ON NEAR-MARTINGALES AND A CLASS OF ANTICIPATING LINEAR SDES

HUI-HSIUNG KUO, PUJAN SHRESTHA*, SUDIP SINHA, AND PADMANABHAN SUNDAR

Abstract. The primary goal of this paper is to prove a near-martingale optional stopping theorem and establish solvability and large deviations for a class of anticipating linear stochastic differential equations. We prove the existence and uniqueness of solutions using two approaches: (1) Ayed–Kuo differential formula using an ansatz, and (2) a novel braiding technique by interpreting the integral in the Skorokhod sense. We establish a Freidlin–Wentzell type large deviations result for solution of such equations.

1. Introduction

Anticipating stochastic calculus has been an active and important research area for several years, and lies at the intersection of probability theory and infinite-dimensional analysis. Enlargement of filtration, Malliavin calculus, and white noise theory provide three distinct methodologies to incorporate anticipation (of future) into classical Itô theory of stochastic integration and differential equations.

It is to the credit of Itô who constructed an anticipating stochastic integral in 1976[6], and laid the foundation for the idea of enlargement of the underlying filtration. Ever since, the method was embraced by several researchers that led to many important works (see articles in [7]). The advent of an integral invented by Skorokhod resulted in an impressive edifice built by Malliavin on stochastic calculus of variations in order to prove Hörmander’s hypoellipticity result by stochastic analysis. Malliavin calculus provided a natural basis for the development and study of anticipative stochastic analysis and differential equations. Around the same time, a systematic study of Hida distributions gave rise to white noise theory and a general framework for stochastic calculus.

Malliavin calculus and white noise theory have vast applicability to the theory of stochastic differential equations with anticipation. However, the results obtained by these theories are primarily abstract though general. A more tractable theory was envisaged by Kuo based on a concrete stochastic integral known as the Ayed–Kuo integral[1]. Under less generality, the latter allows one to obtain results under easily understood, verifiable hypotheses.

In this article, we prove some results about stopped near-martingales, which are generalizations of martingales. We then study existence, uniqueness and large deviation principle for linear stochastic differential equations with anticipating initial conditions and drifts. While we rely mostly on the Ayed–Kuo formalism, other theories are minimally used either out of necessity, or to compare and contrast the conclusions of certain results.

2020 Mathematics Subject Classification. Primary 60H10, 60F10, 60G48, 60G40; Secondary 60H05, 60H07, 60H20.

Key words and phrases. anticipatory integral, stochastic integral, stochastic differential equation, near-martingale, optimal stopping theorem, large deviation principles.

* Corresponding author.
The structure of the paper is as follows. In section 2, we introduce the Ayed–Kuo integral. In section 3, study near-martingales. We show that Ayed–Kuo integrals are near-martingales. We also show that stopped near-martingales are near-martingales, and prove an optional stopping theorem for near-submartingales. In section 4, we study methods for solving anticipating linear stochastic differential equations by interpreting the anticipating stochastic integral from two perspectives. For the Ayed–Kuo formulation, we use the differential formula and an ansatz to derive the solution. For the Skorokhod interpretation, we introduce a novel braiding technique inspired by Trotter’s product formula[12]. We show that the solutions coincide when the assumptions are identical. Finally, in section 5, we derive large deviation principles for the solutions of the class of anticipating linear stochastic differential equations studied in section 4. In this paper, we assume \( t \in [0, 1] \), unless specified otherwise.

2. The Ayed–Kuo anticipating stochastic calculus

Before we define the Ayed–Kuo integral, we need to define instantly independent processes. A stochastic process \( \phi(t) \) is called instantly independent with respect to \( \{F_t\} \) if for each \( t \in [0, 1] \), the random variable \( \phi(t) \) and the \( \sigma \)-field \( F_t \) are independent. Instantly independent processes are the counterpart of adapted processes in this theory.

We refer to [4, section 2] for a detailed definition of the anticipating stochastic integral. In what follows, we highlight the crucial steps in the definition in a concise manner.

**Definition 2.1 ([4, definition 2.3]).** The anticipating integral is defined in following three steps:

(1) Suppose \( f(t) \) is an \( F_t \)-adapted continuous stochastic process and \( \phi(t) \) is a continuous stochastic processes that is instantly independent with respect to \( \{F_t\} \). Then the stochastic integral of \( \Phi(t) = f(t)\phi(t) \) is defined by

\[
\int_0^1 f(t)\phi(t) \, dW_t = \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^n f(t_{j-1}) \phi(t_j) (W_{t_j} - W_{t_{j-1}}),
\]

provided that the limit exists in probability.

(2) For any stochastic process of the form \( \Phi(t) = \sum_{i=1}^n f_i(t)\phi_i(t) \), where \( f_i(t) \) and \( \phi_i(t) \) are given as in step (1), the stochastic integral is defined by

\[
\int_0^1 \Phi(t) \, dW_t = \sum_{i=1}^n \int_0^1 f_i(t) \phi_i(t) \, dW_t.
\]

(3) Let \( \Phi(t) \) be a stochastic process such that there is a sequence \( (\Phi_n(t))_{n=1}^\infty \) of stochastic processes of the form in step (2) satisfying

(a) \( \int_0^1 |\Phi_n(t) - \Phi(t)|^2 \, dt \to 0 \) almost surely as \( n \to \infty \), and

(b) \( \int_0^1 \Phi_n(t) \, dW_t \) converges in probability as \( n \to \infty \).

Then the stochastic integral of \( \Phi(t) \) is defined by

\[
\int_0^1 \Phi(t) \, dW_t = \lim_{n \to \infty} \int_0^1 \Phi_n(t) \, dW_t \text{ in probability.}
\]

This integral is well defined, as demonstrated in [4, lemma 2.1]. In order to use the definition of the integral, we first need to decompose the integrand into a sum of products of
adapted and instantly independent parts. The main idea is to then use the left-endpoints of subintervals to evaluate the adapted parts and the right-endpoints of subintervals to evaluate the instantly independent parts.

3. Near-martingales

3.1. Near-martingale property of the Ayed–Kuo integral. Martingales are an extremely important class of processes that are used to model fair games, and hence find applications not only in probability theory, but also in mathematical finance and numerous other fields. Itô’s integrals are essentially continuous martingale transforms, and therefore retain the martingale nature of the integrator. Since the Ayed–Kuo integral is an extension of the Itô integral, it is natural to ask if Ayed–Kuo integrals are martingale. Unfortunately, they are not. However, we have a very similar property, which gives rise to the idea of near-martingales.

Definition 3.1. An integrable stochastic process $N_t$ is called a near-submartingale with respect to the filtration $\{F_t\}$ if for any $s \leq t$, we have $E(N_t - N_s | F_s) \geq 0$ almost surely. It is called a near-martingale if $E(N_t - N_s | F_s) = 0$ almost surely.

The following result links martingales and near-martingales. In particular, it says that conditioned near-martingales are martingales.

Theorem 3.2 ([5, theorem 2.11]). Let $N_t$ be an integrable stochastic process and let $M_t = E(N_t | F_t)$. Then $N_t$ is a near-martingale if and only if $M_t$ is a martingale.

Ayed–Kuo integrals are near-martingales, as stated by this theorem.

Theorem 3.3. Let $\Theta(x,y)$ be a function that is continuous in both variables such that the stochastic integral

$$N_t = \int_a^t \Theta(W_s, W_b - W_s) dW_s, \quad a \leq t \leq b,$$

exists and $E|N_t| < \infty$ for each $t$ in $[0,1]$. Furthermore, assume that the family of partial sums

$$\sum_{i=1}^n \Theta(W_{t_i}, W_1 - W_{t_{i-1}}) (W_{t_i} - W_{t_{i-1}})$$

are uniformly integrable. Then $N_t$, $a \leq t \leq b$, is a near-martingale with respect to the filtration generated by Brownian motion given by $\{F_t\}$.

Proof. Let $s \leq t$ and consider a partition, $\Delta_n$, of $[s,t]$ with $t_0 = s$ and $t_n = t$. The definition of the Ayed–Kuo stochastic integral in conjunction with the uniform integrability condition
on the partial sums implies
\[
\mathbb{E} [N_t - N_s \mid \mathcal{F}_s] = \mathbb{E} \left[ \int_s^t \Theta(W_v, W_b - W_v) \, dW_v \mid \mathcal{F}_s \right] \\
= \mathbb{E} \left[ \lim_{n \to \infty} \sum_{k=1}^n \Theta(W_{k-1}, W_b - W_k) \Delta W_k \mid \mathcal{F}_s \right] \\
= \lim_{n \to \infty} \sum_{k=1}^n \mathbb{E} [\Theta(W_{k-1}, W_b - W_k) \Delta W_k \mid \mathcal{F}_s].
\] (3.1)

Consider, \( \mathcal{H}^{(b)}_a = \sigma(\mathcal{F}_a \cup \mathcal{G}^{(b)}) \). Then \( \mathcal{F}_s \subseteq \mathcal{F}_{k-1} \subseteq \mathcal{H}^{(k)}_{k-1} \). Using this fact alongside the continuity of \( \Theta \) in both variables, we have that \( \Theta(W_{k-1}, W_b - W_k) \) is \( \mathcal{H}^{(k)}_{k-1} \)-measurable. Furthermore, via the independence of the Brownian increments, \( \Delta W_k \) is independent of \( \mathcal{H}^{(k)}_{k-1} \). Thus,
\[
\mathbb{E} [\Theta(W_{k-1}, W_b - W_k) \Delta W_k \mid \mathcal{F}_s] \\
= \mathbb{E} \left[ \mathbb{E} \left[ \Theta(W_{k-1}, W_b - W_k) \Delta W_k \mid \mathcal{H}^{(k)}_{k-1} \right] \mid \mathcal{F}_s \right] \\
= \mathbb{E} [\Theta(W_{k-1}, W_b - W_k) \mathbb{E} [\Delta W_k \mid \mathcal{F}_s]] \\
= 0.
\]

Using this result for each \( k \) in equation (3.1), we have \( \mathbb{E} [N_t - N_s \mid \mathcal{F}_s] = 0 \), and so \( N_t \) is a near-martingale.

\[\text{Figure 1. A } t\text{-dependence plot of the disjoint increments of } W. \text{ The shaded regions represents the forward and separation } \sigma\text{-field.}\]

3.2. **Stopped near-martingales.** In this section, we show that stopped near-martingales are near-martingales. We also generalize Doob’s optional stopping theorem for near-martingales.

**Definition 3.4.** Let \( (A_n)_{n=0}^{\infty} \) be an adapted process and \( (X_n)_{n=0}^{\infty} \) a discrete time near-submartingale. Then the processes \( (Y_n)_{n=0}^{\infty} \), where \( Y_0 = 0 \) and
\[
Y_n = (A \bullet X)_n = \sum_{i=1}^{n} A_{n-1} (X_n - X_{n-1})
\]
is called the **near-martingale transform** of \( X \) by \( A \).
Near-martingale transforms retain the near-martingale property, as is shown in the following result.

**Proposition 3.5.**  
(1) If $X$ is a near-submartingale and $A$ is a bounded non-negative adapted process, then $(A \bullet X)$ is a near-submartingale.  
(2) If $X$ is a near-martingale and $A$ is a bounded adapted process, then $(A \bullet X)$ is a near-martingale.  
(3) If $X$ and $A$ are both square integrable, then we do not require the boundedness condition in items 1 and 2.

**Proof.** We only prove item 1 because the rest follow the same process. Let $X$ be a near-submartingale and $Y = (A \bullet X)$. Suppose $n$ is an arbitrary time. Note that $(Y_n - Y_{n-1}) = A_{n-1}(X_n - X_{n-1})$, which is integrable since $A$ is bounded. Using the adaptedness of $A$, we get
\[
\mathbb{E}(Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}) = \mathbb{E}(A_{n-1}(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) = A_{n-1}\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) \geq 0,
\]
where the last inequality holds since $A$ is non-negative. □

We now show that stopped near-submartingales are near-submartingales.

**Theorem 3.6.** Suppose $X$ is a discrete time near-submartingale and $\tau$ a stopping time. Then the stopped process $X^\tau$ defined by $X_n^\tau = X_{\tau \wedge n}$ is a (discrete time) near-submartingale.

**Proof.** Let $A_n = 1_{\{n \leq \tau\}}$, so the process $A$ is bounded, non-negative, and adapted. Now, note that $X_n^\tau - X_0 = X_{\tau \wedge n} - X_0 = (A \bullet X)_n$. Therefore, by proposition 3.5, we get that $X^\tau$ is a near-submartingale. □

Now, we show the equivalent result of Doob’s optional stopping theorem for discrete time near-submartingales.

**Theorem 3.7.** Let $X$ be a discrete time near-submartingale. Suppose $\sigma$ and $\tau$ are two bounded stopping times with $\sigma \leq \tau$. Then $X_\sigma$ and $X_\tau$ are integrable, and $\mathbb{E}(X_\tau - X_\sigma \mid \mathcal{F}_\sigma) \geq 0$ almost surely.

**Proof.** Since $\sigma$ and $\tau$ are bounded, there exists $N < \infty$ such that $\sigma \leq \tau \leq N$. Let $Y$ be any near-submartingale. Clearly, $Y_\sigma$ is integrable. Suppose $B \in \mathcal{F}_\sigma$. Then for any $n \leq N$, we have $B \cap \{\sigma = n\} \in \mathcal{F}_n$, and so
\[
\int_{B \cap \{\sigma = n\}} (Y_n - Y_\sigma) \, d\mathbb{P} = \int_{B \cap \{\sigma = n\}} (Y_n - Y_n) \, d\mathbb{P} \geq 0.
\]
Summing over $n$, we get $\int_B (Y_n - Y_\sigma) \, d\mathbb{P} \geq 0$, and so $\mathbb{E}(Y_n - Y_\sigma \mid \mathcal{F}_\sigma) \geq 0$. Finally, let $Y_n = X_n^\tau$ to get
\[
\mathbb{E}(X_\tau^\tau - X_\sigma^\tau \mid \mathcal{F}_\sigma) = \mathbb{E}(X_\tau - X_\sigma \mid \mathcal{F}_\sigma) \geq 0.
\]
□

We need the following definition and lemma to prove the result in continuous time.

**Definition 3.8.** Let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a decreasing sequence of $\sigma$-algebras, and let $X = (X_n)_{n=1}^{\infty}$ be a stochastic process. Then the pair $(X_n, \mathcal{F}_n)_{n=1}^{\infty}$ is called a **backward near-submartingale** if for every $n$,

(1) $X_n$ is integrable and $\mathcal{F}_n$-measurable, and
(2) \( \mathbb{E}(X_n - X_{n+1} \mid \mathcal{F}_{n+1}) \geq 0 \).

Lemma 3.9. Let \((X_n, \mathcal{F}_n)_{n=1}^{\infty}\) be a backward near-submartingale with \( \lim_{n \to \infty} \mathbb{E}(X_n) > -\infty \). If \( X \) is non-negative for every \( n \), then \( X \) is uniformly integrable.

Proof. As \( n \to \infty \), we have \( \mathbb{E}(X_n) \downarrow \lim_{n \to \infty} \mathbb{E}(X_n) = \inf_n \mathbb{E}(X_n) > -\infty \). Fix \( \epsilon > 0 \). By the definition of infimum, there exists a \( K > 0 \) such that for any \( n \geq K \), we have \( \mathbb{E}(X_K) - \lim_{n \to \infty} \mathbb{E}(X_n) < \epsilon \).

For any \( k \geq n \) and \( \delta > 0 \), we have
\[
\mathbb{E}\left(|X_k| \mathbb{1}_{\{|X_k| > \delta\}}\right) = \mathbb{E}\left(X_k \mathbb{1}_{\{|X_k| > \delta\}}\right) + \mathbb{E}\left(X_k \mathbb{1}_{\{|X_k| \leq \delta\}}\right) - \mathbb{E}(X_k).
\]

Moreover, since \( X \) is a backward near-submartingale, \( \mathbb{E}\left(X_k \mathbb{1}_{\{|X_k| \leq \delta\}}\right) \leq \mathbb{E}\left(X_n \mathbb{1}_{\{|X_n| \leq \delta\}}\right) \). Therefore,
\[
\mathbb{E}\left(|X_k| \mathbb{1}_{\{|X_k| > \delta\}}\right) \leq \mathbb{E}\left(X_n \mathbb{1}_{\{|X_n| > \delta\}}\right) + \mathbb{E}\left(X_n \mathbb{1}_{\{|X_n| \leq \delta\}}\right) - \mathbb{E}(X_n) - \epsilon \leq \mathbb{E}\left(|X_n| \mathbb{1}_{\{|X_n| > \delta\}}\right) + \epsilon.
\]

By Markov’s inequality and the non-negativity of \( X \),
\[
\mathbb{P}\{|X_k| > \delta\} \leq \frac{1}{\delta} \mathbb{E}|X_k| = \frac{1}{\delta} \mathbb{E}(X_k) \leq \frac{1}{\delta} \mathbb{E}(X_1) \to 0
\]
as \( \delta \to \infty \). This concludes the proof. \( \square \)

We are now ready to prove the near-martingale optional stopping theorem in continuous time.

Theorem 3.10. Let \( N \) be a near-submartingale with right-continuous sample paths. Suppose \( \sigma \) and \( \tau \) are two bounded stopping times with \( \sigma \leq \tau \). If \( N \) is either non-negative or uniformly integrable, then \( N_\sigma \) and \( N_\tau \) are integrable, and
\[
\mathbb{E}(N_\tau - N_\sigma \mid \mathcal{F}_\sigma) \geq 0 \text{ almost surely.}
\]

Proof. We use a discretization argument to prove the result. Let \( T > 0 \) be a bound for \( \tau \). For every \( n \in \mathbb{N} \), define the discretization function
\[
f_n : [0, \infty) \to \left\{ \frac{k}{n} : k = 0, \ldots, n \right\} : x \mapsto \left\lfloor \frac{2^n x + 1}{2^n} \right\rfloor \land T,
\]
and let \( \sigma_n = f_n(\sigma) \) and \( \tau_n = f_n(\tau) \).

Now, for any \( n \) and \( t \),
\[
\{\tau_n \leq t\} = \{f_n(\tau) \in [0, t]\} = \{\tau \in f_n^{-1}[0, t]\} = \{\tau \in f_n^{-1}\left[0, \frac{2^n t}{2^n}\right]\} \in \mathcal{F}_{\frac{2^n t}{2^n}} \subseteq \mathcal{F}_t,
\]
so \( \tau_n \) is a stopping time. Similarly, \( \sigma_n \) is a stopping time. Moreover, it can be easily seen that \( \sigma_n \leq \tau_n \) for every \( n \), and \( \sigma_n \searrow \sigma \) and \( \tau_n \nearrow \tau \) as \( n \to \infty \). Therefore, by theorem 3.7, we get \( N_{\sigma_n} \) and \( N_{\tau_n} \) are integrable, and \( \mathbb{E}(N_{\tau_n} - N_{\sigma_n} \mid \mathcal{F}_{\sigma_n}) \geq 0 \) almost surely. Furthermore, it is easy to see that \( \mathcal{F}_\sigma = \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n} \subset \mathcal{F}_{\sigma_n} \) for any \( n \). Therefore, \( \mathbb{E}(N_{\tau_n} - N_{\sigma_n} \mid \mathcal{F}_\sigma) \geq 0 \) almost surely for any \( n \).

If \( N \) is non-negative, by construction, \((N_{\sigma_n}, \mathcal{F}_{\sigma_n})_{n=1}^{\infty}\) is a backward near-submartingale such that \( N_{\sigma_n} \geq 0 \) for every \( n \). Therefore, \( \mathbb{E}(N_{\sigma_n}) \searrow \mathbb{E}(N_{\sigma}) > -\infty \) as \( n \to \infty \). Using lemma 3.9, \((N_{\sigma_n})_{n=1}^{\infty}\) is uniformly integrable. Similarly, \((N_{\tau_n})_{n=1}^{\infty}\) is also uniformly integrable. On the other hand, if \( N \) is uniformly integrable, this is trivial.
Using the right continuity of $N$ and the boundedness assumption of $\sigma$ and $\tau$, we get
\[
\lim_{n \to \infty} N_{\sigma_n} = N_{\sigma} \quad \text{and} \quad \lim_{n \to \infty} N_{\tau_n} = N_{\tau}
\]
almost surely. Furthermore, the uniform integrability of $(N_{\sigma_n})_{n=1}^{\infty}$ and $(N_{\tau_n})_{n=1}^{\infty}$ allows us to conclude that $N_{\sigma}$ and $N_{\tau}$ are integrable and that the convergence is also in $L^1$, giving us $\mathbb{E}(N_{\tau} - N_{\sigma} \mid \mathcal{F}_\sigma) \geq 0$ almost surely. □

We highlight the special case of theorem 3.10.

**Corollary 3.11.** Let $N$ be a non-negative near-martingale with right-continuous sample paths and $\tau$ a bounded stopping time. Then $N_{\tau}$ is integrable, and $\mathbb{E}(N_{\tau}) = \mathbb{E}(N_0)$ almost surely.

4. Anticipating linear stochastic differential equations

In our previous works [9, 8], we studied linear stochastic differential equations with anticipating initial conditions, where the stochastic integral is in the Ayed–Kuo sense. In [3], the authors gave examples of linear stochastic differential equations with anticipating diffusion coefficient. In this paper, we focus on anticipating drift.

In particular, we shall be concerned about the solution of a class of anticipating linear stochastic differential equations of the form

\[
\begin{cases}
  dX_t = \sigma_t X_t dW_t + f \left( \int_0^1 \gamma_t dW_t \right) X_t dt, & t \in [0, 1], \\
  X_0 = \xi,
\end{cases}
\]

(4.1)

where $W_t$ is a Brownian motion, $f : \mathbb{R} \to \mathbb{R}$ is bounded function, $\xi$ a random variable, and $\sigma_t$ is an bounded adapted process such that all integrability conditions are satisfied. We choose this class because we want to study linear stochastic differential equations where the anticipation comes from the drift coefficient being Brownian functionals.

4.1. The Ayed–Kuo sense. We look at an extension of Itô’s formula that can account for instantly independent processes. Let $X_t$ and $Y^{(t)}$ be stochastic processes of the form

\[
X_t = X_a + \int_a^t g(s) dB(s) + \int_a^t h(s) ds,
\]

(4.2)

\[
Y^{(t)} = Y^{(b)} + \int_t^b \xi(s) dB(s) + \int_t^b \eta(s) ds,
\]

(4.3)

where $g(t), h(t)$ are adapted (so $X_t$ is an Itô process), and $\xi(t), \eta(t)$ are instantly independent such that $Y^{(t)}$ is also instantly independent.

**Theorem 4.1** ([4, theorem 3.2]). Suppose \( \{X^{(i)}_t\}_{i=1}^n \) and \( \{Y^{(t)}_j\}_{j=1}^m \) are stochastic processes of the form given by equations (4.2) and (4.3), respectively. Suppose \( \theta(t, x_1, \ldots, x_n, y_1, \ldots, y_m) \) is a real-valued function that is $C^1$ in $t$ and $C^2$ in other variables. Then the stochastic
differential of \( \theta(t, X_t^{(1)}, \ldots, X_t^{(n)}, Y_1^{(t)}, \ldots, Y_m^{(t)}) \) is given by
\[
d\theta(t, X_t^{(1)}, \ldots, X_t^{(n)}, Y_1^{(t)}, \ldots, Y_m^{(t)})
= \theta_t \, dt + \sum_{i=1}^n \theta_{x_i} \, dX_t^{(i)} + \sum_{j=1}^m \theta_{y_j} \, dY_j^{(t)}
+ \frac{1}{2} \sum_{i,k=1}^n \theta_{x_i x_k} \, dX_t^{(i)} \, dX_t^{(k)} - \frac{1}{2} \sum_{j,l=1}^m \theta_{y_j y_l} \, dY_j^{(t)} \, dY_l^{(t)}.
\]

The above differential formula allows us to calculate the solutions of anticipating stochastic differential equations. We shall see two instances of its application in section 4.1.

We apply the differential formula to derive a general result for existence of linear stochastic differential equations with anticipating coefficients.

**Theorem 4.2.** Suppose \( \sigma \in L^2_{ad}([0,1] \times \Omega), \gamma \in L^2[0,1], \) and \( \xi \) be a random variable independent of the Wiener process \( W \). Moreover, suppose \( f \in C^2(\mathbb{R}) \) along with \( f, f', f'' \in L^1(\mathbb{R}) \). Then the solution of equation (4.1) in the Ayed–Kuo theory is given by
\[
Z_t = \xi \exp \left[ \int_0^t \sigma_s \, dW_s - \frac{1}{2} \int_0^t \sigma_s^2 \, ds + \int_0^t f \left( \int_0^1 \gamma_u \, dW_u - \int_s^t \gamma_u \, \sigma_u \, du \right) \, ds \right]. \tag{4.4}
\]

**Proof.** We show that equation (4.4) solves equation (4.1). The initial condition is trivially verified.

Note that equation (4.4) can be written as
\[
Z_t = \xi \exp \left[ \int_0^t \sigma_s \, dW_s - \frac{1}{2} \int_0^t \sigma_s^2 \, ds + \int_0^t f \left( \int_0^1 \gamma_u \, dW_u + \int_0^1 \gamma_u \, dW_u - \int_s^t \gamma_u \, \sigma_u \, du \right) \, ds \right].
\]

Motivated by this, we define
\[
\theta(t, x_1, x_2, y) = \xi \exp \left[ x_1 - \frac{1}{2} \int_0^t \sigma_s^2 \, ds + \int_0^t f \left( \gamma_u + y - \int_s^t \gamma_u \, \sigma_u \, du \right) \, ds \right].
\]

Moreover, let
\[
X_t^{(1)} = \int_0^t \sigma_s \, dW_s \quad (\text{so } dX_t^{(1)} = \sigma_t \, dW_t),
\]
\[
X_t^{(2)} = \int_0^t \gamma_s \, dW_s \quad (\text{so } dX_t^{(2)} = \gamma_t \, dW_t),
\]
and
\[
Y^{(t)} = \int_t^1 \gamma_s \, dW_s \quad (\text{so } dY^{(t)} = -\gamma_t \, dW_t).
\]

Then we can write \( Z_t = \theta \left( t, X_t^{(1)}, X_t^{(2)}, Y^{(t)} \right) \).

For conciseness, we denote \( F = f \left( \int_0^1 \gamma_t \, dW_t - \int_s^t \gamma_u \, \sigma_u \, du \right) \), and similarly the derivatives
\[
F' = f' \left( \int_0^1 \gamma_t \, dW_t - \int_s^t \gamma_u \, \sigma_u \, du \right) \quad \text{and} \quad F'' = f'' \left( \int_0^1 \gamma_t \, dW_t - \int_s^t \gamma_u \, \sigma_u \, du \right). \tag{4.8}
\]
derivatives of \( \theta \), we have
\[
\theta_{x_1} = \theta_{x_1x_1} = \theta, \\
\theta_{x_2} = \theta_{x_1x_2} = \theta_y = \theta \cdot \int_0^t F' \, ds, \\
\theta_{x_2x_2} = \theta_{yy} = \theta \cdot \left( \int_0^t F' \, ds \right)^2 + \theta \cdot \int_0^t F'' \, ds, \text{ and} \\
\theta_t = -\frac{1}{2} \theta \sigma^2 + \theta f(x_2 + y) - \gamma_t \sigma_t \theta_y,
\]
where we used the Leibniz integral rule and the second line for the last identity.

Since \( \xi \) is independent of the Wiener process, by theorem 4.1, we get
\[
d\theta = \theta_t \, dt + \theta_{x_1} \, dX_t^{(1)} + \theta_{x_2} \, dX_t^{(2)} + \theta_y \, dY^{(t)} \\
+ \frac{1}{2} \theta_{x_1x_1} \left( dX_t^{(1)} \right)^2 + \frac{1}{2} \theta_{x_2x_2} \left( dX_t^{(2)} \right)^2 + \theta_{x_1x_2} \, dX_t^{(1)} \, dX_t^{(2)} - \frac{1}{2} \theta_{yy} \, (dY^{(t)})^2.
\]
Using the relationships between the derivatives of \( \theta \) and its differential form, we have
\[
d\theta = \theta_t \, dt + \theta \sigma_t \, dW_t + \theta_y \gamma_t \, dW_t - \theta_{x_1} \sigma_t \, dW_t \\
+ \frac{1}{2} \theta \gamma_t^2 \, dt + \frac{1}{2} \theta \gamma_t \sigma_t \, dt - \frac{1}{2} \theta \gamma_t^2 \, dt
\]
\[
= \left( \theta_t + \frac{1}{2} \theta \sigma_t^2 + \theta y \gamma_t \sigma_t \right) \, dt + \theta \sigma_t \, dW_t.
\]

Now,
\[
\theta_t + \frac{1}{2} \theta \sigma_t^2 + \theta y \gamma_t \sigma_t = \left( -\frac{1}{2} \theta \sigma_t^2 + \theta f(x_2 + y) - \theta \gamma_t \sigma_t \right) + \frac{1}{2} \theta \sigma_t^2 + \theta y \gamma_t \sigma_t = \theta f(x_2 + y),
\]
and so
\[
d\theta = f(x_2 + y) \, dt + \theta \sigma_t \, dW_t.
\]
Since \( Z_t = \theta \left( t, X_t^{(1)}, X_t^{(2)}, Y^{(t)} \right) \), we get
\[
dZ_t = f \left( \int_0^1 \gamma_s \, dW_s \right) Z_t \, dt + \sigma_t \, Z_t \, dW_t,
\]
which is exactly equation (4.1).

\[\square\]

Theorem 4.1 is an indispensable tool for analyzing anticipating processes. We show another example by finding the stochastic differential equation corresponding to the square of the above solution.

**Theorem 4.3.** Under the condition of theorem 4.2, the stochastic differential equation
\[
\left\{
\begin{array}{l}
    dV_t = \left[ \sigma_t^2 + f \left( \int_0^1 \gamma_s \, dW_s \right) + 2 \gamma_t \sigma_t \int_0^t f' \left( \int_0^1 \gamma_u \, dW_u - \int_s^t \gamma_u \, \sigma_u \, du \right) \, ds \right] V_t \, dt \\
    + 2 \sigma_t V_t \, dW_t, \\
    V_0 = \xi^2
\end{array}
\right.
\]
is solved by \( Z_t^2 \), where \( Z \) is given by equation (4.4).
Remark 4.4. An interesting feature is that the derivative of $f$ appears in the stochastic differential equation.

Proof. We follow the exact same strategy as the proof of theorem 4.2. The initial condition is trivially true. Let $V_t = Z_t^2$.

Taking the square of both sides of equation (4.4), we get

$$V_t = \xi^2 \exp \left[ \int_0^t 2\sigma_s dW_s - \int_0^t \sigma_s^2 ds + \int_0^t 2f \left( \int_0^t \gamma_u dW_u - \int_s^t \gamma_u \sigma_u du \right) ds \right]$$

We have $V_t = \theta \left( t, X_t^{(1)}, X_t^{(2)}, Y^{(t)} \right)$, where

$$\theta(t, x_1, x_2, y) = \xi^2 \exp \left[ x_1 - \int_0^t \sigma_s^2 ds + \int_0^t 2f \left( x_2 + y - \int_s^t \gamma_u \sigma_u du \right) ds \right],$$

and

$$X_t^{(1)} = \int_0^t 2\sigma_s dW_s \quad \text{(so } dX_t^{(1)} = 2\sigma_t dW_t),$$

$$X_t^{(2)} = \int_0^t \gamma_s dW_s \quad \text{(so } dX_t^{(2)} = \gamma_t dW_t),$$

and $Y^{(t)} = \int_t^1 \gamma_s dW_s \quad \text{(so } dY^{(t)} = -\gamma_t dW_t).$

As before, writing $F = f \left( \int_0^1 \gamma_t dW_t - \int_s^t \gamma_u \sigma_u du \right)$, $F' = f' \left( \int_0^1 \gamma_t dW_t - \int_s^t \gamma_u \sigma_u du \right)$, and $F'' = f'' \left( \int_0^1 \gamma_t dW_t - \int_s^t \gamma_u \sigma_u du \right)$, we get

$$\theta_{x_1} = \theta_{x_1 x_1} = \theta,$$

$$\theta_{x_2} = \theta_{x_1 x_2} = \theta_y = 2\theta \cdot \int_0^t F' ds,$$

$$\theta_{x_2 x_2} = \theta_{y y} = \theta \cdot \left( \int_0^t F' ds \right)^2 + \theta \cdot \int_0^t F'' ds,$$

and

$$\theta_t = -\theta \sigma_t^2 + 2\theta f(x_2 + y) - \gamma_t \sigma_t \theta_y.$$

Using the general Itô formula (theorem 4.1), we get

$$d\theta = \theta_t dt + \theta_{x_1} dX_t^{(1)} + \theta_{x_2} dX_t^{(2)} + \theta_y dY^{(t)}$$

$$+ \frac{1}{2} \theta_{x_1 x_1} \left( dX_t^{(1)} \right)^2 + \frac{1}{2} \theta_{x_2 x_2} \left( dX_t^{(2)} \right)^2 + \theta_{x_1 x_2} dX_t^{(1)} dX_t^{(2)} - \frac{1}{2} \theta_{y y} (dY^{(t)})^2$$

$$= \theta_t dt + 2\theta \sigma_t dW_t + \theta_{x_1} \gamma_t dW_t - \gamma_t \sigma_t dW_t$$

$$+ \theta_{x_2 x_2} dW_t^2 + \frac{1}{2} \theta_{x_1 x_1} \frac{1}{2} dW_t^2 dt + 2\theta \gamma_t \sigma_t dt - \frac{1}{2} \theta_{y y} \gamma_t^2 dt$$

$$= \left( \theta_t + 2\theta \sigma_t^2 + 2\theta \gamma_t \sigma_t \right) dt + 2\theta \sigma_t dW_t$$

$$= \left[ \theta_t + 2\theta f(x_2 + y) + 2\gamma_t \sigma_t \theta \int_0^t F' ds \right] dt + 2\theta \sigma_t dW_t.$$
Finally, using $V_t = \theta \left( t, X_t^{(1)}, X_t^{(2)}, Y(t) \right)$, we get the stochastic differential equation. □

4.2. A novel braiding technique for the Skorokhod sense. In the prior section, we showed the existence of the solution via the Ayed–Kuo differential formula. However, the procedure started with intelligently guessing an ansatz for the solution and applying the differential formula to it. Can a solution be found without this “guessing”? In this section, we use elementary ideas from Malliavin calculus to interpret the stochastic differential equation in the Skorokhod sense. We introduce an iterative “braiding” technique in the spirit of Trotter’s product formula[12] that allows us to construct the solution without needing to know the form of the solution. Note that we expect to arrive at the same solution as in section 4.1 since under the definition of the Ayed–Kuo integral using $L^2(\Omega)$ convergence, the Hitsuda–Skorokhod integral and the Ayed–Kuo integrals are equivalent, as shown in [11, theorem 2.3]. In what follows, we briefly introduce some ideas of Malliavin calculus and Skorokhod integral so that we can introduce our braiding technique.

A well known extension of the Itô integral is Hitsuda–Skorokhod integral. For this text, we shall introduce the Hitsuda–Skorokhod integral as the adjoint of the Gross–Malliavin derivative. Let us first set up the spaces to operate on. Consider the probability space $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is the $\sigma$-field generated by the Brownian motion. Let $\mathbb{H} = L^2[0, 1]$ be the space of square integrable functions defined on the positive reals. For any $h \in \mathbb{H}$, consider the Wiener integral

$$W(h) = \int_0^1 h(t) \, dW_t.$$ 

In particular, if $h = 1_{[0, \frac{1}{2}]} \in \mathbb{H}$ then

$$W\left(1_{[0, \frac{1}{2}]}\right) = \int_0^1 1_{[0, \frac{1}{2}]}(t) \, dW_t = W_{\frac{1}{2}}.$$ 

This Hilbert space $\mathbb{H}$ plays an important role in the definition of the derivative. Let $\mathcal{S}$ be the class of smooth random variables such that $F \in \mathcal{S}$ has the form

$$F = f\left(W(h_1), W(h_2), \ldots, W(h_n)\right), \quad h_i \in \mathbb{H}, \ i \in \{1, 2, \ldots, n\},$$

where $f$ is a real valued $n$-dimensional smooth function whose derivatives have at most polynomial growth.

**Definition 4.5 ([10, definition 1.2.1]).** The Gross–Malliavin derivative of a smooth random variable $F \in \mathcal{S}$ is the real valued random variable given by

$$D_tF = \sum_{i=1}^n \partial_i f\left(W(h_1), W(h_2), \ldots, W(h_n)\right) h_i(t),$$

where $d_i$ is the derivative with respect to the $i$th variable.

We denote $\mathbb{D}^{1,2}$ as the closure of the derivative operator $D$ from $L^2(\Omega)$ to $L^2(\Omega; \mathbb{H})$. In other words, $\mathbb{D}^{1,2}$ is the completion of the class of smooth Brownian functionals with respect to the inner product

$$\langle F, G \rangle_{1,2} = E (FG) + E \left( \langle DF, DG \rangle_\mathbb{H} \right).$$

We now introduce the Skorokhod integral operator $\delta$. 

\[11\]
Definition 4.6 ([10, definition 1.3.1]). We denote by $\delta$ the adjoint of the operator $D$. That is, $\delta$ is an unbounded operator on $L^2(\Omega; \mathbb{H})$ with values in $L^2(\Omega)$ such that:

1. The domain of $\delta$ is the set of $\mathbb{H}$-valued square integrable random variables $u \in L^2(\Omega; \mathbb{H})$ such that for any $F \in \mathcal{D}^{1,2}$, where $c$ is some constant depending on $u$.

$$
\mathbb{E}(\langle DF, u \rangle_{\mathbb{H}}) \leq c \|F\|_2.
$$

2. If $u$ belongs to the domain of $\delta$, then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by

$$
\mathbb{E}(F \delta u) = \mathbb{E}(\langle DF, u \rangle_{\mathbb{H}}).
$$

for any $F \in \mathcal{D}^{1,2}$.

It is natural to ask about the nature of relationship of these two stochastic integral. While that is an open question, we refer to the following result.

Theorem 4.7 ([11, theorem 2.3]). Let $f$ be an adapted $L^2$-continuous stochastic process and $\phi$ be an instantly independent $L^2$-continuous stochastic process such that the sequence

$$
\sum_{i=1}^{n} f(t_i-1)\phi(t_i) (W_{t_i} - W_{t_{i-1}}),
$$

converges strongly in $L^2(\Omega)$ as $\|\Delta_n\| \to 0$. Then the limit $I(f\psi)$ equals the Hitsuda–Skorokhod integral $\delta(f\psi)$ in $\text{Dom}(\delta)$.

Now, we move on to finding the solution of the linear stochastic differential equation when the anticipating integral is taken in the sense of Skorokhod. First, fix the family of translation on the space of continuous functions starting at the origin in the Cameron–Martin direction given by

$$
(A_t(\omega))_s = \omega_s - \int_0^{t\wedge s} \sigma(u) \, du \quad \text{and} \quad (T_t(\omega))_s = \omega_s + \int_0^{t\wedge s} \sigma(u) \, du.
$$

We look at an existence result for stochastic differential equations in the Skorokhod sense.

Lemma 4.8. Suppose $\sigma \in L^2[0,1]$ and $\xi \in L^p(\Omega)$ for some $p > 2$. Then the stochastic differential equation

$$
\begin{cases}
    dZ_t = \sigma(t) \, Z_t \, dW_t \\
    Z_0 = \xi,
\end{cases}
$$

(4.5)

has the unique solution given by

$$
Z_t = (\xi \circ A_t) \, \mathcal{E}_t.
$$

(4.6)

Proof. It is clear that the family $\{(\xi \circ A_t) \, \mathcal{E}_t \mid t \in [0,1]\}$ is $L^r(\Omega)$-bounded for all $r < p$ by Girsanov’s theorem and Hölder’s inequality. Let $G$ be any smooth random variable. Multiply
both sides of equation (4.5) by \( G \). With the process \( X \) given by (4.6),

\[
\mathbb{E}\left( G \int_0^t \sigma(s) \, Z_s \, dW_s \right) = \mathbb{E}\left( \int_0^t \sigma(s) \, Z_s \, D_s G \, ds \right) \\
= \mathbb{E}\left( \xi \int_0^t \sigma(s) \, (D_s G)(T_s) \, ds \right) \quad \text{(using Girsanov theorem)} \\
= \mathbb{E}\left( \xi \int_0^t \frac{d}{ds} G(T_s) \, ds \right) \\
= \mathbb{E}(\xi(G(T_t) - G)) \\
= \mathbb{E}(\xi(A_t) \, \mathcal{E}_t \, G) - \mathbb{E}(\xi \, G) \quad \text{(again by Girsanov theorem)} \\
= \mathbb{E}(Z_t \, G) - \mathbb{E}(\xi \, G).
\]

Thus, a solution of the stochastic equation equation (4.5) is explicitly given by (4.6).

Uniqueness follows since the solution of equation (4.5) started at \( \xi \equiv 0 \) is identically zero at all times. \( \square \)

Now we introduce the braiding technique to solve equation (4.1), where \( \gamma \in L^2[0,1] \) and \( f : \mathbb{R} \to \mathbb{R} \). To simplify notation, define

\[
I_\gamma = \int_0^1 \gamma_s \, dW_s, \\
A_w^u(\omega) = \omega - \int_u^{(\omega \wedge \nu) \vee \nu} \sigma(s) \, ds, \\
E_u^v = \exp \left[ \int_u^v \sigma(s) \, dW_s - \frac{1}{2} \int_u^v \sigma(s)^2 \, ds \right], \text{ and} \\
g_u^v = \exp \left[ f(I_\gamma) \, (v - u) \right].
\]

Directly from the definitions above, for any \( u < v < w \), we get the compositions

\[
A_w^u \circ A_u^v = A_w^v, \\
E_u^v \circ A_v^w = E_u^w, \\
g_u^v \circ A_v^w = \exp \left[ f(I_\gamma \circ A_v^w) \, (v - u) \right],
\]

and the products

\[
E_u^v \cdot E_v^w = E_u^w, \text{ and} \\
g_u^v \cdot g_v^w = g_u^w.
\]

We suppress the dependence on \( \omega \) for notational convenience.

Fix \( t \in [0,1] \), and consider a sequence of partitions \( \Delta_n = \{0 = t_0 < t_1 < \cdots < t_n = t\} \) of \( [0,t] \) such that \( \|\Delta_n\| = \sup \{t_i - t_{i-1} \mid i \in [n]\} \to 0 \). On each subinterval, we

1. solve the equation having only the diffusion with the initial condition as the solution of the previous step, and
2. use the solution obtained in step 1 as the initial condition and solve the equation having only the drift.

13
For the first subinterval, the initial condition of step 1 is taken to be $\xi$. For a visual representation of the idea, see figure 2.

We explicitly demonstrate the process for the first two subintervals. An index $(i)$ in the superscript refers to the $i$th subinterval.

First subinterval.

1. The stochastic differential equation that we want to solve is
\[
\begin{align*}
    dY^{(1)}_u &= \sigma(u) Y^{(1)}_u dW_u, \quad u \in [0, t_1], \\
    Y^{(1)}_0 &= \xi.
\end{align*}
\]

Lemma 4.8 gave us the almost sure unique solution $Y^{(1)}_u = (\xi \circ A_{t_1}^u) E^{\xi}_{t_1}$, so
\[
Y^{(1)}_{t_1} = (\xi \circ A_{t_1}^u) E^{\xi}_{t_1}
\]
on a set $\Omega_1$, where $\mathbb{P}(\Omega_1) = 1$.

2. For each $\omega \in \Omega_1$, we want to solve the ordinary differential equation
\[
\begin{align*}
    dX^{(1)}_u &= f(I_\gamma) X^{(1)}_u du, \quad u \in [0, t_1], \\
    X^{(1)}_0 &= Y^{(1)}_{t_1}.
\end{align*}
\]

By the existence and uniqueness theorem of ordinary differential equations, the unique solution is given by $X^{(1)}_u = Y^{(1)}_{t_1} g^u_0 = (\xi \circ A_{t_1}^u) E^{\xi}_{t_1} g^u_0$, and so
\[
X^{(1)}_{t_1} = (\xi \circ A_{t_1}^u) E^{\xi}_{t_1} g^u_0.
\]

Second subinterval.

1. The stochastic differential equation that we want to solve is
\[
\begin{align*}
    dY^{(2)}_u &= \sigma(u) Y^{(2)}_u dW_u, \quad u \in [t_1, t_2], \\
    Y^{(2)}_{t_1} &= X^{(1)}_{t_1}.
\end{align*}
\]

Lemma 4.8 gives us the almost sure unique solution $Y^{(2)}_u = (X^{(1)}_{t_1} \circ A_{t_1}^u) E^{\xi}_{t_1}$. Now,
\[
Y^{(2)}_u = [(\xi \circ A_{t_1}^u) E^{\xi}_{t_1} g^u_0 \circ A_{t_1}^u] E^{\xi}_{t_1}
= (\xi \circ A_{t_1}^u \circ A_{t_1}^u) E^{\xi}_{t_1} E^{\xi}_{t_1} (g^u_0 \circ A_{t_1}^u)
= (\xi \circ A_{t_1}^u) E^{\xi}_{t_1} (g^u_0 \circ A_{t_1}^u),
\]
where we used the fact that \( E_{t_1}^{t_0} \) is invariant under \( A_{t_1}^u \). This is because, by definition,
\[
A_{t_1}^u(\omega) = \omega - \int_{t_1}^{t_0} \sigma(s) \, ds.
\]
Now, for \( E_{t_1}^{t_0} \), we have \( t \in [0, t_1] \). Therefore,
\[
A_{t_1}^u(\omega) = \omega - \int_{t_1}^{t_0} \sigma(s) \, ds = \omega,
\]
showing the invariance. This gives the motivation behind why we define \( A \) as such, and is a key trick in the method.

Continuing, we get
\[
Y_{t_2}^{(2)}(\omega) = (\xi \circ A_{t_2}^{t_0}) \ E_{t_2}^{t_0} (g_{t_1}^{t_0} \circ A_{t_1}^{t_2})
\]
on a set \( \Omega_2 \subseteq \Omega_1 \), where \( \mathbb{P}(\Omega_2) = 1 \).

(2) For each \( \omega \in \Omega_2 \), we have the ordinary differential equation
\[
\begin{aligned}
\left\{ 
\begin{array}{l}
\frac{dX_u^{(2)}}{du} = f(I_\gamma) X_u^{(2)}, \quad u \in [t_1, t_2], \\
X_{t_1}^{(2)} = Y_{t_2}^{(2)}.
\end{array}
\right.
\end{aligned}
\]
The unique solution is given by
\[
X_{t_2}^{(2)} = \left[ (\xi \circ A_{t_2}^{t_0}) \ E_{t_2}^{t_0} (g_{t_1}^{t_0} \circ A_{t_1}^{t_2}) \right] (g_{t_1}^{t_0} \circ A_{t_2}^{t_2})
= (\xi \circ A_{t_2}^{t_0}) \ E_{t_2}^{t_0} \prod_{i=1}^{2} (g_{t_1}^{t_0} \circ A_{t_2}^{t_2})
\]

It should now become obvious what the pattern is. We prove this using induction in the following lemma.

**Lemma 4.9.** Let \( \xi \in L^p(\Omega) \) for some \( p > 2 \). Consider the \( k \)th subinterval \( u \in [t_{k-1}, t_k] \) for any \( k \in [n] \), and define

(1) the stochastic differential equation
\[
\begin{aligned}
\left\{ 
\begin{array}{l}
\frac{dY_u^{(k)}}{du} = \sigma(u) Y_u^{(k)} \, dW_u, \quad u \in [t_{k-1}, t_k], \\
Y_{t_{k-1}}^{(k)} = X_{t_{k-1}}^{(k-1)},
\end{array}
\right.
\end{aligned}
\]

(2) the ordinary differential equation
\[
\begin{aligned}
\left\{ 
\begin{array}{l}
\frac{dX_u^{(k)}}{du} = f(I_\gamma) X_u^{(k)} \, du, \quad u \in [t_{k-1}, t_k], \\
X_{t_{k-1}}^{(k)} = Y_{t_k}^{(k)}.
\end{array}
\right.
\end{aligned}
\]

Then there exists a set \( \Omega_k \subseteq \Omega \) with \( \mathbb{P}(\Omega_k) = 1 \) such that on \( \Omega_k \), we have
\[
X_{t_k}^{(k)} = (\xi \circ A_{t_k}^{t_0}) \ E_{0}^{t_k} \prod_{i=1}^{k} (g_{t_i}^{t_{i-1}} \circ A_{t_i}^{t_k}).
\]
Proof. **Base cases.** This is true for $k = 1$ and $k = 2$ as shown in the computations above.

**Induction step.** Assume that the result holds for $k = m - 1$. This means that there exists $\Omega_{m-1}$ with $\mathbb{P}(\Omega_{m-1}) = 1$ such that on $\Omega_{m-1}$, we have

$$X^{(m-1)}_{t_{m-1}} = (\xi \circ A^0_{t_{m-1}}) \ E^0_{t_{m-1}} \prod_{i=1}^{m-1} (g^i_{t_{i-1}} \circ A^{t_{m-1}}_{t_{i}}).$$

Using the ideas of computations on the second subinterval, we get that there exists $\Omega_m$ with $\mathbb{P}(\Omega_m) = 1$ such that on $\Omega_m$, we have

$$Y^{(m)}_{t_m} = (\xi \circ A^0_{t_m}) \ E^0_{t_m} \prod_{i=1}^{m-1} (g^i_{t_{i-1}} \circ A^t_{t_{i}}).$$

Since $A^t_{t_m}$ is the identity function, on $\Omega_m$, we have

$$X^{(m)}_{t_m} = Y^{(m)}_{t_m} \ g^m_{t_{m-1}} = (\xi \circ A^0_{t_m}) \ E^0_{t_m} \prod_{i=1}^{m} (g^i_{t_{i-1}} \circ A^t_{t_{i}}).$$

The proof is now complete by mathematical induction. $\square$

We are now able to derive a closed form solution of equation (4.1) in the Skorokhod sense. This is the main theorem of the section.

**Theorem 4.10.** Suppose $\sigma, \gamma \in L^2[0, 1], f : \mathbb{R} \to \mathbb{R}$, and $\xi \in L^p(\Omega)$ for some $p > 2$. Then the unique solution of equation (4.1) in the Skorokhod sense is given by

$$Z_t = (\xi \circ A^t_0) \exp \left[ \int_0^t \sigma(s) \, dW_s - \frac{1}{2} \int_0^t \sigma(s)^2 \, ds + \int_0^t f \left( \int_0^1 \gamma_u \, dW_u - \int_0^t \gamma_u \sigma(u) \, du \right) \, ds \right].$$

**Remark 4.11.** Note that $\xi$ may depend on the Wiener process.

**Proof.** Using lemma 4.9, for any $t \in [0, 1]$, we have

$$X^{(n)}_t = (\xi \circ A^t_0) \ E^t_0 \prod_{i=1}^{k} (g^i_{t_{i-1}} \circ A^t_{t_{i}}).$$

Now,

$$\prod_{i=1}^{k} (g^i_{t_{i-1}} \circ A^t_{t_{i}}) = \prod_{i=1}^{k} \exp \left[ f(I_\gamma \circ A^t_{t_{i}}) (t_i - t_{i-1}) \right]$$

$$= \exp \left[ \sum_{i=1}^{k} f \left( \int_0^1 \gamma_u \, dW_u - \int_0^t \gamma_u \sigma(u) \, du \right) \Delta t_i \right].$$
Finally, taking $n \to \infty$, we get

$$Z_t = \lim_{n \to \infty} X_t^{(n)} = (\xi \circ A_t^0) E_0^t \exp \left[ \int_0^t f \left( \int_0^1 \gamma_u dW_u - \int_0^t \gamma_u \sigma(u) du \right) ds \right],$$

which exactly equals the proposed solution.

The solution exists almost surely, due to the continuity of the measure. Moreover, the solution is unique. For if not, there are two solutions which disagree for the first time on a particular interval, say the $k$th interval. Recall that the solutions obtained using Malliavin calculus and also for ordinary differential equations are unique for each interval of time. Therefore, such a disagreement would violate these uniqueness results. □

5. LARGE DEVIATION PRINCIPLES

The theory of large deviation allow us to find probabilities of rare events that decay exponentially. Our goal is to derive large deviation principles for the solutions of LSDEs that we derived in section 4. But first, we give the formal setting for sample path large deviations.

**Definition 5.1.** Let $(\mathcal{X}, d)$ be a Polish space and $(\mu^\epsilon)_{\epsilon > 0}$ a sequence of Borel probability measures on $\mathcal{X}$. Suppose $I : \mathcal{X} \to \infty$ is a lower semicontinuous functional. Then the sequence $(\mu^\epsilon)_{\epsilon > 0}$ is said to satisfy a large deviation principle on $\mathcal{X}$ with rate function $I$ if and only if

1. (upper bound) for every closed set $F \subseteq \mathcal{X}$,
   $$\overline{\lim}_{\epsilon \to 0} \epsilon \log \mu^\epsilon(F) \leq - \inf_{x \in F} I(x),$$

2. (lower bound) and for every open set $G \subseteq \mathcal{X}$,
   $$\lim_{\epsilon \to 0} \epsilon \log \mu^\epsilon(G) \geq - \inf_{x \in G} I(x).$$

The next result states how large deviation principles are transferred under continuous transformations.

**Theorem 5.2 ([2, theorem 4.2.1]).** Let $\mathcal{X}$ and $\mathcal{Y}$ be two polish spaces, $I$ a rate function on $\mathcal{X}$, and $f$ a continuous function mapping $\mathcal{X}$ to $\mathcal{Y}$. Then the following conclusions hold.

1. For each $y \in \mathcal{Y}$,
   $$J(y) = \inf \left\{ I(x) \middle| x \in f^{-1}(y) \right\}$$
   is a rate function on $\mathcal{Y}$,

2. If $\{X_n\}$ satisfies large deviation principle on $\mathcal{X}$ with rate function $I$, then $\{f(X_n)\}$ satisfies large deviation principle on $\mathcal{Y}$ with rate function $J$.

When are large deviation principles are conserved? To answer this question, we introduce the idea of superexponential closeness.
**Definition 5.3** ([2, definition 4.2.10]). For \( \epsilon > 0 \), let \( X^\epsilon \) and \( Y^\epsilon \) be families of random variables on \((\Omega, \mathcal{F}, P)\) that take values in \( \mathcal{X} \). Then the families \( X^\epsilon \) and \( Y^\epsilon \) (and their corresponding families of laws) are said to be superexponentially close if

\[
\lim_{\epsilon \to 0} \epsilon \log P \{ d(X^\epsilon, Y^\epsilon) > \delta \} = -\infty.
\]

The following theorem says that large deviation principles are preserved for superexponentially close families.

**Theorem 5.4** ([2, theorem 4.2.13]). Suppose \( X^\epsilon \) and \( Y^\epsilon \) be superexponentially close families of random variables on \((\Omega, \mathcal{F}, P)\). Then \( X^\epsilon \) follows large deviation principle with rate function \( I \) if and only if \( Y^\epsilon \) follows large deviation principle with the same rate function \( I \).

Finally, we give an example of large deviation principle. Consider the family of process \( (\sqrt{\epsilon} W^t)_{t > 0} \), where a Wiener process \( W \) is scaled down by a parameter \( \sqrt{\epsilon} \). As \( \epsilon \to 0 \), we have \( \sqrt{\epsilon} W \to 0 \) almost surely. But at what rate does the convergence happen? This is answered by Schilder’s theorem.

**Theorem 5.5** (Schilder [2, theorem 5.2.3]). The sequence of probability measure \( \{ p_\epsilon \} \) as \( \epsilon \to 0 \) follows Large Deviation Principle on \( C_0([0, 1]) \) with rate function \( I(f) \) where

\[
I(f) = \begin{cases} \frac{1}{2} \int_0^1 |f'(t)|^2 dt & \text{if } f \in H^1 \\ \text{otherwise.} \end{cases}
\]

where \( H^1 = \{ f \in C_0([0, 1]) \mid f \text{ is absolutely continuous and } f' \in L^2[0, 1] \} \).

### 5.1. LSDEs with constant initial conditions.

Suppose \( \sigma \) and \( \gamma \) are deterministic functions of bounded variation on \([0, 1]\). Moreover, suppose \( f \in C^2(\mathbb{R}) \) is Lipschitz continuous along with \( f, f', f'' \in L^1(\mathbb{R}) \). For a fixed \( \kappa \in \mathbb{R} \), consider the family of linear stochastic differential equations with parameter \( \epsilon > 0 \) given by

\[
\begin{cases}
    dZ^\epsilon_\kappa(t) = f \left( \sqrt{\epsilon} \int_0^t \gamma_s dW_s \right) Z^\epsilon_\kappa(t) \, dt + \sqrt{\epsilon} \sigma(t) \, Z^\epsilon_\kappa(t) \, dW_t \\
    Z^\epsilon_\kappa(0) = \kappa,
\end{cases}
\tag{5.1}
\]

Using the results from section 4, the unique solutions to equation (5.1) are given by

\[
Z^\epsilon_\kappa(t) = \kappa \exp \left[ \sqrt{\epsilon} \int_0^t \sigma(s) \, dW_s - \frac{\epsilon}{2} \int_0^t \sigma(s)^2 \, ds \right.
\]

\[
+ \int_0^t f \left( \sqrt{\epsilon} \int_0^1 \gamma_u \, dW_u - \epsilon \int_s^t \gamma_u \, \sigma(u) \, du \right) \, ds \right] \tag{5.2}
\]

In order to use the continuity principle (theorem 5.2), we need the following lemma.

**Lemma 5.6.** The function \( \theta : C_0 \to C_\kappa \) defined by

\[
\theta(x) = \kappa \exp \left[ \int_0^t \sigma(s) \, dx(s) - \frac{\epsilon}{2} \int_0^t \sigma(s)^2 \, ds \right.
\]

\[
+ \int_0^t f \left( \int_0^1 \gamma_u \, dx(u) - \epsilon \int_s^t \gamma_u \, \sigma(u) \, du \right) \, ds , \nonumber
\]

is continuous in the topology induced by the canonical suprenorm.
Proof. We can write
\[ \theta(x) = \kappa \exp \left[ \phi(x) - \frac{\epsilon}{2} \int_0^t \sigma_s^2 ds + \psi(x) \right], \]
where \( \phi, \psi : C_0 \rightarrow C_0 \) is given by
\[ \phi(x) = \int_0^t \sigma(s) dx(s) = \sigma(t)x(t) - \int_0^t x(s) d\sigma(s), \]
and
\[ \psi(x) = \int_0^t f \left( \int_0^1 \gamma_u dx(u) - \epsilon \int_s^t \gamma_u \sigma(u) du \right) ds. \]

Using integration by parts,
\[ \phi(x) = \sigma(t)x(t) - \int_0^t x(s) d\sigma(s), \]
and
\[ \psi(x) = \int_0^t f \left( \gamma(1)x(1) - \int_0^1 x(s) d\gamma_s - \epsilon \int_s^t \gamma_u \sigma(u) du \right) ds. \]

Since multiplication by \( \kappa \exp \left( -\frac{\epsilon}{2} \int_0^t \sigma_s^2 ds \right) \) and exp are continuous transformations, continuity of \( \theta \) is guaranteed if we prove continuity of \( \phi \) and \( \psi \). This is what we show below. For \( \phi \), we have
\[
\| \phi(x) - \phi(y) \|_\infty = \left\| \left( \sigma(t)x(t) - \int_0^t x(s) d\sigma(s) \right) - \left( \sigma(t)y(t) - \int_0^t y(s) d\sigma(s) \right) \right\|_\infty
\leq \| \sigma(t)(x(t) - y(t)) \|_\infty + \left\| \int_0^t (x(s) - y(s)) d\sigma(s) \right\|_\infty
\leq \| \sigma \|_\infty \| x - y \|_\infty + | \sigma(t) - \sigma(0) | \| x - y \|_\infty
\leq 3 \| \sigma \|_\infty \| x - y \|_\infty,
\]
so \( \phi \) is continuous.

For \( \psi \), if \( L_f \) is the Lipschitz constant for \( f \), we get
\[
\| \psi(x) - \psi(y) \|_\infty \leq \left\| \int_0^t L_f \left[ \left( \gamma(1)x(1) - \int_0^1 x(s) d\gamma_s - \epsilon \int_s^t \gamma_u \sigma_u du \right) - \left( \gamma(1)y(1) - \int_0^1 y(s) d\gamma_s - \epsilon \int_s^t \gamma_u \sigma_u du \right) \right] ds \right\|_\infty
\leq L_f \left\| \int_0^t \left( \gamma(1)(x(1) - y(1)) - \int_0^1 (x(s) - y(s)) d\gamma_s \right) ds \right\|_\infty
\leq L_f (\| \gamma \|_\infty \| x - y \|_\infty + | \gamma(1) - \gamma(0) | \| x - y \|_\infty)
= 3L_f \| \gamma \|_\infty \| x - y \|_\infty,
\]
which proves the continuity of \( \psi \). \( \square \)

The following result now follows directly from the continuity of \( \theta \) (lemma 5.6), the continuity principle (theorem 5.2), and Schilder’s theorem (theorem 5.5).
Theorem 5.7. The laws of the solutions $Z_{\kappa}^\epsilon$ given by equation (5.2) of the family of stochastic differential equations given by equation (5.1) follow a large deviation principle on $(C_\kappa, \|\cdot\|_\infty)$ with the rate function
\[
J(y) = \inf \{ I \circ \theta^{-1}(y) \},
\]
where $\theta$ is as defined in lemma 5.6, and $I$ is the rate function given by theorem 5.5.

5.2. LSDEs with random initial conditions. Is it necessary for the family of linear stochastic differential equations equation (5.1) to start at a constant point $\kappa \in \mathbb{R}$ in order for it to have a large deviation principle? In this section, we generalize theorem 5.7 and show that we can derive a similar result under a stronger version of exponential equivalence and more restrictive conditions on the functions $f, \sigma, \gamma$.

Suppose $\sigma$ and $\gamma$ are deterministic functions of bounded variation on $[0,1]$. Moreover, suppose $f \in C^2(\mathbb{R})$ is Lipschitz continuous along with $f, f', f'' \in L^1(\mathbb{R})$. Consider the family of linear stochastic differential equations with parameter $\epsilon > 0$ given by
\[
\begin{cases}
    dZ_{\xi}(t) = f \left( \sqrt{\epsilon} \int_0^1 \gamma_s \, dW_s \right) Z_{\xi}(t) \, dt + \sqrt{\epsilon} \sigma(t) Z_{\xi}(t) \, dW_t \\
    Z_{\xi}(0) = \xi,
\end{cases}
\]
where $\xi$'s are random variables independent of the Wiener process $W$. For each $\epsilon$, just as before, the unique solution to equation (5.4) is given by
\[
Z_{\xi}(t) = \xi \exp \left[ \sqrt{\epsilon} \int_0^t \sigma(s) \, dW_s - \frac{\epsilon}{2} \int_0^t \sigma(s)^2 \, ds \right. \\
\left. + \int_0^t f \left( \sqrt{\epsilon} \int_0^1 \gamma_u \, dW_u - \epsilon \int_0^t \gamma_u \sigma(u) \, du \right) \, ds \right]
\]
(5.5)

We now state a more general large deviation principle.

Theorem 5.8. Let $\kappa \in \mathbb{R}$ and consider the family of random variables $\xi^\epsilon$ such that the following hold
\[
\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}[(\xi^\epsilon - \kappa)^2] = -\infty.
\]
Moreover, assume that the functions $f, f', \sigma, \gamma$ are all bounded. Then the laws of the solutions $Z_{\xi}^\epsilon$ given by equation (5.5) of the family of stochastic differential equations given by equation (5.4) follow a large deviation principle on $(C_\kappa, \|\cdot\|_\infty)$ with the rate function given by equation (5.3), where $\theta$ is as defined in lemma 5.6, and $I$ is the rate function given by theorem 5.5.

Proof. Let $V^\epsilon = Z_{\xi}^\epsilon - Z_{\kappa}^\epsilon$. Then $V^\epsilon$ satisfies the stochastic differential equation
\[
\begin{cases}
    dV_t^\epsilon = f \left( \sqrt{\epsilon} \int_0^1 \gamma_s \, dW_s \right) V_t^\epsilon \, dt + \sqrt{\epsilon} \sigma(t) V_t^\epsilon \, dW_t \\
    V_0^\epsilon = \xi - \kappa,
\end{cases}
\]
(5.7)
whose solution is given by
\[
V_t^\varepsilon = (\xi^\varepsilon - \kappa) \exp \left[ \sqrt{\varepsilon} \int_0^t \sigma_s \, dW_s - \frac{\varepsilon}{2} \int_0^t \sigma_s^2 \, ds 
+ \int_0^t f \left( \sqrt{\varepsilon} \int_0^1 \gamma_u \, dW_u - \varepsilon \int_0^t \gamma_u \, du \right) \, ds \right].
\]

Let \( \phi(z) = |z|^2 \) and let \( U^\varepsilon = \phi(V^\varepsilon) \). From theorem 4.3, \( U^\varepsilon \) satisfies the integral equation
\[
U^\varepsilon(t) = (\xi^\varepsilon - \kappa)^2 + 2\sqrt{\varepsilon} \int_0^t \sigma_s \, U_s^\varepsilon \, dW_s
+ \varepsilon \int_0^t \sigma_s^2 \, U_s^\varepsilon \, ds + f \left( \int_0^1 \sqrt{\varepsilon} \, \gamma_s \, dW_s \right) \int_0^t U_s^\varepsilon \, ds
+ 2\varepsilon \int_0^t \gamma_s \, \sigma_s \, U_s^\varepsilon \int_0^s f' \left( \int_0^1 \sqrt{\varepsilon} \, \gamma_v \, dW_v - \varepsilon \int_u^s \gamma_v \, \sigma_v \, dv \right) \, du \, ds.
\]

Fix \( \delta > 0 \) and let \( \tau = \inf \{ t \in [0, 1] : |V_t^\varepsilon| \geq \delta \} \). Taking expectation of the stopped process \( U_{t \wedge \tau}^\varepsilon \), we get
\[
E(U_{t \wedge \tau}^\varepsilon)
= E[(\xi^\varepsilon - \kappa)^2] + 2\sqrt{\varepsilon} E \left[ \int_0^{t \wedge \tau} \sigma_s \, U_{s \wedge \tau}^\varepsilon \, dW_s \right]
+ \varepsilon E \left[ \int_0^{t \wedge \tau} \sigma_s^2 \, U_{s \wedge \tau}^\varepsilon \, ds \right] + E \left[ f \left( \int_0^1 \sqrt{\varepsilon} \, \gamma_s \, dW_s \right) \int_0^{t \wedge \tau} U_{s \wedge \tau}^\varepsilon \, ds \right]
+ 2\varepsilon E \left[ \int_0^{t \wedge \tau} \gamma_s \, \sigma_s \, U_{s \wedge \tau}^\varepsilon \int_0^s f' \left( \int_0^1 \sqrt{\varepsilon} \, \gamma_v \, dW_v - \varepsilon \int_u^s \gamma_v \, \sigma_v \, dv \right) \, du \, ds \right].
\]

The second integral on the right-hand side is a near-martingale by theorem 3.3. Suppose \( f, f', \sigma, \gamma \) are all bounded by some \( M > 1 \). Using non-negativity of \( U^\varepsilon \) and the near-martingale optional stopping theorem (corollary 3.11), we get
\[
E(U_{t \wedge \tau}^\varepsilon) \leq E[(\xi^\varepsilon - \kappa)^2] + 0
+ \varepsilon M^2 E \left[ \int_0^{t \wedge \tau} U_{s \wedge \tau}^\varepsilon \, ds \right] + M E \left[ \int_0^{t \wedge \tau} U_{s \wedge \tau}^\varepsilon \, ds \right]
+ 2\varepsilon M^3 E \left[ \int_0^{t \wedge \tau} U_{s \wedge \tau}^\varepsilon \, ds \right]
\leq E[(\xi^\varepsilon - \kappa)^2] + (M + 2\varepsilon M^3) E \left[ \int_0^{t \wedge \tau} U_{s \wedge \tau}^\varepsilon \, ds \right].
\]

By Gronwall’s inequality, we get
\[
E(U_t^\varepsilon) = E(U_{1 \wedge \tau}^\varepsilon) \leq E[(\xi^\varepsilon - \kappa)^2] \, e^{M + 2\varepsilon M^3}.
\]

Since \( \phi(z) \) is a monotonically increasing non-negative function in \( |z| \), we use Markov’s inequality to get
\[
P\{|V_t^\varepsilon| \geq \delta\} = P\{\phi(V_t^\varepsilon) \geq \phi(\delta)\} \leq \frac{E(\phi(V_t^\varepsilon))}{\phi(\delta)} = \frac{E(U_t^\varepsilon(\tau))}{\phi(\delta)} \leq \frac{E[(\xi^\varepsilon - \kappa)^2]}{\phi(\delta)} \, e^{M + 2\varepsilon M^3}.
\]
Taking log and multiplying by $\epsilon$, we get
\[
\epsilon \log \mathbb{P}\{|V^\epsilon_\tau| \geq \delta\} \leq \epsilon \log \mathbb{E}\left[ (\xi^\epsilon - \kappa)^2 \right] - 2\epsilon \log \delta + \epsilon (M + 2\epsilon M^3).
\]
Finally, taking limit of $\epsilon \to 0$ and using equation (5.6),
\[
\lim_{\epsilon \to 0} \epsilon \log \mathbb{P}\{|V^\epsilon_\tau| \geq \delta\} = -\infty.
\]
This result allows us to say that $Z^\epsilon_\xi$ and $Z^\epsilon_\kappa$ are exponentially equivalent. Since exponentially equivalent families have the same large deviation principle due to theorem 5.4, $Z^\epsilon_\xi$ follows a large deviation principle with the same rate function given by equation (5.3).

References

[1] Wided Ayed and Hui-Hsiung Kuo. “An extension of the Itô integral”. In: Communications on Stochastic Analysis 2.3 (2008), pp. 323–333. DOI: 10.31390/cosa.2.3.05.
[2] Amir Dembo and Ofer Zeitouni. Large Deviations Techniques and Applications. Applications of mathematics. Springer, 1998. ISBN: 9780387984063. DOI: 10.1007/978-3-642-03311-7.
[3] Chii Ruey Hwang, Hui-Hsiung Kuo, and Kimiaki Saitô. “Anticipating Exponential Processes and Stochastic Differential Equations”. In: Communications on Stochastic Analysis 13.3 (2019). DOI: 10.31390/cosa.13.3.09.
[4] Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. “A general Itô formula for adapted and instantly independent stochastic processes”. In: Communications on Stochastic Analysis 10.3 (2016). DOI: 10.31390/cosa.10.3.05.
[5] Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. “Near-martingale Property of Anticipating Stochastic Integration”. In: Communications on Stochastic Analysis 11.4 (2017). DOI: 10.31390/cosa.11.4.06.
[6] Kiyosi Itô. “Extension of stochastic integrals”. In: Proceedings of the International Symposium on Stochastic Differential Equations. Kinokuniya, 1978, pp. 95–109.
[7] Thierry Jeulin and Marc Yor, eds. Grossissements de filtrations: exemples et applications. Séminaire de Calcul Stochastic 1982/83 Université Paris VI. 1st ed. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1985. ISBN: 978-3-540-15210-1. DOI: 10.1007/BFb0075765.
[8] Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. “Anticipating Linear Stochastic Differential Equations with Adapted Coefficients”. In: Journal of Stochastic Analysis 2.2 (2021). DOI: 10.31390/josa.2.2.05.
[9] Hui-Hsiung Kuo, Sudip Sinha, and Jiayu Zhai. “Stochastic Differential Equations with Anticipating Initial Conditions”. In: Communications on Stochastic Analysis 12.4 (2018). DOI: 10.31390/cosa.12.4.06.
[10] David Nualart. The Malliavin Calculus and Related Topics. 2nd ed. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 978-3-540-28329-4. DOI: 10.1007/3-540-28329-3.
[11] Peter Parczewski. “Extensions of the Hitsuda–Skorokhod Integral”. In: Communications on Stochastic Analysis 11.4 (2017). DOI: 10.31390/cosa.11.4.05.
[12] H. F. Trotter. “On the Product of Semi-Groups of Operators”. In: Proceedings of the American Mathematical Society 10.4 (1959), pp. 545–551. ISSN: 00029939, 10886826. DOI: 10.2307/2033649.
