A NEW PROOF FOR THE CONVERGENCE OF PICARD’S FILTER USING PARTIAL MALLIAVIN CALCULUS

HIDEYUKI TANAKA

Abstract. The discrete-time approximation for nonlinear filtering problems is related to both of strong and weak approximations of stochastic differential equations. In this paper, we propose a new method of proof for the convergence of approximate nonlinear filter analyzed by Jean Picard (1984), and show a more general result than the original one. For the proof, we develop an analysis of Hilbert space valued functionals on Wiener space.

1. Introduction

The aim of this paper is to determine the convergence rate of Picard’s filter for nonlinear filtering in a more general condition than that of Picard (1984), and to understand deeply why the scheme can perform with the rate. Although Picard’s filter is based on an Euler-type approximation of stochastic differential equations, the error estimate does not rely on the standard argument of strong and weak convergence of the Euler-type scheme. As seen in the following, the properties of stochastic integrals under a conditional probability make the proof of convergence much more complicated.

Let us first formulate the nonlinear filtering problem with continuous time observations. Consider a stochastic process \((X_t)_{t \geq 0}\) (often called the signal process) defined as the solution of an \(N\)-dimensional stochastic differential equation

\[
X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s
\]

with \(x \in \mathbb{R}^N\) and an \(N\)-dimensional standard Brownian motion \(B = (B_t)_{t \geq 0}\) on a probability space \((\Omega, \mathcal{F}, P)\) with a filtration \((\mathcal{F}_t)_{t \geq 0}\) satisfying the usual conditions. We observe another \(d\)-dimensional process \((Y_t)_{t \geq 0}\) (called the observation process) defined by

\[
Y_t = \int_0^t h(X_s)ds + W_t
\]

where \(W = (W_t)_{t \geq 0}\) is a \(d\)-dimensional standard Brownian motion independent of \(B\). We denote the filtrations associated to \(B\) and \(Y\) with \(P\)-null sets by \((\mathcal{F}_t^B)\) and \((\mathcal{F}_t^Y)\) respectively. The primary goal of nonlinear filtering problem is to investigate the evolution of the conditional distribution of \(X_T\) under the observation \((Y_t)_{0 \leq t \leq T}\). In other words, we are interested in computing the value

\[
E_P[g(X_T)|\mathcal{F}_T^Y].
\]

For this purpose, we consider the new probability measure \(Q\) on \(\mathcal{F}_\infty = \sigma(\cup_{t \geq 0}\mathcal{F}_t)\) under which \((Y_t)\) is a standard Brownian motion independent of \((X_t)\), and \((X_t)\) has the same law under \(P\) and \(Q\). Throughout

Date: May.21, 2014.
2000 Mathematics Subject Classification. 60G35, 60H07, 65C20.
Key words and phrases. Nonlinear filtering, Numerical approximation, Picard’s filter, Malliavin calculus.
the paper, we denote the expectation under $Q$ by $E[\cdot]$. Then the conditional expectation (2) has the expression

$$E^P[g(X_T)|\mathcal{F}^Y_T] = \frac{E[g(X_T)\Phi_T|\mathcal{F}^Y_T]}{E[\Phi_T|\mathcal{F}^Y_T]}$$

with the Radon-Nikodym derivative

$$\Phi_t = \exp\left(\sum_{j=1}^{d} \left( \int_0^t h^j(X_s)dy^j - \frac{1}{2} \int_0^t (h^j)^2(X_s)ds \right) \right).$$

This is called the Kallianpur-Striebel formula (cf. [10], [1]). We need time discretization methods in order to compute $E[g(X_T)\Phi_T|\mathcal{F}^Y_T]$ since the stochastic integral term cannot be computed exactly.

In what follows, we discuss a discrete-time approximation scheme for $\Phi_t$ under the probability measure $Q$. Let us use the notation $\|\cdot\|_P := E[|\cdot|^P]^{1/P}$. Fix $T > 0$ and $\eta(t) = t_i := iT/n$ if $t \in [iT/n, (i+1)T/n)$. We now consider an approximation by a Riemann sum for $\Phi_t$.

Jean Picard showed the following surprising result of $L^2$-convergence of $E[g(X_T)\Phi_T|\mathcal{F}^Y_T]$.

**Theorem 1.1** ([19]). Assume that $g$, $b$ and $\sigma$ are Lipschitz continuous and $h \in C^2_0(\mathbb{R}^N; \mathbb{R}^d)$. Then

$$\left\| E[g(X_T)\Phi_T|\mathcal{F}^Y_T] - E[g(X_T)\tilde{\Phi}_T|\mathcal{F}^Y_T] \right\|_2 \leq \frac{C_T}{n}. \quad (3)$$

**Remark 1.2.** The assumption $\|h\|_\infty < \infty$ can be weakened (see [19], [3]). For example, Picard ([19]) discusses the condition

$$E\left[ \exp \left( (1 + \varepsilon)TH\left( \sup_{0 \leq t \leq T} |X_t| \right) \right) \right] < \infty, \quad \text{for some} \ \varepsilon > 0$$

where

$$H(y) := \sup\left\{ \sum_{j=1}^{d} (h^j)^2(x); |x| \leq y \right\}.$$  

The convergence error (3) is related to both of weak convergence of $\mathcal{F}^Y_T$-measurable random variables and strong convergence of $\mathcal{F}^Y_T$-measurable random variables. Very roughly speaking, the order of convergence of the error is mainly from

$$\int_0^T (h(X_s) - h(X_{\eta(s)}))dY_s.$$  

We notice that the difference $h(X_s) - h(X_{\eta(s)})$ has the weak error of $O(1/n)$, but this is averaged over the trajectory of $(Y_s)$. That is why the rate of convergence is not so obvious. The proof given by Picard is quite complicated since we have to deal carefully with $\int_0^T \cdot dY_s$ under the conditional expectation $E[\cdot|\mathcal{F}^Y_T]$. In this work, we generalize the result ([3]) in terms of the regularity of $g$ (without any ellipticity condition) and $L^p$-estimates with $p > 2$ using several techniques in Malliavin calculus, and however, $h$ is basically assumed to be bounded because of the difficulty in $L^p$-moment estimates for $\Phi_T$ and $\tilde{\Phi}_T$. See the main result in Theorem 2.1 and its proof.

We review here numerical methods required for the simulation of Picard’s filter $E[g(X_T)\Phi_T|\mathcal{F}^Y_T]$. Except in some specific situations the closed-form distribution of $X_t$ is not available, and therefore we need some time discretization schemes applied to $X_t$. Let $X_t$ be a time discretization scheme for $X$, such
A NEW PROOF FOR THE CONVERGENCE OF PICARD’S FILTER

2.1. An extension of Picard’s theorem. Let us fix \( T > 0 \). Throughout the paper, the condition

\[
E^P[\Phi_T^{-1}] = 1
\]

is always assumed to define the probability measure \( Q \) on \( \mathcal{F}_T \), i.e. \( Q(A) := E^P[1_A \Phi_T^{-1}] \) for \( A \in \mathcal{F}_T \). The assumptions (A2)-(A3) introduced below imply the condition \([7]\). See Kallianpur \([8]\), Section 11.3.

We shall extend Picard’s theorem as follows. In Section 2, we state the main result which is an extension of Picard’s theorem, and shall give only the outline of the proof. In Section 3, we show the main part of the proof using infinite dimensional analysis on Wiener space, and in Section 4 we give some remarks on this research.

2. The Main result

2.1. An extension of Picard’s theorem. Let us fix \( T > 0 \). Throughout the paper, the condition

\[
E^P[\Phi_T^{-1}] = 1
\]

is always assumed to define the probability measure \( Q \) on \( \mathcal{F}_T \), i.e. \( Q(A) := E^P[1_A \Phi_T^{-1}] \) for \( A \in \mathcal{F}_T \). The assumptions (A2)-(A3) introduced below imply the condition \([7]\). See Kallianpur \([8]\), Section 11.3.

We shall extend Picard’s theorem as follows.

**Theorem 2.1.** Assume that the following conditions hold:

(A1) The function \( g : \mathbb{R}^N \to \mathbb{R} \) is a measurable function such that \( g(X_T) \in \cap_{p \geq 1} L^p(\Omega, \mathcal{F}_T, Q) \).

(A2) The coefficients \( b \) and \( \sigma \) are Lipschitz continuous.

(A3) The function \( h : \mathbb{R}^N \to \mathbb{R}^d \) is a \( C^2 \)-function of polynomial growth with all derivatives.

(A4) For every \( p \geq 1 \),

\[
\|\Phi_T\|_p + \sup_n \|\Phi_T\|_p \leq K(p, T) < \infty.
\]

Then for every \( p \geq 1 \), there exists a constant \( C = C(p, T) > 0 \) such that

\[
E^P \left[ E^P[g(X_T)\Phi_T|\mathcal{F}_T^Y] - E^P[g(X_T)\Phi_T|\mathcal{F}_T^Y] \right] \leq \frac{C}{n}.
\]

A typical example of (A4) is that \( h \) is bounded. The following corollary for the convergence of the normalized conditional expectation is an immediate consequence of Theorem 2.1.

**Corollary 2.2.** Suppose the assumptions (A1)-(A3) hold, and moreover \( h \) is assumed to be bounded. Then for every \( p \geq 1 \), there exists a constant \( C = C(p, T) > 0 \) such that

\[
E^P \left[ \left( E^P[g(X_T)|\mathcal{F}_T^Y] - \frac{E^P[g(X_T)\Phi_T|\mathcal{F}_T^Y]}{E[\Phi_T|\mathcal{F}_T^Y]} \right) \right]^{1/p} \leq \frac{C}{n}.
\]
Proof. Let $\rho_T(g) := E[g(X_T)\Phi_T | \mathcal{F}_T]$ and $\tilde{\rho}_T(g) := E[g(X_T)\tilde{\Phi}_T | \mathcal{F}_T]$. The error is expressed as

$$\frac{\rho_T(g) - \tilde{\rho}_T(g)}{\rho_T(1)} = \frac{\rho_T(g) - \tilde{\rho}_T(g)}{\rho_T(1)} + \frac{\tilde{\rho}_T(g)}{\rho_T(1)}(\rho_T(1) - \tilde{\rho}_T(1)).$$

It is possible to show from the boundedness of $h$ that the $L^p(\Omega, \mathcal{F}_T, Q)$-norms of $\Phi_T$, $\tilde{\Phi}_T$, $\rho_T(1)^{-1}$ and $\tilde{\rho}_T(1)^{-1}$ are bounded for every $p \geq 1$. Hence we obtain from Cauchy-Schwarz’s inequality

$$E^P\left[\left|\frac{\rho_T(g)}{\rho_T(1)} - \frac{\tilde{\rho}_T(g)}{\rho_T(1)}\right|^p\right]^{1/p} = E\left[\left|\frac{\rho_T(g)}{\rho_T(1)} - \frac{\tilde{\rho}_T(g)}{\rho_T(1)}\right|^p\Phi_T\right]^{1/p} \leq C_1(p, T)\|\rho_T(g) - \tilde{\rho}_T(g)\|_{2p} + C_2(p, T)\|\rho_T(1) - \tilde{\rho}_T(1)\|_{2p},$$

which proves the desired result.

Remark 2.3. For the proof of Theorem 2.1, the probability space $(\Omega, \mathcal{F}_T, Q)$ can be replaced by any other probability space on which $(X_t, Y_t)_{0 \leq t \leq T}$ has the same law. In the following, we fix the probability space so that $(B_t)_{0 \leq t \leq T}$ and $(Y_t)_{0 \leq t \leq T}$ are independent Brownian motions, and $(X_t)_{0 \leq t \leq T}$ is the solution of (1). The probability space will be assumed to be the Wiener space in Section 3.

Remark 2.4. As mentioned in the introduction, the time evolution $\rho_t(g) : t \mapsto E[g(X_t)\Phi_t | \mathcal{F}_T]$, $(g \in C^2_b)$ solves the Zakai equation

$$\rho_t(g) = \rho_0(g) + \int_0^t \rho_s(\mathcal{L}g)ds + \int_0^t \rho_s(gh^T)dY_s$$

where $\rho_0(g) = E[g(X_0)] = g(x)$ and $\mathcal{L}$ is the generator of $X$, i.e.

$$(\mathcal{L}g)(x) = \sum_{i=1}^N h^i(x)\frac{\partial g}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^N (\sigma^i \sigma^j)(x) \frac{\partial^2 g}{\partial x_i \partial x_j}(x).$$

Picard’s filter $\tilde{\rho}_t(g) : t \mapsto E[g(X_t)\tilde{\Phi}_t | \mathcal{F}_T]$ can be understood as a semigroup-type approximation (or Markov chain approximation) in the following sense. Let $X^p_t$ be a stochastic flow of the SDE (1) and $(P_t g)(x) := E[g(X^p_t)]$. Define a parameterized operator $\tilde{P}^p_t$, $y \in \mathbb{R}^d$ by

$$(\tilde{P}^p_t g)(x) := (P_t g)(x) \exp\left(\sum_{j=1}^d \left(h^j(x)y^j - \frac{1}{2}(h^j(x))^2(x)\right)\right).$$

Then we can deduce that for $t_i \leq t < t_{i+1},$

$$\tilde{\rho}_t(g) = \tilde{P}^{Y_{t_i-Y_{t_{i-1}}}}_{t_{i-1}-t_{i-1}} \circ \cdots \circ \tilde{P}^{Y_{t_1-Y_{t_0}}}_{t_{0}-t_{0}} \circ \tilde{P}^{Y_{t-Y_{t_0}}}_{t_0}(g),$$

and $\tilde{P}^{Y_{t-Y_{t_0}}}_{t_{i-1}}(g)(x)$ is a solution of the evolution equation

$$\tilde{P}^{Y_{t-Y_{t_0}}}_{t_{i-1}}(g) = g(x) + \int_{t_i}^t \tilde{P}^{Y_{s-Y_{t_0}}}_{s-t_{i}}(\mathcal{L}g)ds + \int_{t_i}^t \tilde{P}^{Y_{s-Y_{t_0}}}_{s-t_{i}}(g)h^T(x)dY_s,$$

which can be considered as the Zakai equation with the freezing coefficient $h(x)$.

2.2. Outline of proof. The proof of Theorem 2.1 is entirely different from the original one in [19]. Let us compute

$$g(X_T)\Phi_T - g(X_T)\tilde{\Phi}_T = g(X_T)\Gamma_T \sum_{j=1}^d \left( \int_0^T (h^j(X_s) - h^j(X_{\eta(s)}))dY_s^j - \frac{1}{2} \int_0^T ((h^j)^2(X_s) - (h^j)^2(X_{\eta(s)}))ds \right)$$
Proposition 2.5. Under the assumption

\[ C \]

Applying Itô’s formula to \( \zeta(X_s) \) with \( \zeta = h^j \) or \( (h^j)^2 \in C^2 \), we have

\[ \zeta(X_s) - \zeta(X_{\eta(s)}) = \int_{\eta(s)}^{s} \nabla \zeta(X_r) \sigma(X_r) dB_r + \int_{\eta(s)}^{s} (L \zeta)(X_r) dr. \]

So the error \( E[g(X_T)\Phi_T|\mathcal{F}_T^Y] - E[g(X_T)\hat{\Phi}_T|\mathcal{F}_T^Y] \) can be decomposed into four parts \( (E_i)_{1 \leq i \leq 4} \):

\[ E_1 = E \left[ g(X_T) \Gamma_T \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} \nabla (h^j)(X_r) \sigma(X_r) dB_r \right) dY^j_r \bigg| \mathcal{F}_T^Y \right] \]

\[ E_2 = E \left[ g(X_T) \Gamma_T \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} L h^j(X_r) dr \right) dY^j_r \bigg| \mathcal{F}_T^Y \right] \]

\[ E_3 = \frac{1}{2} E \left[ g(X_T) \Gamma_T \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} \nabla ((h^j)^2)(X_r) \sigma(X_r) dB_r \right) ds \bigg| \mathcal{F}_T^Y \right] \]

\[ E_4 = -\frac{1}{2} E \left[ g(X_T) \Gamma_T \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} L (h^j)^2(X_r) dr \right) ds \bigg| \mathcal{F}_T^Y \right]. \]

We are going to prove that

\[ \|E_i\|_p \leq \frac{C(i,p,T)}{n} \]

for \( p \geq 2 \) and \( 1 \leq i \leq 4 \). The estimation for \( E_1 \) is the most difficult task since \( E_1 \) includes both \( dB \) and \( dY \) parts. First, we give the estimates for \( E_2 \) and \( E_4 \).

**Proposition 2.5. Under the assumption (A1)-(A4),** for every \( p \geq 1 \), there exists a constant \( C = C(p,T) > 0 \) such that

\[ \|E_2\|_p + \|E_4\|_p \leq \frac{C}{n}. \]

**Proof.** By the assumption (A4), it holds that

\[ \|\Gamma_T\|_q \leq \|\Phi_T\|_q + \|\hat{\Phi}_T\|_q \leq K(q,T) < \infty \]

for every \( q \geq 1 \). Thus we have easily

\[ \|E_4\|_p \leq \|g(X_T)\Gamma_T\|_{2p} E \left[ \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} L (h^j)^2(X_r) dr \right) ds \right]^{2p} \]

\[ \leq \frac{C_1(p,T)}{n} \sum_{j=1}^{d} E \left[ \sup_{0 \leq r \leq T} |L (h^j)^2(X_r)|^{2p} \right]^{1/2p}. \]

This gives the estimate \( \|E_4\|_p \leq C/n. \)
We next turn to prove $\|E_2\|_p \leq C/n$. Using the Cauchy-Schwarz inequality and Burkholder-Davis-Gundy inequality, we have

$$\|E_2\|_p \leq \|g(X_T)\|_2 \sqrt{\sum_{j=1}^{d} \int_0^T \left( \int_{\eta(s)}^{t} \mathcal{L}h^j(X_r)dr \right)^2 ds} \leq C_2(p, T) \sum_{j=1}^{d} \left( \int_0^T \left( \int_{\eta(s)}^{t} \mathcal{L}h^j(X_r)dr \right)^2 ds \right)^{1/2}.$$

We can finally get the estimate

$$E \left[ \left( \int_0^T \left( \int_{\eta(s)}^{t} \mathcal{L}h^j(X_r)dr \right)^2 ds \right)^{p/2} \right]^{1/2} \leq E \left[ \sup_{0 \leq r \leq T} (|\mathcal{L}h^j(X_r)|^{2p}) \left( \int_0^T (s - \eta(s))^2 ds \right)^{p/2} \right]^{1/2} \leq \frac{C_3(p, T)}{n}.$$

\[ \square \]

3. The estimation via infinite dimensional analysis

This section is devoted to the estimates for $E_1$ and $E_3$ defined in previous. The Malliavin calculus for Hilbert space valued functionals plays an important role in the estimates.

3.1. A brief review of Malliavin calculus and Hilbert space valued martingales. Let $(\Omega, \mathcal{F}, Q)$ be a $d$-dimensional Wiener measure and $(B_t)_{0 \leq t \leq T}$ be the $d$-dimensional canonical Brownian motion on $(\Omega, \mathcal{F}, Q)$. More precisely, $\Omega = C([0, T]; \mathbb{R}^d)$, $\mathcal{F}$ is the Borel $\sigma$-field on $\Omega$, and $Q$ is the Wiener measure under which the coordinate map $t \mapsto B_t$, $B \in \Omega$ becomes a standard Brownian motion.

The Malliavin derivative $D : L^2(\Omega) \supset \text{Dom}(D) \to L^2(\Omega; L^2([0, T]; \mathbb{R}^d))$ is defined as the extension of the following closable operator for smooth Wiener functional $F$:

$$F = f \left( \int_0^T h_1(s)dB_s, \ldots, \int_0^T h_m(s)dB_s \right)$$

where $f : \mathbb{R}^m \to \mathbb{R}$ is a polynomial function and $(h_i) \subset L^2([0, T]; \mathbb{R}^d)$. Then define

$$DF := \sum_{i=1}^{m} (\partial_i f) \left( \int_0^T h_1(s)dB_s, \ldots, \int_0^T h_m(s)dB_s \right) h_i.$$ 

The Skorohod integral $\delta : L^2(\Omega; L^2([0, T]; \mathbb{R}^d)) \supset \text{Dom}(\delta) \to L^2(\Omega)$ is the adjoint operator of $D$. Let $K$ be a real separable Hilbert space. We can similarly define $D$ and $\delta$ for $K$-valued Wiener functionals. The spaces $\mathbb{D}^{1,p}(K) \subset L^p(\Omega; K)$ are defined as the Sobolev spaces induced by the derivative operator $D$ for $K$-valued Wiener functionals. For the details of the precise formulation of Malliavin calculus, we refer to [20] and [17].

We prepare some results for the Skorohod integral $\delta$ (cf. [17]).

**Lemma 3.1.** For $u(\cdot) = \sum_{i=1}^{n} F_i 1_{(t_i, t_{i+1})}(\cdot) \in L^2([0, T]; \mathbb{R}^d)$ with $F_i \in \mathbb{D}^{1,2}(\mathbb{R}^d)$, we have

$$\delta(u) = \sum_{i=1}^{n} F_i \cdot (B_{t_{i+1}} - B_{t_i}) - \sum_{i=1}^{n} \int_{t_i}^{t_{i+1}} \sum_{j=1}^{d} \mathcal{D}^{(j)}(F_i) \cdot F_{i,j} \cdot \delta^{(j)} dr.$$ 

**Lemma 3.2 (Continuity of $\delta$).** Let $p > 1$. There exists $C > 0$ such that

$$\|\delta(u)\|_p \leq C \|u\|_{\mathbb{D}^{1,p}(L^2([0, T]; \mathbb{R}^d))}$$

for every $u \in \mathbb{D}^{1,p}(L^2([0, T]; \mathbb{R}^d))$. 


We will use a kind of Fubini’s theorem below.

**Lemma 3.3.** Let \((u_s)_{0 \leq s \leq T} \in L^2([0, T]; \mathbb{D}^{1,2}_L([0, T]; \mathbb{R}^d)))\), then

\[ \int_0^T \delta(u_s(\cdot)) \, ds = \delta \left( \int_0^T u_s(\cdot) \, ds \right) \quad \text{a.s.} \tag{7} \]

**Proof.** Let \( u^k_s = \sum_{j=1}^{m_k} a^k_j B^j(s) \) with \( a^k_j \in \mathbb{D}^{1,2}_L([0, T]; \mathbb{R}^d) \) and \( B^j \in \mathcal{B}([0, T]) \) such that \( u^k \to u \) in the norm of \( L^2([0, T]; \mathbb{D}^{1,2}_L([0, T]; \mathbb{R}^d))) \) as \( k \to \infty \). Clearly we have

\[ \int_0^T \delta(u^k_s(\cdot)) \, ds = \delta \left( \int_0^T u^k_s(\cdot) \, ds \right). \]

It suffices to check the limit of both sides. By taking \( L^2 \)-norm,

\[
\left\| \int_0^T \delta(u^k_s(\cdot)) \, ds - \int_0^T \delta(u_s(\cdot)) \, ds \right\|_2^2 \leq C_1 \int_0^T \| \delta(u^k_s(\cdot) - u_s(\cdot)) \|_2^2 \, ds
\]

\[
\leq C_2 \int_0^T \| u^k_s(\cdot) - u_s(\cdot) \|_{\mathbb{D}^{1,2}}^2 \, ds
\]

and

\[
\left\| \delta \left( \int_0^T u^k_s(\cdot) \, ds \right) - \delta \left( \int_0^T u_s(\cdot) \, ds \right) \right\|_2^2 \leq C_5 \left\| \int_0^T (u^k_s(\cdot) - u_s(\cdot)) \, ds \right\|_{\mathbb{D}^{1,2}}^2
\]

\[
\leq C_4 \int_0^T \| u^k_s(\cdot) - u_s(\cdot) \|_{\mathbb{D}^{1,2}}^2 \, ds.
\]

Thus we obtain the result (7) as \( k \to \infty \). \( \Box \)

We can derive the following fundamental inequalities for Hilbert space valued martingales.

**Lemma 3.4.** Let \( M_t \) be a continuous \( K \)-valued martingale with respect to a filtration \((\mathcal{F}_t)\) which satisfies the usual conditions. Then for every \( p > 0 \), there exists positive constants \( K_p, c_p < C_p \) such that Doob’s inequality:

\[ E \left[ \sup_{0 \leq t \leq T} |M_t|^p_K \right] \leq K_p E \left[ |M_T|^p_K \right]. \]

Burkholder-Davis-Gundy’s inequality:

\[ c_p E \left[ (M)_T^{p/2} \right] \leq E \left[ \sup_{0 \leq t \leq T} |M_t|^p_K \right] \leq C_p E \left[ (M)_T^{p/2} \right]. \]

**Proof.** See e.g. [20, Theorem 3.1]. \( \Box \)

**Lemma 3.5.** If \( F \in L^p(\mathcal{F}_T^B; K) \) for some \( p \geq 2 \), then there exists a unique process \( f_s = (f^1_s, \ldots, f^d_s) \) such that \( f^i_s \) are \( K \)-valued progressively measurable processes satisfying

\[ F = E[F] + \int_0^T f_s dB_s, \]

and

\[ E \left[ \left( \int_0^T \sum_{i=1}^d |f^i_s|_K^2 \, ds \right)^{p/2} \right] \leq C_p E[|F|_K^p]. \tag{8} \]

In particular, if \( F \in \mathbb{D}^{1,2}(\mathcal{F}_T^B; K) \), then we have the so-called Clark-Ocone formula

\[ f_s(\omega) = E[D_s F|\mathcal{F}_s^B](\omega) \text{ a.e. } (s, \omega) \in [0, T] \times \Omega. \]
Proposition 3.7. Under the assumption $(K_3, Lemma 21.2)$. We can apply Lemma 3.4 with $F$ as the representation $F$ by Lemma 3.5. We obtain the representation $F = C_1(p)E[F - E[F]]_{K_3}^{p/2}$.

Proof. We check only the inequality (8) using the inequalities in Lemma 3.4:

$E\left[\left(\int_0^T \sum_{i=1}^d |f_i|_K^2 ds\right)^{p/2}\right] \leq C_1(p) E\left[\int_0^T f_s dB_s \right]^{p/2}$

$= C_1(p) E[F - E[F]]_{K_3}^{p/2}$

$\leq C_2(p) E[|F|_{K_3}^p]$. 

$\square$

3.2. Infinite dimensional Itô calculus for $E_3$. Let us define two Wiener spaces $(W_B, B(W_B), P^{WB})$ and $(W_Y, B(W_Y), P^{WY})$ on which $(B_t)_{0 \leq t \leq T}$ and $(Y_t)_{0 \leq t \leq T}$ are canonical Brownian motions respectively. From now on we specify

$(\Omega, \mathcal{F}, Q) = (W_B, B(W_B), P^{WB}) \times (W_Y, B(W_Y), P^{WY})$.

We denote by $E^{WB}$ and $E^{WY}$ the expectations under $P^{WB}$ and $P^{WY}$ respectively. Since $B$ and $Y$ are independent, we notice that $E[\cdot | \mathcal{F}_T] = E^{WB}[\cdot]$.

We now return to prove $\|E_3\|_p = O(1/n)$. The fundamental idea to get the order of convergence is as follows (see also [2]): Let $F \in L^2(W_B \times W_Y; \mathbb{R})$ and $\theta_s$ be a $\mathcal{F}_s^{W_B}$-adapted process with finite moments. We are going to give the error estimates for the type of formula

$\int_0^T f_s dB_s$.

Applying this representation to $E^{WB}[F \int_{t_i}^{t_{i+1}} \theta_s dB_s]$, we obtain a conditional duality formula

$E^{WB}[F \int_{t_i}^{t_{i+1}} \theta_s dB_s] = E^{WB}[\int_{t_i}^{t_{i+1}} f_s \theta_s ds] \in L^2(W_Y; \mathbb{R})$.

This means that it is possible to prove the convergence of $O(1/n)$ from the term $\int_{t_i}^{t_{i+1}} \cdot ds$ if $(f_s)$ has good moment estimates.

Lemma 3.6. Let $p \geq 2$ and suppose $F \in L^p(W_B \times W_Y; \mathbb{R})$ has the representation $F = E^{WB}[F] + \int_0^T f_s dB_s$ (in Lemma 3.4), then there exists a constant $C = C(p) > 0$ such that

$E\left[\left(\int_0^T |f_s|^2 ds\right)^{p/2}\right] \leq CE[|F|^p]$.

Proof. Recall that $|\cdot|$ is the norm on $\mathbb{R}^d$. We can consider the $L^2(W_Y; \mathbb{R})$-valued martingale $\int_0^T f_s dB_s$ as the $\mathbb{R}$-valued stochastic integral for the $\mathbb{R}^d$-valued process $f_s$ which is progressively measurable with respect to the enlarged filtration $\mathcal{F}_s^{W_B} \vee \mathcal{F}_s^{W_Y}$ on $(\Omega, \mathcal{F}, Q)$ through usual approximation arguments (see e.g. [3] Lemma 21.2). We can apply Lemma 3.4 with $K = \mathbb{R}$ to it. 

$\square$

Proposition 3.7. Under the assumption (A1)-(A4), for every $p \geq 1$, there exists a constant $C = C(p, T) > 0$ such that

$\|E_3\|_p \leq \frac{C}{n}$.

Proof. We prove only the one dimensional case. Let $\theta_r = \frac{1}{r} (h^2)'(X_r) \sigma(X_r)$. Applying Lemma 3.4 and 3.6 to $g(X_T)\Gamma_T$, we have a representation

$g(X_T)\Gamma_T = E^{WB}[g(X_T)\Gamma_T] + \int_0^T f_s dB_s$.
Using Itô’s formula for stochastic integrals with respect to \( B_t \), we can deduce that
\[
E^{W_B}[E^{W_B}[g(X_T)\Gamma_T]] = \int_0^T \int_{\eta(s)}^s \theta_r dB_r ds = 0
\]
and
\[
E^{W_B} \left[ \int_0^T f_s dB_s \int_0^s \int_{\eta(s)}^s \theta_r dB_r ds \right] = \int_0^T E^{W_B} \left[ \int_0^T f_s dB_s \int_{\eta(s)}^s \theta_r dB_r ds \right] dr ds = \int_0^T \int_{\eta(s)}^s f_s \theta_r dr ds ds.
\]
We notice that
\[
|E^{W_B}[f_s \theta_r]| \leq E^{W_B}[|f_r|^2]^{1/2} \sup_{0 \leq t \leq T} E^{W_B}[|\theta_r|^2]^{1/2}.
\]
Therefore the estimate \( (9) \) in Lemma 3.6 implies
\[
\left\| E^{W_B} \left[ \int_0^T \int_{\eta(s)}^s f_s \theta_r dr ds \right] \right\|_p^p \leq C_1 \left( \frac{T}{n} \right)^p E \left[ \left( \int_0^T |f_r|^{2p} dr \right)^{p/2} \right] \leq C_2 \left( \frac{T}{n} \right)^p \left\| g(X_T) \Gamma_T \right\|_p^p
\]
for some constant \( C_2 = C_2(p, T) \).

### 3.3. Partial Malliavin calculus for \( E_1 \)

In order to analyze the \( E_1 \) term, we again use the representation
\[
g(X_T) \Gamma_T = E^{W_B}[g(X_T)\Gamma_T] + \int_0^T f_s dB_s.
\]
We can then obtain
\[
E_1 = E \left[ g(X_T) \Gamma_T \sum_{j=1}^d \int_0^T \left( \int_{\eta(s)}^s \nabla (h^j)(X_r) \sigma(X_r) dB_r \right) dY_s^j \right]_{\mathcal{F}_T^Y} \]
\[
= E \left[ \int_0^T f_s dB_s \sum_{j=1}^d \int_0^T \left( \int_{\eta(s)}^s \nabla (h^j)(X_r) \sigma(X_r) dB_r \right) dY_s^j \right]_{\mathcal{F}_T^Y}.
\]
We should mention that it is impossible to apply Itô calculus to the inside of the conditional expectation since \( f_s \) is not adapted to \( \mathcal{F}_s^B \lor \mathcal{F}_s^Y \).

For this reason, instead of Itô calculus, we review partial Malliavin calculus introduced in [18]. Consider Malliavin calculus for each space of \( (W_B, B(W_B), P^{W_B}) \) and \( (W_Y, B(W_Y), P^{W_Y}) \). Let us denote the Sobolev spaces, the Malliavin derivative, and the Skorohod integral on \( (W_B, B(W_B), P^{W_B}) \) by \( \mathbb{D}^{k,p}_B, D^B_t, \delta_B \), and on \( (W_Y, B(W_Y), P^{W_Y}) \) by \( \mathbb{D}^{k,p}_Y, D^Y_t, \delta_Y \). We note that \( D^B \) and \( D^Y \) are naturally extended to \((N+d)\)-dimensional Wiener space \( (\Omega, \mathcal{F}, Q) \), and the pair \( (D^B, D^Y) \) coincides with the standard Malliavin derivative \( D : \Omega \to L^2([0,T]; \mathbb{R}^{N+d}) \) in the following sense: Let us consider an orthogonal decomposition
\[
L^2([0,T]; \mathbb{R}^{N+d}) = L^2_B \oplus L^2_Y
\]
Let \( \Pi_B \) and \( \Pi_Y \) be the projections from \( L^2([0, T]; \mathbb{R}^{N+d}) \) to \( L^2_B \) and \( L^2_Y \) respectively. Then we can define \( D^B := \Pi_B \circ D \) and \( D^Y := \Pi_Y \circ D \) on the \((N+d)\)-dimensional Wiener space \((\Omega, F, Q)\). This formulation is called the “partial” Malliavin calculus \([13, 18]\).

In this section, we realize partial Malliavin calculus using a “Sobolev space valued” Sobolev space \( \mathbb{D}^1,N^2(\mathbb{R}) \). Let us start the detailed formulation. Let \( K \) be a real separable Hilbert space and \( G \in L^2(W_B; K) \). We define \( J^B \) the projection so that \( G = E^{W_B}[G] + \int_0^T J^B(G)dB_s \). In particular, if we take \( K = \mathbb{D}^1,N^2(\mathbb{R}) \) and

\[
G \in \mathbb{D}^1,N^2 \left( \mathbb{D}^1,N^2(\mathbb{R}) \right) \subset L^2(W_B; \mathbb{D}^1,N^2(\mathbb{R})),
\]

we have by the Clark-Ocone formula

\[
J^B(G) = E^{W_B}[D^BG|\mathcal{F}_s] \in \mathbb{D}^1,N^2(\mathbb{R}).
\]

We note that \( \mathbb{D}^1,N^2(\mathbb{D}^1,N^2(\mathbb{R})) \neq \mathbb{D}^1,N^2(\mathbb{R}) \) where \( \mathbb{D}^1,N^2(\mathbb{R}) \) is the usual Sobolev space on \( W_B \times W_Y \). One notices that the space \( \mathbb{D}^1,N^2(\mathbb{D}^1,N^2(\mathbb{R})) \) is spanned by products of smooth functionals:

\[
F = f \left( \int_0^T h_1(s)dB_s, \ldots, \int_0^T h_m(s)dB_s \right) g \left( \int_0^T \theta_1(s)DY_s, \ldots, \int_0^T \theta_k(s)DY_s \right)
\]

\[
\in L^2(W_B \times W_Y; \mathbb{R}) \cong L^2(W_B; L^2(W_Y; \mathbb{R}))
\]

with \( \{h_i\}_{1 \leq i \leq m} \subset L^2([0, T]; \mathbb{R}^N), \{\theta_i\}_{1 \leq i \leq k} \subset L^2([0, T]; \mathbb{R}^d) \), real-valued \( C^1 \)-functions \( f \) and \( g \).

Let us first present auxiliary lemma which will be used in later computations.

**Lemma 3.8.** (i): For \( G \in L^2(W_B; \mathbb{D}^1,N^2(\mathbb{R})) \),

\[
D^Y E^{W_B}[G] = E^{W_B}[D^YG] \quad a.s.
\]

(ii): If \( \xi \in \mathbb{D}^1,p(B^p(L^2([0, T]; \mathbb{R}^d))) \) with some \( p \geq 2 \), then \( \int_0^T \xi_s dy_s \in \mathbb{D}^1,p(\mathbb{D}^1,N^2(\mathbb{R})) \) and

\[
D^B \left( \int_0^T \xi_s dy_s \right) = \int_0^T (D^B\xi_s)dy_s,
\]

\[
D^Y \left( \int_0^T \xi_s dy_s \right) = \xi,
\]

\[
D^Y D^B \left( \int_0^T \xi_s dy_s \right) = D^B D^Y \left( \int_0^T \xi_s dy_s \right) = D^B \xi.
\]

**Proof.** (i): We choose an approximation sequence \( (G_k) \) of the form \( G_k = \sum_{i=1}^m S_i A_i \), \( S_i \in \mathbb{D}^1,N^2(\mathbb{R}) \) and \( A_i \in B(W_B) \). For each \( k \), \( G_k \) clearly satisfies the desired equality. Thus we obtain the result using the continuity of \( D \). (ii): This is a version of the proof of \([19, \text{Proposition 1.3.8]}\), recall that \( D^B(Y_1) = 0 \). \( \square \)

For the proof of the estimate \( \|E_1\|_p \leq C/n \), we will take an approximation sequence \( (Z_\ell)_{\ell} \subset \mathbb{D}^1,N^2(\mathbb{R}) \) such that \( Z_\ell \to g(X_T) \) in \( L^{2p}(W_B) \) as \( \ell \to \infty \). The following lemma plays a key role for the estimate of \( E_1 \).

**Lemma 3.9.** Let \( p \geq 2 \) and \( Z \in \mathbb{D}^1,p(\mathbb{R}) \). Then under the assumptions \( (A2)-(A4) \), \( Z G_T(\rho) \in \mathbb{D}^1,p(\mathbb{D}^1,N^2(\mathbb{R})) \). Moreover, let \( (\theta_s) \) be a \( \mathbb{R}^d \)-valued continuous \( \mathcal{F}_s \)-progressively measurable process with
We can show by Jensen’s inequality and Lemma 3.6 that
\[ E^{W_\theta} \left( \int_0^T \sup_{0 \leq s \leq T} |D_s^B (Z_{T}(\rho)) \cdot \theta_s|^2 ds \right)^{p/2} \leq M p C \| Z_{T}(\rho) \|^p \]

Proof. We can check that \( X_t \in \mathbb{R}^1 \) under Assumption (A2). Using the chain rule of Malliavin derivative, we obtain from Lemma [38] and Assumption (A4)
\[ \Gamma_T(\rho, k) := \sum_{l=0}^k \left( \frac{\log(\Gamma_T(\rho))}{l} \right)^l \in \bigcap_{p \geq 1} \mathbb{D}_B^p(\mathbb{D}_Y^2(\mathbb{R})). \]
Thus taking the limit \( k \to \infty \), we can show that
\[ \Gamma_T(\rho) \in \bigcap_{p \geq 1} \mathbb{D}_B^p(\mathbb{D}_Y^2(\mathbb{R})), \]
which implies \( Z_{\Gamma_T(\rho)} \in \mathbb{D}_B^1(\mathbb{D}_Y^2(\mathbb{R})). \)

We now start to prove the desired inequality (12). Applying the Clark-Ocone formula (11) to \( Z_{\Gamma_T(\rho)} \), we deduce that
\[ D_s^B (Z_{\Gamma_T(\rho)}) = D_s^B (\mathbb{F}_T(\rho)) = E^{W_\theta} \left( D_s^B (Z_{\Gamma_T(\rho)}) | \mathcal{F}_T \right) \]
almost every \( (r, s, \omega) \in [0, T]^2 \times \Omega \). We notice that
\[ D_s^B ( \rho (X_r) + (1 - \rho) h(X_{\eta(r)})) \]
and then
\[ D_s^B (Z_{\Gamma_T(\rho)}) = D_s^B (\rho (X_r) + (1 - \rho) h(X_{\eta(r)})) + Z_{\Gamma_T(\rho)} D_s^B (\rho (X_r) + (1 - \rho) h(X_{\eta(r)})). \]
This formula and the Cauchy-Schwarz inequality for the conditional expectation \( E[|\mathcal{F}_r^p] \) imply
\[ E^{W_\theta} \left( |D_s^B (Z_{\Gamma_T(\rho)}) \cdot \theta_s|^2 \right) \leq 2 E^{W_\theta} \left( |D_s^B (Z_{\Gamma_T(\rho)})|^2 \right) E^{W_\theta} \left( (\rho (X_r) + (1 - \rho) h(X_{\eta(r)})) \cdot \theta_s \right)^2 \]
\[ + 2 E^{W_\theta} \left( |Z_{\Gamma_T(\rho)}|^2 \right) E^{W_\theta} \left( |D_s^B (\rho (X_r) + (1 - \rho) h(X_{\eta(r)})) \cdot \theta_s|^2 \right). \]
We refer for the reader to the basic estimate (17); for any \( q \geq 1 \),
\[ E^{W_\theta} \left( \sup_{0 \leq t \leq T} |X_t|^q \right) + E^{W_\theta} \left( \sup_{0 \leq s \leq T} |D_s^B X_t|^q \right) \leq C_1 (q, T) < \infty. \]
The above inequality allows us to show that
\[ E^{W_\theta} \left( |D_s^B (Z_{\Gamma_T(\rho)}) \cdot \theta_s|^2 \right) \leq C_2 (p, T) (E^{W_\theta} \left( |D_s^B (Z_{\Gamma_T(\rho)})|^2 \right) + E^{W_\theta} \left( |Z_{\Gamma_T(\rho)}|^2 \right)). \]
We can show by Jensen’s inequality and Lemma 3.6 that
\[ E^{W_\theta} \left( \left( \int_0^T |D_s^B (Z_{\Gamma_T(\rho)})|^2 ds \right)^{p/2} \right) \leq E \left( \left( \int_0^T |J_s^B (Z_{\Gamma_T(\rho)})|^2 ds \right)^{p/2} \right) \]
\[ \leq C_3 (p) E \left( \| Z_{\Gamma_T(\rho)} \|^p \right). \]
Using these inequalities, we obtain the constant \( C \) in the assertion. \( \square \)

We now finish the proof of the main theorem.

**Proposition 3.10.** Let the assumptions (A1)-(A4) hold. Then for every \( p \geq 2 \), there exists a constant \( C = C(p, T) > 0 \) such that
\[ \| E_1 \|_p \leq \frac{C}{n}. \]
Proof. We first define

$$E_1(\rho) := E \left[ g(X_T) \Gamma_T(\rho) \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} \nabla(h^j)(X_r) \sigma(X_r) dB_r \right) dY_s \right| \mathcal{F}_T^Y ]$$

and then

$$\|E_1\|_p \leq \int_{0}^{1} \|E_1(\rho)\|_p d\rho \leq \sup_{0 \leq \rho \leq 1} \|E_1(\rho)\|_p.$$

So it suffices to give an estimate for $\|E_1(\rho)\|_p$.

Let us define for $E \in \mathbb{D}_B^{1,2p}(\mathbb{R})$

$$E_1(\rho, Z) := E \left[ Z \Gamma_T(\rho) \sum_{j=1}^{d} \int_{0}^{T} \left( \int_{\eta(s)}^{s} \nabla(h^j)(X_r) \sigma(X_r) dB_r \right) dY_s \right| \mathcal{F}_T^Y ].$$

We shall show that

$$\|E_1(\rho, Z)\|_p \leq \frac{C}{n} \|Z \Gamma_T(\rho)\|_p,$$

and then taking an approximation sequence $(Z_\ell)_\ell \in \mathbb{D}_B^{1,2p}(\mathbb{R})$ such that $Z_\ell \to g(X_T)$ in $L^{2p}$, we have

$$\|E_1(\rho)\|_p \leq \frac{C}{n} \|g(X_T) \Gamma_T(\rho)\|_p \leq \frac{\tilde{C}(p,T)}{n},$$

which is what we want to prove.

For notational simplicity, we prove (14) only the case where $B$ and $Y$ are one dimensional Brownian motions. Let $\theta_r = (h^j)(X_r) \sigma(X_r)$. By Itô’s formula,

$$\int_{0}^{T} \int_{\eta(s)}^{s} \theta_r dB_r dY_s = \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} \theta_s dB_s \right) (Y_{t_{i+1}} - Y_{t_i}) - \int_{t_i}^{t_{i+1}} (Y_s - Y_{t_i}) \theta_s dB_s.$$ 

Set $f_s = f_s(\rho, Z) := J_s^B(\Gamma_T(\rho))$. We can deduce that

$$E^{W_B} \left[ E^{W_B} [Z \Gamma_T(\rho)] \int_{0}^{T} \int_{\eta(s)}^{s} \theta_r dB_r dY_s \right] = 0$$

and

$$E^{W_B} \left[ \int_{0}^{T} f_s dB_s \int_{0}^{T} \int_{\eta(s)}^{s} \theta_r dB_r dY_s \right]$$

$$= E^{W_B} \left[ \int_{0}^{T} f_s dB_s \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} \theta_s dB_s \right) (Y_{t_{i+1}} - Y_{t_i}) - \int_{t_i}^{t_{i+1}} (Y_s - Y_{t_i}) \theta_s dB_s \right]$$

$$= E^{W_B} \left[ \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} f_s \theta_s ds \right) (Y_{t_{i+1}} - Y_{t_i}) - \int_{t_i}^{t_{i+1}} (Y_s - Y_{t_i}) f_s \theta_s ds \right]$$

$$= \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} E^{W_B} [f_s \theta_s ds] \right) (Y_{t_{i+1}} - Y_{t_i}) - \int_{t_i}^{t_{i+1}} (Y_s - Y_{t_i}) E^{W_B} [f_s \theta_s ds].$$
By using Lemma 3.1 and the fact that $D^Y E^{W_B} [\cdot] = E^{W_B} [D^Y \cdot]$ in Lemma 3.8 it holds that

$$\sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} E^{W_B} [f_s \theta_s] ds \right) (Y_{t_{i+1}} - Y_{t_i})$$

$$= \delta_Y \left( \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} E^{W_B} [f_s \theta_s] ds \right) 1_{(t_i, t_{i+1})} (\cdot) \right) + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t_{i+1}} E^{W_B} [(D^Y f_s) \theta_s] ds dr,$$

and

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (Y_s - Y_{t_i}) E^{W_B} [f_s \theta_s] ds$$

$$= \int_0^T \left( \delta_Y (E^{W_B} [f_s \theta_s] 1_{[\eta(s), \cdot)} (\cdot)) + \int_{\eta(s)}^s E^{W_B} [(D^Y f_s) \theta_s] ds \right) dr ds$$

$$= \delta_Y \left( \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} E^{W_B} [f_s \theta_s] ds \right) 1_{(t_i, t_{i+1})} (\cdot) \right) + \int_0^T \int_{\eta(s)}^\infty E^{W_B} [(D^Y f_s) \theta_s] ds dr.$$

Here we used Lemma 3.3 in the second equality. Consequently we derive the formula

$$E^{W_B} \left[ \int_0^T f_s dB_s \int_0^T \int_{\eta(s)}^\infty \theta_s dB_s dY_s \right]$$

$$= \delta_Y \left( \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} E^{W_B} [f_s \theta_s] ds \right) 1_{(t_i, t_{i+1})} (\cdot) \right) + \int_0^T \int_{\eta(s)}^\infty E^{W_B} [(D^Y f_s) \theta_s] ds dr.$$

Using the above formula and Lemma 3.2 we finally get the estimate

$$\|E^1 (\rho, Z)\|_p \leq \frac{C_1}{n^p} \left( \int_0^T |f_s|^2 ds \right)^{p/2} + \frac{C_2}{n^p} / \left( \int_0^T \int_{\eta(s)}^\infty |E^{W_B} [(D^Y f_s) \theta_s]|^2 ds dr \right)^{p/2}$$

$$\leq \frac{C_3}{n^p} \|Z^T (\rho)\|_p^p + \frac{C_4}{n^p} \left( \int_0^T \text{ess sup}_{0 \leq r \leq T} |E^{W_B} [(D^Y f_s) \theta_s]|^2 ds \right)^{p/2}$$

Applying Lemma 3.3 to the last term, we obtain the result (14). This finishes the proof. □

4. Conclusion and some remarks on further research

The generalization discussed in the present paper consists of two parts. The first one is to determine the rate of convergence even if $g$ is irregular, and the analysis relies on the duality of stochastic integrals in Section 3.2 and a sharp estimate via partial Malliavin calculus in Section 3.3. The second one is the estimate by $L^p$-norm with $p > 2$, which is derived from the computation of the Skorohod integral and its continuity by means of Lemma 3.2.

We finally remark three problems which should be take into account in future research.

i) The author expects that the method of proof works as well if $X$ is the solution of a Lévy-driven stochastic differential equation (independent of $W$). Of course, we need several techniques on Wiener-Poisson space such as the Clark-Ocone formula and its moment estimates. In addition to the duality of the form

$$E \left[ \int_0^T f_s dB_s \int_{t_i}^{t_{i+1}} \theta_s dB_s \right] = E \left[ \int_{t_i}^{t_{i+1}} f_s \theta_s ds \right],$$
we will also use the duality for a Poisson random measure $N(dx, dt)$ of the form

$$E\left[ \int_0^T f_s(x)\tilde{N}(dx, ds) \int_{t_i}^{t_{i+1}} \theta_s(x)\tilde{N}(dx, ds) \right] = E\left[ \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}^N} f_s(x)\nu(x)dxds \right]$$

where $\tilde{N}$ is the compensated Poisson random measure and $\nu$ is the Lévy measure associated with $N(dx, ds)$. The detailed discussion is left for future work.

ii) If $B$ and $W$ are not independent (more generally, $X$ depends on $W$), we cannot apply the procedure of our proof to the error estimates. To begin with, the rate of convergence is not clear ($n^{-1/2}$ or $n^{-1}$) in that case. Similarly, the case where the coefficient $h$ depends on $Y$ is also quite complicated situation to determine the rate of convergence.

iii) Another subject of interest in this field is an asymptotic limit (central limit theorem) with rate $1/n^\alpha$ by means of

$$n^\alpha \left( E[g(X_T)\Phi_T|\mathcal{F}_T] - E[g(X_T)\tilde{\Phi}_T|\mathcal{F}_T] \right) \rightarrow G \neq 0 \text{ in law},$$

which implies the optimal rate of convergence of the conditional expectation. The result in Theorem 2.1 is not sufficient for this purpose since we merely can take $\alpha = 1 - \epsilon$ (any $\epsilon > 0$) with $G = 0$.

ACKNOWLEDGEMENT

The author would like to thank Professor Masatoshi Fujisaki for giving him the opportunity to study nonlinear filtering and motivating this research through helpful discussion. This work was supported by JSPS KAKENHI Grant Number 12J03138.

REFERENCES

[1] Bain, A., Crisan, D., Fundamentals of stochastic filtering, Springer, 2009.
[2] Clément, D., Kohatsu-Higa, A., Lamberton, D., A duality approach for the weak approximation of stochastic differential equations, Ann. Appl. Probab. 16, 2006, 1124-1154.
[3] Crisan, D., Discretizing the continuous-time filtering problem: order of convergence, In The Oxford Handbook of Nonlinear Filtering, Oxford University Press, 2011, 572-597.
[4] Crisan, D., Ghazali, S., On the convergence rates of a general class of weak approximations of SDEs, Stochastic differential equations: theory and applications, 221-248, Interdiscip. Math. Sci., 2, World Sci. Publ., Hackensack, NJ, 2007.
[5] Crisan, D., Ortiz-Latorre, S., A KL V Filter, preprint, 2012.
[6] Douset, A., Johansen, A.M., A tutorial on particle filtering and smoothing: fifteen years later, In The Oxford Handbook of Nonlinear Filtering, Oxford University Press, 2011, 656-704.
[7] Gobet, E., Pagès, G., Pham, H., Printemps, J., Discretization and simulation of Zakai equation, SIAM Journal on Numerical Analysis, 44, 2006, 2505-2538.
[8] Gordon, N.J., Salmond, D.J., Smith, A.F.M., Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proceedings F (Radar and Signal Processing), 140, 1993, 107-113.
[9] Kallianpur, G., Stochastic Filtering Theory, Springer, 1980.
[10] Kallianpur, G., Striebel, C., Estimation of stochastic systems: Arbitrary system process with additive white noise observation errors, Ann. Math. Statist. 39, 1968, 785-801.
[11] Kitagawa, G., Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models, Journal of Computational and Graphical Statistics, 5, 1996, 1-25.
[12] Kunita, H., Nonlinear filtering problems II. associated equations, In The Oxford Handbook of Nonlinear Filtering, Oxford University Press, 2011, 55-94.
[13] Kusuoka, S., Stroock, D., The partial Malliavin calculus and its application to non-linear filtering, Stochastics, 12, 1984, 83-142.
[14] Lyons, T., Victoir, N., Cubature on Wiener space, Proc. R. Soc. Lond. Ser. A 460, 2004, 169-198.
[15] Milstein, G.N., Tret’yakov, M.V., Monte Carlo methods for backward equations in nonlinear filtering, Adv. in Appl. Probab. 41, 2009, 63-100.
[16] Ninomiya, S., Victoir, N., Weak approximation of stochastic differential equations and application to derivative pricing, Appl. Math. Finance 15, 2008, 107-121.
[17] Nualart, D., Malliavin calculus and related topics, Springer, Berlin, 2006.
A NEW PROOF FOR THE CONVERGENCE OF PICARD’S FILTER

[18] Nualart, D., Zakai, M., The partial Malliavin calculus, Séminaire de Probabilités XXIII. Lecture Notes in Math. 1372, 362-381, Springer, Berlin, 1989.

[19] Picard, J., Approximation of nonlinear filtering problems and order of convergence, In Filtering and Control of Random Processes (Lecture Notes Control Inform. Sci. 61), Springer, Berlin, 1984, 219-236.

[20] Shigekawa, I., Stochastic Analysis, Translations of Mathematical Monographs Vol. 224, American Mathematical Society, Providence, RI, 2004.

[21] Talay, D., Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications, In Filtering and Control of Random Processes (Lecture Notes Control Inform. Sci. 61), Springer, Berlin, 1984, 294-313.

Department of Mathematical Sciences, Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

E-mail address: hitanaka@fc.ritsumei.ac.jp