

The Bismut–Elworthy–Li formula for mean-field stochastic differential equations

David Baños

Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, N-0316 Oslo, Norway. E-mail: davidru@math.uio.no

Received 26 October 2015; revised 5 October 2016; accepted 11 October 2016

Abstract. We generalise the so-called Bismut–Elworthy–Li formula to a class of stochastic differential equations whose coefficients might depend on the law of the solution. We give some examples of where this formula can be applied to in the context of finance and the computation of Greeks and provide a simple but rather illustrative simulation experiment showing that the use of the Bismut–Elworthy–Li formula, also known as Malliavin method, is more efficient compared to the finite difference method.

Résumé. Nous généralisons la formule dite Bismut–Elworthy–Li à une classe d'équations différentielles stochastiques dont les coefficients pourrait dépendre de la loi de la solution. Nous donnons quelques exemples où cette formule peut être appliquée dans le contexte de la finance et le calcul des Grecs et de fournir une expérience de simulation simple mais significative montrant que l'utilisation de la formule Bismut–Elworthy–Li, également connu comme méthode de Malliavin, est plus efficace que la méthode des différences finies.

MSC: 60H07; 60H10; 60J60; 65C05

Keywords: Bismut–Elworthy–Li formula; Malliavin calculus; Monte Carlo methods; Stochastic differential equations; Integration by parts formulas

1. Introduction

It is known that the spatial derivative of the solution to the (backward) Kolmogorov equation can be represented as an expectation of a functional of the solution of an SDE with some weight, namely the so-called Bismut–Elworthy–Li (BEL) formula as shown in [2] and extended in [5]. In [6] the authors use techniques from Malliavin calculus to prove BEL formula and employ it for the computation of sensitivities of financial options, also known as *Greeks*. In the classical setting where no dependence on the law of the process takes place, also the authors in their series of papers [8–10] have made substantial contributions to the derivation of integration by parts formula using techniques from Malliavin calculus and used them to study regularity of the associated semigroup.

In many applications, it is very natural to expect that the coefficients of a stochastic differential equation (SDE) may depend on properties of the law of the solution, such as dependence on its moments. In this direction, the author in [14] has studied such problems in detail and derived integration by parts formula for the case of so-called de-coupled equations when the initial condition is taken as a random variable and the derivative is with respect to a measure. Here, we want to extend the formula to mean-field type SDEs following the essence of [6] for equations where the dependence on the law of the solution also depends on its initial condition and show that such generalisation is actually non-trivial, requiring more regularity of the solution in the sense of Malliavin which is not needed in the classical setting of [6]. First, we give a relationship between the Malliavin derivative and the spatial derivative of the solution with respect to the initial condition. Already here we see that such generalisation involves an extra factor which is no longer adapted, thus requiring more (Malliavin) regularity on the solution which is not immediate. Fortunately, if the

coefficients are continuously differentiable with bounded Lipschitz derivatives, then the solution is twice Malliavin differentiable and hence a formula using the Skorokhod integral may be expected. Using such relation one can find the BEL formula in this context. Some merely illustrative examples are provided in order to give a better insight on the effect of mean-field SDEs in the BEL formula. In the last examples we carry out some simulations to show that the Malliavin method is more efficient compared to a finite difference method, especially when the function involved is discontinuous.

The paper is organised as follows: In Section 2 we collect some summarised basic facts on Malliavin Calculus needed for the derivation of the main results of the paper. In Section 3 we include all intermediate steps towards the main result which is the Bismut–Elworthy–Li formula in the context of mean-field SDEs. Finally, Section 4 is devoted to provide some illustrative examples of this generalised Bismut–Elworthy–Li formula with simulations. The findings are similar to those in [6], the use of Bismut–Elworthy–Li’s formula serves as a much more efficient method to compute sensitivities with respect to initial data, especially when the function involved in the expectation is highly irregular.

Notations. Let \mathbb{R}_+ denote the non-negative real numbers. Denote by $|\cdot|$ the Euclidean norm in \mathbb{R}^d , $d \geq 1$. Given a Banach space E , denote by $\|\cdot\|_E$ its associated norm. Let $k, p \geq 0$ integers and $\mathbb{D}^{k,p}$ be the space of k times Malliavin differentiable random variables with all p -moments. Denote by $D_s, s \geq 0$ the Malliavin derivative as introduced in [15, Chapter 1, Section 1.2.1] and δ its dual operator (Skorokhod integral). Denote by $\text{Dom } \delta$ the domain of δ (Skorokhod integrable processes). Denote the trace of a matrix $M \in \mathbb{R}^{d \times d}$ by $\text{Tr}(M) := \sum_{j=1}^d M_{j,j}$ and by M^* its transpose. For a (weakly) differentiable function $f : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $(x, y) \mapsto f(x, y)$, denote by ∂_1 , respectively by ∂_2 , (weak) differentiation with respect to the first (space) variable $x \in \mathbb{R}^d$, respectively the second (space) variable $y \in \mathbb{R}^d$.

2. Framework

Our main results centrally rely on tools from Malliavin calculus. We here provide a concise introduction to the main concepts in this area. For deeper information on Malliavin calculus the reader is referred to e.g. [4,11,12,15].

2.1. Malliavin calculus

Let $W = \{W_t, t \in [0, T]\}$ be a standard Wiener process on some complete filtered probability space $(\Omega, \mathfrak{A}, P)$ where $\mathcal{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$ is the P -augmented natural filtration. Denote by \mathcal{S} the set of simple random variables $F \in L^2(\Omega)$ in the form

$$F = f\left(\int_0^T h_1(s) dW_s, \dots, \int_0^T h_n(s) dW_s\right), \quad h_1, \dots, h_n \in L^2([0, T]), f \in C_0^\infty(\mathbb{R}^n).$$

The Malliavin derivative operator D acting on such simple random variables is the process $DF = \{D_t F, t \in [0, T]\}$ in $L^2(\Omega \times [0, T])$ defined by

$$D_t F = \sum_{i=1}^n \partial_i f\left(\int_0^T h_1(s) dW_s, \dots, \int_0^T h_n(s) dW_s\right) h_i(t).$$

Define the following norm on \mathcal{S} :

$$\|F\|_{1,2} := \|F\|_{L^2(\Omega)} + \|DF\|_{L^2(\Omega; L^2([0, T]))} = E[|F|^2]^{1/2} + E\left[\int_0^T |D_t F|^2 dt\right]^{1/2}. \tag{1}$$

We denote by $\mathbb{D}^{1,2}$ the closure of the family of simple random variables \mathcal{S} with respect to the norm given in (1) and we will refer to this space as the space of Malliavin differentiable random variables in $L^2(\Omega)$ with Malliavin derivative belonging to $L^2(\Omega)$.

In the derivation of the probabilistic representation for the Delta, the following chain rule for the Malliavin derivative will be essential:

Lemma 2.1. *Let $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$ be continuously differentiable with bounded partial derivatives. Further, suppose that $F = (F_1, \dots, F_m)$ is a random vector whose components are in $\mathbb{D}^{1,2}$. Then $\varphi(F) \in \mathbb{D}^{1,2}$ and*

$$D_t \varphi(F) = \sum_{i=1}^m \partial_i \varphi(F) D_t F_i, \quad P\text{-a.s.}, t \in [0, T].$$

The Malliavin derivative operator $D : \mathbb{D}^{1,2} \rightarrow L^2(\Omega \times [0, T])$ admits an adjoint operator $\delta = D^* : \text{Dom}(\delta) \rightarrow L^2(\Omega)$ where the domain $\text{Dom}(\delta)$ is characterised by all $u \in L^2(\Omega \times [0, T])$ such that for all $F \in \mathbb{D}^{1,2}$ we have

$$E \left[\int_0^T D_t F u(t) dt \right] \leq C \|F\|_{1,2},$$

where C is some constant depending on u .

For a stochastic process $u \in \text{Dom}(\delta)$ (not necessarily adapted to \mathcal{F}) we denote by

$$\delta(u) := \int_0^T u(t) \delta W_t \tag{2}$$

the action of δ on u . The above expression (2) is known as the Skorokhod integral of u and it is an anticipative stochastic integral. It turns out that all \mathcal{F} -adapted processes in $L^2(\Omega \times [0, T])$ are in the domain of δ and for such processes u we have

$$\delta(u) = \int_0^T u(t) dW_t,$$

i.e. the Skorokhod and Itô integrals coincide. In this sense, the Skorokhod integral can be considered to be an extension of the Itô integral to non-adapted integrands.

As for the Itô integral, there is also a corresponding isometry property for the Skorokhod integral. The proof of this can e.g. be found in [4, Theorem 6.17].

Theorem 2.2. *Let u be a process such that $u(s) \in \mathbb{D}^{1,2}$ for a.e. $s \in [0, T]$ and*

$$E \left[\int_0^T u(t)^2 dt + \int_0^T \int_0^T |D_t u(s) D_s u(t)| ds dt \right] < \infty.$$

Then u is Skorokhod integrable and

$$E \left[\left(\int_0^T u(t) \delta W_t \right)^2 \right] = E \left[\int_0^T u(t)^2 dt + \int_0^T \int_0^T D_t u(s) D_s u(t) ds dt \right].$$

The dual relation between the Malliavin derivative and the Skorokhod integral implies the following important formula:

Theorem 2.3 (Duality formula). *Let $F \in \mathbb{D}^{1,2}$ be \mathcal{F}_T -measurable and $u \in \text{Dom}(\delta)$. Then*

$$E \left[F \int_0^T u(t) \delta W_t \right] = E \left[\int_0^T u(t) D_t F dt \right]. \tag{3}$$

The following is the corresponding integration by parts formula for the Skorokhod integral. See e.g. [4, Theorem 3.15].

Theorem 2.4 (Integration by parts). *Let $u \in \text{Dom}(\delta)$ and $F \in \mathbb{D}^{1,2}$ such that $Fu \in \text{Dom}(\delta)$. Then*

$$F \int_0^T u(t) \delta W(t) = \int_0^T Fu(t) \delta W(t) + \int_0^T u(t) D_t F dt.$$

3. The (mean-field) Bismut–Elworthy–Li formula

The object of study is a *mean-field type stochastic differential equation* (SDE) of the form

$$dX_t^x = b(t, X_t^x, \rho_t^x) dt + \sum_{k=1}^m \sigma_k(t, X_t^x, \pi_t^x) dW_t^k, \quad X_0 = x \in \mathbb{R}^d, \quad (4)$$

$$t \in [0, T], \rho_t^x := E[\varphi(X_t^x)], \pi_t^x := E[\psi(X_t^x)],$$

where $T \in \mathbb{R}$, $T > 0$, $b : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma_k : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $k = 1, \dots, m$, $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ are measurable functions and $W = \{W_t, t \in [0, T]\}$ is an m -dimensional Brownian motion on some probability space $(\Omega, \mathfrak{A}, P)$ equipped with the natural filtration augmented by all P -null sets, denoted by $\mathcal{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$.

We will usually consider the solution as a function of x , therefore we write X_t^x to stress this fact. Moreover, we will assume the following conditions as in [3]

Hypotheses (H).

- (i) *the functions $(t, x, y) \mapsto b(t, x, y)$ and $(t, x, y) \mapsto \sigma_k(t, x, y)$, $k = 1, \dots, m$ are continuously differentiable with bounded Lipschitz derivatives.*
- (ii) *Assume $d \leq m$ and the matrix $(\sigma_1, \dots, \sigma_m)$ admits a right-inverse.*
- (iii) *The functions φ and ψ are continuously differentiable with bounded Lipschitz derivatives.*

The above hypotheses will be assumed throughout the whole paper. Under these conditions, we know that there exists a pathwise unique global solution $X^x = \{X_t^x, t \in [0, T]\}$ for every initial condition $x \in \mathbb{R}^d$. We will henceforward denote by X^x the process solution.

By a result of Buckdahn, Li, Peng and Rainer in [3] we know that, under the hypotheses above, equation (4) admits a stochastic flow of diffeomorphisms, in particular, the function $x \mapsto X_t^x$ is P -a.s. classically differentiable. This property is crucial to define the first variation process and relate it to the Malliavin derivative of the solution.

It is known that when the coefficients b and σ are globally Lipschitz continuous with linear growth, then the solution X_t^x belongs to $\mathbb{D}^{1, \infty}$ for any $t \in [0, T]$ and $x \in \mathbb{R}^d$, see e.g. [15, Theorem 2.2.1]. Also, if b and σ are infinitely often differentiable with all derivatives being bounded, then X_t^x falls into the space $\bigcap_{k \geq 1} \bigcap_{p \geq 1} \mathbb{D}^{k, p}$ for any $t \in [0, T]$ and $x \in \mathbb{R}^d$, see e.g. [15, Theorem 2.2.2]. The proof of the latter result is actually proved by induction which then implies the following straight-forward regularity of our solution X under the hypotheses given in (H).

Proposition 3.1. *The random variable $\omega \mapsto X_t^x(\omega)$ belongs to the space $\bigcap_{p \geq 1} \mathbb{D}^{2, p}$ for any $t \in [0, T]$ and $x \in \mathbb{R}^d$.*

Proof. See [15, Theorem 2.2.2]. □

To relax notation let us introduce the following stochastic processes which will be used extensively throughout the rest of this note. Let $A^x = \{A_t^x, t \in [0, T]\}$, $\alpha^x = \{\alpha_t^x, t \in [0, T]\}$, $B^{x, k} = \{B_t^{x, k}, t \in [0, T]\}$, $\beta^{x, k} = \{\beta_t^{x, k}, t \in [0, T]\}$, $k = 1, \dots, m$ be matrix-valued processes defined as:

$$A_t^x := \partial_1 b(t, X_t^x, \rho_t^x), \quad \alpha_t^x := \partial_2 b(t, X_t^x, \rho_t^x) \frac{\partial}{\partial x} \rho_t^x$$

$$B_t^{x, k} := \partial_1 \sigma_k(t, X_t^x, \pi_t^x), \quad \beta_t^{x, k} := \partial_2 \sigma_k(t, X_t^x, \pi_t^x) \frac{\partial}{\partial x} \pi_t^x \quad (5)$$

for $k = 1, \dots, m$.

Next proposition shows that the first variation process is invertible for every $t > 0$.

Proposition 3.2. *Let $Y^x = \{Y_t^x, t \in [0, T]\}$ be the solution to the following matrix-valued linear SDE*

$$dY_t^x = A_t^x Y_t^x dt + \sum_{k=1}^m B_t^{x, k} Y_t^x dW_t^k, \quad Y_0^x = I, \quad t \in [0, T].$$

Then

$$(\det Y_t^x)^{-1} \in \bigcap_{p \geq 1} L^p(\Omega).$$

As a consequence, Y_t^x is P -a.s. invertible for every $t \in [0, T]$.

Proof. We want to show that

$$E[|\det Y_t^x|^{-p}] < \infty$$

for every integer $p \geq 1$. Indeed, in virtue of (stochastic) Liouville's formula, see e.g. [13, p.92], one has

$$\det Y_t^x = \exp \left\{ \int_0^t \left(\text{Tr} A_u^x - \frac{1}{2} \sum_{k=1}^m \text{Tr}[(B_u^{x,k})^2] \right) du + \sum_{k=1}^m \int_0^t \text{Tr} B_u^{x,k} dW_u^k \right\}, \quad P\text{-a.s.}$$

Observe that since the processes $\text{Tr} B^{x,k}$ are in $L^2([0, T])$ and \mathcal{F} -adapted, the stochastic integrals appearing in the exponent are martingales. Hence,

$$M_t := \exp \left\{ \sum_{k=1}^m \int_0^t \text{Tr} B_u^{x,k} dW_u^k - \frac{1}{2} \sum_{k=1}^m \int_0^t (\text{Tr} B_u^{x,k})^2 du \right\}$$

is an \mathcal{F} -martingale with $E[M_t] = E[M_0] = 1$. Thus, using Cauchy–Schwarz' inequality we have for every $p \geq 1$,

$$\begin{aligned} E[|\det Y_t^x|^{-p}] &\leq E \left[\exp \left\{ \frac{1}{2} p \sum_{k=1}^m \int_0^t \text{Tr}[(B_u^{x,k})^2] du + 2p \sum_{k=1}^m \int_0^t (\text{Tr} B_u^{x,k})^2 du - 2p \int_0^t \text{Tr} A_u^x du \right\} \right]. \end{aligned}$$

In particular, a sufficient condition to prove the claim is reduced to showing that

$$\sup_{t \in [0, T]} \left| E \left[\exp \left\{ \lambda \int_0^t \text{Tr} A_u^x du \right\} \right] \right| < \infty,$$

and that

$$\sup_{t \in [0, T]} \left| E \left[\exp \left\{ \lambda \int_0^t \sum_{k=1}^m (\text{Tr} B_u^{x,k})^2 du \right\} \right] \right| + \sup_{t \in [0, T]} \left| E \left[\exp \left\{ \lambda \int_0^t \sum_{k=1}^m \text{Tr}[(B_u^{x,k})^2] du \right\} \right] \right| < \infty$$

for every $\lambda \in \mathbb{R}$ which clearly holds since A^x and $B^{x,k}$, $k = 1, \dots, m$ are uniformly bounded. \square

The following statement is one of the main observations for the derivation of the Bismut–Elworthy–Li formula in the mean-field context. It can be seen as a generalisation of the well-known relation between the first variation process Y^x in the non mean-field context and the Malliavin derivative, see e.g. [15, Ch.2, Section 2.3.1].

Theorem 3.3. *For every $s, t \in [0, T]$, $s \leq t$ one has the following relationship between the spatial derivative and the Malliavin derivative of X_t^x*

$$\frac{\partial}{\partial x} X_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t), \quad s \leq t, \quad (6)$$

where

$$u^x(t) := \left(I + \int_0^t (Y_u^x)^{-1} \left(\alpha_u^x - \sum_{k=1}^m B_u^{x,k} \beta_u^{x,k} \right) du + \sum_{k=1}^m \int_0^t (Y_u^x)^{-1} \beta_u^{x,k} dW_u^k \right),$$

σ^{-1} denotes the right-inverse of σ , $Y^x = \{Y_t^x, t \in [0, T]\}$ is the $d \times d$ fundamental matrix satisfying

$$dY_t^x = A_t^x Y_t^x dt + \sum_{k=1}^m B_t^{x,k} Y_t^x dW_t^k, \quad Y_0^x = I, \quad t \in [0, T]$$

and where $A^x = \{A_t^x, t \in [0, T]\}$, $\alpha^x = \{\alpha_t^x, t \in [0, T]\}$, $B^{x,k} = \{B_t^{x,k}, t \in [0, T]\}$, $\beta^{x,k} = \{\beta_t^{x,k}, t \in [0, T]\}$, $k = 1, \dots, m$ are the matrix-valued processes defined as in (5).

Proof. Differentiating with respect to $x \in \mathbb{R}^d$ we have that $\frac{\partial}{\partial x} X_t^x$ satisfies the following matrix-valued linear equation

$$\begin{aligned} \frac{\partial}{\partial x} X_t^x &= I + \int_0^t \left(\partial_1 b(u, X_u^x, \rho_u^x) \frac{\partial}{\partial x} X_u^x + \partial_2 b(u, X_u^x, \rho_u^x) \frac{\partial}{\partial x} \rho_u^x \right) du \\ &\quad + \sum_{k=1}^m \int_0^t \left(\partial_1 \sigma_k(u, X_u^x, \pi_u^x) \frac{\partial}{\partial x} X_u^x + \partial_2 \sigma_k(u, X_u^x, \pi_u^x) \frac{\partial}{\partial x} \pi_u^x \right) dW_u^k, \end{aligned} \tag{7}$$

where we have used the regularity of the integrands to pull the differentiation operator inside the integrals. For the Lebesgue integral the usual differentiation under the integral sign theorem follows, as for the Itô integral, which is not a path-wise integral, we refer to [7, Theorem 2.2] to justify that this can be done.

Using the notations in the statement of the theorem, we can solve (7) and express $\frac{\partial}{\partial x} X_t^x$ as

$$\frac{\partial}{\partial x} X_t^x = Y_t^x \left(I + \int_0^t (Y_u^x)^{-1} \left(\alpha_u^x - \sum_{k=1}^m B_u^{x,k} \beta_u^{x,k} \right) du + \sum_{k=1}^m \int_0^t (Y_u^x)^{-1} \beta_u^{x,k} dW_u^k \right).$$

By the well-known classical relation, see e.g. [15, Chapter 2, Section 2.3.1], it is true that,

$$Y_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x, \quad s \leq t,$$

where σ^{-1} denotes the right-inverse of σ and hence the relation follows. □

Remark 3.4. For the relation $Y_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x$, $s \leq t$ to hold in the mean-field setting one also needs the property that $x \mapsto X_t^x$ defines a stochastic semiflow. In the mean-field case we point out that the fact that b , σ , φ and ψ are continuously differentiable with bounded Lipschitz derivative implies this fact in virtue of [3].

Remark 3.5. It is shown in [1] that SDE (4) is twice Malliavin differentiable when the vector field b is globally Lipschitz continuous with linear growth and does not depend on the law of X^x and one has additive noise. Nevertheless, using the same method one can prove the same result since the dependence on $E[\varphi(X_t^x)]$ and $E[\psi(X_t^x)]$ does not bring stochasticity to the equation. In the sense that the Malliavin derivative of X_t^x for every fixed $t \in [0, T]$ takes the same form as in the usual linear setting. This fact alludes that relation (6) still holds whenever $\sigma \equiv 1$ and b is only globally Lipschitz with linear growth.

Proposition 3.6. Let $t \in [0, T]$ and define the random variables

$$F := \int_0^t (Y_u^x)^{-1} \left(\alpha_u^x - \sum_{k=1}^m B_u^{x,k} \beta_u^{x,k} \right) du, \quad G := \sum_{k=1}^m \int_0^t (Y_u^x)^{-1} \beta_u^{x,k} dW_u^k,$$

then $F, G \in \text{Dom } \delta$.

Proof. Since $\mathbb{D}^{1,2} \subset \text{Dom } \delta$, see [15, Proposition 1.3.1.], it is enough show that $F, G \in \mathbb{D}^{1,2}$. In both cases it suffices to show that the integrands are Malliavin differentiable. For the Lebesgue integral it is immediate as for the Itô integral one may justify this by [15, Lemma 1.3.4.].

Now, since $Y_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x$, $s \leq t$ and $X_t^x \in \mathbb{D}^{2,2}$ for every $t \in [0, T]$ we have together with Proposition 3.2 that $(Y_t^x)^{-1} \in \mathbb{D}^{1,2}$ for every $t \in [0, T]$. The result follows since the processes α^x , $B^{x,k}$ and $\beta^{x,k}$ have Malliavin differentiable marginals. That is, α_t^x , $B_t^{x,k}$ and $\beta_t^{x,k}$ are Malliavin differentiable for every $t \in [0, T]$. This is a consequence of the fact that b and σ_k , $k = 1, \dots, m$ are continuously differentiable with bounded Lipschitz derivatives and hence, since X_t^x has no atoms one can apply the chain rule for the Malliavin derivative, see [15, p.29]. \square

Corollary 3.7. *The process*

$$s \mapsto \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x \left(I + \int_0^t (Y_u^x)^{-1} \left(\alpha_u^x - \sum_{k=1}^m B_u^{x,k} \beta_u^{x,k} \right) du + \sum_{k=1}^m \int_0^t (Y_u^x)^{-1} \beta_u^{x,k} dW_u^k \right),$$

$0 \leq s \leq t$, is Skorokhod integrable.

Proof. Indeed, it is the product of an adapted process, hence Skorokhod integrable and by Proposition 3.6 a Skorokhod integrable random variable. \square

Theorem 3.8 (Bismut–Elworthy–Li formula). *Let $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}_+$ be a measurable function such that $\Phi(X_t^x) \in L^2(\Omega)$. Define the function*

$$v(x) := E[\Phi(X_t^x)].$$

Then v is continuously differentiable and its derivative admits the following representation

$$\frac{\partial}{\partial x} v(x) = E \left[\Phi(X_t^x) \int_0^t a(s) [\sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t)]^* \delta W_s \right]^*, \quad (8)$$

where $$ denotes transposition, Y^x is the fundamental matrix obtained in Theorem 3.3 and*

$$u^x(t) := I + \int_0^t (Y_s^x)^{-1} \left(\alpha_s^x - \sum_{k=1}^m B_s^{x,k} \beta_s^{x,k} \right) ds + \sum_{k=1}^m \int_0^t (Y_s^x)^{-1} \beta_s^{x,k} dW_s^k,$$

here $a : [0, T] \rightarrow \mathbb{R}$ is a bounded integrable function such that $\int_0^t a(s) ds = 1$ and $a \equiv 0$ on $[t, T]$. The processes α^x , $B^{x,k}$, $\beta^{x,k}$, $k = 1, \dots, m$ are defined as in Theorem 3.3.

Proof. We will carry out the proof in four steps. First, we will show the formula for smooth functions Φ with compact support. Then extend it to any continuous and bounded function Φ by using a limit argument. We then get rid of the continuity by employing a monotone class argument. Finally, we consider any general function with the property that $\Phi(X_t^x) \in L^2(\Omega)$.

Step I: Assume first that Φ is infinitely differentiable with compact support. By Theorem 3.3 we have

$$\frac{\partial}{\partial x} X_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t), \quad s \leq t, P\text{-a.s.}$$

Then multiplying both sides by the function a and integrating over $s \in [0, t]$ we have

$$\frac{\partial}{\partial x} X_t^x = \int_0^t a(s) D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t) ds, \quad P\text{-a.s.} \quad (9)$$

As a consequence,

$$\begin{aligned} \frac{\partial}{\partial x} v(x) &= E \left[\Phi'(X_t^x) \frac{\partial}{\partial x} X_t^x \right] \\ &= E \left[\Phi'(X_t^x) \int_0^t a(s) D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t) ds \right] \\ &= E \left[\int_0^t a(s) D_s \Phi(X_t^x) \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t) ds \right] \\ &= E \left[\Phi(X_t^x) \int_0^t a(s) [\sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t)]^* \delta W_s \right]^*, \end{aligned}$$

where we have used relation (9), the chain rule for the Malliavin derivative (backwards) and the duality formula for the Malliavin derivative which is justified by Corollary 3.7.

Step 2: Assume Φ is bounded and continuous, in particular, $\Phi(X_t^x) \in L^2(\Omega)$. We can approximate Φ by a sequence of smooth functions $\{\Phi_n\}_{n \geq 0}$ with compact support such that $\Phi_n(x) \rightarrow \Phi(x)$ for every $x \in \mathbb{R}^d$ as $n \rightarrow \infty$. Define

$$\bar{v}(x) := E \left[\Phi(X_t^x) \int_0^t a(s) [\sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t)]^* \delta W_s \right]^*.$$

To make reading clearer introduce the notation $C := E[|\Phi(X_t^x)|^2]^{1/2}$ and the matrix-valued process $\xi_s := a(s) \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x u^x(t)$, $0 \leq s \leq t$. Then the objects $v(x) := E[\Phi(X_t^x)]$ and $\bar{v}(x)$ are well-defined since $\Phi(X_t^x) \in L^2(\Omega)$ and using Cauchy–Schwarz' inequality we have

$$|\bar{v}(x)| \leq CE \left[\int_0^t \text{Tr}[\xi_s \xi_s^*] ds + \int_0^t \int_0^t D_s \xi_r D_r \xi_s dr ds \right]^{1/2},$$

where we used Itô's isometry property for Skorokhod integrals, see Theorem 2.2 or e.g. [4, Theorem 6.17.]. Observe that the first term is bounded since a and σ^{-1} are uniformly bounded and u^x has integrable trajectories. The second term is bounded since ξ_s is Malliavin differentiable for every $s \in [0, t]$ because $Y_s^x u^x(t)$ is Malliavin differentiable for every $s \in [0, T]$ in virtue of Proposition 3.2 in connection with Proposition 3.1 as for $u^x(t)$, due to Corollary 3.7.

Now, we approximate v by $v_n(x) := E[\Phi_n(X_t^x)]$. It is clear that $v_n \rightarrow v$ a.e. and now we can use the Bismut–Elworthy–Li formula on $\frac{\partial}{\partial x} v_n$ in order to estimate $|\frac{\partial}{\partial x} v_n(x) - \bar{v}(x)|$. Indeed, again by Cauchy–Schwarz inequality and Itô's isometry for the Skorokhod integral we have

$$\left| \frac{\partial}{\partial x} v_n(x) - \bar{v}(x) \right| \leq E[|\Phi_n(X_t^x) - \Phi(X_t^x)|^2]^{1/2} E[|w(t)|^2]^{1/2},$$

where $w(t) := \int_0^t \xi_s^* \delta W_s$ denotes the Malliavin weight. Now since Φ_n and Φ are continuous and bounded we have for every compact subset $K \subset \mathbb{R}^d$

$$\lim_{n \rightarrow \infty} \sup_{x \in K} \left| \frac{\partial}{\partial x} v_n(x) - \bar{v}(x) \right| = 0.$$

Hence, v is continuously differentiable with $\frac{\partial}{\partial x} v = \bar{v}$.

Step 3: Let us denote

$$\mathcal{G} := \{ \Phi : \mathbb{R}^d \rightarrow \mathbb{R}_+ \text{ continuous and bounded} \}.$$

It is clear that \mathcal{G} is a multiplicative class, i.e. $\psi_1, \psi_2 \in \mathcal{G}$ then $\psi_1 \psi_2 \in \mathcal{G}$. Further, let \mathcal{H} be the class of functions $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}_+$ for which (8) holds. From Step 2 we have $\mathcal{G} \subset \mathcal{H}$. Then \mathcal{H} is a monotone vector space on \mathbb{R}^d , see e.g. [16, p.23] for definitions. Indeed, from dominated convergence we have monotonicity. In fact, if $\{\Phi_n\}_{n \geq 0} \subset \mathcal{H}$ such that $0 \leq \Phi_1 \leq \dots \leq \Phi_n \leq \dots$ with $\lim_n \Phi_n = \Phi$ and Φ is bounded then $\Phi \in \mathcal{H}$. Furthermore, denote by

$\sigma(\mathcal{G}) := \{f^{-1}(B), B \in \mathcal{B}(\mathbb{R}_+), f \in \mathcal{G}\}$ where $\mathcal{B}(\mathbb{R}_+)$ denotes the Borel σ -algebra in \mathbb{R}_+ . Then we are able to apply the monotone class theorem, see e.g. [16, Theorem 8] and conclude that \mathcal{H} contains all bounded and $\sigma(\mathcal{G})$ -measurable functions $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}_+$. Nevertheless, $\sigma(\mathcal{G})$ coincides with the Borel σ -algebra of \mathbb{R}^d since \mathcal{G} contains all continuous bounded functions. So we conclude that \mathcal{H} contains all bounded Borel measurable functions on \mathbb{R}^d .

Step 4: The last step is then to approximate any $\mathcal{B}(\mathbb{R}^d)$ -measurable function $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}_+$ such that $\Phi(X_t^x) \in L^2(\Omega)$ by a sequence $\{\Phi_n\}_{n \geq 0}$ of bounded $\mathcal{B}(\mathbb{R}^d)$ -measurable functions. For example,

$$\Phi_n(x) = \Phi(x)1_{\{\Phi(x) \leq n\}}, \quad x \in \mathbb{R}^d, n \geq 0.$$

Then $\Phi_n \in \mathcal{H}$ for each $n \geq 0$. Define $\tilde{v}(x) := E[\Phi(X_t^x)w(t)]$. Then by Cauchy–Schwarz’ inequality and Itô’s isometry we know

$$\sup_{x \in K} \left| \frac{\partial}{\partial x} v_n(x) - \tilde{v}(x) \right| \leq C \sup_{x \in K} E[|\Phi_n(X_t^x) - \Phi(X_t^x)|]^2,$$

for any compactum $K \subset \mathbb{R}^d$ and some finite constant $C > 0$. Finally, observe that clearly one has

$$\sup_{x \in K} E[|\Phi_n(X_t^x) - \Phi(X_t^x)|]^2 \xrightarrow{n \rightarrow \infty} 0$$

thus proving the result. \square

The following is an immediate corollary of Theorem 3.8 which provides a representation of the gradient of v in terms of an Itô integral.

Corollary 3.9. *Under the hypotheses of Theorem 3.8 we have the following representation for the derivative of v ,*

$$\begin{aligned} \frac{\partial}{\partial x} v(x) = E \left[\Phi(X_t^x) \left(\int_0^t a(s) [\sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x]^* dW_s u^x(t) \right. \right. \\ \left. \left. - \int_0^t a(s) [\sigma^{-1}(s, X_s^x, \pi_s^x) Y_s^x D_s u^x(t)]^* ds \right) \right]. \end{aligned} \quad (10)$$

Proof. Indeed, the proof follows by observing that $u^x(t) \in \mathbb{D}^{1,2}$. Then integration by parts for the Skorokhod integral, see Theorem 2.4 or [4, Theorem 3.15], yields the result. \square

4. Applications

In this section we wish to give a rather simple but illustrative example of how the dependence on the expectation of the solution may give rise to more complicated terms when deriving the Bismut–Elworthy–Li formula. In one of the examples we adopt the context of finance where the formula has a broad use for the computation of the so-called Delta sensitivities which, in short, is the sensitivity of prices of contracts with respect to the initial value of the price of the stock taken into consideration. We will consider the price of an option written on a stock whose dynamics depend on the expectation of the price process. Then we provide two numerical examples in order to demonstrate that the Bismut–Elworthy–Li formula, or the so-called Malliavin method for computing the Delta is numerically more efficient than the usual finite difference method even when the function Φ is discontinuous.

Example 4.1 (Black–Scholes model with continuous dividend payments). Let $S^x = \{S_t^x, t \in [0, T]\}$ represent the price dynamics of some asset with initial price $x > 0$ governed by the following SDE

$$\frac{dS_t^x}{S_t^x} = (\mu - q\rho_t^x) dt + \sigma dW_t, \quad \rho_t^x := E[S_t^x], t \in [0, T], S_0^x = x > 0, \quad (11)$$

where $\mu, q, \sigma \in \mathbb{R}$ and $\sigma > 0$. Let $S_t^0 = e^{rt}$, $t \in [0, T]$, $r \in \mathbb{R}$ with $r > 0$ be the risk-less asset and $\Phi : \mathbb{R} \rightarrow [0, \infty)$ a pay-off function. Here, q represents the dividend yield, and the philosophy behind this model is that, the higher the mean-price, the higher the dividend yield will get and thus, implying a decrease of speed in the drift and as a by-product of the mean-price of the process S^x . This gives an endogenous way of including dividend yields into the model depending on the average price of the stock.

Then the price of a European option at current time with maturity $T > 0$ (under the risk-neutral valuation approach) is given by

$$p_T(x) = e^{-rT} E_{\tilde{P}}[\Phi(S_T^x)], \quad (12)$$

where \tilde{P} is the risk-neutral measure, i.e.

$$\left. \frac{d\tilde{P}}{dP} \right|_{\mathcal{F}_t} = M_t^x := e^{-\int_0^t \theta_u^x dW_u - \frac{1}{2} \int_0^t (\theta_u^x)^2 du}, \quad t \in [0, T],$$

where

$$\theta_t^x := \frac{\mu - r - q\rho_t^x}{\sigma}, \quad t \in [0, T]$$

is the market price of risk process.

It follows that $\rho_t^x = \frac{x\mu e^{\mu t}}{qx e^{\mu t} + \mu - qx}$ obtained as the solution of a Riccati equation. Also, we have $\frac{\partial}{\partial x} \rho_t^x = \frac{e^{-\mu t}}{x^2} (\rho_t^x)^2$ and $\frac{\partial}{\partial x} \theta_t^x = -\frac{q}{\sigma} \frac{e^{-\mu t}}{x^2} (\rho_t^x)^2$.

The sensitivity of the price (12) with respect to the current value of the stock, also known as *Delta* denoted by Δ , is then given by $p'(x)$. This quantity gives the variation of the future contract price with respect to today's value, hence, it gives a measure of risk of the valuation of the contract, since a high value of Δ yields a higher uncertainty of the price of the contract.

Nevertheless, observe that we are no longer under the setting of Theorem 3.8 since the expectation is taken under a different measure which, actually, depends on x . Hence, we need to make the following observation

$$\frac{\partial}{\partial x} E_{\tilde{P}}[\Phi(S_T^x)] = \frac{\partial}{\partial x} E[M_T^x \Phi(S_T^x)] = E\left[\frac{\partial}{\partial x} M_T^x \Phi(S_T^x)\right] + E\left[M_T^x \frac{\partial}{\partial x} \Phi(S_T^x)\right].$$

The aim is to obtain an expression independent of Φ' . In the first term this is not an issue whereas in the second term Theorem 3.8 can now be applied, the only difference is that M_T^x goes into the Skorokhod integral appearing in the Malliavin weight. As a consequence, the Δ -sensitivity of an option Φ written on S_T^x is given by

$$\Delta = e^{-rT} E\left[\Phi(S_T^x) \left(\frac{\partial}{\partial x} M_T^x + \frac{1}{x^2 \sigma} e^{-\mu T} \rho_T^x \int_0^T a(s) M_T^x \delta W_s\right)\right].$$

Let us find a simpler expression for the stochastic integral. Using the integration by parts formula for the Skorokhod integral, see Theorem 2.4, we find that

$$\int_0^T M_T^x \delta W_s = \left(W(T) - \int_0^T \theta_s^x ds\right) M_T^x$$

and hence, taking $a \equiv \frac{1}{T}$ we find that under the risk-neutral measure \tilde{P} , the Δ -sensitivity is given by

$$\Delta = e^{-rT} E_{\tilde{P}}[\Phi(S_T^x) Z_T^x]$$

with Malliavin weight

$$Z_T^x := \frac{1}{x^2 \sigma} \left(q \int_0^T e^{-\mu s} (\rho_s^x)^2 dW_s + q \int_0^T \theta_s^x e^{-\mu s} (\rho_s^x)^2 ds + \frac{e^{-\mu T} \rho_T^x}{T} \left(W(T) - \int_0^T \theta_s^x ds\right)\right).$$

Finally, observe that if we ignore the dependence on $E[S_t^x]$, e.g. taking $q = 0$ then we obtain

$$\Delta = e^{-rT} E_{\tilde{P}} \left[\Phi(S_T^x) \frac{W(T) - \int_0^T \theta_s^x ds}{Tx\sigma} \right] = e^{-rT} E_P \left[\Phi(S_T^x) \frac{\hat{W}(T)}{Tx\sigma} \right],$$

where \hat{W} is a standard Brownian motion under P and hence the Δ coincides with the classical one.

Example 4.2. Consider now the following SDE,

$$\frac{dX_t^x}{X_t^x} = f(\rho_t^x) dt + \sigma dW_t, \quad \rho_t^x := E[X_t^x], t \in [0, T], X_0^x = x > 0, \quad (13)$$

for some suitable continuously differentiable function $f : \mathbb{R} \rightarrow \mathbb{R}$. Let us compute the derivative of $v(x) := E[\Phi(X_T^x)]$ for an irregular function Φ .

Using Theorem 3.8 with $a(s) \equiv 1/T$ we have

$$v'(x) = E \left[\Phi(X_T^x) \frac{1}{T} \int_0^T \frac{1}{\sigma X_s} \frac{X_s}{x} u^x(T) \delta W_s \right] = \frac{1}{\sigma x T} u^x(T) E[\Phi(X_T^x) W_T],$$

where

$$u^x(T) = 1 + x \int_0^T f'(\rho_s^x) \frac{\partial}{\partial x} \rho_s^x ds.$$

This special “geometric-type” case shows that whenever u is deterministic then the delta v' is an rescaled version of the classical delta and they coincide when f is constant, indeed.

Let us then see what happens when we make the volatility coefficient depend on the expectation, which leads to a stochastic u .

Example 4.3. Consider now

$$\frac{dX_t^x}{X_t^x} = \mu dt + \sigma \rho_t^x dW_t, \quad \rho_t^x := E[X_t^x], t \in [0, T], X_0^x = x > 0. \quad (14)$$

In this case according to Theorem 3.3 in connection with Theorem 3.8 the Malliavin weight, denoted by w_T^x , becomes

$$w_T^x := \frac{1}{\sigma x T} \int_0^T \frac{1}{\rho_s^x} \left(1 - \sigma^2 \int_0^T (\rho_s^x)^2 ds + \sigma \int_0^T \rho_s^x dW_s \right) \delta W_s.$$

Fortunately we can rewrite the above expression in terms of an Itô integral using integration by parts. Namely,

$$w_T^x = \left(1 - \sigma^2 \int_0^T (\rho_s^x)^2 ds + \sigma \int_0^T \rho_s^x dW_s \right) \int_0^T \frac{1}{\rho_s^x} dW_s - \sigma T.$$

Denote the random variables $F^x := \int_0^T \rho_s^x dW_s$ and $G^x := \int_0^T \frac{1}{\rho_s^x} dW_s$. Then the vector (F^x, G^x) is normally distributed with zero mean and covariance matrix

$$\Sigma := \begin{pmatrix} \int_0^T (\rho_s^x)^2 ds & T \\ T & \int_0^T \frac{1}{(\rho_s^x)^2} ds \end{pmatrix}.$$

Altogether we obtain

$$v'(x) = \frac{1}{\sigma x T} E \left[\Phi(X_T^x) \left(1 - \sigma^2 \int_0^T (\rho_s^x)^2 ds + \sigma F \right) G \right] - \frac{1}{x} E[\Phi(X_T^x)].$$

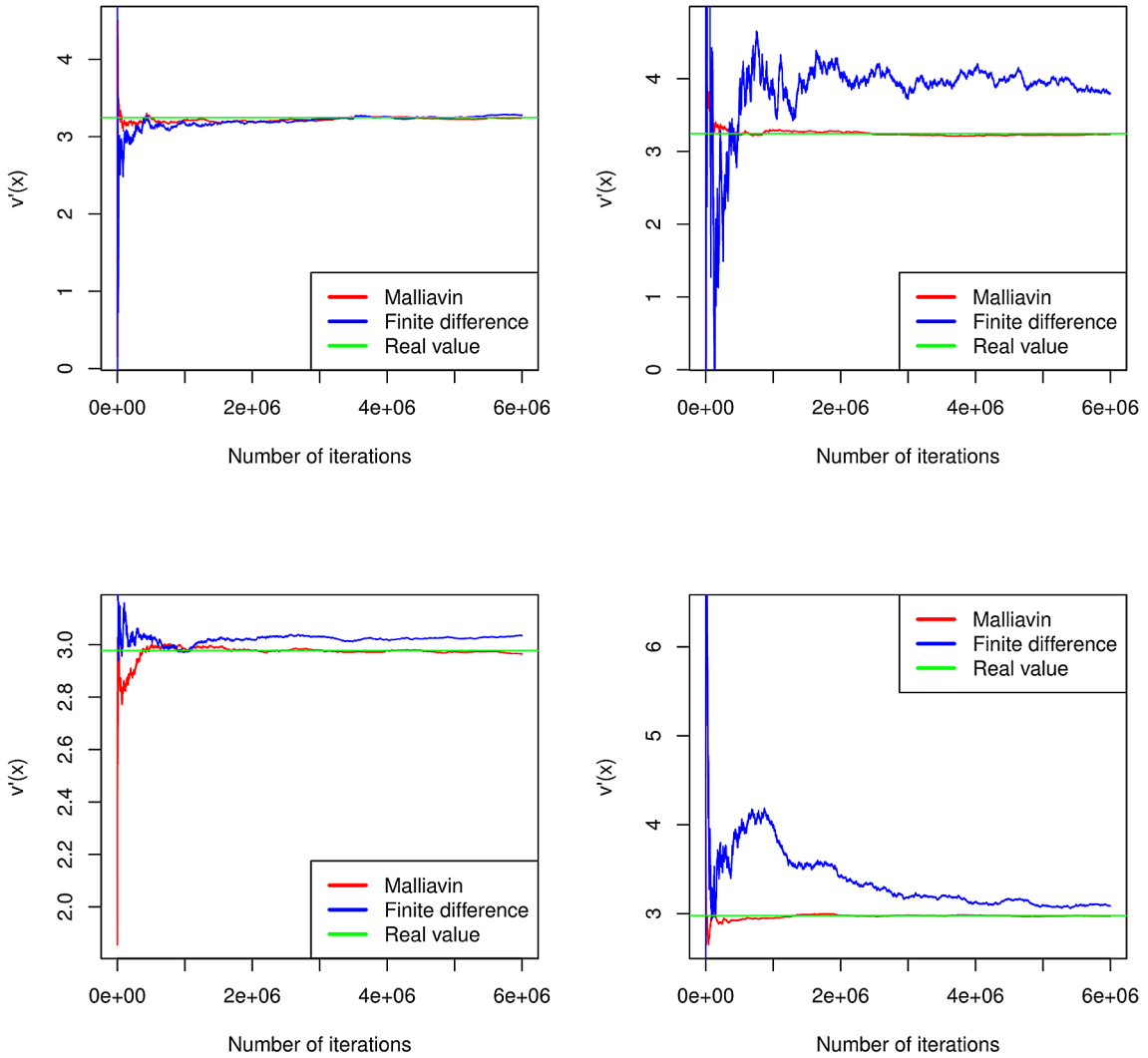


Fig. 1. Approximation of a Call Option: On top parameters set to $\sigma = 0.8$, $\mu = 1$, $x = 1$, $K = 2$ and $T = 1$ with $h = 0.1$ to the right and $h = 0.01$ to the left. On bottom parameters set to $\sigma = 1.2$, $\mu = 1$, $x = 0.5$, $K = 0.7$ and $T = 1$ with $h = 0.1$ to the right and $h = 0.01$ to the left.

Let us, for instance, take $\Phi(x) = (x - K)1_{\{x \geq K\}}$ for some fixed $K > 0$, also known as a *European call option* in the context of finance. Then we use a Monte Carlo method to compute the above expression and compare it to the following finite difference method scheme hereunder

$$v'(x) \approx \frac{E[\Phi(X_T^{x+h})] - E[\Phi(X_T^x)]}{h}, \quad h \approx 0.$$

In Figure 1 (upper left picture), $h = 0.1$ for the finite difference method and the two methods are seemingly giving similar accurate results, although the Malliavin method is more efficient in number of iterations. If one wishes to decrease h in order to gain precision we can see how the finite difference method becomes unstable (Figure 1, upper right and lower right pictures). In conclusion, the integration by parts formula seems to be a much more efficient tool for the computation of sensitivities for mean-field SDEs, at least, in this setting.

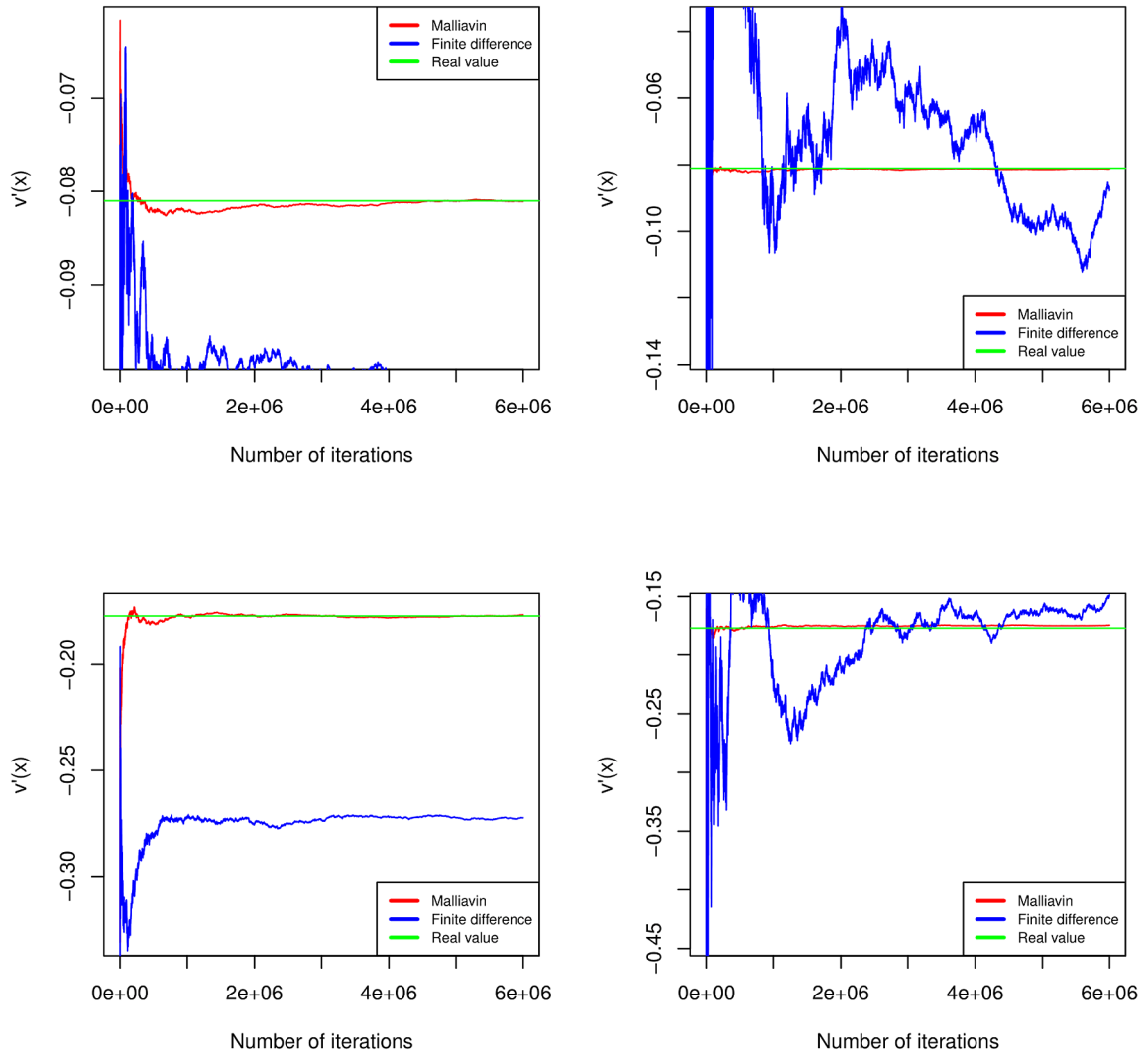


Fig. 2. Approximation of a Digital Option: On top parameters set to $\sigma = 0.8, \mu = 1, x = 1, K = 2$ and $T = 1$ with $h = 0.1$ to the right and $h = 0.01$ to the left. On bottom parameters set to $\sigma = 1.2, \mu = 1, x = 0.5, K = 0.7$ and $T = 1$ with $h = 0.1$ to the right and $h = 0.01$ to the left.

Let us now try a more irregular function Φ , namely $\Phi(x) = 1_{\{x \geq K\}}$ which has a discontinuity at $x = K$, also known as a *European digital option* in the context of finance. Denote

$$d(x) := \frac{\log \frac{K}{x} - \mu T + \frac{1}{2} \sigma^2 \int_0^T (\rho_s^x)^2 ds}{\sigma}.$$

Then

$$v'(x) = \frac{1}{\sigma x T} E \left[1_{\{F^x \geq d(x)\}} \left(\left(1 - \frac{\sigma^2 x^2}{2\mu} (e^{2\mu T} - 1) + \sigma F^x \right) G^x - \sigma T \right) \right].$$

We compare again the Malliavin method with a finite difference scheme (see Figure 2).

The conclusions here are clear. The regularity of the function Φ plays an important role. We see that the bias in the finite difference method seems high and it becomes unstable when decreasing the values of h . On the con-

rary, the Bismut–Elworthy–Li formula gives a better approximation of the sensitivity, even when the function Φ is discontinuous.

Acknowledgements

The author would like to thank the unknown referees for very helpful comments.

References

- [1] D. Baños and T. Nilssen. Malliavin and flow regularity of SDEs. Application to the study of densities and the stochastic transport equation. *Stochastics* **88** (4) (2016) 540–566. [MR3473850](#)
- [2] J. M. Bismut. *Large Deviation and Malliavin Calculus*. *Progress in Mathematics* **45**. Birkhäuser, 1984. [MR0755001](#)
- [3] R. Buckdahn, J. Li, P. Peng and C. Rainer. Mean-field stochastic differential equations and associated PDEs. Available at [arXiv:1407.1215 \[math.PR\]](#).
- [4] G. Di Nunno, B. Øksendal and F. Proske. *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer, 2008.
- [5] K. D. Elworthy and X.-M. Li. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** (1994) 252–286. [MR1297021](#)
- [6] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance Stoch.* **3** (4) (1999) 391–412. [MR1842285](#)
- [7] J. E. Hutton and P. I. Nelson. Interchanging the order of differentiation and stochastic integration. *Stochastic Process. Appl.* **18** (1984) 371–377. [MR0770201](#)
- [8] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. I. In *Stochastic Analysis (Katata/Kyoto, 1982)* 271–306. *North-Holland Math. Library* **32**. North-Holland, Amsterdam, 1984. [MR0780762](#)
- [9] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci., Univ. Tokyo, Sect. 1A, Math.* **32** (1) (1985) 1–76. [MR0783181](#)
- [10] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci., Univ. Tokyo, Sect. 1A, Math.* **34** (2) (1987) 391–442. [MR0914028](#)
- [11] P. Malliavin. Stochastic calculus of variations and hypoelliptic operators. In *Proc. Inter. Symp. on Stoch. Diff. Equations (Kyoto, 1976)* 195–263. Wiley, 1978. [MR0536013](#)
- [12] P. Malliavin. *Stochastic Analysis*. Springer, 1997. [MR1450093](#)
- [13] X. Mao. *Stochastic Differential Equations and Applications*. Woodhead Publishing, Glasgow, 2007. [MR2380366](#)
- [14] E. McMurray. Regularity of McKean-Vlasov Stochastic Differential Equations and Applications. Ph.D. thesis, Imperial College, London.
- [15] D. Nualart. *The Malliavin Calculus and Related Topics*, 2nd edition. Springer, 2010. [MR2200233](#)
- [16] P. E. Protter. *Stochastic Integration and Differential Equations*, 2nd edition. *Applications of Mathematics (New York)* **21**. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2004. [MR2020294](#)