A Representation Theorem for Smooth Brownian Martingales

Sixian Jin[1], **Qidi Peng**[2], **Henry Schellhorn**[3]

Abstract: We show that, under certain smoothness conditions, a Brownian martingale, when evaluated at a fixed time, can be represented via an exponential formula. The time-dependent generator of this exponential operator only depends on the second order Malliavin derivative operator evaluated along a "frozen path". The exponential operator can be expanded explicitly to a series representation, which resembles the Dyson series of quantum mechanics. Our continuous-time martingale representation result can be proven independently by two different methods. In the first method, one constructs a time-evolution equation, by passage to the limit of a special case of a *backward Taylor expansion* of an approximating discrete time martingale. The exponential formula is a solution of the time-evolution equation, but we emphasize in our article that the time-evolution equation is a separate result of independent interest. In the second method, which we only highlight in this article, we use the property of denseness of exponential functions. We provide several applications of the exponential formula, and briefly highlight numerical applications of the backward Taylor expansion.
Keywords: Continuous martingales, Malliavin calculus.
MSC 2010: 60G15 ; 60H07

1 Introduction

And many more

The problem of representing Brownian martingales has a long and distinguished history. Dambis [3] and Dubins-Schwarz [5] showed that continuous martingales can be represented in terms of time-changed Brownian motions. Doob [6], Wiener and Itô developed what is often called Itô's martingale representation theorem: every local Brownian martingale has a version which can be written as an Itô integral plus a constant. In this article, we consider a special kind of martingales which are conditional expectations of a $\mathcal{F}_T$-measurable random variable F. Recall that, when the random variable F is Malliavin differentiable, the Clark-Ocone formula ([2, 14]) states that the integrand in Itô's martingale representation theorem is equal to the conditional expectation of the Malliavin derivative of F. We focus on a less general case, where the Brownian martingale is assumed to be "infinitely smooth". Namely, the target random variable F is infinitely differentiable in the sense of Malliavin. We show that such a Brownian martingale, when evaluated at time $t \leq T$, $E[F|\mathcal{F}_t]$, can be represented as an exponential operator of its value at the later time T.

While smoothness is a limitation to our result, our representation formula opens the way to new numerical schemes, and some analytical asymptotic calculations, because the exponential operator can be calculated explicitly in a series expansion, which resembles the

[1]Institute of Mathematical Sciences, Claremont Graduate University, Sixian.Jin@cgu.edu
[2]Institute of Mathematical Sciences, Claremont Graduate University, Qidi.Peng@cgu.edu
[3]Corresponding author, Institute of Mathematical Sciences, Claremont Graduate University, Henry.Schellhorn@cgu.edu

Dyson series of quantum mechanics. Although we still call our martingale's expansion Dyson series, there are two main differences between our martingale representation and the Dyson formula for the initial value problem in quantum mechanics. First, in the case of martingales, time flows backward. Secondly, the time-evolution operator is equal to one half of the second-order Malliavin derivative evaluated along a constant path, while for the initial value problem in quantum mechanics the time-evolution operator is equal to $-2\pi i$ times the time-dependent Hamiltonian divided by the Planck constant.

Our continuous-time martingale representation result can be proved using two different methods: by discrete time approximation and by approximation from a dense subset. In the first method, the key idea is to construct the *backward Taylor expansion* (BTE) of an approximating discrete-time martingale. The BTE was introduced in Schellhorn and Morris [17], and applied to price American options numerically. The idea in that paper was to use the BTE to approximate, over one time-step, the conditional expectation of the option value at the next time-step. While not ideal to price American options because of the lack of differentiability of the payoff, the BTE is better suited to the numerical calculation of the solution of smooth backward stochastic differential equations (BSDE). In a related paper, Hu et al. [7] introduce a numerical scheme to solve a BSDE with drift using Malliavin calculus. Their scheme can be viewed as a Taylor expansion carried out until the first order. Our BTE can be seen as a generalization to higher order of that idea, where the Malliavin derivatives are calculated at the future time-step rather than at the current time-step.

The time-evolution equation results then by a passage to the limit, when the time increment goes to zero, of the BTE, following the "frozen path". The exponential formula is then a solution of the time-evolution equation, under certain conditions. We stress the fact that both the BTE and the time-evolution equation are interesting results in their own right. Since the time-evolution equation is obtained from the BTE only along a particular path, we conjecture that there might be other types of equations that smooth Brownian martingales satisfy in continuous time. The time-evolution equation can also be seen as a more general result than the exponential formula, in the same way that the semi-group theory of partial differential equations does not subsume the theory of partial differential equations. For instance, other types of series expansion, like the Magnus expansion [16] can be used to solve a time-evolution equation.

We also sketch an alternate method, which we call the *density method* of proof of the exponential formula, which uses the denseness of stochastic exponentials in $L^2(\Omega)$. The complete proof [4] goes along the lines of the proof of the exponential formula for fractional Brownian motion (fBm) with Hurst parameter $H > 1/2$, which we present in a separate paper [10]. We emphasize that it is most likely nontrivial to obtain the exponential formula in the Brownian case by a simple passage to the limit of the exponential formula for fBm when H tends to $1/2$ from above. We mention three main differences between Brownian motion and fBm in our context. First, by the Markovian nature of Brownian motion, the backward Taylor expansion leads easily in the Brownian case to a numerical scheme. Second, there is a time-evolution equation in the Brownian case, but probably not in the fBm case, so that the BTE method of proof is unavailable. Third, the fractional conditional expectation (which is defined only for $H > 1/2$ in [1]) in general does not coincide with the conditional

[4] This proof is available from the authors, upon request.

expectation.

The structure of this paper is the following. We first expose the discrete time result, namely the Backward Taylor Expansion (BTE) for functionals of discrete Brownian sample path. We then move to continuous time, and present the time-evolution equation and exponential formula. We then sketch the density method of proof. Four explicit examples are given, which show the usefulness of the Dyson series in analytic calculations. Example 4 is about the Cox-Ingersoll-Ross model with time-varying parameters, which, as far as we know, is a new result. The proofs of the theorems are relegated to the appendix.

2 Martingale Representation

2.1 Preliminaries and notation

This section reviews some basic Malliavin calculus and introduces some definitions that are used in our article. We denote by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ a complete filtered probability space, where the filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions, i.e., it is the usual augmentation of the filtration generated by Brownian motion W on $\mathbb{R}$ (most results can be easily generalized to Brownian motion on $\mathbb{R}^d$). Unless stated otherwise all equations involving random variables are to be regarded to hold P-almost surely.

Following by [15], we say that a real function $g : [0, T] \to \mathbb{R}^n$ is symmetric if:

$$g(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = g(x_1, \dots, x_n),$$

for all permutations σ on $(1, 2, .., n)$. If in addition, $g \in L^2([0, T]^n)$, i.e.,

$$||g||^2_{L^2([0,T]^n)} = \int_0^T \dots \int_0^T g^2(x_1, \dots, x_n)\, \mathrm{d}x_1 \dots \mathrm{d}x_n < +\infty,$$

then we say g belongs to $\hat{L}^2([0, T]^n)$, the space of symmetric square-integrable functions on $[0, T]^n$. Denote by $L^2(\Omega)$ the space of random variables with finite variance, i.e., the norm of $F \in L^2(\Omega)$ is

$$\|F\|_{L^2(\Omega)} = \sqrt{E[F^2]} < +\infty.$$

The Wiener chaos expansion of $F \in L^2(\Omega)$, is defined by:

$$F = \sum_{m=0}^{\infty} I_m(f_m) \quad \text{in } L^2(\Omega), \tag{2.1}$$

where $\{f_m\}_{m=0}^{\infty}$ is a uniquely determined sequence of deterministic functions (m-dimensional kernels) in $\hat{L}^2([0, T]^n)$, and the operator $I_m : \ \hat{L}^2([0, T]^n) \to L^2(\Omega)$ is defined as

$$\begin{cases} I_m(f_m) &= m! \int_0^T \int_0^{t_m} ... \int_0^{t_2} f(t_1, ..., t_m)\, \mathrm{d}W(t_1)\, \mathrm{d}W(t_2) ... \,\mathrm{d}W(t_m), \quad \text{if } m > 0, \\ I_0(f_0) &= f_0. \end{cases}$$

For some stochastic process u, we use $\int_0^T u(s)\delta W(s)$ to denote its Skorohod integral, which is considered as the adjoint of the Malliavin derivative operator. It can be defined this way: for all $t \in [0, T]$, if the Wiener chaos expansion of $u(t)$ is:

$$u(t) = \sum_{m=0}^{\infty} I_m(f_m(., t)) \text{ in } L^2(\Omega),$$

where f_m is some $(m+1)$-dimensional kernel, then we define (when $L^2(\Omega)$-convergence holds):

$$\int_0^T u(s)\delta W(s) = \sum_{m=0}^{\infty} I_{m+1}(\tilde{f}_m) \text{ in } L^2(\Omega),$$

where $\tilde{f}_m$ denotes the symmetrization of the $(m+1)$-dimensional kernel f_m with respect to its $(m+1)$th argument (see Proposition 1.3.1 in [15] for details):

$$\tilde{f}_m(t_1, \ldots, t_m, t) = \frac{1}{m+1}\left(f_m(t_1, \ldots, t_m, t) + \sum_{i=1}^{m} f_m(t_1, \ldots, t_{i-1}, t, t_{i+1}, \ldots, t_m, t_i)\right).$$

Following Lemma 4.16 in [15], the Malliavin derivative $D_t F$ of F (when it exists) satisfies:

$$D_t F = m \sum_{m=0}^{\infty} I_{m-1}(f_m(\cdot, t)) \quad \text{in } L^2(\Omega).$$

We denote the Malliavin derivative of order l of F at time t by $D_t^l F$, as a shorthand notation for $\underbrace{D_t \ldots D_t}_{\imath \text{ times}} F$. We call $\mathbb{D}_\infty([0, T])$ the set of random variables which are infinitely Malliavin differentiable and $\mathcal{F}_T$-measurable. A random variable is said to be infinitely Malliavin differentiable if for any integer n:

$$E\left[\left(\sup_{s_1,\ldots,s_n \in [0,T]} \left|D_{s_n} \ldots D_{s_1} F\right|\right)^2\right] < +\infty. \tag{2.2}$$

In particular, we denote by $\mathbb{D}^N([0, T])$ the space of all random variables F which satisfy (2.2) for all $n \leq N$.

Given $\omega \in \Omega$, let the freezing operator ω^t [5] be defined as:

$$W(s, \omega^t(\omega)) = \begin{cases} W(s, \omega) \text{ if } s \leq t \\ W(t, \omega) \text{ if } t \leq s \leq T. \end{cases} \tag{2.3}$$

When Brownian motion is defined as the coordinate mapping process (see [9]), then $\omega^t(\omega)$ represents obviously a "frozen path" - a particular path where the corresponding Brownian motion becomes constant after time t. Thus the freezing operator ω^t is a mapping from Ω to Ω. We could view ω^t as a left-operator such that, for any $\mathcal{F}_T$-measurable random variable F:

$$(\omega^t \circ F)(\omega) = F(\omega^t(\omega)).$$

[5] The introduction of this operator ω^t on the path may seem awkward, but it facilitates the proof of Theorem 2.4.

This reminds us that, generally speaking,

$$(D_s F)(\omega^t(\omega)) \neq D_s \left(F(\omega^t(\omega)) \right).$$

For instance, if $t \leq T$, then $\frac{1}{2}(D_s W(T)^2)(\omega^t(\omega)) = \chi_{[s \leq T]} W(T, \omega^t(\omega)) = \chi_{[s \leq T]} W(t, \omega)$, with χ being the indicator function. In the remainder of text, whenever possible, we only write ω^t for convenience, instead of $\omega^t(\omega)$ and F instead of $F(\omega)$. We give in the remarks below some examples of applications of the freezing path operator, which can also be viewed as a constructive definition of the operator.

Remark 2.1. *We show hereafter the freezing path transformation of some random variables with particular forms. Let $t \leq T$:*

1. *For $p \in \mathcal{P}$, space of polynomials, suppose $F = p(W(s_1), \ldots, W(s_n))$, then $F(\omega^t) = p(W(s_1 \wedge t), \ldots, W(s_n \wedge t))$.*

2. *The following equations hold:*

$$\left(\int_0^T f(s) \, \mathrm{d}W(s) \right) (\omega^t) = \int_0^t f(s) \, \mathrm{d}W(s);$$
$$\left(\int_0^T W(s) \, \mathrm{d}s \right) (\omega^t) = \int_0^t W(s) \, \mathrm{d}s + W(t)(T - t).$$

   These results can be explained by representing the integrals as Riemann sums and using Theorem 2.2.

3. *For a general Itô integral, there is not yet a satisfactory or general approach to compute its frozen path so far. We have to first evaluate the integral and then apply the freezing operator to it. For example,*

$$\left(\int_0^T W(s) \, \mathrm{d}W(s) \right) (\omega^t) = \left(\frac{W(T)^2 - T}{2} \right) (\omega^t) = \frac{W(t)^2 - T}{2}.$$

2.2 Backward Taylor Expansion (BTE)

Through this subsection, as well as in its proofs, we assume that $F \in \mathbb{D}_\infty([0, T])$ is cylindrical. In other words, it can be written as $G(W(\Delta), W(2\Delta), ..., W(T))$ where $T = M\Delta$ and $G : \mathbb{R}^M \to \mathbb{R}$ is a deterministic infinitely differentiable function.

We now present the backward Taylor expansion of the Brownian martingale evaluated at time $m\Delta$. First we define h_n, the Hermite polynomial of degree $n \geq 0$, by $h_0 = 1$ and for $n \geq 1$,

$$h_n(x) = (-1)^n \exp\left(\frac{x^2}{2} \right) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \exp\left(-\frac{x^2}{2} \right) \quad \text{for all } x \in \mathbb{R}. \tag{2.4}$$

We denotes $\lfloor x \rfloor$ as the floor number of real-value x, which is defined as

$$\lfloor x \rfloor := \max\{m \in \mathbb{Z}; m \leq x\}. \tag{2.5}$$

Theorem 2.1. *If F satisfies, for each $m = 0, ..., M-1$, the following condition:*

$$\sum_{i=0}^{L} E\left[\left(D^{2L-i}_{(m+1)\Delta}F\right)^2\right]\binom{L}{i}^4 \frac{i!}{(L!)^2}\Delta^{2L-i} \xrightarrow[L\to+\infty]{} 0, \tag{2.6}$$

then for each m, the following series converges in $L^2(\Omega)$:

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{l=0}^{\infty}\gamma(m,l)E[D^l_{(m+1)\Delta}F|\mathcal{F}_{(m+1)\Delta}], \tag{2.7}$$

where $\gamma(m,0) = 1$ and otherwise,

$$\gamma(m,l) = (-1)^l\Delta^{l/2}\sum_{j=0}^{\lfloor l/2\rfloor}\frac{1}{(l-2j)!}h_{l-2j}\left(\frac{W_{(m+1)\Delta}-W_{m\Delta}}{\sqrt{\Delta}}\right). \tag{2.8}$$

Important Remark 1: A quite large range of random variables fits Condition (2.6). For instance, by applying Stirling's approximation to the factorials, we can show that:

$$\sum_{i=0}^{L}\binom{L}{i}^4\frac{i!}{(L!)^2}\Delta^{2L-i} \le \frac{\alpha^L}{L^L}$$

for some constant $\alpha > 0$, which does not depend on L. Thus Condition (2.6) is satisfied by F, if there is some constant $c > 0$ such that for any integers L and any $m = 0, 1, \ldots, M-1$,

$$E\left[\left(D^L_{(m+1)\Delta}F\right)^2\right] \le c^L. \tag{2.9}$$

A simple example that satisfies (2.9) is $F = e^{c'W_T}$, where c' can be any constant.

Important Remark 2: Clearly, the BTE (2.7) is a finite sum of L terms when all the Malliavin derivatives of F of order $l \ge L$ vanish.

We provide an intuitive example to show how to apply Theorem 2.1:

Example: Let $F = W((m+1)\Delta)^2$, then $E[F|\mathcal{F}_{m\Delta}] = W(m\Delta)^2 + \Delta$. From (2.8) we see that:

$$\begin{aligned}
\gamma(m,1) &= -W((m+1)\Delta) + W(m\Delta);\\
\gamma(m,2) &= \frac{\Delta}{2} + \frac{1}{2}(W((m+1)\Delta) - W(m\Delta))^2.
\end{aligned}$$

By Theorem 2.1 and some algebraic computations, we get:

$$\begin{aligned}
&E[W((m+1)\Delta)^2|\mathcal{F}_{m\Delta}]\\
&= \gamma(m,0)W((m+1)\Delta)^2 + \gamma(m,1)E[2W((m+1)\Delta)|\mathcal{F}_{(m+1)\Delta}] + 2\gamma(m,2)\\
&= W(m\Delta)^2 + \Delta.
\end{aligned}$$

Note that (2.7) is a one-step backward time equation. A complete expression of $E[F|\mathcal{F}_{m\Delta}]$ can be derived by applying (2.7) recursively:

Corollary 2.1. *Let F satisfy the assumption (2.6), then we have*

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{j_{m+1}=0}^{\infty} \dots \sum_{j_M=0}^{\infty} \prod_{k=m+1}^{M} \gamma(k-1, j_k) D_{(m+1)\Delta}^{j_{m+1}} \dots D_{M\Delta}^{j_M} F. \tag{2.10}$$

A non-intuitive feature of the BTE is that *any path* from t to T can be chosen to approximate conditional expectations evaluated at time t, as opposed to Monte Carlo simulation where many paths are needed. In the next subsection we will choose the frozen paths:

$$W((m+1)\Delta, \omega^t(\omega)) = W(m\Delta, \omega^t(\omega)).$$

for $t \leq m\Delta$ to derive our main result. We proceed now to discuss two applications of the BTE. Since space is limited, we describe these informally.

Application to solving FBSDEs

Consider a forward backward stochastic differential equation(FBSDE), where the problem is to find a triplet of adapted processes $(X(t), Y(t), Z(t))$ such that:

$$\begin{cases} \mathrm{d}X(t) &= b(t, X(t))\,\mathrm{d}t + \sigma(t, X(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X(t), Y(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x \\ Y(T) &= F(X), \end{cases}$$

where F is a function of the path $X(., \omega)$ from 0 to T. This problem is at the same time more general (because of the path-dependency of F) and less general than a standard FBSDE, because the coefficients of the diffusion X do not depend on Y or Z. We define:

$$U(t) = Y(t) - \int_0^t h(s, X(s), Y(s))\,\mathrm{d}s. \tag{2.11}$$

We select a path ω along which we simulate (an approximation of) $X(t, \omega)$. Then, along the same path, we use (2.10) to compute:

$$U(t, \omega) = E[F(X)|\mathcal{F}_{m\Delta}](\omega)$$

and then compute $Y(t, \omega)$ by discretizing the integral (2.11). The main numerical difficulty is to evaluate the Malliavin derivatives. One can apply the change of variables defined in [4] to calculate the Malliavin derivatives of X, and then (mutatis mutandis) the Faa di Bruno formula for the composition of F with X. One must of course calculate finite sums instead of infinite ones in (2.10). One could then imagine a scheme where, at each step, the "optimal path" is chosen so as to minimize the global truncation error. We leave all these considerations for future research.

Application to Pricing Bermudan Options

Casual observation of (2.10) shows that the number of calculations grows exponentially with time. This shortcoming of the BTE does not occur in the problem of Monte Carlo pricing Bermudan options, where the BTE can be competitive as we hint now. Most of the computational burden in Bermudan option pricing consists in evaluating:

$$C(mk\Delta) = E[\max(C((m+1)k\Delta), h((m+1)k\Delta))|\mathcal{F}_{m\Delta}], \tag{2.12}$$

where $C(m\Delta)$ and $h(m\Delta)$ are respectively the continuation value and the exercise value of the option, and k is the number of time units between each exercise. The data in this problem consists in the exercise value at all times and the continuation value $C(M\Delta)$ at expiration. The conditional expectation $C(m\Delta)$ must be evaluated at all times $mk\Delta$, with $0 \leq mk < M$ and along every scenario. In regression-based algorithms, such as [11] and [20], the continuation value is regressed at each time $mk\Delta$ on a basis of functions of the state variables, so that $C(mk\Delta)$ can be expressed as a formula. The formula is generally a polynomial function of the previous values of the state variable. This important fact, that $C(mk\Delta)$ is available formulaically rather than numerically, makes possible the use of symbolic Malliavin differentiation.[6] Since the formula is a polynomial, there is no truncation error in (2.10) if the state variables are Brownian motion.

2.3 The Time-Evolution Equation

For notational simplicity we define:

$$M^F(t) = E[F|\mathcal{F}_t].$$

Define the set of random variables $\mathcal{M}(t) := \left\{M^F(t) : F \in L^2(\Omega) \text{ and } \mathcal{F}_T\text{-measurable}\right\}$. Let the time-evolution operator P_s be the conditional expectation operator, restricted to the following subspaces. For any time $s \in [0, T]$:

$$\begin{aligned} P_s \ : \ \bigcup_{\tau \in [s,T]} \mathcal{M}(\tau) \ &\longrightarrow \ \mathcal{M}(s) \\ M^F(\tau) \ &\longmapsto \ M^F(s). \end{aligned} \tag{2.13}$$

We also define the time-derivative of the time-evolution operator as: for $0 < s \leq \tau \leq T$, provided the limit exists in $L^2(\Omega)$:

$$\frac{\mathrm{d}P_s}{\mathrm{d}s}(M^F(\tau)) = \lim_{h \downarrow 0} \frac{P_s(M^F(\tau)) - P_{s-h}(M^F(\tau))}{h} \text{ in } L^2(\Omega). \tag{2.14}$$

Before we state the time-evolution equation, we first give an important property of continuity of the freezing path operator ω^t defined in (2.3), which plays a key role in the proof of Theorem 2.3 below.

Theorem 2.2. *Let $F^{(M)}$ and F belong to $L^2(\Omega)$. If $F^{(M)} \to F$ as $M \to \infty$ in $L^2(\Omega)$, then for any $t \geq 0$,*

$$F^{(M)}(\omega^t) \xrightarrow[M\to\infty]{L^2(\Omega)} F(\omega^t).$$

[6] One remaining problem is that, in (2.12), the maximum of C and h is not Malliavin differentiable. We will show in another article how one can use the conditional expectation $E[C((m+1)k\Delta|\mathcal{F}_{m\Delta}]$ as a control variate, where the control variate is calculated using the BTE.

Theorem 2.3. *Let $t < s$. Suppose $F \in \mathbb{D}^6([0,T])$, then the operator P_s satisfies the following equation, whenever the right hand-side is an element in $L^2(\Omega)$:*

$$\left(\frac{\mathrm{d}P_sF}{\mathrm{d}s}\right)(\omega^t) = -\frac{1}{2}\left(D_s^2 P_s F\right)(\omega^t). \tag{2.15}$$

Note that this equality holds for each $s \in (t,T]$ in $L^2(\Omega)$.

We can see the analogy between our time-evolution operator P_s and the one in quantum mechanics. The difference is that in quantum mechanics $-(1/2)D^2$ is replaced by the Hamiltonian divided by $-i\hbar$. While the next theorem will provide a Dyson series solution to Equation (2.15). We remind the reader that other series expansions are sometimes more useful in mathematical physics. For instance, we believe that the Magnus expansion [16] is an easier approach to solve for the bond price in the Cox-Ingersoll-Ross model (and thus, the Riccati equation) than the Dyson series (see Section 3).

2.4 Exponential Formula

For esthetical reasons we introduce a "chronological operator". In this we follow Zeidler [21]. Let $\{H(t)\}_t$ be a collection of operators. The chronological operator $\mathcal{T}$ is defined by:

$$\mathcal{T}(H(t_1)H(t_2)\ldots H(t_n)) := H(t_{1'})H(t_{2'})\ldots H(t_{n'}),$$

where $t_{1'},\ldots,t_{n'}$ is a permutation of $t_1,\ldots,t_n$ such that $t_{1'} \geq t_{2'} \geq \ldots \geq t_{n'}$.

For example, it is showed in Zeidler [21] on Page 44-45 that:

$$\int_0^t \int_0^{t_2} H(t_1)H(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 = \frac{1}{2!}\int_0^t \int_0^t \mathcal{T}(H(t_1)H(t_2))\,\mathrm{d}t_1\,\mathrm{d}t_2.$$

This will be the only property of the chronological operator we will use in this article.

Definition 2.1. *The exponential operator of a time-dependent generator H is:*

$$\mathcal{T}\exp\left(\int_t^T H(s)\,\mathrm{d}s\right) = \sum_{k=0}^{\infty}\frac{1}{k!}\int_t^T \cdots \int_t^T \mathcal{T}(H(\tau_1),\ldots,H(\tau_k))\,\mathrm{d}\tau_1\ldots\mathrm{d}\tau_k. \tag{2.16}$$

In quantum field theory, the series on the right hand-side of (2.16) is called a *Dyson series* (see for instance [21]).

Theorem 2.4. *Suppose $F \in \mathbb{D}_\infty([0,T])$ satisfies the following condition:*

$$\frac{(T-t)^{2n}}{(2^n n!)^2} E\left[\left(\sup_{u_1,\ldots,u_n\in[t,T]}\left|(D_{u_n}^2 \ldots D_{u_1}^2 F)(\omega^t)\right|\right)^2\right] \xrightarrow[n\to\infty]{} 0 \tag{2.17}$$

for fixed $t \in [0,T]$. Then

$$E[F|\mathcal{F}_t] = \left(\mathcal{T}\exp\left(\frac{1}{2}\int_t^T D_s^2\,\mathrm{d}s\right)F\right)(\omega^t) \text{ in } L^2(\Omega). \tag{2.18}$$

Example: Let $F = W(T)^2$, then for $t \le s \le T$:

$$\begin{aligned} F(\omega^t) &= W(t)^2; \\ (D_s^2 F)(\omega^t) &= 2. \end{aligned}$$

It follows from Theorem 2.4 and Theorem 2.2 that

$$E[F|\mathcal{F}_t] = F(\omega^t) + \frac{1}{2}\int_t^T (D_s^2 F)(\omega^t)\,\mathrm{d}s = W(t)^2 + T - t.$$

The importance of the exponential formula (2.18) stems from the Dyson series representation (2.16), which we rewrite hereafter in a more convenient way:

$$\begin{aligned} E\left[F|\mathcal{F}_t\right] &= F(\omega^t) + \frac{1}{2}\int_t^T D_s^2 F(\omega^t)\,\mathrm{d}s + \frac{1}{4}\int_t^T \int_{s_1}^T D_{s_1}^2 D_{s_2}^2 F(\omega^t)\,\mathrm{d}s_2\,\mathrm{d}s_1 + \ldots \\ &= \sum_{i=0}^{+\infty} \frac{1}{2^i} \int_{t \le s_1 \le \ldots \le s_i \le T} (D_{s_i}^2 \ldots D_{s_1}^2 F)(\omega^t)\,\mathrm{d}s_i \ldots \mathrm{d}s_1 \\ &= \sum_{i=0}^{+\infty} \frac{1}{2^i i!} \int_{[t,T]^i} (D_{s_i}^2 \ldots D_{s_1}^2 F)(\omega^t)\,\mathrm{d}s_i \ldots \mathrm{d}s_1. \end{aligned} \tag{2.19}$$

We will use this formula for analytical calculations, as we show in the next subsection. The analytical calculations become quickly nontrivial, though, and, for numerical applications, one may want to develop an automatic tool that performs symbolic Malliavin differentiation (see the earlier discussion, on the implementation of the BTE).

Remark: One may wonder how to evaluate the condition 2.17. By closability of the Malliavin derivative and Theorem 2.2, condition 2.17 holds if the following condition holds:

$$\frac{(T-t)^{2n}}{(2^n n!)^2} E\left[\sup_{u_1,\ldots,u_n \in [t,T]} \left|(D_{u_n}^2 \ldots D_{u_1}^2 F)\right|^2\right] \xrightarrow[n\to\infty]{} 0. \tag{2.20}$$

Important Remark: As mentioned in the introduction, there is another way of proving the exponential formula, by the so-called density method (the denseness of the exponential functions). Here we just sketch out the idea. Let f be a deterministic function, and define the exponential function of f as

$$\varepsilon(f) = \exp\left(\int_0^T f(s)\,\mathrm{d}W(s)\right).$$

We can see that $\varepsilon(f)$ has an exponential formula representation. On the one hand, since $\int_0^T f(s)\,\mathrm{d}W(s)$ is Gaussian, the left hand-side of (2.19) is

$$E[\varepsilon(f)|\mathcal{F}_t] = \exp\left(\int_0^t f(s)\,\mathrm{d}W(s) + \frac{1}{2}\int_t^T f(s)^2\,\mathrm{d}s\right).$$

On the other hand, we have:

$$\begin{aligned}\varepsilon(f)(\omega^t) &= \exp\left(\int_0^t f(s)\,\mathrm{d}W(s)\right);\\ D_s^2\varepsilon(f) &= (f(s))^2\varepsilon(f) \text{ for } s\in[0,T].\end{aligned}$$

Thus on the right hand-side of (2.19),

$$\begin{aligned}\frac{1}{2^i i!}\int_{[t,T]^i}(D_{s_i}^2\dots D_{s_1}^2\varepsilon(f))(\omega^t)\,\mathrm{d}s_i\dots\mathrm{d}s_1 &= \varepsilon(f)(\omega^t)\frac{1}{2^i i!}\int_{[t,T]^i}(f(s_i)\dots f(s_1))^2\,\mathrm{d}s_i\dots\mathrm{d}s_1\\ &= \exp\left(\int_0^t f(s)\,\mathrm{d}W(s)\right)\frac{1}{i!}\left(\frac{1}{2}\int_t^T f(s)^2\,\mathrm{d}s\right)^i.\end{aligned}$$

And finally (2.19) holds for $F=\varepsilon(f)$:

$$\begin{aligned}\left(\exp\left(\frac{1}{2}\int_t^T D_s^2\right)\varepsilon(f)\right)(\omega^t) &= \exp\left(\int_0^t f(s)\,\mathrm{d}W(s)\right)\sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{1}{2}\int_t^T f(s)^2\,\mathrm{d}s\right)^i\\ &= \exp\left(\int_0^t f(s)\,\mathrm{d}W(s)+\frac{1}{2}\int_t^T f(s)^2\,\mathrm{d}s\right)\\ &= E[\varepsilon(f)|\mathcal{F}_t].\end{aligned}$$

For general $F\in L^2(\Omega)$, the proof of (2.19) can be completed by using the fact that the linear span of the exponential functions is dense in $L^2(\Omega)$.

3 Solution of Some Problems by Dyson Series

We provide four different examples. The first example is a very well-known example, but it illustrates nicely the computation of Dyson series in case the random variable F (seen as a functional of Brownian motion) is not path-dependent. In the second example, the functional F is path-dependent. The third example is path-dependent again and illustrates the problem of convergence of the Dyson series. The fourth example shows a new representation of the price of a discount bond in the Cox-Ingersoll-Ross model; see [18] for a discussion of the problem.

3.1 The Heat Kernel

We consider the random variable

$$F=f(W(T),T),$$

where f is some heat kernel satisfying the following differential equation

$$\frac{\partial^{2n}f}{\partial x^{2n}}=(-2)^n\frac{\partial^n f}{\partial T^n}.$$

For example, for $\tau > T$,

$$f_\tau: \ (x,T) \mapsto \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{x^2}{2(\tau - T)}\right)$$

is such a heat kernel.

For $t \le s_1, \ldots, s_i \le T$,

$$(D^2_{s_i} \ldots D^2_{s_1} f(W(T),T))(\omega^t) = \left.\frac{\partial^{2i} f(x,T)}{\partial x^{2i}}\right|_{x=W(T)} (\omega^t) = (-2)^i \frac{\partial^i f(W(T),T)}{\partial T^i}(\omega^t) = (-2)^i \frac{\partial^i f(W(t),T)}{\partial T^i}.$$

It follows that

$$\frac{1}{2^i i!} \int_{[t,T]^i} (D^2_{s_i} \ldots D^2_{s_1} f(W(T),T))(\omega^t)\, \mathrm{d}s_i \ldots \mathrm{d}s_1 = \frac{(t-T)^i}{i!} \frac{\partial^i f(W(t),T)}{\partial T^i}.$$

Then by (2.19), the Dyson series expansion of $E\left[F|\mathcal{F}_t\right]$ $(t \le T)$ coincides with the Taylor expansion of $t \mapsto f(x,t)$ around $t = T$, evaluated at $x = W(t)$:

$$\begin{aligned} E\left[F|\mathcal{F}_t\right] &= \sum_{i=0}^{\infty} \frac{(t-T)^i}{i!} \frac{\partial^i f(W(t),T)}{\partial T^i} \\ &= f(W(t),t). \end{aligned}$$

As a particular case, when taking

$$F = \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W(T)^2}{2(\tau - T)}\right),$$

one has

$$E\left[F|\mathcal{F}_t\right] = \frac{1}{\sqrt{\tau - t}} \exp\left(-\frac{W^2(t)}{2(\tau - t)}\right).$$

Observation: We deliberately took $\tau > T$ so that the functional F would be infinitely Malliavin differentiable. It remains to be seen whether proper convergence results can be obtained when $\tau \downarrow T$.

This example is not new, in the sense that it could have been obtained by applying $\exp\left(\frac{1}{2}(T-t)\frac{\partial^2}{\partial x^2}\right)$, i.e., the time-evolution operator of the heat equation, to the function $f\left(x,T\right) = \frac{1}{\sqrt{\tau - T}} \exp(-\frac{x^2}{2(\tau - T)})$ (see [8] page 162). This example hints to the fact that our time-evolution operator is a generalization of the time-evolution operators for the heat kernel, the latter being applicable to path-independent problems, and the former being applicable to path-dependent problems.

3.2 The Merton Interest Rate Model

Let $F = \exp(-\int_0^T W(u)\, \mathrm{d}u)$. By Itô's formula we have:

$$\int_0^T W(u)\, \mathrm{d}u = \int_0^T (T-u)\, \mathrm{d}W(u).$$

We observe that, for $t \leq s_1, \ldots, s_i \leq T$,

$$D^2_{s_i} \ldots D^2_{s_1} F = (T - s_i)^2 \ldots (T - s_1)^2 F$$

and

$$F(\omega^t) = \exp\left(- \int_0^t W(u)\,\mathrm{d}u - W(t)(T-t) \right).$$

Therefore by (2.19), the Dyson series expansion of $E\left[F|\mathcal{F}_t\right]$ is:

$$\begin{aligned} E\left[F|\mathcal{F}_t\right] &= F(\omega^t) \sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{1}{2} \int_t^T (T-s)^2\,\mathrm{d}s \right)^i \\ &= \exp\left(- \int_0^t W(u)\,\mathrm{d}u - W(t)(T-t) + \frac{1}{6}(T-t)^3 \right). \end{aligned}$$

3.3 Moment Generating Function of Geometric Brownian Motion

Let

$$X(T) = e^{M - \sigma W(T)} \text{ and } F = e^{-X(T)}.$$

Then $E[F]$ is the moment generating function of the lognormal random variable $X(T)$ evaluated at the parameter $z = -1$. Now we are going to compute $E[F|\mathcal{F}_t]$ for $t \leq T$. We have, for all $t \leq s_1, ..., s_n < T$:

$$\begin{aligned} D_{s_1} F &= F \sigma e^{M - \sigma W(T)}, \\ D_{s_1} D_{s_2} F &= F \sigma^2 (e^{2M - 2\sigma W(T)} - e^{M - \sigma W(T)}). \end{aligned}$$

By induction, we obtain:

$$D^2_{s_n} \ldots D^2_{s_1} F = F \sigma^{2n} \sum_{i=1}^{2n} (-1)^i \left\{ {2n \atop i} \right\} e^{i(M - \sigma W(T))},$$

where $\left\{ {j \atop i} \right\}$ denotes the Stirling number of the second kind, with convention $\left\{ {0 \atop 0} \right\} = 1$, and $\left\{ {j \atop 0} \right\} = 0$ for any $j \neq 0$.

So by the exponential formula, for $t \leq T$,

$$\frac{1}{2^n n!} \int_{[t,T]^n} (D^2_{s_n} \ldots D^2_{s_1} F)(\omega^t)\,\mathrm{d}s_n \ldots \mathrm{d}s_1 = e^{-e^{M - \sigma W(t)}} \frac{(T-t)^n \sigma^{2n}}{2^n n!} \sum_{i=0}^{2n} (-1)^i \left\{ {2n \atop i} \right\} e^{i(M - \sigma W(t))},$$

and

$$E[F|\mathcal{F}_t] = e^{-e^{M - \sigma W(t)}} \sum_{n=0}^{\infty} \sum_{i=0}^{2n} \frac{(\frac{\sigma^2 (T-t)}{2})^n (-1)^i}{n!} \left\{ {2n \atop i} \right\} e^{i(M - \sigma W(t))}. \tag{3.1}$$

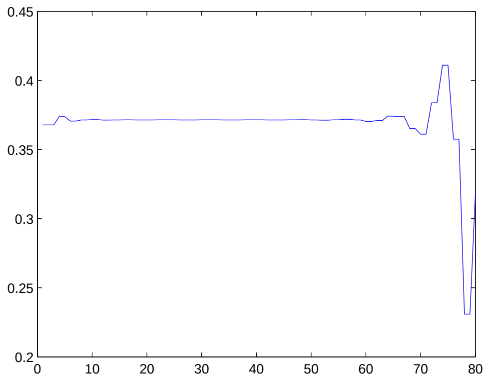

Figure 1: The first 80 terms of Dyson series (3.1)

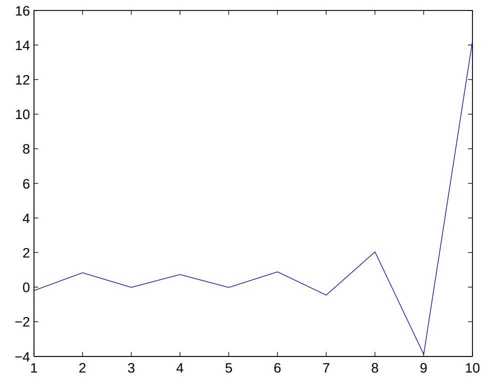

Figure 2: The first 10 terms of Taylor series (3.2)

This example shows that, although the MGF of GBM is well-defined, the Dyson series does not converge. Also it is indeed well-known in [19] that the Taylor series of the MGF of GBM, which is:

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} e^{nM + \frac{n^2\sigma^2 T}{2}} \tag{3.2}$$

does not converge either. An interesting phenomenon arises, the analysis of which is probably quite difficult[7]. While both series (3.1) and (3.2) are divergent, the Taylor series (3.2) diverges much earlier and has higher deviation than the Dyson series. We show in Fig 1,2 and Table 1 the results of our numerical experiments. We calculate (3.1) and (3.2) as a function of the number of terms in the series and plug in the following parameters: $t = 0$, $T = 1$, and $\sigma = 0.6$. From Fig 1 and 2, we observe that the first terms of series (3.1) are more stable than the first terms of series (3.2). Also in Table 1, we list the first eleven terms of the two series, which we compare to a straightforward calculation of $E[F]$ by numerical integration with sample size from 2^{10} to 2^{20}, reported in the third column: every subsequent row corresponds to a finer grid. We also observed in other experiments that the Dyson series has a better performance than the Taylor series (3.2) when approximating the expectation of F as σ grows.

Also it is interesting to see that (3.1) resembles the MGF for a Poisson random variable, which can be expressed as:

$$E[\exp(zN)] = \sum_{n=0}^{\infty} \sum_{i=0}^{n} \frac{z^n}{n!} \left\{ \begin{matrix} n \\ i \end{matrix} \right\} \lambda^i,$$

where N is a Poisson random variable with mean parameter λ.

[7] According to Wikipedia, the question of convergence of the moment generating function of the lognormal distribution has been referred to as the "Matterhorn" of probability theory.

Computation of (3.1)	Computation of (3.2)	Approximation of $E[e^{-e^{-0.6W(1)}}]$
0.3679	1	0.3696
0.3679	-0.1972	0.3727
0.3679	0.8300	0.3724
0.3679	-0.0122	0.3719
0.3738	0.7301	0.3717
0.3706	-0.0201	0.3718
0.3706	0.8855	0.3716
0.3714	-0.4575	0.3718
0.3714	2.0403	0.3717
0.3717	-3.8787	0.3717

Table 1: Comparison of Taylor expansion and Dyson series

3.4 Bond Pricing in the Extended Cox-Ingersoll-Ross Model

Assume that the interest rate is given by:

$$\mathrm{d}r(s) = \left(-2br(s) + d\sigma(s)^2\right)\,\mathrm{d}s + 2\sigma(s)\sqrt{r(s)}\,\mathrm{d}W(s) \tag{3.3}$$

where $r(0) = r_0$, $\sigma(s)$ is a deterministic function, b is a real number and d is a positive integer.

Let $\{W_i(s) : i = 1, .., d\}$ be a sequence of independent Brownian motions. Each W_i generates a filtration $\{\mathcal{F}_t^i\}$. By Itô's lemma and Levy's theorem (see, e.g. [18]) the interest rate $r(t)$ can be written as follows:

$$r(s) = \sum_{i=1}^{d} X_i^2(s) \tag{3.4}$$

where $X_i(t)$ is the Ornstein-Uhlenbeck process defined by:

$$\mathrm{d}X_i(s) = -bX_i(s)\,\mathrm{d}s + \sigma(s)\,\mathrm{d}W_i(s)$$

with $X_i(0) = \sqrt{\frac{r_0}{d}}$ for all i. Let

$$F = \exp\left(-\int_t^T r(s)\,\mathrm{d}s\right).$$

Our goal is to find the bond price $E[F|\mathcal{F}_t]$. Since F can be written as $F = \prod_{i=1}^d F_i$ where $F_i := \exp\left(-\int_t^T X_i(s)^2\,\mathrm{d}s\right)$. we can compute each conditional expectation separately:

$$E[F|\mathcal{F}_t] = \prod_{i=1}^{d} E[F_i|\mathcal{F}_t^i]. \tag{3.5}$$

We consider separately the case $b = 0$ and the case $b > 0$. When $b = 0$:

$$X_i(s) = X_i(0) + \int_0^s \sigma(u)\,\mathrm{d}W_i(u). \tag{3.6}$$

When $s \in (t,T), u \geq s$ for any i, we obtain:

$$\begin{array}{rcl} X_i(s)(\omega^t) & = & X_i(t); \\ (D_s X_i(u))(\omega^t) & = & 2\sigma(s)\sqrt{X_i(t)}; \\ (D_s^2 X_i(u))(\omega^t) & = & 2\sigma^2(s). \end{array}$$

By (3.5) and Theorem 2.4, the first terms of the Dyson series are explicitly given:

$$\begin{aligned} E[F|\mathcal{F}_t] = e^{-(T-t)r(t)} \prod_{i=1}^{d} \Bigg\{ & 1 + \frac{1}{2} \int_t^T \sigma^2(s_1)[4(T-s_1)^2 X_i(t)^2 - 2(T-s_1)] \, \mathrm{d}s_1 \\ & + \frac{1}{2^2} \int_t^T \int_{s_1}^T \sigma^2(s_1)\sigma^2(s_2) \times \Big\{ 16(T-s_1)^2(T-s_2)^2 X_i(t)^4 \\ & - \left(8(T-s_1)^2(T-s_2) + 40(T-s_1)(T-s_2)^2\right) X_i(t)^2 + 8(T-s_2)^2 \\ & + 4(T-s_1)(T-s_2) \Big\} \, \mathrm{d}s_1 \, \mathrm{d}s_2 + \ldots \Bigg\}. \end{aligned} \tag{3.7}$$

We respectively denote by A_0, A_1, A_2 the deterministic coefficients of $1, X_i(t)^2, X_i(t)^4$ in (3.7). Then we see that

$$\begin{aligned} A_0 &= 1 - \int_t^T (T-s_1)\sigma^2(s_1) \, \mathrm{d}s_1 + \int_t^T \int_{s_1}^T \{2(T-s_2)^2 + (T-s_1)(T-s_2)\} \sigma^2(s_1)\sigma^2(s_2) \, \mathrm{d}s_2 \, \mathrm{d}s_1 + \ldots; \\ A_1 &= \int_t^T 2(T-s_1)^2 \sigma^2(s_1) \, \mathrm{d}s_1 \\ &\quad + \int_t^T \int_{s_1}^T (10(T-s_1)(T-s_2)^2 - 2(T-s_1)^2(T-s_2)) \sigma^2(s_1)\sigma^2(s_2) \, \mathrm{d}s_2 \, \mathrm{d}s_1 + \ldots; \\ A_2 &= \int_t^T \int_{s_1}^T 4(T-s_1)^2(T-s_2)^2 \sigma^2(s_1)\sigma^2(s_2) \, \mathrm{d}s_2 \, \mathrm{d}s_1. \end{aligned}$$

Then according to the relation (3.4) between $r(t)$ and $X_i(t)$, the representation formula can be probably reorganized into:

$$E[F|\mathcal{F}_t] = e^{-(T-t)r(t)}(A_0^d + A_0^{d-1} A_1 r(t) + A_0^{d-1} A_2 r(t)^2 + \ldots). \tag{3.8}$$

Remark: To get (3.8), the relation $A_0 A_2 = \frac{1}{2} A_1^2$ has to hold. However, we can not yet mathematically prove it because the total structures of all coefficients A_i, $i \geq 1$ remain unknown or too complicated to be known. An alternative testing method is to check this result with several first terms which we have computed out. To this end, we recall that, in the particular case $\sigma \equiv 1$, there is an existing analytical formula (see [18] again) which agrees with ours:

$$E[F|\mathcal{F}_t] = \left(\operatorname{sech}\left(\sqrt{2}\tau\right)\right)^{\frac{d}{2}} e^{\frac{r(t)}{\sqrt{2}} \tanh \sqrt{2}\tau}, \tag{3.9}$$

where $\tau = T - t$. Using Taylor expansion, (3.9) becomes

$$E[F|\mathcal{F}_t] = e^{-(T-t)r(t)} \times \left\{ \left(1 - \frac{1}{2}\tau^2 + \frac{7}{24}\tau^4 + \dots\right) + r(t)\left(\frac{2}{3}\tau^3 - \frac{8}{15}\tau^5 + \dots\right) + r(t)^2\left(\frac{2}{9}\tau^6 + \dots\right) + \dots \right\}. \tag{3.10}$$

By taking $\sigma \equiv 1$ in (3.8), we see that the terms we computed coincide with those in (3.10).

We now consider the case $b > 0$. Then for all $1 \le i \le d$:

$$X_i(s) = X_i(0)e^{-bs} + e^{-bs}\int_0^s e^{br}\sigma(r)\,\mathrm{d}W_i(r).$$

We apply Theorem 2.3 separately to each F_i in (3.5) as follows:

$$D_u^2 F_i = F_i \left(D_u \int_t^T X_i(s)^2\,\mathrm{d}s\right)^2 - F_i D_u^2 \int_t^T X_i(s)^2\,\mathrm{d}s$$

in which each term can be computed out as:

$$\begin{aligned}
& D_u \int_t^T X_i(s)^2\,\mathrm{d}s \\
&= 2X_i(0)e^{bu}\sigma(u)\int_u^T e^{-2bs}\,\mathrm{d}s + 2e^{bu}\sigma(u)\int_u^T \left(e^{-2bs}\int_0^s e^{br}\sigma(r)\,\mathrm{d}W_i(r)\right)\mathrm{d}s \\
&= X_i(0)\frac{e^{-bu} - e^{-b(2T-u)}}{b}\sigma(u) + 2e^{bu}\sigma(u)\int_u^T \left(e^{-2bs}\int_0^s e^{br}\sigma(r)\,\mathrm{d}W_i(r)\right)\mathrm{d}s
\end{aligned}$$

and

$$D_u^2 \int_t^T X_i(s)^2\,\mathrm{d}s = 2e^{2bu}\sigma(u)^2 \int_u^T e^{-2bs}\,\mathrm{d}s = \frac{1 - e^{-2b(T-u)}}{b}\sigma(u)^2.$$

Then by acting on the freezing path operator ω^t,

$$F_i(\omega^t) = \exp\left(\frac{e^{-2b(T-t)} - 1}{2b} X_i(t)^2\right)$$

as well as

$$\begin{aligned}
\left(D_u \int_t^T X_i(s)^2\,\mathrm{d}s\right)(\omega^t) &= \frac{e^{-bu} - e^{-b(2T-u)}}{b}\sigma(u)\left(X_i(0) + \int_0^t e^{br}\sigma(r)\,\mathrm{d}W_i(r)\right); \\
\left(D_u^2 \int_t^T X_i(s)^2\,\mathrm{d}s\right)(\omega^t) &= \frac{1 - e^{-2b(T-u)}}{b}\sigma(u)^2.
\end{aligned}$$

So we obtain the first two terms in the exponential formula:

$$E[F_i|\mathcal{F}_t^{(i)}] = \exp\left(\frac{e^{-2b(T-t)} - 1}{2b} X_i(t)^2\right) \times \left(1 + \frac{1}{2}\int_t^T \left(X_i(t)^2\left(\frac{e^{-b(u-t)} - e^{-b(2T-u-t)}}{b}\right)^2 - \frac{1 - e^{-2b(T-u)}}{b}\right)\sigma(u)^2\,\mathrm{d}u + \dots\right)$$

By (3.5), the first terms of the Dyson series are:

$$
\begin{aligned}
E[F|\mathcal{F}_t] &= \exp\left(\frac{e^{-2b(T-t)}-1}{2b}r(t)\right) \\
&\times \prod_{i=1}^{d}\left(1+\frac{1}{2}\int_t^T \left(X_i(t)^2\left(\frac{e^{-b(u-t)}-e^{-b(2T-u-t)}}{b}\right)^2 - \frac{1-e^{-2b(T-u)}}{b}\right)\sigma(u)^2\,\mathrm{d}u+\ldots\right).
\end{aligned}
$$

By similar method as in writing (3.8) this formula can be also rearranged into:

$$
E[F|\mathcal{F}_t] = \exp\left(\frac{e^{-2b(T-t)}-1}{2b}r(t)\right)\left(A_0(b)^d + A_0(b)^{d-1}A_1(b)r(t)+\ldots\right) \tag{3.11}
$$

where

$$
\begin{aligned}
A_0(b) &= 1-\frac{1}{2}\int_t^T \frac{1-e^{-2b(T-u)}}{b}\sigma^2(u)\,\mathrm{d}u+\ldots; \\
A_1(b) &= \frac{1}{2}\int_t^T \left(\frac{e^{-b(u-t)}-e^{-b(2T-u-t)}}{b}\right)^2 \sigma^2(u)\,\mathrm{d}u+\ldots.
\end{aligned}
$$

It is well-known (see [18] again) that the bond price is still affine in $r(t)$ and satisfies:

$$
E[F|\mathcal{F}_t] = \exp(-C(t,T)r(t) - A(t,T)) \tag{3.12}
$$

where $C(t,T)$ solves the time dependent Riccati equation (3.13):

$$
C'(t,T) = 2b(t)C(t,T) + 2\sigma(t)^2C^2(t,T) - 1, \tag{3.13}
$$

and $A'(t,T) = -d\sigma(t)^2C(t,T)$.

As mentioned in the introduction, a Magnus expansion may be an easier tool than the Dyson expansion to calculate all the terms in (3.11) when σ is not constant. Such an expansion would be a remarkable result, since it would provide a new solution to (3.13).

4 Conclusion and Future Work

For future work, we intend to design and analyze new numerical schemes that implement the Dyson series to solve BSDEs. The main weakness of Theorem 2.4 is that it currently requires the functional F to be infinitely Malliavin differentiable. It is unknown at this point whether this smoothness requirement can be relaxed. Theorem 2.4 can certainly be extended to a filtration generated by several Brownian motions, and probably to Lévy processes. A generalization from representation of martingales to representation of semi-martingales would also be interesting.

5 Appendix

5.1 Proof of Theorem 2.1

Let $t = m\Delta$, we just prove the case $m = M - 1$. The general case obtains similarly. We remind the reader of Proposition 5.6 in [15], namely that, if both F, $E[F|\mathcal{F}_s] \in \mathbb{D}^{1,2}$ then for $t \leq s$:

$$D_t(E[F|\mathcal{F}_s]) = E[D_t F|\mathcal{F}_s]. \tag{5.1}$$

Using (5.1) and the Clark-Ocone formula (see, e.g. [13]), we get, for $t \leq T$ and any positive integer l:

$$\begin{aligned}
E[D_T^l F|\mathcal{F}_t] &= E[D_T^l F] + \int_0^t E[D_s E[D_T^l F|\mathcal{F}_t]|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&= E[D_T^l F] + \int_0^t E[E[D_s D_T^l F|\mathcal{F}_t]|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&= E[D_T^l F] + \int_0^t E[D_s D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&= E[D_T^l F] + \int_0^T E[D_s D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&\quad - \int_t^T E[D_s D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&= E[D_T^l F|\mathcal{F}_T] - \int_t^T E[D_s D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s).
\end{aligned} \tag{5.2}$$

Now we are going to show

$$\int_t^T E[D_s D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s) = \int_t^T E[D_T^{l+1} F|\mathcal{F}_s] \,\mathrm{d}W(s). \tag{5.3}$$

Since F is assumed discrete, for $s \in (t, T]$ we have:

$$D_s D_T^l F = D_T^{l+1} F. \tag{5.4}$$

Since this part is crucial, we provide the following example to illustrate it: if $l = 0$ and $F = W^2(T)$, then $D_s F = D_T F = 2W(T)$.

By (5.2) and (5.4),

$$\begin{aligned}
E[D_T^l F|\mathcal{F}_t] = E[D_T^l F|\mathcal{F}_T] &- \int_t^T E[D_T D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s) \\
&+ \int_t^{t+0^+} E[(D_T - D_s) D_T^l F|\mathcal{F}_s] \,\mathrm{d}W(s).
\end{aligned} \tag{5.5}$$

It remains to verify

$$\int_t^{t+0^+} E\left[(D_T - D_s) D_T^l F|\mathcal{F}_s\right] \,\mathrm{d}W(s) = 0.$$

For any $\varepsilon > 0$ arbitrarily small, denote by

$$Y(t,\varepsilon) = \int_t^{t+\varepsilon} E[(D_T - D_s)D_T^l F|\mathcal{F}_s]\, \mathrm{d}W(s),$$

we show that:

$$\lim_{\varepsilon\to 0} Y(t,\varepsilon) = 0 \quad a.s.. \tag{5.6}$$

Observe that

$$E\left[Y^2\left(t,\frac{1}{n^2}\right)\right] = \int_t^{t+\frac{1}{n^2}} E[E[(D_T - D_s)D_T^l F|\mathcal{F}_s]]^2\, \mathrm{d}s \le C\frac{1}{n^2}, \tag{5.7}$$

where $C > 0$ is a constant and (5.7) follows because Malliavin derivatives are bounded. Then by following Chebyshev inequality and the Borel-Cantelli lemma we get (5.6). Then (5.3) follows. We thus obtain, for $l = 0$,

$$E[F|\mathcal{F}_t] = E[F|\mathcal{F}_T] - \int_t^T E[D_T F|\mathcal{F}_{s_1}]\, \mathrm{d}W(s_1) \tag{5.8}$$

$$= E[F|\mathcal{F}_T] - \int_t^T E[D_T F|\mathcal{F}_T]\delta W(s_1) + \int_t^T \int_{s_1}^T E[D_T^2 F|\mathcal{F}_{s_2}]\delta W(s_2)\delta W(s_1), \tag{5.9}$$

where (5.8) follows from (5.5) with $l = 0$, and (5.9) follows from (5.5) with $l = 1$. We continue the expansion iteratively:

$$\begin{aligned} E[F|\mathcal{F}_t] &= E[F|\mathcal{F}_T] - \int_t^T E[D_T F|\mathcal{F}_T]\delta W(s_1) + \dots \\ &+(-1)^{L-1} \int_t^T \int_{s_1}^T \dots \int_{s_{L-2}}^T E[D_T^L F|\mathcal{F}_T]\delta W(s_{L-1})\dots\delta W(s_1) \\ &+(-1)^L \int_t^T \int_{s_1}^T \dots \int_{s_{L-1}}^T E[D_T^{L+1} F|\mathcal{F}_{s_L}]\delta W(s_L)\dots\delta W(s_1) \end{aligned} \tag{5.10}$$

To show the series converges is equivalent to showing the following remainder converges to 0 in $L^2(\Omega)$:

$$R_{[t,T]}^L := \int_t^T \int_{s_1}^T \dots \int_{s_{L-1}}^T E[D_T^{L+1} F|\mathcal{F}_{s_L}]\delta W(s_L)\dots\delta W(s_1). \tag{5.11}$$

To simplify notation, we denote the iterated Skorohod integral of the multiparameter stochastic process X by:

$$\delta^L(X) := \int_{\mathbb{R}^L} X(s_1,\dots,s_n)\,(\delta W(s))^{\otimes L}$$

where $(\delta W(s))^{\otimes L} := \delta W(s_L)\dots\delta W(s_1)$ and similarly at below we denote $(\,\mathrm{d}s)^{\otimes L} := \mathrm{d}s_L \dots \mathrm{d}s_1$.

Lemma 5.1. *For any multiparameter stochastic process X, we have*

$$E[\delta^L(X)^2] = \sum_{i=0}^L \binom{L}{i}^2 i! E\left[\left\|D^{L-i}X\right\|^2_{H^{\otimes(2L-i)}}\right],$$

where the norm of the tensor product of Hilbert space $H^{\otimes(2L-i)}$, $\left\|D^{L-i}X\right\|_{H^{\otimes(2L-i)}}$, is defined by

$$\left\|D^{L-i}X\right\|^2_{H^{\otimes(2L-i)}} := \int_{\mathbb{R}^{2L-i}} D^{L-i}_{x_1,\ldots,x_{L-i}} X(s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L)(\mathrm{d}x)^{\otimes(L-i)}$$
$$\times\left(D^{L-i}_{s_1,\ldots,s_{L-i}} X(x_1,\ldots,x_{L-i},s_{L-i+1},\ldots,s_L)\right)(\mathrm{d}s)^{\otimes L-i}(\mathrm{d}s_{L-i+1}\ldots\mathrm{d}s_L).$$

This lemma is referred to (2.12) in [12]. Since their paper did not include a proof of this statement, we came up with a proof, which is available upon request.

Lemma 5.2.

$$E\left[\left(R^L_{[t,T]}\right)^2\right] \le \sum_{i=0}^{L} E\left[\left(D^{2L-i}_T F\right)^2\right]\binom{L}{i}^4 \frac{i!}{(L!)^2}(T-t)^{2L-i} \xrightarrow[L\to\infty]{} 0. \tag{5.12}$$

Proof. To compute the $L^2(\Omega)$-norm of the remainder $R^L_{[t,T]}$, we calculate its symmetrization:

$$R^L_{[t,T]} = \frac{(-1)^L}{L!}\int_{\mathbb{R}^L} H_L(s_1,\ldots,s_L)\left(\delta W(s)\right)^{\otimes L}, \tag{5.13}$$

where

$$H_L(s_1,\ldots,s_L) := \sum_{\sigma\in S_L} E[D^L_T F|\mathcal{F}_{s_{\sigma(L)}}]\chi_{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T}(s_1,\ldots,s_L) \tag{5.14}$$

is a symmetric function with S_L being the set of all permutations of $\{1,\ldots,L\}$. Then according to Lemma 5.1, and Fubini's theorem, we obtain:

$$E[\delta^L(H_L)^2] = \sum_{i=0}^{L}\binom{L}{i}^2 i!E\left[\left\|D^{L-i}H_L\right\|^2_{H^{\otimes(2L-i)}}\right]$$
$$= \sum_{i=0}^{L}\binom{L}{i}^2 i!E\Big[\int_{\mathbb{R}^{2L-i}}\left[D^{L-i}_{x_1,\ldots,x_{L-i}}H_L(s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L)(\,\mathrm{d}x)^{\otimes(L-i)}\right]$$
$$\times\left[D^{L-i}_{s_1,\ldots,s_{L-i}}H_L(x_1,\ldots,x_{L-i},s_{L-i+1},\ldots,s_L)(\mathrm{d}s)^{\otimes(L-i)}\right](\mathrm{d}s_{L-i+1}...\mathrm{d}s_L)\Big]. \tag{5.15}$$

By definition of H_L in (5.14), we obtain:

$$D^{L-i}_{x_1,\ldots,x_{L-i}}H_L(s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L)$$
$$= \sum_{\sigma\in S_L} E[D^{L-i}_{x_1,\ldots,x_{L-i}}D^L_T F|\mathcal{F}_{s_{\sigma(L)}}]\chi_{[t,s_{\sigma(L)}]}(x_1,\ldots,x_{L-i})\chi_{\{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T\}}(s_1,\ldots,s_L)$$

and with $r_l = x_l\chi_{\{l\le L-i\}} + s_l\chi_{\{l>L-i\}}$, l is a positive integer which is no larger than L, we obtain also:

$$D^{L-i}_{s_1,\ldots,s_{L-i}}H_L(x_1,\ldots,x_{L-i},s_{L-i+1},\ldots,s_L)$$
$$= \sum_{\sigma'\in S_L} E[D^{L-i}_{s_1,\ldots,s_{L-i}}D^L_T F|\mathcal{F}_{s_{\sigma(L)}}]\chi_{[t,s_{\sigma(L)}]}(s_1,\ldots,s_{L-i})\chi_{\{t\le r_{\sigma'(1)}\le\ldots\le r_{\sigma'(L)}\le T\}}(r_1,\ldots,r_L)$$

Then by Cauchy-Schwarz inequality and the discreteness of F, we obtain:

$$\begin{aligned}
& E\left[D^{L-i}_{x_1,\ldots,x_{L-i}} H_L(s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L) D^{L-i}_{s_1,\ldots,s_{L-i}} H_L(x_1,\ldots,x_{L-i},s_{L-i+1},\ldots,s_L)\right] \\
= & E\left[\sum_{\sigma\in S_L}\sum_{\sigma'\in S'_L} E[D^{L-i}_{x_1,\ldots,x_{L-i}} D^L_T F|\mathcal{F}_{s_{\sigma(L)}}] E[D^{L-i}_{s_1,\ldots,s_{L-i}} D^L_T F|\mathcal{F}_{r_{\sigma'(L)}}]\right] \\
& \chi_{[t,s_{\sigma(L)}]}(x_1,\ldots,x_{L-i}) \chi_{\{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T\}}(s_1,\ldots,s_L) \\
& \chi_{[t,s_{\sigma'(L)}]}(s_1,\ldots,s_{L-i}) \chi_{\{t\le r_{\sigma'(1)}\le\ldots\le r_{\sigma'(L)}\le T\}}(r_1,\ldots,r_L) \\
\le & E\left[\left(D^{2L-i}_T F\right)^2\right] \sum_{\sigma\in S_L}\sum_{\sigma'\in S'_L} \chi_{[t,s_{\sigma(L)}]}(x_1,\ldots,x_{L-i}) \chi_{[t,s_{\sigma'(L)}]}(s_1,\ldots,s_{L-i}) \\
& \chi_{\{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T\}}(s_1,\ldots,s_L) \chi_{\{t\le r_{\sigma'(1)}\le\ldots\le r_{\sigma'(L)}\le T\}}(r_1,\ldots,r_L) \\
\le & E\left[\left(D^{2L-i}_T F\right)^2\right] \sum_{\sigma\in S_L}\sum_{\sigma'\in S'_L} \chi_{[t,T]}(x_1,\ldots,x_{L-i},s_1,\ldots,s_{L-i}) \\
& \chi_{\{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T\}}(s_1,\ldots,s_L) \chi_{\{t\le r_{\sigma'(1)}\le\ldots\le r_{\sigma'(L)}\le T\}}(r_1,\ldots,r_L). \qquad (5.16)
\end{aligned}$$

Then by observing that the function

$$\begin{aligned}
& f(x_1,\ldots,x_{L-i},s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L) = \sum_{\sigma\in S_L}\sum_{\sigma'\in S'_L} \chi_{[t,T]}(x_1,\ldots,x_{L-i},s_1,\ldots,s_{L-i}) \\
& \times \chi_{\{t\le s_{\sigma(1)}\le\ldots\le s_{\sigma(L)}\le T\}}(s_1,\ldots,s_L) \chi_{\{t\le r_{\sigma'(1)}\le\ldots\le r_{\sigma'(L)}\le T\}}(r_1,\ldots,r_L)
\end{aligned}$$

is symmetric for each of the three groups of variables: $\{x_1,\ldots,x_{L-i}\}, \{s_1,\ldots,s_{L-i}\}$ and $\{s_{L-i+1},\ldots,s_L\}$, we obtain

$$\begin{aligned}
& \int_{[t,T]^{2L-i}} f(x_1,\ldots,x_{L-i},s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L)(\,\mathrm{d}x)^{\otimes(L-i)}(\mathrm{d}s)^{\otimes(L-i)}(\mathrm{d}s_{L-i+1}\ldots\mathrm{d}s_L) \\
= & ((L-i)!)^2\, i! \int_A f(x_1,\ldots,x_{L-i},s_1,\ldots,s_{L-i},s_{L-i+1},\ldots,s_L)(\,\mathrm{d}x)^{\otimes(L-i)}(\mathrm{d}s)^{\otimes(L-i)}(\mathrm{d}s_{L-i+1}\ldots\mathrm{d}s_L)
\end{aligned}$$

where $A=\{x_m, s_n\ m\le L-i,\ n\le L : x_1\le\ldots\le x_{L-i};\ s_1\le\ldots\le s_{L-i};\ s_{L-i+1}\le\ldots\le s_L\}$. Meanwhile by basic calculation the following property holds:

$$\int_{t\le x_1\le\ldots\le x_n\le T}(\,\mathrm{d}x)^{\otimes n} \le \int_{t\le x_{\alpha_1}\le\ldots\le x_{\alpha_{n-i}}\le T, t\le x_{\beta_1}\le\ldots\le x_{\beta_i}\le T}(\,\mathrm{d}x_{i+1}\ldots\mathrm{d}x_n)^{\otimes(n-i)}(\,\mathrm{d}x)^{\otimes i}$$

where $\{\alpha_m, m=1,\dots,n-i\}\cup\{\beta_m, m=1,\dots,i]\}=\{1,\dots,n\}$. Then we finally obtain:

$$
\begin{aligned}
& E[\delta^L(H_L)^2] \\
\leq\ & \sum_{i=0}^{n} E\left[\left(D_T^{2L-i}F\right)^2\right]\binom{L}{i}^2 i! \\
& \int_{[t,T]^{2L-i}} f(x_1,\dots,x_{L-i},s_1,\dots,s_{L-i},s_{L-i+1},\dots,s_L)(\,\mathrm{d}x)^{\otimes(L-i)}(\mathrm{d}s)^{\otimes(L-i)}(\mathrm{d}s_{L-i+1}\dots\mathrm{d}s_L) \\
\leq\ & \sum_{i=0}^{L} E\left[\left(D_T^{2L-i}F\right)^2\right]\binom{L}{i}^2 i!\,((L-i)!)^2\, i!\left(\frac{L!}{i!(L-i)!}\right)^2 \\
& \int_{t\leq x_1\leq\ldots\leq x_{L-i}\leq T,\ t\leq s_1\leq\ldots\leq s_{L-i}\leq T,\ t\leq s_{L-i+1}\leq\ldots\leq s_L\leq T}(\,\mathrm{d}x)^{\otimes(L-i)}(\mathrm{d}s)^{\otimes(L-i)}(\mathrm{d}s_{L-i+1}\dots\mathrm{d}s_L) \\
=\ & \sum_{i=0}^{L} E\left[\left(D_T^{2L-i}F\right)^2\right]\binom{L}{i}^2 i!\,((L-i)!)^2\, i!\left(\frac{L!}{i!(L-i)!}\right)^2\frac{(T-t)^{2L-i}}{((L-i)!)^2\, i!} \\
=\ & \sum_{i=0}^{L} E\left[\left(D_T^{2L-i}F\right)^2\right]\binom{L}{i}^4 i!(T-t)^{2L-i}.
\end{aligned}
\tag{5.17}
$$

Thus combine (5.17) with (5.13) as well as the theorem's assumption, the remainder tends to zero:

$$
E\left[\left(R_{[t,T]}^L\right)^2\right]\leq\sum_{i=0}^{L} E\left[\left(D_T^{2L-i}F\right)^2\right]\binom{L}{i}^4\frac{i!}{(L!)^2}(T-t)^{2L-i}\xrightarrow[L\to\infty]{}0. \tag{5.18}
$$

□

In order to establish (2.7), we first prove in Lemma 5.3 that the coefficients $\gamma(m,l)$ are independent of F.

Lemma 5.3. *There exists random variables $\gamma(m,l)$, $l=1,\dots,L$ and $m=0,\dots,M-1$, which are independent of F, such that: For any discrete random variable F with the Malliavin derivatives of order higher than L vanish, we have:*

$$
E[F|\mathcal{F}_t]=\sum_{l=0}^{L}\gamma(m,l)E[D_T^lF|\mathcal{F}_T]. \tag{5.19}
$$

Proof. We remind the reader again that we have assumed $t=(M-1)\Delta$ and $T=M\Delta$ throughout this long proof. The general cases of $m\geq 0$ are similar. We define, for any $s\in[t,T]$ and $g>0$:

$$
\begin{aligned}
M_{l,0}(s) &= E_T[D_T^lF|\mathcal{F}_T]; \\
M_{l,g}(s) &= \int_s^T\dots\int_{s_{g-2}}^T\int_{s_{g-1}}^T E[D_T^lF|\mathcal{F}_T]\delta W(s_g)\delta W(s_{g-1})\dots\delta W(s_1).
\end{aligned}
$$

We first claim that there exists stochastic processes $\delta(.;h,g)$ independent of F such that, for $s \in [t,T]$:

$$M_{l,g}(s) = \sum_{h=0}^{g} \delta(s;h,g) E[D_T^{l+h} F|\mathcal{F}_T] \tag{5.20}$$

From the formula for the Skorohod integral of a process multiplied by a random variable (see Equation (1.48) in [13]), we calculate indeed:

$$\begin{aligned} M_{l,1}(s) &= \int_s^T E[D_T^l F|\mathcal{F}_T]\delta W(u) \\ &= E[D_T^l F|\mathcal{F}_T](W(T) - W(s)) - E[D_T^{l+1} F|\mathcal{F}_T](T-s), \end{aligned}$$

so that:

$$\begin{aligned} \delta(s;0,1) &= W(T) - W(s); \\ \delta(s;1,1) &= s - T. \end{aligned}$$

We suppose by induction that (5.20) holds for $M_{l,g}$. To prove that it holds for $M_{l,g+1}$, we reapply Equation (1.48) in [13], and obtain, for $s \in [t,T]$:

$$\begin{aligned} &M_{l,g+1}(s) = \int_s^T M_{l,g}(u)\delta W(u) = \sum_{h=0}^{g} \int_s^T \delta(u;h,g) E[D_T^{l+h} F|\mathcal{F}_T]\delta W(u) \\ &= \sum_{h=0}^{g} \left(E[D_T^{l+h} F|\mathcal{F}_T] \int_s^T \delta(u;h,g)\delta W(u) - E[D_T^{l+h+1} F|\mathcal{F}_T] \int_s^T \delta(u;h,g)\,\mathrm{d}u \right) \\ &= \sum_{h=0}^{g+1} E[D_T^{l+h} F|\mathcal{F}_T]\delta(s;h,g+1), \end{aligned}$$

where

$$\begin{aligned} \delta(s;0,g+1) &= \int_s^T \delta(u;0,g)\delta W(u); \\ \delta(s;h,g+1) &= \int_s^T \delta(u;h,g)\delta W(u) - \int_s^T \delta(u;h-1,g)\,\mathrm{d}u \quad \text{for} \quad 1 \le h \le g; \\ \delta(s;g+1,g+1) &= -\int_s^T \delta(u;h,g)\,\mathrm{d}u. \end{aligned}$$

Thus (5.20) holds. By (5.10), we know that for any discrete F such that the Malliavin derivatives of order higher than L vanish:

$$\begin{aligned} E[F|\mathcal{F}_t] &= \sum_{l=0}^{L} (-1)^l M_{l,l}(t) \\ &= \sum_{l=0}^{L} (-1)^l \sum_{h=0}^{l} \delta(t;h,l) E[D_T^{l+h} F|\mathcal{F}_T]. \end{aligned} \tag{5.21}$$

Thus, we can set:

$$\gamma(M-1,l)=\sum_{i=0}^{\lfloor l/2\rfloor}(-1)^{l-i}\delta(t;i,l-i). \tag{5.22}$$

This proves that $\gamma(M-1,l)$ is independent of F. $\square$

It is possible, but rather complicated, to calculate $\gamma(m,l)$ from (5.22). However, since (5.19) holds for any differentiable F, we resort to a simpler strategy. Our strategy is to vary F to determine recursively $\gamma(m,l)$. For simplicity, we suppose $m=0$, i.e., $t=0$. Clearly the first coefficient (take F=constant) is:

$$\gamma(0,0)=1.$$

To determine the second coefficient, $\gamma(0,L)$, with $L=1$, we choose a function F such that $D^{L+1}F=0$. The only such function is a linear function of $W(T)$. We thus put in (5.19) $F=W(T)$ and calculate:

$$E[F]=\sum_{l=0}^{1}\gamma(0,l)D_T^lF.$$

In other terms:

$$0=\gamma(0,0)W(T)+\gamma(0,1)\times 1$$

Thus:

$$\gamma(0,L)=-W(T).$$

The general structure of the recursion is then:

$$\gamma(0,L)=\frac{E[F]-\sum_{l=0}^{L-1}\gamma(0,l)D_T^lF}{D_T^LF}. \tag{5.23}$$

Clearly, formula (5.23) applies for any coefficient L. To calculate $\gamma(0,L)$ for $L=2$ we thus pick $F=W^2(T)$ (so that $D^{L+1}F=0$) and obtain:

$$\begin{aligned}\gamma(0,2) &= \frac{E[W^2(T)]-\sum_{l=0}^{1}\gamma(0,l)D_T^lW^2(T)}{D_T^2W^2(T)}\\ &= \frac{T-W^2(T)+W(T)\times 2W(T)}{2}\\ &= \frac{W^2(T)+T}{2}.\end{aligned}$$

We complete the proof by induction. Suppose that, for

$$\gamma(0,L-1)=\chi_{\{L\text{ even}\}}\left(\frac{\Delta}{2}\right)^{L/2}\frac{1}{(L/2)!}-\sum_{l=0}^{L-2}\gamma(0,l)\frac{W(T)^{L-1-l}}{(L-1-l)!}, \tag{5.24}$$

we select $F=W^L(T)$, which we insert together with (5.24) in (5.23) to arrive at

$$\gamma(0,L)=\chi_{\{L\text{ even}\}}\left(\frac{\Delta}{2}\right)^{L/2}\frac{1}{(L/2)!}-\sum_{l=0}^{L-1}\gamma(0,l)\frac{W(T)^{L-l}}{(L-l)!}. \tag{5.25}$$

The formula in the case where $0 < m < M$ obtains by replacing $W(T)$ by $W((m+1)\Delta) - W(m\Delta)$, that is choosing for test functions F above the successive powers of $W((m+1)\Delta) - W(m\Delta)$. We obtain then:

$$\gamma(m, L) = \chi_{\{L \text{ even}\}} \left(\frac{\Delta}{2}\right)^{L/2} \frac{1}{(L/2)!} - \sum_{l=0}^{L-1} \gamma(m, l) \frac{(W((m+1)\Delta) - W(m\Delta))^{L-l}}{(L-l)!}. \tag{5.26}$$

It is then trivial to verify by induction that (2.8) satisfies (5.26). Finally by using lemma 5.3 in (5.10), the theorem follows. □

5.2 Proof of Theorem 2.2

We prove this theorem following the notation in [13] Section 1.1.1.. Without loss of generality, we assume $(F^{(M)})$ and F are all $\mathcal{F}_T$-measurable, with some $T \geq 0$. Let $f \in L^2([0,T])$, we firstly claim that for an exponential function $\epsilon(f) = \exp(\int_0^T f(s)\,\mathrm{d}W(s) - \frac{1}{2}\|f\|^2_{L^2([0,T])})$, by definition, for $0 \leq t \leq T$ (the case for $t > T$ is trivial),

$$\epsilon(f)(\omega^t) = \epsilon(f\chi_{[0,t]}).$$

Then by using Theorem 1.1.1 and Proposition 1.1.1 in [13], there exists a sequence of real numbers $\{c_\alpha\}_{\alpha \in \mathcal{J}}$, such that for all $0 \leq t \leq T$,

$$\begin{aligned}\epsilon(f)(\omega^t) &= \left(\sum_{\alpha \in \mathcal{J}} c_\alpha \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^T e_i(s)\,\mathrm{d}W(s)\right)\right)(\omega^t) \\ &= \sum_{\alpha \in \mathcal{J}} c_\alpha \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right).\end{aligned}$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)$, $\alpha! = \alpha_1! \ldots \alpha_m!$, $\{e_i\}_i$ is a basis of $L^2([0,T])$ and $\mathcal{J} := \{(\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m : m \geq 1\}$ denotes the set of all finite sequence of nonnegative integers. Since the linear span of $\epsilon(f)$, for $f \in L^2([0,T])$ is a total subset of $L^2(\Omega)$, then by arguing with convergence in $L^2(\Omega)$, we have for $F \in L^2(\Omega)$, there exists a sequence of real numbers $\{c_\alpha\}_{\alpha \in \mathcal{J}}$, such that for all $t \geq 0$,

$$F(\omega^t) = \sum_{\alpha \in \mathcal{J}} c_\alpha \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right).$$

Thus we have proved that

$$\|F(\omega^t)\|^2_{L^2(\Omega)} = E\left[\left(\sum_{\alpha \in \mathcal{J}} c_\alpha \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right)\right)^2\right] < \infty \tag{5.27}$$

It is well-known that

$$\|F\|^2_{L^2(\Omega)} = \sum_{\alpha \in \mathcal{J}} \alpha! c_\alpha^2. \tag{5.28}$$

Here we notice that $\left\{e_i\chi_{[0,t]}\right\}_{i\geq 0}$ is no longer a orthonormal basis thus we can not have an isometry like formula (1.7) in [13] .

Suppose that $F^{(M)} - F \xrightarrow[M\to\infty]{L^2(\Omega)} 0$, and let $F^{(M)} - F$ has the following representation:

$$F^{(M)} - F = \sum_{\alpha\in\mathcal{J}} c_\alpha^{(M)} \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^T e_i(s)\,\mathrm{d}W(s)\right).$$

We know by (5.28) that:

$$\left\|F^{(M)} - F\right\|_{L^2(\Omega)}^2 = \sum_{\alpha\in\mathcal{J}} \alpha! \left(c_\alpha^{(M)}\right)^2 \xrightarrow[M\to\infty]{} 0,$$

which implies $c_\alpha^{(M)} \to 0$. Then according to (5.27), we obtain

$$\left\|\left(F^{(M)} - F\right)(\omega^t)\right\|_{L^2(\Omega)}^2 = E\left[\left(\sum_{\alpha\in\mathcal{J}} c_\alpha^{(M)} \sqrt{\alpha!} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right)\right)^2\right] < \infty$$

and by the fact that $c_\alpha^{(M)} \to 0$, we know

$$\lim_{M\to\infty}\left|c_\alpha^{(M)} c_\beta^{(M)} \prod_{i=1}^{\infty} h_{\alpha_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right) \prod_{i=1}^{\infty} h_{\beta_i}\left(\int_0^t e_i(s)\,\mathrm{d}W(s)\right)\right| = 0,$$

then by dominated convergence theorem, we obtain:

$$\left\|\left(F^{(M)} - F\right)(\omega^t)\right\|_{L^2(\Omega)} \xrightarrow[M\to\infty]{} 0,$$

and then the theorem is proven.

5.3 Proof of Theorem 2.3

We first show there exists a discrete approximation $F^{(M)}$ of F in $L^2(\Omega)$. Assume that $F \in L^2(\Omega)$ is $\mathcal{F}_T$-measurable, then there exists a sequence of polynomial functions (see for instance [13])

$$p_M\left(\int_0^T h_1(s)\,\mathrm{d}W(s), \ldots, \int_0^T h_M(s)\,\mathrm{d}W(s)\right) \xrightarrow[M\to+\infty]{L^2(\Omega)} F,$$

where $h_1, \ldots, h_M \in L^2([0,T])$. Also observe that, for each Wiener integral $\int_0^T h_i(s)\,\mathrm{d}W(s)$, by definition, there exists a sequence of polynomial functions

$$J_M^{(i)}\left(W(\frac{T}{M}), W(\frac{2T}{M}), \ldots, W(T)\right) \xrightarrow[M\to+\infty]{L^2(\Omega)} \int_0^T h_i(s)\,\mathrm{d}W(s).$$

Therefore, by the continuity of the Brownian trajectory and polynomials, one has

$$F^{(M)} = \left(p_M \circ \left(J_M^{(1)}, \ldots, J_M^{(M)}\right)\right)\left(W(\frac{T}{M}), W(\frac{2T}{M}), \ldots, W(T)\right) \xrightarrow[M\to+\infty]{L^2(\Omega)} F.$$

We now wish to prove the following result: for $t \leq s - \delta$ with some $0 < \delta < \Delta$,

$$\frac{E[F|\mathcal{F}_{s-\delta}](\omega^t) - E[F|\mathcal{F}_s](\omega^t)}{\delta} \xrightarrow[\delta \to 0]{L^2(\Omega)} \frac{1}{2}\left(D_s^2 E[F|\mathcal{F}_s]\right)(\omega^t). \tag{5.29}$$

To this effect, we prove that (5.29) is true for discrete $F^{(M)}$. The main difficulty in the proof of the next lemma is that, unlike in the BTE, the time s belongs to the real line, and is not necessarily aligned on the grid points mT/M.

Lemma 5.4. *For any* $0 \leq t < s < T$*:*

$$\frac{E[F^{(M)}|\mathcal{F}_{s-\delta}](\omega^t) - E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta} \xrightarrow[\delta \to 0]{L^2(\Omega)} \frac{1}{2}\left(D_s^2 E[F^{(M)}|\mathcal{F}_s]\right)(\omega^t). \tag{5.30}$$

Proof. Fix s, and let $\Delta = T/M$. We construct a sequence of numbers $\{\delta_k = \Delta/k\}$, for k integer and larger than 1. Let $m = \lfloor s/\Delta \rfloor$ is the floor number of s/Δ. Thus $s \in [m\Delta, (m+1)\Delta)$. First, suppose that $s \in (m\Delta, (m+1)\Delta)$. Then there exists a value of K large enough, so that, for all $k \geq K$, $s - \delta_k \in (m\Delta, (m+1)\Delta)$. We compute $E[F^{(M)}|\mathcal{F}_{s-\delta_k}]$ similarly as in (5.10):

$$\begin{aligned} & E[F^{(M)}|\mathcal{F}_{s-\delta_k}] \\ = \ & E[F^{(M)}|\mathcal{F}_s] - \int_{s-\delta_k}^s E[D_s F^{(M)}|\mathcal{F}_s]\delta W(s_1) + \int_{s-\delta_k}^s \int_{s_1}^s E[D_s^2 F^{(M)}|\mathcal{F}_s]\delta W(s_2)\delta W(s_1) \\ & -R^3_{[s-\delta_k, s]}, \end{aligned} \tag{5.31}$$

where:

$$R^3_{[s-\delta_k,s]} = \int_{s-\delta_k}^s \int_{s_1}^s \int_{s_2}^s E[D_s^3 F^{(M)}|\mathcal{F}_{s_3}]\delta W(s_3)\delta W(s_2)\delta W(s_1).$$

On one hand, by Lemma 5.2, we obtain:

$$E\left[\left(R^3_{[s-\delta_k,s]}\right)^2\right] \leq \sum_{i=0}^{3} E\left[\left(D_s^{6-i} F\right)^2\right] \binom{3}{i}^4 \frac{i!}{(3!)^2}\delta_k^{6-i}. \tag{5.32}$$

On the other hand, in (5.31) we can compute:

$$\left(-\int_{s-\delta_k}^s E[D_s F^{(M)}|\mathcal{F}_s]\delta W(s_1)\right)(\omega^t) = \delta_k E[D_s^2 F^{(M)}|\mathcal{F}_s](\omega^t); \tag{5.33}$$

$$\left(\int_{s-\delta_k}^s \int_{s_1}^s E[D_s^2 F^{(M)}|\mathcal{F}_s]\delta W(s_2)\delta W(s_1)\right)(\omega^t) = \left(-\frac{\delta_k}{2} E[D_s^2 F^{(M)}|\mathcal{F}_s] + \frac{\delta_k^2}{2} E[D_s^4 F^{(M)}|\mathcal{F}_s]\right)(\omega^t). \tag{5.34}$$

Thus combining (5.31)-(5.34) as well as the assumption $F \in \mathbb{D}^6([0,T])$ and Theorem 2.2, we have:

$$\frac{E[F^{(M)}|\mathcal{F}_{s-\delta_k}](\omega^t) - E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta_k} - \frac{1}{2} D_s^2 E[F^{(M)}|\mathcal{F}_s](\omega^t)] \xrightarrow[\delta_k \to 0]{L^2(\Omega)} 0.$$

Suppose now that $s = m\Delta$. Then similarly we can choose K such that when $k > K$, $s - \delta_k \in ((m-1)\Delta, m\Delta)$ and then clearly (5.31)-(5.34) also hold and the proof is completed. $\square$

To show (5.29) we need the following relations.

$$E[(E[F-F^{(M)}|\mathcal{F}_s](\omega^t))^2] \xrightarrow[M\to\infty]{} 0; \tag{5.35}$$

$$E[(E[D_s^2F-D_s^2F^{(M)}|\mathcal{F}_s](\omega^t))^2] \xrightarrow[M\to\infty]{} 0; \tag{5.36}$$

Relations (5.35) and (5.36) result from applying Theorem 2.2, the closability of the operator D^2 on $[0,+\infty)$, the definition of $F^{(M)}$, and the property that for any stochastic process X_s and any time $t\le s$, $E[E[X_s|\mathcal{F}_t]^2]\le E[X_s^2]$. As in the proof of the previous lemma we use a sequence $\{\delta_{k,M}\}$, where $\delta_{k,M}=\frac{T}{Mk}$. We apply the triangle inequality, (5.35), (5.36) in (5.30) to obtain for $t\le s-\delta_{k,M}$:

$$\begin{aligned}
&E\left[\left(\frac{E[F|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}-\frac{1}{2}\left(D_s^2E[F|\mathcal{F}_s]\right)(\omega^t)\right)^2\right]^{1/2}\\
&=E\left[\left(\frac{E[F|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}-\frac{E[F^{(M)}|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}\right.\right.\\
&\quad+\frac{E[F^{(M)}|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}-\frac{1}{2}\left(D_s^2E[F^{(M)}|\mathcal{F}_s]\right)(\omega^t)\\
&\quad\left.\left.+\frac{1}{2}\left(D_s^2E[F^{(M)}|\mathcal{F}_s]\right)(\omega^t)-\frac{1}{2}\left(D_s^2E[F|\mathcal{F}_s]\right)(\omega^t)\right)^2\right]^{1/2}\\
&\le E\left[\left(\frac{E[F|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F^{(M)}|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)}{\delta_{k,M}}\right)^2\right]^{1/2}\\
&\quad+E\left[\left(\frac{E[F|\mathcal{F}_s](\omega^t)-E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}\right)^2\right]^{1/2}\\
&\quad+E\left[\left(\frac{E[F^{(M)}|\mathcal{F}_{s-\delta_{k,M}}](\omega^t)-E[F^{(M)}|\mathcal{F}_s](\omega^t)}{\delta_{k,M}}-\frac{1}{2}\left(D_s^2E[F^{(M)}|\mathcal{F}_s]\right)(\omega^t)\right)^2\right]^{1/2}\\
&\quad+E\left[\left(\frac{1}{2}\left(D_s^2E[F^{(M)}|\mathcal{F}_s]\right)(\omega^t)-\frac{1}{2}D_s^2E[F|\mathcal{F}_s](\omega^t)\right)^2\right]^{1/2}\xrightarrow[k,M\to\infty]{} 0.
\end{aligned} \tag{5.37}$$

Remembering the definitions (2.13) and (2.14), relation (5.37) becomes then, for $t\le s-\Delta$:

$$E\left[\left(\left(\frac{P_{s-\delta_{k,M}}-P_s}{\delta_{k,M}}F\right)(\omega^t)-\frac{1}{2}\left((D_s^2\circ P_s)F\right)(\omega^t)\right)^2\right]\xrightarrow[k,M\to\infty]{} 0.\ \square$$

5.4 Proof of Theorem 2.4

In this proof, all equalities are in $L^2(\Omega)$ and $s\ge t$. From (5.29), we obtain:

$$P_sF(\omega^t)=F(\omega^t)+\frac{1}{2}\int_s^T(D_u^2\circ P_u)F(\omega^t)\,\mathrm{d}u. \tag{5.38}$$

Then for positive integer n we define the operator $T_s^{(n)}$ by:

$$T_s^{(n)} F := \sum_{i=0}^{n} \mathcal{A}_{i,s} F,$$

where

$$\mathcal{A}_{i,s} F := \int_s^T \cdots \int_{s_{i-1}}^T \frac{1}{2^i} D_{s_1}^2 \dots D_{s_i}^2 F \, \mathrm{d}s_i \dots \mathrm{d}s_1.$$

Then by iterating (5.38) we obtain: for $n > 0$

$$P_s F(\omega^t) = T_s^{(n-1)} F(\omega^t) + \frac{1}{2^n} \int_s^T \cdots \int_{u_{n-1}}^T (D_{u_n}^2 \dots D_{u_1}^2 \circ P_{u_n}) F(\omega^t) \, \mathrm{d}u_n \dots \mathrm{d}u_1.$$

Thus according to assumption (2.17):

$$E\left[\left((P_s - T_s^{(n-1)}) F(\omega^t)\right)^2\right] = E\left[\left(\frac{1}{2^n} \int_s^T \cdots \int_{u_{n-1}}^T (D_{u_n}^2 \dots D_{u_1}^2 \circ P_{u_n}) F(\omega^t) \, \mathrm{d}u_n \dots \mathrm{d}u_1\right)^2\right]$$

$$\leq \frac{(T-s)^{2n}}{(2^n n!)^2} E\left[\left(\sup_{u_1,\dots,u_n \in [0,T]} \left|D_{u_n}^2 \dots D_{u_1}^2 F\right|\right)^2\right] \xrightarrow[n\to\infty]{} 0, \tag{5.39}$$

We now take $s = t$ and obtain:

$$E[F|\mathcal{F}_t] = P_t F = T_t^{(\infty)} F(\omega^t) = \sum_{n=0}^{\infty} \frac{1}{2^n} \int_t^T \cdots \int_{s_{n-1}}^T D_{s_n}^2 \dots D_{s_1}^2 F(\omega^t) \, \mathrm{d}s_n \dots \mathrm{d}s_1. \quad \square$$

Acknowledgements

We thank *Josep Vives* and *Hank Krieger* for reviewing our manuscript. All errors are ours.

References

[1] F. Biagini, Y. Hu, B. Øksendal and T. Zhang T, *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer, London, 2008.

[2] J. M. C. Clark, *The representation of functionals of Brownian motion by stochastic integrals*, Ann. Math. Statist. 41 (1970), pp. 1282-1295.

[3] K. E. Dambis, *On the decomposition of continuous martingales*, Theor. Prob. Appl. 10 (1965), pp. 401-410.

[4] J. Detemple, R. Garcia and M. Rindisbacher, *Representation formulas for Malliavin derivatives of diffusion processes*, Finance and Stochastics 9(3) (2005), pp. 349-367.

[5] L. Dubins, L. and G. Schwarz, *On continuous martingales*, Proc. Nat. Acad. Sci. USA 53 (1965), pp. 913-916.

[6] J. L. Doob, Stochastic Processes, Wiley, New York, 1953.

[7] Y. Hu, D. Nualart and X. Song, *Malliavin calculus for backward stochastic differential equations and application to numerical solutions*, The Annals of Applied Probability 21, 6 (2011), pp. 2379-2423.

[8] J. Hunter and B. Nachtergaele, Applied Analysis, World Scientific, 2007.

[9] I. Karatzas and S. Shreve, Brownian Motion and Stochastic Calculus, Springer, 1991.

[10] S. Jin, Q. Peng and H. Schellhorn, *Fractional Hida Malliavin derivatives and series representations of fractional conditional expectations*, preprint (2015), submitted for publication. Available at arXiv: 1406.1538.

[11] F. Longstaff and E. Schwartz, *Valuing American options by simulation: a simple least-squares approach*, Review of Financial Studies 14 (2001), pp. 113-147.

[12] I. Nourdin and D. Nualart, (2010) *Central limit theorem for multiple Skorohod integrals*, J. Theor. Probab. 23 (2010), pp. 39-64.

[13] D. Nualart, The Malliavin Calculus and Related Topics, Springer, 1995.

[14] D. Ocone, *Malliavin calculus and stochastic integral representation of diffusion processes*, Stochastics 12 (1984), pp. 161-185.

[15] B. Øksendal, *An introduction to Malliavin calculus with applications to economics*, Norwegian School of Economics and Business Administration (1997).

[16] J. Oteo and J. Ros, *From time-ordered products to Magnus expansion*, Journal of Mathematical Physics 41,5 (2000), pp. 3268-3277.

[17] H. Schellhorn and H. Morris, *A Differential Tree Approach to Price Path-dependent American Options Using Malliavin Calculus*, IAENG Transactions on Engineering Technologies Volume II, American Institute of Physics, 2009.

[18] S. Shreve, Stochastic Calculus for Finance II: Continuous-Time Models, Springer, 2004.

[19] C. Tellambura, *Accurate computation of the MGF of the lognormal distribution and its application to sums of lognormal*, IEEE Transactions on Communication 58,5 (2010), pp. 1568-1577.

[20] J. Tsitsiklis and B. Van Roy, *Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives*, IEEE Transactions on Automatic Control 44 (1999), pp. 1850-1851.

[21] E. Zeidler, Quantum Field Theory 1: Basics in Mathematics and Physics, Springer, 2006.