Finance Stochast. 9, 349–367 (2005)
DOI: 10.1007/s00780-004-0151-6

Finance and
Stochastics

Representation formulas for Malliavin derivatives of diffusion processes

Jérôme Detemple[1], **René Garcia**[2], **Marcel Rindisbacher**[3]

[1] Boston University School of Management, Boston, MA, 02215, USA, and CIRANO, Montréal H3A2A5, Canada (e-mail: detemple@bu.edu)
[2] Université de Montréal and CIREQ, Montréal H3C3J7, Canada, and CIRANO, Montréal H3A2A5, Canada (e-mail: rene.garcia@umontreal.ca)
[3] J.L. Rotman School of Management, Toronto M5S3E6, Canada, and CIRANO, Montréal H3A2A5, Canada (e-mail: rindisbm@rotman.utoronto.ca)

Abstract. We provide new representation formulas for Malliavin derivatives of diffusions, based on a transformation of the underlying processes. Both the univariate and the multivariate cases are considered. First order as well as higher order Malliavin derivatives are characterized. Numerical illustrations of the benefits of the transformation are provided.

Key words: Malliavin derivatives, Doss transformation, multivariate diffusions

JEL Classification: G11, G12, G13

Mathematics Subject Classification (2000): 60H07, 60J60, 65C05

1 Introduction

Consider a stochastic process Y that satisfies the stochastic differential equation

$$dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t; \quad Y_0 \quad \text{given} \tag{1.1}$$

where W is a Brownian motion process and the coefficients $\mu(t, y) : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ and $\sigma(t, y) : [0, T] \times \mathbb{R} \mapsto \mathbb{R}_+$ satisfy growth and Lipschitz conditions. Also assume that the coefficients are infinitely differentiable functions of y with bounded derivatives of all orders greater than or equal to one, and that $\mu(t, 0)$ and $\sigma(t, 0)$ are bounded. Under these conditions, Nualart (1995), Theorem 2.2.2, shows that

Financial support from the Network of Centers of Excellence (MITACS) is gratefully acknowledged. The second author is grateful to Hydro-Québec and the Bank of Canada for financial support.

Correspondence to: René Garcia.

Maunuscript received: February 2002; final version received: November 2004

$Y(t) \in \mathbb{D}^\infty$ for all $t \in [0, T]$. Moreover, the (first order) Malliavin derivative of the process, denoted $\mathcal{D}_t Y_s$, satisfies the linear stochastic differential equation

$$d\mathcal{D}_t Y_s = \partial_2 \mu(s, Y_s) \mathcal{D}_t Y_s ds + \partial_2 \sigma(s, Y_s) \mathcal{D}_t Y_s dW_s \tag{1.2}$$

for $s \geq t$, where $\partial_2 f(s, y) \equiv \partial f(s, y)/\partial y$, subject to the boundary condition $\lim_{s \to t} \mathcal{D}_t Y_s = \sigma(t, Y_t)$. Differentiating (1.2) repeatedly provides stochastic differential equations for higher order Malliavin derivatives, $\mathcal{D}_t^j Y_s, j \geq 2$. Since the equation for $\mathcal{D}_t^j Y_s$ is linear, the solution can be written in terms of lower order derivatives and of the underlying process (1.1).

The characterization of $\mathcal{D}_t^j Y_s$ as the solution of an SDE is particularly useful for applications since it is amenable to numerical computations. Standard methods developed to simulate the solution of an SDE (see for example Kloeden and Platen 1997) can be used for that purpose. The convergence rate of numerical estimates is determined by the convergence rate of the martingale part of the SDE, which is slower than the convergence rate of the drift.

In this paper we derive an alternative characterization of Malliavin derivatives of diffusion processes involving the solutions of ordinary differential equations. Specifically we show that the Malliavin derivative $\mathcal{D}_t Y_s$ can be written as $\mathcal{D}_t Y_s = \sigma(s, Y_s) \mathcal{D}_t Z_s$ where the process $\mathcal{D}_t Z_s$ solves an ODE of the form

$$d\mathcal{D}_t Z_s = A^{(1)}(s, Z_s) \mathcal{D}_t Z_s ds \tag{1.3}$$

for $s \geq t$, for some suitable function $A^{(1)}(s, z)$, subject to the boundary condition $\lim_{s \to t} \mathcal{D}_t Z_s = 1$. The key to this formula is a change of variables, from Y to Z, which reduces the volatility of the transformed process, Z, to one. This change of variables is closely related to the transformation developed by Doss (1977) to convert a stochastic differential equation into an ordinary differential equation. It has also been used by other authors in univariate settings (e.g. Lamperti 1964; Gihman and Skorohod 1972). Straightforward application of the rules of Malliavin calculus to the SDE satisfied by Z leads to (1.3). We establish this formula for univariate as well as multivariate processes and provide characterizations for higher order derivatives. In all cases the Malliavin derivatives of the underlying processes can be written as functions of the Malliavin derivatives of transformed processes which satisfy an ODE in the univariate case and a matrix ODE in the multivariate case. These characterizations are useful for numerical applications since they improve the stability and convergence properties of numerical estimates of $\mathcal{D}_t Y_s$ to their true values.

A rapidly growing literature applies Malliavin calculus to resolve various problems of interest to financial economists. For instance, Fournié et al. (1999) use Malliavin calculus to compute the Greeks for discontinuous path-dependent payoffs written on multivariate diffusions. Fournié et al. (2001) propose an application to the global pricing and hedging of a European option. Detemple et al. (2003) compute optimal portfolios based on the Ocone and Karatzas (1991) formula. Cvitanić et al. (2003) deal with the efficient computation of hedges for options with discontinuous payoffs. Asymptotic properties of Monte Carlo estimators of diffusions are studied by Detemple et al. (2004), while issues regarding variance reduction are considered by Bouchard et al. (2004). All these applications use Malliavin derivatives of first and higher order and are concerned, in some way, with increasing the convergence speed of Monte Carlo schemes. The transformation we propose offers a solution that improves the rate of convergence of numerical estimates of Malliavin derivatives.

The next section of the paper details the univariate case. Section 3 considers multivariate diffusions. Section 4 summarizes asymptotic properties of estimators and provides a numerical illustration of the benefits of the transformation.

2 Malliavin derivatives of univariate diffusions

Consider a stochastic process Y that satisfies the stochastic differential equation

$$dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t; \quad Y_0 \text{ given} \tag{2.1}$$

where W is a Brownian motion process and the coefficients $\mu(t, y) : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ and $\sigma(t, y) : [0, T] \times \mathbb{R} \mapsto \mathbb{R}_+$ satisfy growth and Lipschitz conditions for the existence of a unique strong solution to (2.1). We also assume that σ is bounded away from 0. Following Lamperti (1964), Gihman and Skorohod (1972) and Doss (1977) we consider a new state variable $Z_t = F(t, Y_t)$ where the function $F : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ is selected so that $\partial_2 F = 1/\sigma$, where $\partial_2 F \equiv \partial F/\partial Y$. Since $\sigma(t, y) > 0$ we can define the inverse transformation $y = G(t, z)$. With this definition the relation $Y_t = G(t, Z_t)$ holds.

Using $\partial_{22} F = \partial_2 (1/\sigma) = -(\partial_2 \sigma)/\sigma^2$ along with Itô's lemma shows that Z satisfies the equation

$$dZ_t = \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](t, G(t, Z_t))dt + dW_t, \tag{2.2}$$

which has unit volatility coefficient.

The representation of Y as a function of a diffusion Z with unit volatility is closely related to a construction proposed by Doss (1977). Doss establishes the existence of a function h, solving $\partial_2 h = \sigma(t, h)$, and a process D satisfying another ODE such that $Y_t = h(t, D_t, W_t)$. It can be shown that $h(t, d, w) = G(t, g(d, w))$ where G solves $\partial_2 G = \sigma(t, G)$ and where $g(d, w) = d + w$. Clearly, with $Z_t \equiv D_t + W_t$, we obtain $Y_t = G(t, Z_t)$ where Z solves an SDE with identity volatility coefficient.

Theorem 2.1 *Suppose that the following conditions hold: (i) differentiability of drift, $\mu \in \mathcal{C}^1([0, T] \times \mathbb{R})$, (ii) differentiability of volatility, $\sigma \in \mathcal{C}^2([0, T] \times \mathbb{R})$, (iii) growth condition, $\mu(t, 0)$ and $\sigma(t, 0)$ are bounded for all $t \in [0, T]$. Then, for $s \geq t$,*

$$\mathcal{D}_t Y_s = \sigma(s, Y_s) \exp\left[\int_t^s \left[\partial_2\mu - \frac{\mu\partial_2\sigma}{\sigma} - \frac{1}{2}(\partial_{22}\sigma)\sigma - \frac{\partial_1\sigma}{\sigma}\right](v, Y_v)dv\right]. \tag{2.3}$$

Note that (2.3) is expressed entirely in terms of Riemann-Stieltjes integrals of first and second derivatives of the coefficients of Y. Formula (2.3) is therefore easily computed using standard methods to approximate the Riemann integrals involved.

Proof Let $F : [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ such that $\partial_2 F = 1/\sigma$. Using $\partial_{22} F = \partial_2 (1/\sigma) = -(\partial_2 \sigma)/\sigma^2$ and Itô's lemma implies

$$dF(t, Y_t) = \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](t, G(t, Z_t))dt + dW_t \tag{2.4}$$

where G is the inverse function $G(t, F(t, y)) = y$ and $Z_t \equiv F(t, Y_t)$. Applying the chain rule of Malliavin calculus shows that $\mathcal{D}_t Z_s$ solves the stochastic differential equation

$$d\mathcal{D}_t Z_s = \partial_2 \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](s, G(s, Z_s))\partial_2 G(s, Z_s)\mathcal{D}_t Z_s ds \tag{2.5}$$

for $s \geq t$, subject to the boundary condition $\lim_{s \longrightarrow t} \mathcal{D}_t Z_s = 1$.

But $Y_s = G(s, Z_s)$, so that

$$\mathcal{D}_t Y_s = \partial_2 G(s, Z_s)\mathcal{D}_t Z_s = \sigma(s, Y_s)\mathcal{D}_t Z_s \tag{2.6}$$

where the equality on the right hand side follows from

$$\partial_2 G(s, F(s, Y_s))\partial_2 F(s, Y_s) = 1. \tag{2.7}$$

Solving (2.5) for $\mathcal{D}_t Z_s$ and substituting in (2.6) gives the expression in the theorem. □

We now provide a characterization of higher order derivatives. The j^{th} order Malliavin derivative operator with respect to $W_{t_1}, \dots, W_{t_j}$ will be denoted by $\mathcal{D}^j_{\{t_1,\dots,t_j\}} Y_s$. To state the result we introduce the following notation. First, define a function $A : [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ such that

$$A(t, z) \equiv \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](t, G(t, z)). \tag{2.8}$$

For differentiable functions we write $A^{(j)}(t, z)$ and $G^{(j)}(t, z)$ for the j^{th} derivative with respect to z. Second, for a set with j elements $\{t_1, \dots, t_j\}$ and a number $1 \leq k \leq j$ let $\mathcal{P}(\{t_1, \dots, t_j\})(k)$ be the set of partitions of $\{t_1, \dots, t_j\}$ into k nonempty subsets $\{I_1, ..., I_k\}$. We have

Theorem 2.2 *Suppose that the following conditions hold: (i) infinite differentiability of drift,* $\mu \in \mathcal{C}^\infty([0,T] \times \mathbb{R})$*, (ii) infinite differentiability of volatility,* $\sigma \in \mathcal{C}^\infty([0,T] \times \mathbb{R})$*, (iii) growth condition,* $\mu(t, 0)$ *and* $\sigma(t, 0)$ *are bounded for all* $t \in [0,T]$*. Then, for* $j \geq 2$*,*

$$\mathcal{D}^j_{\{t_1,\dots,t_j\}} Y_s = \sum_{k=1}^{j} G^{(k)}(s, Z_s) \sum_{\{I_1,\dots,I_k\} \in \mathcal{P}(\{t_1,\dots,t_j\})(k)} \prod_{l=1}^{k} \mathcal{D}^{card(I_l)}_{I_l} Z_s \tag{2.9}$$

where $card(I_j)$ *is the number of elements in the subset* I_j *and*

$$d\mathcal{D}^j_{\{t_1,\dots,t_j\}} Z_s = \sum_{k=1}^{j} A^{(k)}(s, Z_s) \sum_{\{I_1,\dots,I_k\}\in\mathcal{P}(\{t_1,\dots,t_j\})(k)} \prod_{l=1}^{k} \mathcal{D}^{card(I_l)}_{I_l} Z_s ds \tag{2.10}$$

for all $s \geq t_j \geq \ldots \geq t_1$, *subject to the boundary condition*

$$\lim_{s\longrightarrow t_j} \mathcal{D}^j_{\{t_1,\dots,t_j\}} Z_s = 0. \tag{2.11}$$

Proof Theorem 2.1 showed that $\mathcal{D}^1_{t_1} Y_s = G^{(1)}(s, Z_s)\mathcal{D}^1_{t_1} Z_s$ where $d\mathcal{D}^1_{t_1} Z_s = A^{(1)}(s, Z_s)\mathcal{D}^1_{t_1} Z_s ds$ subject to $\lim_{s\longrightarrow t_1} \mathcal{D}^1_{t_1} Z_s = 1$. Using the chain rule of Malliavin calculus and assumptions (i)–(ii) gives, for $j = 2$,

$$\mathcal{D}^2_{\{t_1,t_2\}} Y_s = G^{(2)}(s, Z_s) \left(\mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^1_{t_2} Z_s\right) + G^{(1)}(s, Z_s)\mathcal{D}^2_{\{t_1,t_2\}} Z_s \tag{2.12}$$

where $\partial_2 G^{(1)}(s, Z_s) \equiv G^{(2)}(s, Z_s)$ and

$$d\mathcal{D}^2_{\{t_1,t_2\}} Z_s = \left(A^{(2)}(s, Z_s) \left(\mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^1_{t_2} Z_s\right) + A^{(1)}(s, Z_s)\mathcal{D}^2_{\{t_1,t_2\}} Z_s\right) ds \tag{2.13}$$

for $s \geq t_2 \geq t_1$, subject to the boundary condition $\lim_{s\longrightarrow t_2} \mathcal{D}^2_{\{t_1,t_2\}} Z_s = 0$.

For $j = 3$, we get

$$\begin{aligned}\mathcal{D}^3_{\{t_1,t_2,t_3\}} Y_s = {} & G^{(3)}(s, Z_s) \left(\mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^1_{t_2} Z_s \times \mathcal{D}^1_{t_3} Z_s\right) \\ & + G^{(2)}(s, Z_s) \left(\mathcal{D}^2_{\{t_1,t_3\}} Z_s \times \mathcal{D}^1_{t_2} Z_s + \mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^2_{\{t_2,t_3\}} Z_s\right) \\ & + G^{(2)}(s, Z_s) \left(\mathcal{D}^2_{\{t_1,t_2\}} Z_s \times \mathcal{D}^1_{t_3} Z_s\right) \\ & + G^{(1)}(s, Z_s)\mathcal{D}^3_{\{t_1,t_2,t_3\}} Z_s\end{aligned} \tag{2.14}$$

where $\mathcal{D}^2_{\{t_1,t_2\}} Z_s$ is given above, $\mathcal{D}^2_{\{t_1,t_3\}} Z_s$ and $\mathcal{D}^2_{\{t_2,t_3\}} Z_s$ satisfy similar equations, and the third derivative $\mathcal{D}^3_{\{t_1,t_2,t_3\}} Z_s$ solves

$$\begin{aligned}d\mathcal{D}^3_{\{t_1,t_2,t_3\}} Z_s = {} & A^{(3)}(s, Z_s) \left(\mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^1_{t_2} Z_s \times \mathcal{D}^1_{t_3} Z_s\right) ds \\ & + A^{(2)}(s, Z_s) \left(\mathcal{D}^2_{\{t_1,t_3\}} Z_s \times \mathcal{D}^1_{t_2} Z_s\right) ds \\ & + A^{(2)}(s, Z_s) \left(\mathcal{D}^1_{t_1} Z_s \times \mathcal{D}^2_{\{t_2,t_3\}} Z_s\right) ds \\ & + A^{(2)}(s, Z_s) \left(\mathcal{D}^2_{\{t_1,t_2\}} Z_s \times \mathcal{D}^1_{t_3} Z_s\right) ds \\ & + A^{(1)}(s, Z_s)\mathcal{D}^3_{\{t_1,t_2,t_3\}} Z_s ds.\end{aligned} \tag{2.15}$$

Collecting terms and using the definition of $\mathcal{P}(\{t_1,\dots,t_j\})(k)$ leads to the expressions stated in the theorem. Repeated differentiation gives (2.9) and (2.10) for the general case. □

For the special case $t_1 = \ldots = t_j$ the inside sums can be simplified. With the notation

$$C^j_{i_1,\ldots,i_k} \equiv \frac{j!}{i_1! \times \ldots \times i_k!} \tag{2.16}$$

where $j = i_1 + \ldots + i_k$ we can write

Corollary 2.1 *Suppose that the following conditions hold: (i) infinite differentiability of drift,* $\mu \in \mathcal{C}^\infty([0,T] \times \mathbb{R})$*, (ii) infinite differentiability of volatility,* $\sigma \in \mathcal{C}^\infty([0,T] \times \mathbb{R})$*, (iii) growth condition,* $\mu(t,0)$ *and* $\sigma(t,0)$ *are bounded for all* $t \in [0,T]$*. Then, for* $j \geq 2$*,*

$$\mathcal{D}^j_t Y_s = \sum_{k=1}^{j} G^{(k)}(s, Z_s) \sum_{\substack{1 \leq i_1 \leq \cdots \leq i_k \leq j \\ j = i_1 + \ldots + i_k}} C^j_{i_1,\ldots,i_k} \prod_{l=1}^{k} \mathcal{D}^{i_l}_t Z_s \tag{2.17}$$

where

$$d\mathcal{D}^j_t Z_s = \sum_{k=1}^{j} \left(A^{(k)}(s, Z_s) \sum_{\substack{1 \leq i_1 \leq \cdots \leq i_k \leq j \\ j = i_1 + \ldots + i_k}} C^j_{i_1,\ldots,i_k} \prod_{l=1}^{k} \mathcal{D}^{i_l}_t Z_s \right) ds \tag{2.18}$$

for all $s \geq t$*, subject to the boundary condition* $\lim_{s \longrightarrow t} \mathcal{D}^j_t Z_s = 0$.

Proof Consider the case $j = 3$. When $k = 1$ the set of partitions

$$\mathcal{P}(\{t_1, t_2, t_3\})(1) = \mathcal{P}(\{t, t, t\})(1) \tag{2.19}$$

has one element, $I_1 = \{t, t, t\}$, of cardinality $card(I_1) = 3$. Accordingly, $C^3_{i_1} = C^3_3 = 1$. For $k = 2$ the set of partitions is

$$\begin{aligned} \mathcal{P}(\{t_1, t_2, t_3\})(2) &= \{\{(t_1), (t_2, t_3)\},\ \{(t_2), (t_1, t_3)\}, \{(t_3), (t_1, t_2)\}\} \\ &= \{\{(t), (t, t)\}, \{(t), (t, t)\}, \{(t), (t, t)\}\}. \end{aligned} \tag{2.20}$$

Each of these partitions has two sets with respectively 1 and 2 elements. Here $C^3_{i_1,i_2} = C^3_{1,2} = 3$ and there is only one set of indices i_1, i_2 such that $3 = i_1 + i_2$ and $1 \leq i_1 \leq i_2 \leq 3$. This is $i_1 = 1$ and $i_2 = 2$. The inside sum then equals

$$\begin{aligned} \sum_{\substack{1 \leq i_1 \leq \cdots \leq i_k \leq j \\ j = i_1 + \ldots + i_k}} C^j_{i_1,\ldots,i_k} \prod_{l=1}^{k} \mathcal{D}^{i_l}_t Z_s &= \sum_{\substack{1 \leq i_1 \leq i_2 \leq 3 \\ j = i_1 + i_2}} C^3_{i_1,i_2} \left(\mathcal{D}^{i_1}_t Z_s \times \mathcal{D}^{i_2}_t Z_s \right) \\ &= 3\mathcal{D}^1_t Z_s \times \mathcal{D}^2_t Z_s \end{aligned} \tag{2.21}$$

as it should be. The formula can be verified for $j > 3$ in a similar manner. □

3 Multivariate diffusions

This section extends the results to a multivariate setup. We consider a d-dimensional process Y that satisfies the system of SDEs

$$dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t; \quad Y_0 \text{ given} \tag{3.1}$$

where W is a p-dimensional Brownian motion process and the coefficients $\mu(t, y) : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma(t, y) : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^p$ satisfy growth and Lipschitz conditions.

For any $d \times 1$ vector of differentiable functions $f(t, Y)$ let $\partial_1 f$ represent the $d \times 1$ vector of first derivatives relative to time and $\partial_2 f$ the $d \times d$ matrix whose rows are composed of the gradients relative to Y of the elements of f. For any function $G : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and a differentiable matrix $\sigma(t, y) : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^p$ define the operator

$$\mathcal{A}_t G \equiv \mu(t, G) - \partial_1 G - \frac{1}{2} \sum_{j=1}^{d} \partial_2 \sigma_j(t, G) \sigma_j(t, G) \tag{3.2}$$

where $\sigma_j(t, G)$ is the j^{th} column of the matrix $\sigma(t, G)$.

The Malliavin derivative of Y has the following representation.

Theorem 3.1 *Let $p = d$ and suppose that the following conditions hold:*

(i) *differentiability of drift: $\mu \in \mathcal{C}^1([0, T] \times \mathbb{R}^d)$*
(ii) *differentiability of volatility: $\sigma_j \in \mathcal{C}^2([0, T] \times \mathbb{R}^d)$ for all $j = 1, \ldots, p$*
(iii) *growth condition: $\mu(t, 0)$ and $\sigma(t, 0)$ are bounded for all $t \in [0, T]$*
(iv) *volatility conditions:*
  - *(iv-a)* *commutativity: the Lie algebra of the vector fields generated by the columns of σ, $\mathcal{L}\{\sigma_1, ..., \sigma_d\}$, is Abelian, i.e. $(\partial_2 \sigma_i)\sigma_j = (\partial_2 \sigma_j)\sigma_i$*
  - *(iv-b)* *rank condition: $rank(\sigma) = d$ a.e.*

Under these conditions there exists an invertible function $G : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$, solving the total differential equation

$$\partial_2 G(t, z) = \sigma(t, G(t, z)); \quad G(t, 0) = 0 \ \textit{for all} \ t \in [0, T] \tag{3.3}$$

and a d-dimensional process Z that satisfies

$$dZ_s = A(s, Z_s)ds + dW_s; \quad G(0, Z_0) = Y_0 \tag{3.4}$$

with drift function

$$A(t, z) = \sigma(t, G(t, z))^{-1}[\mathcal{A}_t G](t, z), \tag{3.5}$$

where the operator $\mathcal{A}_t$ is defined in (3.2). For all $s \geq t$ the Malliavin derivative of Y is given by

$$\mathcal{D}_t Y_s = \sigma(s, G(s, Z_s))\mathcal{D}_t Z_s, \tag{3.6}$$

where $\mathcal{D}_t Z_s$ solves the stochastic differential equation

$$d\mathcal{D}_t Z_s = \partial_2 A(s, Z_s)\mathcal{D}_t Z_s ds; \quad \lim_{s \to t} \mathcal{D}_t Z_s = I_d \tag{3.7}$$

with I_d the d-dimensional identity matrix.

The commutativity condition in assumption (iv-a) corresponds to the Frobenius condition in Doss (1977) (see the proof of his Lemma 17), which guarantees that the total differential Eq. (3.3) is completely integrable (see Hartman 1982, p. 118, for a definition). Given that σ is Lipschitz by assumption, the commutativity condition on the columns of the volatility matrix is necessary and sufficient for the existence of G (see Doss 1977, Lemma 17). As σ has both full row and column rank, G corresponds to the inverse of a function F that solves $\partial_2 F = \sigma^{-1}$. If the volatility coefficient satisfies the rank condition, the commutativity condition is necessary and sufficient for the existence of F as well. In addition, the rank condition implies that there exists a unique Z_0 such that $G(0, Z_0) = Y_0$. The commutativity assumption is automatically satisfied if the state variables do not interact with each other, i.e. if $\sigma_j(t, Y_t) = \sigma_j(t, Y_t^j)$ for $j = 1, ..., d$. The one dimensional case treated before falls in this category.

Proof The Lipschitz condition on σ and the commutativity assumption (iv-a) imply that there exists a function G which solves (3.3) (see Doss 1977, Lemma 17). The rank condition (iv-b) implies that $\partial_2 G(0, Z_0)$ is a linear isomorphism, i.e. $det(\partial_2 G(0, Z_0)) = det(\sigma(0, G(0, Z_0))) \neq 0$. Thus, by the inverse function theorem, there exists a unique Z_0 such that $G(0, Z_0) = Y_0$. To complete the proof it remains to establish the existence of a drift function $A : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ for the diffusion Z, such that $G(t, Z_t) = Y_t$ for all $t > 0$.

Applying Itô's rule on both sides of $G(t, Z_t) = Y_t$ yields

$$\partial_1 G(t, Z_t)dt + \partial_2 G(t, Z_t)dZ_t + \frac{1}{2}\sum_{i,j=1}^{d} \partial_{ij}^2 G(t, Z_t)d[Z^i, Z^j]_t = dY_t \tag{3.8}$$

where $\partial_{ij}^2 \equiv \partial^2/\partial z_i \partial z_j$. Invoking again (3.3) and the representation (3.4) enables us to write

$$\begin{aligned} dY_t = {} & \partial_1 G(t, Z_t)dt + \sigma(t, G(t, Z_t))dZ_t \\ & + \frac{1}{2}\sum_{j=1}^{d} \partial_2 \sigma_j(t, G(t, Z_t))\sigma_j(t, G(t, Z_t))dt \end{aligned} \tag{3.9}$$

where $\sigma_j(t, G(t, Z_t))$ is the j^{th} column of the matrix $\sigma(t, G(t, Z_t))$. For this equation to hold A must solve the equation

$$\sigma(t, G(t, z))A(t, z) = [\mathcal{A}_t G](t, z) \tag{3.10}$$

where the operator $\mathcal{A}_t$ is given in (3.2). The rank condition on σ then gives (3.5), and a straightforward application of the rules of Malliavin calculus establishes (3.6)–(3.7). □

Remark 3.1 The proof of Theorem 3.1 shows that the rank condition (iv-b) is used to find a unique drift, given by (3.5), for the process Z and a unique initial value Z_0 solving $G(t, Z_0) = Y_0$. A version of the theorem can also be established when the solution of (3.10) exists, but is not unique, and the set

$$L_G(Y_0) \equiv \{z \in \mathbb{R}^d : G(0, z) = Y_0\}, \tag{3.11}$$

is not empty. To see this note that a drift function exists if and only if $[\mathcal{A}_t G](t,z) \in \mathcal{M}(\sigma(t,G(t,z)))$, where the operator $\mathcal{A}_t$ is given in (3.2) and where $\mathcal{M}(\sigma) \equiv \{y : y = \sigma x,\ x \in \mathbb{R}^d\}$ is the vector space spanned by the columns of σ. Since $[\mathcal{A}_t G](t,z) \in \mathcal{M}(\sigma(t,G(t,z)))$ if and only if

$$\sigma(t,G(t,z))\sigma(t,G(t,z))^+[\mathcal{A}_t G](t,z) = [\mathcal{A}_t G](t,z), \tag{3.12}$$

where σ^+ is the Moore-Penrose (MP) generalized inverse of σ (see Magnus and Neudecker 1988, p. 32, for a definition), we obtain

$$\begin{aligned} A^{(Q)}(t,z) = \sigma(t,G(t,z))^+[\mathcal{A}_t G](t,z) + Q(t,z) \\ -\sigma(t,G(t,z))^+\sigma(t,G(t,z))Q(t,z) \end{aligned} \tag{3.13}$$

for some arbitrary function $Q : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$.

To establish that $G(0,Z_0) = Y_0$ for some Z_0 we rely on the implicit function theorem which states that the level set (3.11), which describes the set of admissible initial conditions Z_0, is a non-empty manifold if the linear mapping $\partial_2 G(t,Z_0)$ is onto. Therefore, if $L_G(Y_0) \neq \emptyset$, but the rank condition fails, the initial value of the diffusion $Z_0 \in L_G(Y_0)$ and thus Z itself is not unique. If $\partial_2 G$ is not onto we can proceed as in Remark 3.3 below.

Our next remarks briefly discuss extensions of the results to cases where the volatility matrix fails to satisfy the conditions in (iv) of Theorem 3.1.

Remark 3.2 Let $d < p$ and suppose that $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d\times\mathbb{R}^p$ has full row rank ($rank(\sigma) = d < p$, $a.e.$) and is commutative. Remark 3.1 shows that the solution set $L_G(Y_0)$ is no longer a singleton, and (3.10) has an uncountable number of solutions $A^{(Q)} : [0,T] \times \mathbb{R}^p \mapsto \mathbb{R}^p$ that can be constructed with $\sigma^+ = \sigma'(\sigma\sigma')^{-1}$. The process Z is a p-dimensional diffusion of the form (3.4) with drift $A^{(Q)}$ and initial value $Z_0 \in L_G(Y_0)$. The Malliavin derivative of Y is given by the formulas in Theorem 3.1 with A replaced by $A^{(Q)}$.

Remark 3.3 Let $p < d$ and suppose that $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^p$ has full column rank ($rank(\sigma) = p < d$, $a.e.$) and is commutative. Equation (3.10) and the nonlinear equation $G(0,Z_0) = Y_0$ are overdetermined and therefore, if (3.12) fails and $\partial_2 G$ is not onto, both equations have no solution. In this case, we can embed Y in $X' \equiv [Y', V']$, where V is a q-dimensional diffusion, $q = p - d$, with coefficients (μ^V, σ^V) such that $(\sigma^X)' = [\sigma', (\sigma^V)']$ satisfies assumption (iv) of Theorem 3.1. An application of Theorem 3.1 to X shows that there exists a p-dimensional diffusion Z with drift $A^{(V)}$ and volatility coefficient I_p, such that $X_t = G^{(V)}(t,Z_t)$. Taking $G(t,z) \equiv [G_i^{(V)}(t,Z_t)]_{i=1,\dots,d}$ to be the first d components of $G^{(V)}$ we obtain $Y_t = G(t,Z_t)$ where

$$dZ_t = A^{(V)}(t,Z_t)dt + dW_t; \quad G(0,Z_0) = Y_0. \tag{3.14}$$

Hence,

$$\mathcal{D}_t Y_s = \partial_2 G(t,Z_t)\mathcal{D}_t Z_s \tag{3.15}$$

for $s \geq t$, where

$$d\mathcal{D}_t Z_s = \partial_2 A^{(V)}(s, Z_s)\mathcal{D}_t Z_s ds \tag{3.16}$$

subject to the boundary condition $\lim_{s \to t} \mathcal{D}_t Z_s = I_p$. With the particular selection of the $q-$dimensional Brownian motion $V' \equiv [W_{d+1}, \ldots, W_p]$ as an auxiliary process we obtain

$$A^{(V)}(t, z) = \begin{bmatrix} (\sigma^1(t, G(t, z)))^{-1}[\mathcal{A}_t G](t, z) \\ \mathbf{0}_{q,1} \end{bmatrix}, \tag{3.17}$$

where $\sigma = [\sigma^1, \sigma^2]$ with $\sigma^1 = [\sigma_1, \ldots, \sigma_d]$ and $\sigma^2 = [\sigma_{d+1}, \ldots, \sigma_p]$.

Remark 3.4 Let $p = d$ and suppose that $\sigma : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^d$ has full rank but fails to be commutative. Also suppose that (i) there exists a function $B : [0, T] \mapsto \mathbb{R}^d$ such that $\sigma(t, y)B(t)^+ B(t) = \sigma(t, y)$ and that (ii) the rotated volatility matrix $\sigma^{(B)}(t, y) \equiv \sigma(t, y)B(t)^+$ is commutative, i.e. $((B_j^+)' \otimes I_d)(\partial_2 \sigma)\sigma B_i^+ = ((B_i^+)' \otimes I_d)(\partial_2 \sigma)\sigma B_j^+$ for all $i, j = 1, \ldots, d$ (see Magnus and Neudecker 1988 for a definition of the Kronecker product $\otimes$), and satisfies

$$[\sigma^{(B)}(\sigma^{(B)})^+](t, G^{(B)}(t, z))[\mathcal{A}_t G^{(B)}](t, z)) = [\mathcal{A}_t G^{(B)}](t, z), \tag{3.18}$$

where $G^{(B)}$ solves (3.3) with $\sigma^{(B)}$ in place of σ. Then, if $L_{G^{(B)}}(Y_0)$, defined for Y_0 and $G^{(B)}$ as in (3.11), is non-empty, Theorem 3.1 applies to ensure the existence of a function $G^{(B)}$ such that $G^{(B)}(t, Z_t) = Y_t$, for all $t \in [0, T]$, where Z is a diffusion process with initial value $Z_0 \in L_{G^{(B)}}(Y_0)$, deterministic volatility function $B(t)$ and drift function

$$\begin{aligned} A^{(B,Q)}(t, z) = \sigma^{(B)}(t, G^{(B)}(t, z))^+[\mathcal{A}_t G^{(B)}](t, z) + Q(t, z) \\ -[(\sigma^{(B)})^+ \sigma^{(B)}](t, G^{(B)}(t, z))Q(t, z) \end{aligned} \tag{3.19}$$

for some function Q. (Note that if B is of full rank $\sigma^{(B)}$ is of full rank as well. In this case $A^{(B,Q)} = A^{(B)}$ where

$$A^{(B)}(t, z) = \sigma^{(B)}(t, G^{(B)}(t, z))^{-1}[\mathcal{A}_t G^{(B)}](t, z), \tag{3.20}$$

and $L_{G^{(B)}}(Y_0)$ is a singleton). The Malliavin derivative of Y is given by the formulas in Theorem 3.1 with (G, A) replaced by $(G^{(B)}, A^{(B,Q)})$ and the limiting condition $\lim_{s \to t} \mathcal{D}_t Z_s = B(t)$ in (3.7). Alternatively, if (3.18) is not satisfied, we can proceed as in Remark 3.3 with σ replaced by $\sigma^{(B)}$.

We now present results for higher order derivatives. The j^{th} order partial Malliavin derivative operator with respect to Wiener processes $W_{l_1}, \ldots, W_{l_j}$ will be denoted by $\mathcal{D}^j_{\{(t_1,l_1),\ldots,(t_j,l_j)\}}$. For differentiable functions we write $(A_i)^{(j)}_{\{h_1,\ldots,h_j\}}(t, z)$ and $(G_i)^{(j)}_{\{h_1,\ldots,h_j\}}(t, z)$ for the j^{th} derivative with respect to $\{z_{h_1}, \ldots, z_{h_j}\}$.

Theorem 3.2 *Suppose that the following conditions hold: (i) infinite differentiability of drift, $\mu \in \mathcal{C}^{\infty}([0,T] \times \mathbb{R}^d)$, (ii) infinite differentiability of volatility, $\sigma \in \mathcal{C}^{\infty}([0,T] \times \mathbb{R}^d)$, (iii) growth condition, $\mu(t,0)$ and $\sigma(t,0)$ are bounded for all $t \in [0,T]$, and condition (iv) of theorem 4. Then, for $j \geq 2$,*

$$\mathcal{D}^{j}_{H(j)} Y_{i,s} = \sum_{k=1}^{j} \sum_{h_1,\ldots,h_k=1}^{d} (G_i)^{(k)}_{\{h_1,\ldots,h_k\}}(s,Z_s) \times \sum_{\{I_1,\ldots,I_k\} \in \mathcal{P}(H(j))(k)} \prod_{l=1}^{k} \mathcal{D}^{card(I_l)}_{I_l} Z_{h_{\iota(I_l,k)},s} \tag{3.21}$$

where $\mathcal{P}(H(j))(k)$ is the set of all partitions of the set $H(j) = \{K_1,\ldots,K_j\}$ with $K_m = (t_m,l_m)$ into k nonempty subsets $\{I_1,...,I_k\}$, where $\iota(I_m,k) \equiv$ $\min\{\min\{n : K_n \in I_m\}, k\}$, and where

$$d\mathcal{D}^{j}_{H(j)} Z_{i,s} = \sum_{k=1}^{j} \sum_{h_1,\ldots,h_k=1}^{d} (A_i)^{(k)}_{\{h_1,\ldots,h_k\}}(s,Z_s) \times \sum_{\{I_1,\ldots,I_k\} \in \mathcal{P}(H(j))(k)} \prod_{l=1}^{k} \mathcal{D}^{card(I_l)}_{I_l} Z_{h_{\iota(I_l,k)},s} ds \tag{3.22}$$

subject to the boundary condition

$$\lim_{s \longrightarrow t^{\star}} \mathcal{D}^{j}_{H(j)} Z_{i,s} = 0 \tag{3.23}$$

for all $t^{\star} \leq \max\{t_1,\ldots,t_j\}$.

Proof Theorem 3.1 showed that

$$\mathcal{D}^{1}_{(t_1,l_1)} Y_{i,s} = \sum_{h_1=1}^{d} (G_i)^{(1)}_{h_1}(s,Z_s) \mathcal{D}^{1}_{(t_1,l_1)} Z_{h_1,s} \tag{3.24}$$

where

$$d\mathcal{D}^{1}_{(t_1,l_1)} Z_{i,s} = \sum_{h_1=1}^{d} (A_i)^{(1)}_{h_1}(s,Z_s) \mathcal{D}^{1}_{(t_1,l_1)} Z_{h_1,s} ds \tag{3.25}$$

subject to the boundary condition

$$\lim_{s \longrightarrow t_1} \mathcal{D}^{1}_{(t_1,l_1)} Z_{h_1,s} = 1 \tag{3.26}$$

for all $h_1 = 1,\ldots,d$. Using the chain rule of Malliavin calculus and assumptions (i)–(ii) gives

$$\mathcal{D}^{2}_{\{(t_1,l_1),(t_2,l_2)\}} Z_{i,s} = \sum_{h_1,h_2=1}^{d} (G_i)^{(2)}_{\{h_1,h_2\}}(s,Z_s) \times \left(\mathcal{D}^{1}_{(t_1,l_1)} Z_{h_1,s} \times \mathcal{D}^{1}_{(t_2,l_2)} Z_{h_2,s} \right) + \sum_{h_1=1}^{d} (G_i)^{(1)}_{h_1}(s,Z_s) \mathcal{D}^{2}_{\{(t_1,l_1),(t_2,l_2)\}} Z_{h_1,s} \tag{3.27}$$

where $\partial_{h_2}(G_i)^{(1)}_{h_1}(s, Z_s) \equiv (G_i)^{(2)}_{\{h_1,h_2\}}(s, Z_s)$ and

$$d\mathcal{D}^2_{\{(t_1,l_1),(t_2,l_2)\}}Z_{i,s} = \sum_{h_1,h_2=1}^{d} (A_i)^{(2)}_{\{h_1,h_2\}}(s, Z_s) \times \left(\mathcal{D}^1_{(t_1,l_1)}Z_{h_1,s} \times \mathcal{D}^1_{(t_2,l_2)}Z_{h_2,s}\right) ds + \sum_{h_1=1}^{d} (A_i)^{(1)}_{h_1}(s, Z_s)\mathcal{D}^2_{\{(t_1,l_1),(t_2,l_2)\}}Z_{h_1,s}ds \tag{3.28}$$

subject to the boundary condition $\lim_{s\longrightarrow t^\star} \mathcal{D}^2_{\{(t_1,l_1),(t_2,l_2)\}}Z_{h_1,s} = 0$ for any $t^\star \leq \max\{t_1, t_2\}$, and for all $h_1 = 1, \ldots, d$.

For $j = 3$ we get

$$\mathcal{D}^3_{\{(t_1,l_1),(t_2,l_2),(t_3,l_3)\}}Y_{i,s} = \sum_{h_1,h_2,h_3=1}^{d} (G_i)^{(3)}_{\{h_1,h_2,h_3\}}(s, Z_s)K_{h_1,h_2,h_3,s} + \sum_{h_1,h_2=1}^{d} (G_i)^{(2)}_{\{h_1,h_2\}}(s, Z_s)K_{h_1,h_2,s} + \sum_{h_1=1}^{d} (G_i)^{(1)}_{\{h_1\}}(s, Z_s) \times \left(\mathcal{D}^3_{\{(t_1,l_1),(t_1,l_2),(t_3,l_3)\}}Z_{h_1,s}\right) \tag{3.29}$$

and

$$d\mathcal{D}^3_{\{(t_1,l_1),(t_2,l_2),(t_3,l_3)\}}Z_{i,s} = \sum_{h_1,h_2,h_3=1}^{d} (A_i)^{(3)}_{\{h_1,h_2,h_3\}}(s, Z_s)K_{h_1,h_2,h_3,s}ds + \sum_{h_1,h_2=1}^{d} (A_i)^{(2)}_{\{h_1,h_2\}}(s, Z_s)K_{h_1,h_2,s}ds + \sum_{h_1=1}^{d} (A_i)^{(1)}_{\{h_1\}}(s, Z_s) \times \left(\mathcal{D}^3_{\{(t_1,l_1),(t_1,l_2),(t_3,l_3)\}}Z_{h_1,s}\right) ds \tag{3.30}$$

where $K_{h_1,h_2,h_3,s} \equiv \mathcal{D}^1_{(t_1,l_1)}Z_{h_1,s} \times \mathcal{D}^1_{(t_2,l_2)}Z_{h_2,s} \times \mathcal{D}^1_{(t_3,l_3)}Z_{h_3,s}$ and

$$K_{h_1,h_2,s} \equiv \mathcal{D}^1_{(t_1,l_1)}Z_{h_1,s} \times \mathcal{D}^2_{\{(t_2,l_2),(t_3,l_3)\}}Z_{h_2,s} + \mathcal{D}^1_{(t_2,l_2)}Z_{h_2,s} \times \mathcal{D}^2_{\{(t_1,l_1),(t_3,l_3)\}}Z_{h_1,s} + \mathcal{D}^1_{(t_3,l_3)}Z_{h_2,s} \times \mathcal{D}^2_{\{(t_1,l_1),(t_2,l_2)\}}Z_{h_1,s}. \tag{3.31}$$

Repeated differentiation leads to the general expressions in the theorem. □

4 Asymptotic properties and numerical results

The usefulness of the change of variable and of the characterizations provided stems from the properties of simulation-based estimators of $(Y, \mathcal{D}Y)$. This section briefly reviews the convergence properties of those estimators and provides numerical illustrations of the results.

4.1 Error properties

Consider the d-dimensional random variable Y_T given by the terminal value of the solution of the SDE (3.1). To simplify the notation we assume that the coefficients are time independent. The Euler approximation of Y_T is

$$Y_T^N = Y_0 + \sum_{n=0}^{N-1} \mu(Y_{nh}^N)h + \sum_{n=0}^{N-1} \sum_{j=1}^{d} \sigma_j(Y_{nh}^N) \Delta W_{nh}^j \tag{4.1}$$

where $\sigma_j(y)$ is the j^{th} column of $\sigma(y)$, $h = T/N$, $\Delta W_{nh}^j = W_{(n+1)h}^j - W_{nh}^j$ and N is the number of discretization points selected. The asymptotic error distribution follows from Rootzén (1980), Kurtz and Protter (1991) and Jacod and Protter (1998).

Theorem 4.1 (Kurtz and Protter 1991) *The error $Y_T^N - Y_T$ satisfies $\sqrt{N}(Y_T^N - Y_T) \Rightarrow e_T^Y$ where*

$$e_T^Y = -\frac{1}{\sqrt{2}} \Omega_T \int_0^T \Omega_v^{-1} \sum_{h,j=1}^{d} [\partial\sigma_j \sigma_h](Y_v) dB_v^{h,j} \tag{4.2}$$

with $[B^{h,j}]_{h,j\in\{k,\dots,d\}}$ a $d^2 \times 1$ standard Brownian motion independent of W and $\partial\sigma_j$ a $d \times d$ matrix of derivatives of σ_j with respect to Y. The matrix Ω is given by

$$\Omega_v = \exp\left(\int_0^v [\partial\mu - \frac{1}{2} \sum_{j=1}^{d} \partial\sigma_j (\partial\sigma_j)'](Y_s) ds + \sum_{j=1}^{d} \int_0^v \partial\sigma_j(Y_s) dW_s^j \right) \tag{4.3}$$

where $\partial\mu$ is the $d \times d$ matrix of derivatives of the vector μ with respect to the elements of Y.

Theorem 4.1 shows that the error converges in law, at the rate $\sqrt{N}$, to the random variable e_T^Y as the number of discretization points N becomes large. The asymptotic error e_T^Y depends on the coefficients of the SDE and on their derivatives. It also involves a new Brownian motion, $[B^{h,j}]$, which is orthogonal to the original Brownian motion process W.

The transformed process $Z_t = F(Y_t)$ satisfies $dZ_t = A(Z_t)dt + dW_t; Z_0 = F(Y_0)$ and has Euler approximation

$$Z_T^N = Z_0 + \sum_{n=0}^{N-1} A(Z_{nh}^N)h + \sum_{n=0}^{N-1} \Delta W_{nh}. \tag{4.4}$$

The asymptotic error distribution is given in Detemple et al. (2004).

Theorem 4.2 (Detemple et al. 2004) *The error $Z_T^N - Z_T$ satisfies $N(Z_T^N - Z_T) \Rightarrow e_T^Z$ where*

$$e_T^Z = -\widehat{\Omega}_T \int_0^T \widehat{\Omega}_v^{-1} \partial A(Z_v) \left(\frac{1}{2} dZ_v + \frac{1}{\sqrt{12}} d\hat{B}_v + \frac{1}{2} \sum_{k,l=1}^{d} \partial_{lk} A(Z_s) ds \right). \tag{4.5}$$

Here $\hat{B}$ is a $d \times 1$ standard Brownian motion independent of W and $[B^{h,j}]$, $\partial A(Z) = [\partial_1 A(Z), ..., \partial_d A(Z)]$ is the $d \times d$ matrix with columns given by the derivatives of the vector $A(Z)$, and $\partial_{lk} A(Z)$ the $d \times 1$ vector of cross derivatives of $A(Z)$ with respect to arguments l, k. The $d \times d$ matrix $\widehat{\Omega}_v$ is

$$\widehat{\Omega}_v = \exp\left(\int_0^v \partial A(Z_s) ds \right). \tag{4.6}$$

Applying the inverse transformation gives

$$N(G(Z_T^N) - G(Z_T)) \Rightarrow \sigma(Y_T) e_T^Z. \tag{4.7}$$

Theorem 4.2 shows that the speed of convergence increases after application of the change of variables. Also, the limit law is different and involves exponentials of a bounded variation process instead of a stochastic integral. The convergence rate $1/N$ attained by $G(Z_T^N)$ corresponds to the convergence rate of the Euler scheme applied to an ordinary differential equation. This is the best rate that can be attained with this scheme.

4.2 A numerical example

We now illustrate the convergence results in Theorems 4.1 and 4.2 with an example. Suppose that Y follows a process with Constant Elasticity of Variance (CEV), $dY_t = \mu Y_t dt + \sigma Y_t^\gamma dW_t$ where $Y_0 > 0$ is given. The Malliavin derivative $\mathcal{D}Y$ solves the linear equation $d\mathcal{D}_t Y_s = \mu(\mathcal{D}_t Y_s) ds + \sigma\gamma Y_s^{\gamma-1} (\mathcal{D}_t Y_s) dW_s$ subject to $\lim_{s \to t} \mathcal{D}_t Y_s = \sigma Y_t^\gamma$.

The transformed process $Z = Y^{1-\gamma}$ solves

$$dZ_t = (1-\gamma)\left(\mu Z_t - \frac{1}{2} \gamma \sigma^2 Z_t^{-1} \right) dt + \sigma(1-\gamma) dW_t; \quad Z_0 = Y_0^{1-\gamma} \tag{4.8}$$

and its Malliavin derivative satisfies

$$d\mathcal{D}_t Z_s = (1-\gamma)\mathcal{D}_t Z_s \left(\mu + \frac{1}{2} \gamma \sigma^2 Z_t^{-2} \right) ds \tag{4.9}$$

subject to $\lim_{s \to t} \mathcal{D}_t Z_s = \sigma(1-\gamma)$. Applying the inverse transformation gives $Y_s = Z_s^{1/(1-\gamma)}$ and $\mathcal{D}_t Y_s = (1/(1-\gamma)) Z_s^{\gamma/(1-\gamma)} \mathcal{D}_t Z_s$ for all $s \in [0, T]$.

The numerical experiments compare estimates of Y and $\mathcal{D}_t Y$ based on the original processes to those obtained by the transformation. Distributions of errors

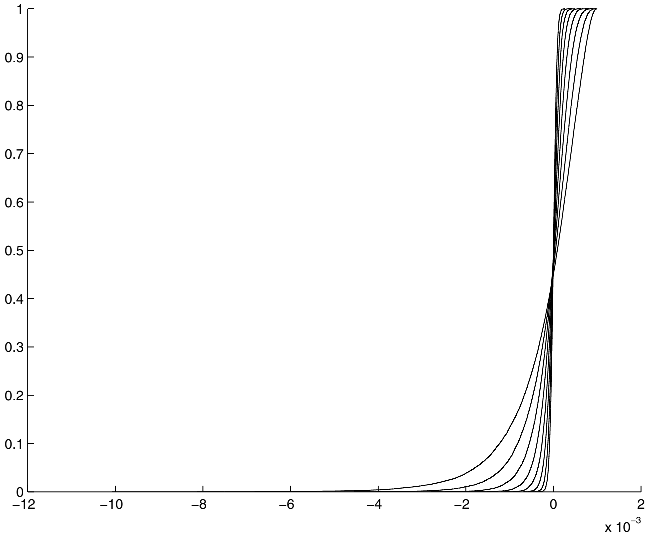

Fig. 1. Distribution of the error $Y_T^N - Y_T$ using the Euler scheme without transformation

and error statistics (e.g. Root Mean-Squared Error and Mean Absolute Error) are computed for different discretizations N of the time interval $[0, T]$, using a Monte Carlo (Euler) scheme. For example the estimate of the Mean Absolute Error for the Malliavin derivative is

$$\epsilon(N, M) = \widehat{E}^M \left|\mathcal{D}_0^N Y_T - \mathcal{D}_0 Y_T\right| = \frac{1}{M} \sum_{i=1}^{M} \left|\mathcal{D}_0^{N,i} Y_T - \mathcal{D}_0^i Y_T\right| \tag{4.10}$$

where $\mathcal{D}_0 Y_T$ is the true value of the derivative, $\mathcal{D}_0^N Y_T$ its approximation based on N discretization points, and $\hat{E}^M$ is the empirical mean (for the evaluation of numerical estimates of Y we use (4.10) with Y_T^N and Y_T in place of $\mathcal{D}_0^N Y_T$ and $\mathcal{D}_0 Y_T$). The computation of $\epsilon(N, M)$ requires the true value of the derivative which is usually not known. Fortunately, Monte Carlo simulation can be used for that purpose as well since estimates with and without transformation both converge to the true value as N becomes large. Accordingly, we compute the benchmark true value from a very fine discretization. We then calculate the difference between this benchmark and the Euler estimates, with and without the transformation, for different values of N. Estimates of Root Mean Square Errors are computed using a similar procedure.

Parameter values were selected as $\mu = 0.10, \sigma = 0.20, \gamma = 1.20, Y_0 = 0.10$ and $T = 1$. For Monte Carlo estimates of expectations we used $M = 25,000$ trajectories. The benchmark true values were computed using the method without transformation with $N = 2^{14}$.

Figures 1–4 and Table 1 illustrate the results. The figures graph the empirical distribution of $25,000$ replications of the error estimates $Y_T^N - Y_T$ and $\mathcal{D}_0^N Y_T - \mathcal{D}_0 Y_T$, for $N = 2^x$ with $x = 1, \ldots, 6$, based on the methods without and with transformation.

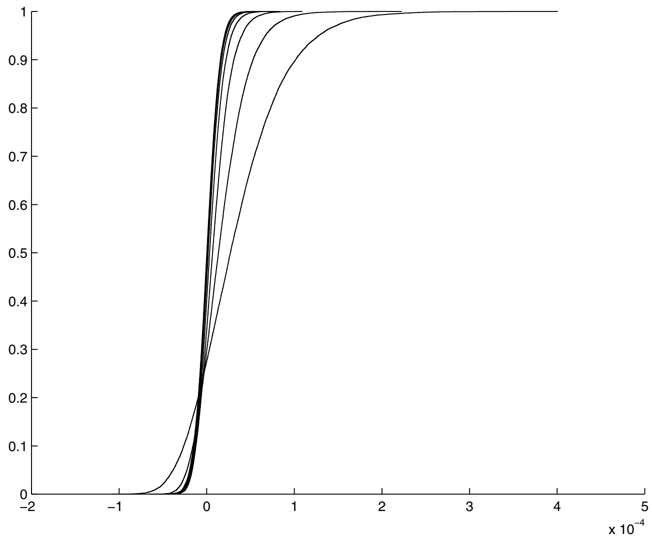

Fig. 2. Distribution of the error $Y_T^N - Y_T$ using the Euler scheme with transformation

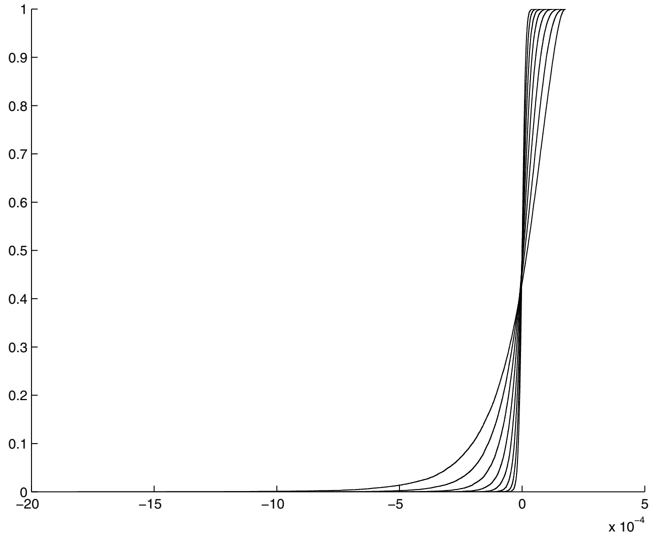

Fig. 3. Distribution of the error $\mathcal{D}_0^N Y_T - \mathcal{D}_0 Y_T$ using the Euler scheme without transformation

The table reports the Mean Absolute Error, the Root Mean-Squared Error and the Probability that Absolute Error exceeds a threshold equal to 0.1%. In all cases the error is displayed as a function of the parameter N. The results show the higher convergence speed of estimates based on the transformation. For instance columns 2 and 4 of the table show that the speed of convergence of the Euler scheme (without transformation) is roughly of order $1/\sqrt{N}$. Columns 3 and 5 exhibit a convergence speed of about $1/N$ when the transformation is used. This numerical

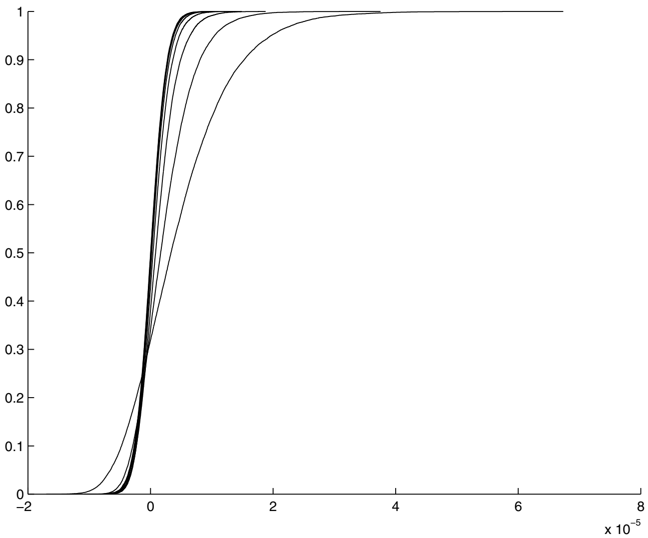

Fig. 4. Distribution of the error $\mathcal{D}_0^N Y_T - \mathcal{D}_0 Y_T$ using the Euler scheme with transformation

example illustrates the theory and confirms the increased convergence speed when the transformation is applied.

5 Conclusion

In this paper we developed representation formulas expressing Malliavin derivatives of diffusions in terms of solutions of ODEs. Characterizations of this type are useful for various purposes, in particular for numerical implementation. Computations based on our formulas can benefit from the faster convergence speed of numerical methods for ODEs. Numerical experiments conducted illustrate the theoretical convergence results for the new representation.

Efficiency is a central issue for simulation-based methods. Approximations of Malliavin derivatives have previously been calculated using a Euler scheme for SDEs. The resulting estimates converge at a slower rate than those based on the representation proposed in this paper. Moreover, the control of errors for numerical solvers of ODEs is well understood and software packages for higher order methods, like Runge-Kutta methods, are readily available.

Malliavin derivatives of first and higher order are increasingly used in financial applications such as risk management and asset allocation. Indeed, the use of Malliavin calculus permits the derivation of appealing formulas for the "Greeks" of options or the optimal portfolio shares in asset allocation problems. Implementation of these formulas often becomes a simple matter of applying Monte Carlo methods, even for large scale problems with nonlinear dynamics and significant number of state variables. Additional applications that have been developed in the recent literature include the representation of densities and conditional expectations as well as the numerical resolution of backward stochastic differential equations. In

Table 1. Error estimates for a CEV process and its Malliavin derivative. Estimates are reported for the method without (columns 2 and 4) and with transformation (columns 3 and 5). The top panel gives the Root Mean Square Error, the middle panel the Mean Absolute Error and the bottom panel the Absolute Error Probability (probability of an absolute error larger than 0.1% of initial value)

Error estimates for $N = 2^x$				
x	Y_T	$G(Z_T)$	$\mathcal{D}_0 Y_T$	$\mathcal{D}_0 G(Z_T)$
Root mean-squared error				
2	8.8117114e-04	6.1619880e-05	1.5111781e-04	9.3893084e-06
3	5.9067315e-04	3.2494544e-05	1.0282300e-04	5.0442477e-06
4	4.0364020e-04	1.9422254e-05	7.0938771e-05	3.1474082e-06
5	2.7945076e-04	1.4372872e-05	4.9363312e-05	2.4505407e-06
6	1.9644969e-04	1.2732531e-05	3.4797045e-05	2.2318379e-06
Mean absolute error				
2	6.1439138e-04	4.6034409e-05	1.0605937e-04	6.8902757e-06
3	4.3348976e-04	2.4345276e-05	7.5672567e-05	3.7306049e-06
4	3.0577428e-04	1.4723597e-05	5.3709545e-05	2.3736093e-06
5	2.1533142e-04	1.1110502e-05	3.7966469e-05	1.8906215e-06
6	1.5238578e-04	9.9678655e-06	2.6910408e-05	1.7414161e-06
Absolute error probability (0.1% *level)*				
2	8.9428000e-01	1.0412000e-01	9.2384000e-01	1.4972000e-01
3	8.4980000e-01	9.2000000e-03	8.9248000e-01	2.6160000e-02
4	7.9108000e-01	2.0000000e-04	8.4924000e-01	1.9200000e-03
5	7.0204000e-01	0.0000000e+00	7.8808000e-01	1.6000000e-04
6	5.9176000e-01	0.0000000e+00	6.9424000e-01	8.0000000e-05

all these applications accurate methods for the computation of Malliavin derivatives are essential.

References

Bouchard, B., Ekeland, I., Touzi, N.: On the Malliavin approach to Monte Carlo approximation of conditional expectations. Finance Stochast. **8**, 45–71 (2004)

Cvitanić, J., Ma, J., Zhang, J.: Efficient computation of hedging portfolios for options with discontinuous payoffs. Math. Finance **13**, 135–151 (2003)

Detemple, J.B., Garcia, R., Rindisbacher, M.: A Monte-Carlo method for optimal portfolios. J. Finance **58**, 401–446 (2003)

Detemple, J.B., Garcia, R., Rindisbacher, M.: Asymptotic properties for Monte Carlo estimators of diffusions. Working Paper CIRANO, 2004. J. Econ. (forthcoming)

Doss, H.: Liens entre équations différentielles stochastiques et ordinaires. Ann. l'Institut Henri Poincaré **13**, 99–125 (1977)

Fournié, E., Lasry, J-M., Lebuchoux, J., Lions, P-L., Touzi, N.: Applications of Malliavin calculus to Monte Carlo methods in finance. Finance Stochast. **3**, 391–412 (1999)

Fournié, E., Lasry, J-M., Lebuchoux, J., Lions, P-L.: Applications of Malliavin calculus to Monte-Carlo methods in finance, II. Finance Stochast. **5**, 201–236 (2001)

Gihman, I.I., Skorohod, A.V.: Stochastic differential equations. Berlin, Heidelberg, New York: Springer 1972

Hartman, P.: Ordinary differential equations. Boston: Birkhäuser 1982

Jacod, J., Protter, P.: Asymptotic error distributions for the Euler method for stochastic differential equations. Ann. Probab. **26**, 267–307 (1998)

Kurtz, T.G., Protter, P.: Wong-Zakai corrections, random evolutions and numerical schemes for SDE's (Stochastic Analysis). New York: Academic Press 1991, pp. 331–346

Kloeden, P., Platen, E.: Numerical solution of stochastic differential equations. Berlin, Heidelberg, New York: Springer 1997

Lamperti, J.: A simple construction of certain diffusion processes. J. Math. Kyoto University **4**, 161–170 (1964)

Magnus, J.R., Neudecker, H.: Matrix differential calculus with applications in statistics and econometrics. New York: Wiley 1988

Nualart, D.: The Malliavin calculus and related topics. Berlin, Heidelberg, New York: Springer 1995

Ocone, D., Karatzas, I.: A generalized Clark representation formula, with application to optimal portfolios. Stochast. Stochast. Reports **34**, 187–220 (1991)

Rootzen, H.: Limit distributions for the error in approximations of stochastic integrals. Ann. Probab. **8**, 241–251 (1980)