mathematics

MDPI

Article

Donsker-Type Theorem for Numerical Schemes of Backward Stochastic Differential Equations

Yi Guo [1] and Naiqi Liu [2,*]

[1] Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan 250100, China
[2] School of Mathematics, Shandong University, Jinan 250100, China
* Correspondence: 201711879@mail.sdu.edu.cn

Abstract: This article studies the theoretical properties of the numerical scheme for backward stochastic differential equations, extending the relevant results of Briand et al. with more general assumptions. To be more precise, the Brown motion will be approximated using the sum of a sequence of martingale differences or a sequence of i.i.d. Gaussian variables instead of the i.i.d. Bernoulli sequence. We cope with an adaptation problem of Y^n by defining a new process $\hat{Y}^n$; then, we can obtain the Donsker-type theorem for numerical solutions using a similar method to Briand et al.

Keywords: Donsker-type theorem; backward stochastic differential equations; numerical schemes

MSC: 60F17; 60H35

1. Introduction

In this article we consider the following backward stochastic differential equation (BSDE for short):

$$Y_t = \xi + \int_t^T f(Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \le t \le T, \tag{1}$$

where W is standard Brownian motion on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$, $f(y,z)$ is Lipschitz continuous with y and z, and ξ is measurable with $\mathcal{F}_T$.

Research on backward stochastic differential equations can be traced back to 1973, Bismut [1] first studied linear backward stochastic differential equations. Pardoux and Peng [2] first introduced nonlinear backward stochastic differential equations, and the existence and uniqueness of the solutions under the standard Lipschitz condition were presented. For over thirty years, based on the works of Pardoux and Peng [2], backward stochastic differential equations have developed rapidly in the fields of financial mathematics, stochastic control, differential games, and numerical analysis; see El Karoui et al. [3] for details.

Unlike forward stochastic differential equations, the solution of backward stochastic differential equations is a pair of adapted processes (Y, Z); this also poses certain difficulties in numerically solving backward stochastic differential equations. Early studies mostly used stochastic differential equations to study numerical solutions for forward and backward stochastic differential equations. The system of differential equations consists of a portion of forward stochastic differential equations and a portion of backward stochastic differential equations. Ma et al. [4] proposed a four-step solution method for forward

check for updates

Academic Editor: Leonid Piterbarg

Received: 22 January 2025
Revised: 6 February 2025
Accepted: 11 February 2025
Published: 19 February 2025

Citation: Guo, Y.; Liu, N. Donsker-Type Theorem for Numerical Schemes of Backward Stochastic Differential Equations. *Mathematics* **2025**, *13*, 684. https://doi.org/10.3390/math13040684

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

and backward stochastic differential equations. Douglas et al. [5] provided a numerical solution for forward and backward stochastic differential equations using approximation techniques for partial differential equations and stochastic differential equations. Zhang [6] also provided the L^2 convergence rate of numerical schemes for forward and backward stochastic differential equations. Chevance [7] assumed that when f is irrelevant with z, by discretizing conditional expectations, the following can be obtained:

$$Y_t = \mathbb{E}[\xi + \int_t^T f(Y_s, Z_s)ds|\mathcal{F}_t]$$

This provided a numerical solution for Y and proved the convergence property. Coquet et al. [8] studied the convergence of numerical schemes for backward stochastic differential equations in the sense of function space using the convergence of filtration when f is irrelevant with z.

When f is relevant with z, Briand et al. [9] studied the Donsker-type theorem to develop a numerical scheme of backward stochastic differential equations, and investigated the following numerical scheme:

$$y_k = y_{k+1} + hf(y_k, z_k) - \sqrt{h}z_k\varepsilon_{k+1}, \quad k = n-1, \cdots, 0, \quad y_n = \xi^n,$$

where $h = T/n$, $\{\varepsilon_k\}_{1\le k\le n}$ is a set of independent and identically distributed random variables, and

$$\mathbb{P}(\varepsilon_1 = 1) = \mathbb{P}(\varepsilon_1 = -1) = \frac{1}{2}. \tag{2}$$

Here, we note $\mathcal{G}_k = \sigma(\varepsilon_1, \cdots, \varepsilon_k)$, $1 \le k \le n$ and set ξ^n as a $\mathcal{G}_n-$ measurable random variable. Meanwhile, set

$$z_k = h^{-\frac{1}{2}}\mathbb{E}[y_{k+1}\varepsilon_{k+1}|\mathcal{G}_k],$$

For $0 \le t \le T$, define

$$Y_t^n = y_{[t/h]}, \quad Z_t^n = z_{\lfloor t/h \rfloor},$$

where $\lfloor x \rfloor = (x-1)^+$ when x is an integer, and $\lfloor x \rfloor = [x]$ when x is not an integer.

Briand et al. [9] found that (Y^n, Z^n) converge to (Y, Z) in a certain sense. This is a type of Donsker-type theorem that was first used to describe the weak convergence of partial sum processes to Brownian motion on $C[0, T]$ (see Donsker [10]).

We can see that the Bernoulli distribution is critical for obtaining Donsker-type theorem for numerical solutions. The key contribution of this article is to extend the Bernoulli condition (2) to a more general condition: $\{\varepsilon_k\}_{1\le k\le n}$ is a sequence of martingale difference or i.i.d. Gaussian variables. In the former case, we need an additional assumption (see Assumption 6 below) to deal with the absence of independent and identical distributions. In the latter case, we will use an estimation Lemma 5 to obtain the UT property, which is necessary to prove the Theorem 3. Moreover, we can see in Remark 1 that the original solution Y^n might not be adapted to the filtration. Therefore, we introduce an improved process $\hat{Y}^n$ to insure the adaptation. The structure of this article is as follows: Section 2 introduces the main results of this article. The Section 3 introduces a result for the convergence of filtration, which is our tool to prove the main results of this paper. In Section 4, we provide proof of the main results of this article. We also introduce some lemmas that could be used in Section 5.

2. Main Result

The discussion in this article is based on the complete probability space with filtration $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$; W is a standard Brownian motion on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$; and fil-

tration $\{\mathcal{F}_t\}_{0\le t\le T}$, is the natural filtration generated by W, assumed to be right continuous. Consider random variable $\{\varepsilon_k^n\}_{1\le k\le n}$ defined on the probability space above, and define

$$W_t^n = \sqrt{h}\sum_{k=1}^{[t/h]}\varepsilon_k^n, \quad 0\le t\le T, \quad h=\frac{T}{n}. \tag{3}$$

Similar to the settings of Briand et al. [9], we set $y_n^n = \xi^n$, and for $k = n-1, \cdots, 0$,

$$y_k^n = y_{k+1}^n + hf(y_k^n, z_k^n) - \sqrt{h}z_k^n\varepsilon_{k+1}^n, \tag{4}$$

and

$$z_k^n = h^{-\frac{1}{2}}\mathbb{E}[y_{k+1}^n\varepsilon_{k+1}^n|\mathcal{G}_k^n],$$

where $\mathcal{G}_k^n = \sigma(\varepsilon_1^n, \ldots, \varepsilon_k^n)$.

Remark 1. *In fact, (4) is equivalent to the following equation:*

$$y_k^n = \xi^n + h\sum_{i=k}^{n-1} f(y_i^n, z_i^n) - \sum_{i=k}^{n-1} z_i^n\varepsilon_{i+1}^n. \tag{5}$$

Let $k = 0$ and subtract it from (5); we obtain

$$y_k^n = y_0^n + h\sum_{i=0}^{k-1} f(y_i^n, z_i^n) - \sum_{i=0}^{k-1} z_i^n\varepsilon_{i+1}^n.$$

However, this could not yield that y^n is adapted to the filtration $\{\mathcal{G}_k^n\}_{0\le k\le n}$ because we know nothing about y_0.

To adapt y_k^n to filtration $\mathcal{G}_k^n$, we further define $\hat{y}_k^n$ as a modified version of y_k^n:

$$\hat{y}_k^n = \mathbb{E}[y_k^n|\mathcal{G}_k^n], \tag{6}$$

and for $0\le t\le T$,

$$Y_t^n = y_{[t/h]}^n, \quad Z_t^n = z_{\lfloor t/h\rfloor}^n, \quad \hat{Y}_t^n = \hat{y}_{[t/h]}^n. \tag{7}$$

Remark 2. *Because of the definition of $[\,\cdot\,]$ and $\lfloor\,\cdot\,\rfloor$ in the previous section, Y^n and $\hat{Y}^n$ are càdlàg processes, and Z^n is a càglàd process.*

This article investigates the weak convergence problem of $(\hat{Y}, Z)$ in space $D([0,T])$. The topology of $D([0,T])$ is J_1. Related concepts are referenced in the work of Jacod and Shiryaev [11]. Here, we present a few symbols required for this article. $M^n \Rightarrow M^\infty$ means M^n converge in law under the J_1 topology. $M^n \xrightarrow{u.c.p.} M^\infty$ represent $\sup_{s\le t}|M_s^n - M_s^\infty| \xrightarrow{\mathbb{P}} 0$ for all $0\le t\le T$. If $\{\mathcal{G}_t^n\}_{0\le t\le T}$, $\{\mathcal{G}_t\}_{0\le t\le T}$ are two filtrations, then $\{\mathcal{G}_t^n\}_{0\le t\le T} \xrightarrow{w} \{\mathcal{G}_t\}_{0\le t\le T}$ means that for any $B\in\mathcal{G}_T$, $\mathbb{E}[1_B|\mathcal{G}_\cdot^n]$ probably converges to $\mathbb{E}[1_B|\mathcal{G}_\cdot]$ under topology J_1. We denote the natural filtration generated by W_n in (3) as $\{\mathcal{F}_t^n\}_{0\le t\le T}$, which is assumed to be right continuous. In the following, we recall the UT condition, which is crucial for the convergence of the stochastic integrals.

Definition 1. *A sequence of continuous $\mathbb{R}^d$-valued semimartingales $(X^n)_{n\ge1}$ is said to have the UT condition if for each $n\ge1$, the decomposition $X^n = M^n + A^n$ has*

$$\sup_n \mathbb{E}[\langle M^n, M^n\rangle_T + \int_0^T |dA_s^n|] < \infty,$$

where M_n is a local martingale and A_n is a local finite variation process.

Now, we provide the assumptions of this article.

Assumption 1. *$f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz, which means there exists a constant $K \geq 0$, such that*

$$\forall (y,z), (y',z') \in \mathbb{R}^2, \quad |f(y,z) - f(y',z')| \leq K(|y-y'| + |z-z'|);$$

Assumption 2. *ξ is $\mathcal{F}_T-$ measurable. For $n \geq 1$, ξ^n is $\mathcal{G}_n^n-$ measurable, and*

$$\mathbb{E}\left[\xi^2\right] + \sup_n \mathbb{E}\left[(\xi^n)^2\right] < \infty;$$

Assumption 3.

$$\lim_{n\to\infty} \mathbb{E}[|\xi^n - \xi|] = 0;$$

Assumption 4. $W^n \xrightarrow{u.c.p.} W$

Assumption 5. *for $0 \leq k \leq n$, $\mathbb{E}[\varepsilon_k^n | \mathcal{G}_{k-1}^n] = 0$, $\mathbb{P}(\varepsilon_k^n = 1) = \mathbb{P}(\varepsilon_k^n = -1) = \frac{1}{2}$.*

Assumption 6. *For any bounded continuous function g,*

$$\mathbb{E}[g(W_t^n)|W_s^n] \xrightarrow{\mathbb{P}} \mathbb{E}[g(W_t)|W_s], \quad 0 \leq s < t < T. \tag{8}$$

Remark 3. *At first glance, Assumption 6 may seem quite strange. However, it is a technical condition to apply Lemma 1 to proof the convergence of filtration if $\{\varepsilon_k^n\}_{1\leq k\leq n}$ is not an i.i.d. sequence.*

Assumption 7. *$\{\varepsilon_k^n\}_{1\leq k\leq n}$ is a family of independent and identically distributed Gaussian random variables, and $\mathbb{E}[\varepsilon_k^n] = 0$, $\mathbb{E}[(\varepsilon_k^n)^2] = 1$.*

Remark 4. *Actually, under Assumptions 5 and 6 or Assumption 7, we can obtain $W^n \Rightarrow W$ through Lemma 4. Since W is continuous, we also obtain that W^n converges under local uniform topology in law because of the property of the J_1 topology.*

Now, we present the main results of this article. We should note that the results are based on Assumption 5 or Assumption 7, which cannot coexist, so the conditions for the main results are divided into two independent assumptions.

Theorem 1. *Under Assumptions 1–4, with 5 and 6 or 7, we have*

$$\left(\hat{Y}^n, \int_0^\cdot Z_s^n ds\right) \Rightarrow \left(Y, \int_0^\cdot Z_s ds\right), \tag{9}$$

Furthermore,

$$\sup_{0\leq t\leq T} \left|\hat{Y}_t^n - Y_t\right|^2 + \int_0^T |Z_s^n - Z_s|^2 ds \xrightarrow{\mathbb{P}} 0. \tag{10}$$

Now, we provide an example by applying Theorem 1.

Theorem 2. *Let S^n be defined as follows:*

$$S_t^n = \sqrt{h}\sum_{k=1}^{[t/h]} \varepsilon_k, \quad 0 \le t \le T, \quad h = \frac{T}{n},$$

where ε_k are in Assumptions 5 and 6 or Assumption 7 without superscript n. Let $\{Y_t, Z_t\}_{0\le t\le T}$ be the solution of the BSDE (1) with $\xi = g(W_T)$ and $\{\hat{Y}_t^n, Z_t^n\}_{0\le t\le T}$ be defined by (7) with $\xi^n = g(S_T^n)$. If $g : \mathbb{R} \to \mathbb{R}$ is a bounded continuous function, then $\hat{Y}^n \xrightarrow{u.c.p.} Y$.

Proof. From the Donsker's theorem and Skorokhod representation theorem, we obtain that there exists a probability space, with a Brown motion W and a sequence of ε_k^n satisfying Assumptions 5 and 6 or Assumption 7 and

$$W_t^n = \sqrt{h}\sum_{k=1}^{[t/h]} \varepsilon_k^n, \quad 0 \le t \le T, \quad h = \frac{T}{n}$$

satisfy

$$\sup_{0\le t\le T} |W_t^n - W_t| \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \to \infty.$$

As g is a bounded continuous function, Assumptions 2 and 3 are trivially satisfied. Then, we could apply Theorem 1 to come to the conclusion. □

3. Convergence of Filtration

Before proving our main results, we first provide the conclusion of the convergence of filtration. This conclusion is crucial for us to prove the main results of this article.

We denote the natural filtration generated by W^n as $\{\mathcal{F}_t^n\}_{0\le t\le T}$ in (3). We assume $\{\mathcal{F}_t^n\}_{0\le t\le T}$ is right continuous. We first provide a necessary and sufficient condition for the convergence of filtration.

Lemma 1 (Coquet et al. [12])**.** *S is an adaptive continuous process on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$; $\mathcal{F}^S$ is the natural filtration generated by that, assumed to be right continuous. The following statements are equivalent:*

(1). $\mathcal{F}^n \xrightarrow{w} \mathcal{F}^S$;

(2). For all $(c_1, \ldots, c_m) \in \mathbb{R}^m$, $\left(\lambda_1^l, \ldots \lambda_m^l\right) \in \mathbb{R}^m$, and non-negative real numbers $t_1, \ldots t_k$, in J_1 topology,

$$\mathbb{E}\left[\sum_{j=1}^m c_j \exp\left\{i\sum_{l=1}^k \lambda_j^l S_{t_l}\right\} \mid \mathcal{F}_\cdot^n\right] \xrightarrow{\mathbb{P}} \mathbb{E}\left[\sum_{j=1}^m c_j \exp\left\{i\sum_{l=1}^k \lambda_j^l S_{t_l}\right\} \mid \mathcal{F}_\cdot^S\right].$$

Based on the above results, we have

Theorem 3. *Set X^n as $\mathcal{F}_T^n-$ a measurable integrable random variable; X is $\mathcal{F}_T-$ a measurable integrable random variable:*

$$M_t^n = \mathbb{E}[X^n|\mathcal{F}_t^n], \quad M_t = \mathbb{E}[X|\mathcal{F}_t]. \tag{11}$$

Assume

(a)

$$\mathbb{E}\left[X^2\right] + \sup_n \mathbb{E}\left[(X^n)^2\right] < \infty,$$

(b)

$$\lim_{n\to\infty} \mathbb{E}[|X^n - X|] \to 0.$$

Under Assumptions 4–6 or 4 and 7, we have

$$(W^n, M^n, [M^n, M^n], [M^n, W^n]) \xrightarrow{u.c.p.} (W, M, [M, M], [M, W]).$$

Proof. From Remark 1.(2) in Coquet et al. [12], if we can prove $\mathcal{F}^n \xrightarrow{w} \mathcal{F}$, together with condition (b), we could obtain $M^n \xrightarrow{u.c.p.} M$. Hence, now we prove $\mathcal{F}^n \xrightarrow{w} \mathcal{F}$. According to Assumptions 4–6, we first prove for fix $t > 0$,

$$\mathbb{E}\left[\sum_{j=1}^{m} c_j \exp\left\{i\sum_{l=1}^{k}\lambda_j^l W_{t_l}\right\} \mid \mathcal{F}_t^n\right] \xrightarrow{\mathbb{P}} \mathbb{E}\left[\sum_{j=1}^{m} c_j \exp\left\{i\sum_{l=1}^{k}\lambda_j^l W_{t_l}\right\} \mid \mathcal{F}_t\right].$$

According to Assumption 4, we only need to prove

$$\mathbb{E}\left[\sum_{j=1}^{m} c_j \exp\left\{i\sum_{l=1}^{k}\lambda_j^l W_{t_l}^n\right\} \mid \mathcal{F}_t^n\right] \xrightarrow{\mathbb{P}} \mathbb{E}\left[\sum_{j=1}^{m} c_j \exp\left\{i\sum_{l=1}^{k}\lambda_j^l W_{t_l}\right\} \mid \mathcal{F}_t\right].$$

For $t \in [t_{p-1}, t_p)$,

$$\begin{aligned} F_t &:= \mathbb{E}\left[\sum_{j=1}^{m} c_j \exp\left(i\sum_{l=1}^{p-1}\lambda_j^l W_{t_l} + i\sum_{l=p}^{k}\lambda_j^l W_{t_l}\right) \mid \mathcal{F}_t\right] \\ &= \sum_{j=1}^{m} c_j \prod_{l<p} \exp\left(i\lambda_j^l W_{t_l}\right) \cdot \prod_{l\geq p} \mathbb{E}\left[\exp\left(i\lambda_j^l W_{t_l}\right) \mid \mathcal{F}_t\right], \\ F_t^n &:= \sum_{j=1}^{m} c_j \prod_{l<p} \exp\left(i\lambda_j W_{t_i}^n\right) \cdot \prod_{l\geq p} \mathbb{E}\left[\exp\left(i\lambda_j^l W_{t_i}^n\right) \mid \mathcal{F}_t^n\right]. \end{aligned}$$

According to Assumptions 4 and 6, we have $F_t^n \xrightarrow{\mathbb{P}} F_t$. Therefore, we have proved the pointwise convergence in probability of F_t^n, which is uniformly integrable for fix t to F_t, which is a continuous martingale. Then, using Lemma 1 and Proposition 1.2 in Aldous [13], we can obtain

$$\mathcal{F}^n \xrightarrow{w} \mathcal{F}. \tag{12}$$

If Assumptions 4 and 7 hold, the convergence of filtration is a direct corollary of Proposition 2 of [12] since $\{W^n\}_{n\geq 1}$ is a sequence of processes with independent increments. Therefore we have $M^n \xrightarrow{u.c.p.} M$.

Next, we proceed to prove $[M^n, M^n] \xrightarrow{u.c.p.} [M, M]$ and $[W^n, W^n] \xrightarrow{u.c.p.} [W, W]$. Under Assumptions 4–6 or 4 and 7, taking advantage of the Doob inequality and condition (a),

$$\sup_n \mathbb{E}\left[\sup_{0\leq t\leq T} \Delta|M_t^n|\right] < \infty. \tag{13}$$

From Lemma 7, we could obtain

$$[M^n, M^n] \xrightarrow{u.c.p.} [M, M]. \tag{14}$$

For W^n, if Assumptions 4–6 hold, this trivially satisfies

$$\sup_n \mathbb{E}\left[\sup_{0\leq t\leq T} \Delta|W_t^n|\right] < \infty. \tag{15}$$

On the other hand, if Assumptions 4 and 7 hold, noting that $\sqrt{\log n/n}$ is bounded and

$$\sup_{0\le t\le T}\Delta|W_t^n| = \sup_{1\le k\le n}\sqrt{h}\varepsilon_k^n = \sqrt{\frac{T}{n}}\sup_{1\le k\le n}\varepsilon_k^n,$$

it is easy to verify (15) using Lemma 5. Hence, using the same proof, we have

$$[W^n, W^n] \xrightarrow{u.c.p.} [W, W].$$

Combining this with (14), we have

$$[M^n, W^n] \xrightarrow{u.c.p.} [M, W].$$

Since the convergence likely has addictive properties, we have proved the theorem. □

Using the above theorem, we can obtain the following theorem, similarly to Corollary 3.2 in the work of Briand et al. [9]. The proof is the same as that corollary, so we present only the theorem, without the proof.

Theorem 4. *Under the condition of Theorem 3, a predictable process exists,* $(Z_t^n)_{0\le t\le T}$ *and* $(Z_t)_{0\le t\le T}$*, such that*

$$\forall t\in[0,T],\quad M_t^n = \mathbb{E}[X^n] + \int_0^t Z_s^n dW_s^n,\quad M_t = \mathbb{E}[X] + \int_0^t Z_s dW_s,$$

and

$$\int_0^T (Z_t^n - Z_t)^2 dt \xrightarrow{\mathbb{P}} 0.$$

4. Proof of Theorem 1

Equation (9) is a direct corollary of (10); therefore, we proceed to prove (10). Set $(Y_t^{\infty,p+1}, Z_t^{\infty,p+1})$ as the solution of the backward stochastic differential equation:

$$Y_t^{\infty,p+1} = \xi + \int_t^T f\left(Y_s^{\infty,p}, Z_s^{\infty,p}\right)ds - \int_t^T Z_s^{\infty,p+1}dW_s,\quad 0\le t\le T, \tag{16}$$

and define

$$y_k^{n,p+1} = y_{k+1}^{n,p+1} + hf\left(y_k^{n,p}, z_k^{n,p}\right) - \sqrt{h}z_k^{n,p+1}\varepsilon_{k+1}^n,\quad k = n-1,\cdots,0,\quad y_n^{n,p+1} = \xi^n. \tag{17}$$

Set $Y^{\infty,0}=0, Z^{\infty,0}=0, y^{n,0}=0, z^{n,0}=0$ we have

$$\hat{Y}^n - Y = (\hat{Y}^n - Y^n) + (Y^n - Y^{n,p}) + (Y^{n,p} - Y^{\infty,p}) + (Y^{\infty,p} - Y), \tag{18}$$

$$Z^n - Z = (Z^n - Z^{n,p}) + (Z^{n,p} - Z^{\infty,p}) + (Z^{\infty,p} - Z), \tag{19}$$

where $0\le t\le T$, $Y_t^{n,p} = y_{[t/h]}^{n,p}$ and $Z_t^{n,p} = z_{\lfloor t/h\rfloor}^{n,p}$.

Using the Picard iteration, we can obtain $(Y^{\infty,p}, Z^{\infty,p}) \to (Y, Z)$ in the sense of (10), which is a classic conclusion and can be found in the work of Zhang [14]. Therefore, only the convergence of other terms needs to be considered. Hence, now we need a only few lemmas.

Lemma 2. *Under Assumptions 5 and 6 or 7, there exists* n_0*, such that when* $n > n_0$*,* $p\to\infty$*,*

$$\mathbb{E}[\sup_{0\le t\le T}|Y_t^n - Y_t^{n,p}|^2 + \int_0^T (Z_t^n - Z_t^{n,p})^2 dt] \to 0.$$

Proof. First assume we have Assumptions 5 and 6; then, we need to prove that there exists $\alpha > 1$, $n_0 > 1$ such that $n > n_0$, for all $p \geq 1$,

$$\| \left(y^{n,p+1} - y^{n,p}, z^{n,p+1} - z^{n,p} \right) \|_\alpha^2 \leq \frac{2}{3} \| \left(y^{n,p} - y^{n,p-1}, z^{n,p} - z^{n,p-1} \right) \|_\alpha^2,$$

where

$$\| \left(y^{n,p+1} - y^{n,p}, z^{n,p+1} - z^{n,p} \right) \|_\alpha^2 = \mathbb{E} \left[\sup_{0 \leq k \leq n} \alpha^{kh} \left| y_k^{n,p+1} - y_k^{n,p} \right|^2 + h \sum_{k=0}^{n-1} \alpha^{kh} \left| z_k^{n,p+1} - z_k^{n,p} \right|^2 \right].$$

For the convenience of narration, we use y, z instead of $y^{n,p+1} - y^{n,p}, z^{n,p+1} - z^{n,p}$, and use u, v instead of $y^{n,p} - y^{n,p-1}, z^{n,p} - z^{n,p-1}$. Set $\beta > 1$; since $y_n = 0$, we have $k = 0, \ldots, n-1$,

$$\beta^k y_k^2 = \sum_{i=k}^{n-1} \beta^i y_i^2 - \beta^{i+1} y_{i+1}^2 = (1-\beta) \sum_{i=k}^{n-1} \beta^i y_i^2 + \beta \sum_{i=k}^{n-1} \beta^i \left(y_i^2 - y_{i+1}^2 \right). \tag{20}$$

Set

$$y_i^2 - y_{i+1}^2 = 2y_i(y_i - y_{i+1}) - (y_i - y_{i+1})^2, \tag{21}$$

considering relation (17), we obtain

$$y_i - y_{i+1} = h \left\{ f\left(y_i^{n,p}, z_i^{n,p} \right) - f\left(y_i^{n,p-1}, z_i^{n,p-1} \right) \right\} - \sqrt{h} z_i \varepsilon_{i+1}^n. \tag{22}$$

Since f is Lipschitz continuous, for $\nu > 0$,

$$2y_i \left\{ f\left(y_i^{n,p}, z_i^{n,p} \right) - f\left(y_i^{n,p-1}, z_i^{n,p-1} \right) \right\} \leq 2K|y_i|(|u_i| + |v_i|) \leq 2\left(K^2/\nu \right) y_i^2 + \nu \left(u_i^2 + v_i^2 \right). \tag{23}$$

Furthermore, squaring the last term in (22) and using Assumption 5, we derive

$$h z_i^2 \leq 2(y_i - y_{i+1})^2 + 4K^2 h^2 \left(u_i^2 + v_i^2 \right). \tag{24}$$

From (22) and (23), it is easy to obtain

$$2 \sum_{i=k}^{n-1} \beta^i y_i (y_i - y_{i+1}) \leq 2K^2 (h/\nu) \sum_{i=k}^{n-1} \beta^i y_i^2 + \nu h \sum_{i=k}^{n-1} \beta^i \left(u_i^2 + v_i^2 \right) - 2\sqrt{h} \sum_{i=k}^{n-1} \beta^i y_i z_i \varepsilon_{i+1}^n. \tag{25}$$

On the other hand, using (24), we know

$$- \sum_{i=k}^{n-1} \beta^i (y_i - y_{i+1})^2 \leq -(h/2) \sum_{i=k}^{n-1} \beta^i z_i^2 + 2K^2 h^2 \sum_{i=k}^{n-1} \beta^i \left(u_i^2 + v_i^2 \right).$$

Set $\rho = \left(\nu + 2K^2 h \right) \beta h$; according to (20) and (21) and the two inequalities shown above, we have

$$\begin{aligned}
&\beta^k y_k^2 + \beta(h/2) \sum_{i=k}^{n-1} \beta^i z_i^2 \\
&= (1-\beta) \sum_{i=k}^{n-1} \beta^i y_i^2 + 2\beta \sum_{i=k}^{n-1} \beta^i y_i (y_i - y_{i+1}) - \beta \sum_{i=k}^{n-1} \beta^i (y_i - y_{i+1})^2 + \beta(h/2) \sum_{i=k}^{n-1} \beta^i z_i^2 \\
&\leq \left(1 - \beta + 2K^2 h \beta / \nu \right) \sum_{i=k}^{n-1} \beta^i y_i^2 - 2\beta\sqrt{h} \sum_{i=k}^{n-1} \beta^i y_i z_i \varepsilon_{i+1}^n + \rho \sum_{i=k}^{n-1} \beta^i \left(u_i^2 + v_i^2 \right).
\end{aligned}$$

Therefore, if $1-\beta+2K^2h\beta/\nu \le 0$, the above inequality implies, for $k=0,\dots,n-1$,

$$\beta^k y_k^2+\beta(h/2)\sum_{i=k}^{n-1}\beta^i z_i^2 \le \rho\sum_{i=0}^{n-1}\beta^i\left(u_i^2+v_i^2\right)-2\beta\sqrt{h}\sum_{i=k}^{n-1}\beta^i y_i z_i \varepsilon_{i+1}^n. \tag{26}$$

Since ε_k^n is a martingale difference sequence, using the expectation, we have

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i z_i^2\right] \le 2\left(\nu+2K^2h\right)\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2+v_i^2\right)\right]. \tag{27}$$

Furthermore, using (26), we can obtain

$$\sup_{0\le k\le n}\beta^k y_k^2 \le \rho\sum_{i=0}^{n-1}\beta^i\left(u_i^2+v_i^2\right)+4\beta\sqrt{h}\sup_{0\le k\le n-1}\left|\sum_{i=0}^{k}\beta^i y_i z_i\varepsilon_{i+1}^n\right|,$$

Because of Assumption 5, we can easily verify that $\{\sum_{i=0}^{k}\beta^i y_i z_i\varepsilon_{i+1}^n\}_{0\le k\le n-1}$ is a martingale. Using the Burkholder–Davis–Gundy inequality, $\sqrt{ab}\le\frac{a+b}{2}$, it follows that

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le k\le n}\beta^k y_k^2\right] &\le \rho\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2+v_i^2\right)\right]+C\sqrt{h}\beta\mathbb{E}\left[\left(\sum_{i=0}^{n-1}\beta^{2i}y_i^2z_i^2\right)^{1/2}\right]\\ &\le \rho\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2+v_i^2\right)\right]+C^2\beta^2(h/2)\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i z_i^2\right]+\frac{1}{2}\mathbb{E}\left[\sup_{0\le k\le n}\beta^k y_k^2\right].\end{aligned}$$

Using (27), we have

$$\mathbb{E}\left[\sup_{0\le k\le n}\beta^k y_k^2+h\sum_{i=0}^{n-1}\beta^i z_i^2\right] \le \lambda\mathbb{E}\left[\sup_{0\le k\le n}\beta^k u_k^2+h\sum_{i=0}^{n-1}\beta^i v_i^2\right],$$

where $\lambda = 2\left(\nu+2K^2h\right)\left(1+\beta+C^2\beta^2\right)(1\vee T)$, set $1-\beta+2K^2h\beta/\nu \le 0$. Choose ν, such that $2\nu(2+C^2)(1\vee T)=1/2$. Make $\beta=\alpha^h$, where $\alpha\ge 1$; we hope that $1-\beta+2K^2h\beta/\nu\le 0$, so α should be greater than or equal to $\exp(-h^{-1}\log(1-2K^2h/\nu))$. When $n\to\infty$, this equation tends towards $\exp(2K^2/\nu)$. Hence, we set $\alpha=\exp(2K^2/\nu)$. On the other hand, when $n\to\infty$, $\lambda\to 2\nu(2+C^2)(1\vee T)=1/2$. Based on the analysis above, there exists n_0, such that when $n\ge n_0$,

$$\mathbb{E}\left[\sup_{0\le k\le n}\beta^k y_k^2+h\sum_{i=0}^{n-1}\beta^i z_i^2\right] \le \frac{2}{3}\mathbb{E}\left[\sup_{0\le k\le n}\beta^k u_k^2+h\sum_{i=0}^{n-1}\beta^i v_i^2\right],$$

Now, we prove that for $n\ge n_0$,

$$\mathbb{E}[\sup_{0\le k\le n-1}|y_k^{n,1}|^2+h\sum_{i=0}^{n-1}|z_i^{n,1}|^2]<\infty.$$

In fact, since $y^{n,0}=0, z^{n,0}=0$,

$$y_k^{n,1}=\xi^n+h\sum_{i=k}^{n-1}f(0,0)-\sqrt{h}\sum_{i=k}^{n-1}z_i^{n,1}\varepsilon_{i+1}^n,$$

then,

$$\sup_{0\le k\le n-1}|y_k^{n,1}|^2 \le C[(\xi^n)^2+T^2f^2(0,0)-h\sum_{i=0}^{n}(z_i^{n,1})^2]$$

$$\leq (C \vee 1)[(\xi^n)^2 + T^2 f^2(0,0) - h\sum_{i=0}^{n}(z_i^{n,1})^2],$$

so we have

$$\mathbb{E}[\sup_{0\leq k\leq n-1}|y_k^{n,1}|^2 + (C\vee 1)h\sum_{i=0}^{n}(z_i^{n,1})^2] \leq (C\vee 1)[(\xi^n)^2 + T^2f^2(0,0)] < \infty.$$

According to Assumption 7, in fact, (20)–(23) and (25) are similar to what can be obtained. Meanwhile, it should be noted that (24) can be estimated as follows:

$$hz_i^2(\varepsilon_{i+1}^n)^2 \leq 2(y_i - y_{i+1})^2 + 4K^2h^2\left(u_i^2 + v_i^2\right).$$

Hence,

$$-\sum_{i=k}^{n-1}\beta^i(y_i - y_{i+1})^2 \leq -(h/2)\sum_{i=k}^{n-1}\beta^i z_i^2(\varepsilon_{i+1}^n)^2 + 2K^2h^2\sum_{i=k}^{n-1}\beta^i\left(u_i^2 + v_i^2\right).$$

Combining (25), we have

$$\beta^k y_k^2 + \beta(h/2)\sum_{i=k}^{n-1}\beta^i z_i^2(\varepsilon_{i+1}^n)^2 \leq \left(1 - \beta + 2K^2h\beta/\nu\right)\sum_{i=k}^{n-1}\beta^i y_i^2 - 2\beta\sqrt{h}\sum_{i=k}^{n-1}\beta^i y_i z_i\varepsilon_{i+1}^n + \rho\sum_{i=k}^{n-1}\beta^i\left(u_i^2 + v_i^2\right). \tag{28}$$

Taking the mathematical expectations from both sides, for $k = 0$, since ε_i^n is independent and identically distributed, we can also obtain (27). From (28), we have

$$\mathbb{E}[\sup_{0\leq k\leq n}\beta^k y_k^2] \leq \mathbb{E}\left[\rho\sum_{i=0}^{n-1}\beta^i\left(u_i^2 + v_i^2\right) + 4\beta\sqrt{h}\sup_{0\leq k\leq n-1}\left|\sum_{i=0}^{k}\beta^i y_i z_i\varepsilon_{i+1}^n\right|\right],$$

According to the Burkholder–Davis–Gundy inequality, $\sqrt{ab} \leq \frac{a+b}{2}$ and Assumption 7, we have:

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq k\leq n}\beta^k y_k^2\right] &\leq \rho\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2 + v_i^2\right)\right] + C\sqrt{h}\beta\mathbb{E}\left[\left(\sum_{i=0}^{n-1}\beta^{2i}y_i^2z_i^2(\varepsilon_{i+1}^n)^2\right)^{1/2}\right] \\
&\leq \rho\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2 + v_i^2\right)\right] + C^2\beta^2(h/2)\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i z_i^2\right] + \frac{1}{2}\mathbb{E}\left[\sup_{0\leq k\leq n}\beta^k y_k^2(\varepsilon_{i+1}^n)^2\right] \\
&= \rho\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i\left(u_i^2 + v_i^2\right)\right] + C^2\beta^2(h/2)\mathbb{E}\left[\sum_{i=0}^{n-1}\beta^i z_i^2\right] + \frac{1}{2}\mathbb{E}\left[\sup_{0\leq k\leq n}\beta^k y_k^2\right]
\end{aligned}$$

The remaining parts are similar and can be obtained. □

Lemma 3. *Under Assumptions 1–4, with 5 and 6 or 7, for $p > 0$, when $n \to \infty$, we have*

$$(Y^{n,p}, Z^{n,p}) \to (Y^{\infty,p}, Z^{\infty,p}).$$

Proof. We will use induction to prove that if $(Y^{n,p}, Z^{n,p}) \to (Y^{\infty,p}, Z^{\infty,p})$,, then we have $(Y^{n,p+1}, Z^{n,p+1}) \to (Y^{\infty,p+1}, Z^{\infty,p+1})$. For convenience of notation, we omit the p and consider everything in a continuous setting; hence, by virtue of Remark 2, the situation in (16) and (17) changes to

$$\begin{aligned}
Y_t' &= \xi + \int_t^T f(Y_s, Z_s)ds - \int_t^T Z_s' dW_s, \quad 0 \leq t \leq T, \\
Y_t'^n &= \xi^n + \int_t^T f(Y_{s-}^n, Z_s^n)dA_s^n - \int_t^T Z_s'^n dW_s^n, \quad 0 \leq t \leq T,
\end{aligned}$$

where $A_s^n = [s/h]h$. Suppose we already know that $\{Y_t^n, Z_t^n\}_{0\le t\le T}$ converges to $\{Y_t, Z_t\}_{0\le t\le T}$ and we only need to prove that $\{Y_t'^n, Z_t'^n\}_{0\le t\le T}$ converges to $\{Y_t', Z_t'\}_{0\le t\le T}$. Now, we can use induction on p. For $p = 0$, $Y = Z = Y^n = Z^n = 0$, set

$$M_t^n = Y_t'^n + \int_0^t f(0,0)dA_s^n, \quad 0 \le t \le T,$$

which satisfies

$$M_t^n = M_0^n + \int_0^t Z_s'^n dW_s^n,$$

thus, M^n is a $\mathcal{F}^n$-martingale. Since $Y_T'^n = \xi^n$, we have

$$M_t^n = \mathbb{E}(M_T^n \mid \mathcal{F}_t^n), \quad M_T^n = \xi^n + \int_0^T f(0,0)dA_s^n.$$

To use Theorems 3 and 4, we need to prove the convergence of M_T^n under L^1. Since we have

$$\begin{aligned}\left|M_T^n - Y_T' - \int_0^T f(0,0)ds\right| &= \left|\xi^n - \xi + \int_0^T f(0,0)dA_s^n - \int_0^T f(0,0)ds\right| \\ &\le |\xi^n - \xi| + \left|\int_0^T f(0,0)dA_s^n - \int_0^T f(0,0)ds\right|,\end{aligned}$$

from Assumption 3, the equation above converges under L^1. Applying Theorems 3 and 4, we have $M^n \xrightarrow{u.c.p.} M$, where

$$M_t = Y_t' + \int_0^t f(0,0)ds$$

and

$$\int_0^T (Z_s'^n - Z_t')^2 dt \xrightarrow{\mathbb{P}} 0.$$

Therefore, we can obtain

$$\sup_{0\le t\le T}|M_t^n - M_t| + \int_0^T |Z_s'^n - Z_s'|^2 ds \xrightarrow{\mathbb{P}} 0.$$

Our task is to prove

$$\sup_{0\le t\le T}|Y_t'^n - Y_t'| + \int_0^T |Z_s'^n - Z_s'|^2 ds \xrightarrow{\mathbb{P}} 0,$$

Hence, we only need to prove

$$\sup_{0\le t\le T}\left|\int_0^t f(0,0)dA_s^n - \int_0^t f(0,0)ds\right| \xrightarrow{\mathbb{P}} 0,$$

where $A_s^n = [s/h]h$, and the result clearly follows from the Riemann integration. For the case of $p \ge 1$, using a similar analysis to Briand et al. [9], set

$$M_t^n = Y_t'^n + \int_0^t f(Y_{s-}^n, Z_s^n)dA_s^n, \quad 0 \le t \le T,$$

which satisfies

$$M_t^n = M_0^n + \int_0^t Z_s'^n dW_s^n,$$

Hence, M^n is a $\mathcal{F}^n$-martingale. Since $Y_T^n = Y_T'^n = \xi^n$, we have

$$M_t^n = \mathbb{E}(M_T^n \mid \mathcal{F}_t^n), \quad M_T^n = Y_T'^n + \int_0^T f(Y_{s-}^n, Z_s^n) dA_s^n.$$

To use Theorems 3 and 4, we need to prove the convergence of M_T^n under L^1. Noting that Y^n and Z^n are piecewise constant and Y, Z are continous, we have

$$\begin{aligned} \left| M_T^n - Y_T - \int_0^T f(Y_s, Z_s) ds \right| &\leq |Y_T'^n - Y_T| + \int_0^T |f(Y_{s-}^n, Z_s^n) - f(Y_s, Z_s)| ds \\ &\leq (1 + KT) \sup_{0 \leq t \leq T} |Y_t^n - Y_t| + K \int_0^T |Z_s^n - Z_s| ds. \end{aligned} \tag{29}$$

The equation above converges under L^1 because of the boundedness of L^2; e refer to [14], Theorem 4.2.1. Thanks to Theorems 3 and 4, using the same step as used in $p = 0$, we again have

$$\sup_{0 \leq t \leq T} |M_t^n - M_t| + \int_0^T \left| Z_s'^n - Z_s' \right|^2 ds \xrightarrow{\mathbb{P}} 0.$$

where

$$M_t = Y_t' + \int_0^t f(Y_t, Z_t) ds$$

Now, we proceed to show

$$\sup_{0 \leq t \leq T} \left| Y_t'^n - Y_t' \right| + \int_0^T \left| Z_s'^n - Z_s' \right|^2 ds \xrightarrow{\mathbb{P}} 0,$$

so we only need to prove

$$\sup_{0 \leq t \leq T} \left| \int_0^t f(Y_{s-}^n, Z_s^n) dA_s^n - \int_0^t f(Y_s, Z_s) ds \right| \xrightarrow{\mathbb{P}} 0,$$

However, we have just proved this in (29), which finishes the proof. □

Now, we prove Theorem 1. Due to the two lemmas above, we only need to show

$$\sup_{0 \leq t \leq T} |\hat{Y}_t^n - Y_t^n| \xrightarrow{\mathbb{P}} 0,$$

This can be used to show

$$\sup_{0 \leq t \leq T} |\mathbb{E}[Y_t^n | \mathcal{F}_t^n] - Y_t^n| \xrightarrow{\mathbb{P}} 0.$$

In fact,

$$\begin{aligned} \sup_{0 \leq t \leq T} |\mathbb{E}[Y_t^n | \mathcal{F}_t^n] - Y_t + Y_t - Y_t^n| &\leq \sup_{0 \leq t \leq T} |\mathbb{E}[Y_t^n | \mathcal{F}_t^n] - Y_t| + \sup_{0 \leq t \leq T} |Y_t - Y_t^n| \\ &\leq \sup_{0 \leq t \leq T} |\mathbb{E}[Y_t^n - Y_t | \mathcal{F}_t^n] + \sup_{0 \leq t \leq T} |\mathbb{E}[Y_t | \mathcal{F}_t^n] - Y_t| \\ &+ \sup_{0 \leq t \leq T} |Y_t - Y_t^n|. \end{aligned}$$

From Lemmas 2 and 3, the third term converged in probability. For the first term, thanks to Doob's inequality, we have

$$\lim_{n\to\infty} \mathbb{P}(\sup_{0\le t\le T} |\mathbb{E}[Y_t^n - Y_t|\mathcal{F}_t^n] \ge \varepsilon) \le \lim_{n\to\infty} \mathbb{P}(\sup_{0\le s\le T} \mathbb{E}[\sup_{0\le t\le T} |Y_t^n - Y_t|\mathcal{F}_s^n] \ge \varepsilon)$$
$$\le \lim_{n\to\infty} \varepsilon^{-1}\mathbb{E}[\sup_{0\le t\le T} |Y_t^n - Y_t|] = 0,$$

where the last equality is because of the L^2 boundedness of $\sup_{0\le t\le T} |Y_t^n - Y_t|$.

Furthermore, we proved $\mathcal{F}^n \xrightarrow{w} \mathcal{F}$ in Theorem 3; combining this with

$$\mathbb{E}[\sup_{0\le t\le T} |Y_t|] < \infty,$$

where Y is the solution of (1), we can prove the second convergence using Lemma 6. Hence, we, at length, complete the proof.

5. Some Lemmas

Lemma 4. *Under Assumptions 5 and 6 or Assumption 7, it holds that* $W^n \Rightarrow W$.

Proof. We follow the notations in the book by Jacod and Shiryaev [11], which states that if the Assumptions 5 and 6 hold, then W^n's predictable characteristics "without truncation" are as follows:

$$B_t'^n = \sum_{k=1}^{[t/h]} \mathbb{E}[\sqrt{h}\varepsilon_k^n|\mathcal{G}_{k-1}^n] = 0,$$
$$C_t'^n = \sum_{k=1}^{[t/h]} \mathbb{E}[h(\varepsilon_k^n)^2|\mathcal{G}_{k-1}^n] = h[t/h],$$
$$g * \nu_t'^n = \sum_{k=1}^{[t/h]} \mathbb{E}[g(\sqrt{h}\varepsilon_k^n)|\mathcal{G}_{k-1}^n], \quad g \in C_1(\mathbb{R}).$$

Remember, the characteristics of the standard Brown motion are $B' = 0$, $C'^t = t$, $\nu = 0$. Furthermore, for some $C, a > 0$, we have $|g(x)| \le C|x|^2 1_{\{|x|>a\}}$ because of $g \in C_1(\mathbb{R})$, implying

$$g * \nu_t'^n \le \sum_{k=1}^{[t/h]} \mathbb{E}[Ch(\varepsilon_k^n)^2 1_{\{|\sqrt{h}\varepsilon_k^n|>a\}}] = [t/h]Ch1_{\{\sqrt{h}>a\}},$$

which tend toward 0 as $n \to \infty$. Additionally, $|x|^2 1_{\{|x|>a\}} * \nu_t'^n = [t/h]h1_{\{\sqrt{h}>a\}} \to 0$ as $n \to \infty$. Combining the above, thanks to [11] VII.3.7, we can obtain $W^n \Rightarrow W$.

Now, if the Assumption 7 holds, the proof is the same as the above. □

Lemma 5. *Let* η_i, $1 \le i \le n$ *be a sequence of random variables satisfying the standard normal distribution* $N(0,1)$. *Then, we have*

$$\mathbb{E}[\max_{1\le i\le n} \eta_i] \sim O(\sqrt{\log n}).$$

Consequently, it holds that

$$\mathbb{E}[\max_{1\le i\le n} |\eta_i|] \sim O(\sqrt{\log n}).$$

Proof. On the one hand, let $Z_n = \sum_{i=1}^{n} \exp(\beta\eta_i)$ for $\beta > 0$. We have

$$\beta \max_{1\le i\le n} \eta_i = \log\left(\exp\left(\max_{1\le i\le n} \beta\eta_i\right)\right) \le \log Z_n.$$

Using Jensen's inequality and taking the expectations on both sides of the above inequality, we could obtain

$$\beta\mathbb{E}[\max_{1\le i\le n} \eta_i] \le \mathbb{E}[\log Z_n] \le \log \mathbb{E}[Z_n] = \log n + \beta^2/2.$$

Taking $\beta = \sqrt{\log n}$, we can obtain $\mathbb{E}[\max_{1\le i\le n} \eta_i] \le \frac{3}{2}\sqrt{\log n}$.

On the other hand, it is obvious that $\{\max_{1\le i\le n} \eta_i > t\} = \cup_{i=1}^{n}\{\eta_i > t\}$ for all $t > 0$. Since η_i is independent, we can obtain

$$\begin{aligned}\mathbb{P}\left\{\max_{1\le i\le n} \eta_i > t\right\} &= 1 - \mathbb{P}\left\{\bigcup_{i=1}^{n}\{\eta_i \le t\}\right\} \\ &= 1 - \prod_{i=1}^{n}(1 - \mathbb{P}\{\eta_i > t\})\end{aligned}$$

For $0 < x < 1$, since $1 - x < \exp(-x)$, we have

$$\mathbb{P}\left\{\max_{1\le i\le n} \eta_i > t\right\} > 1 - \exp\left(-\sum_{i=1}^{n}\mathbb{P}\{\eta_i > t\}\right). \tag{30}$$

Set $A = \cap_{i\ge 1}\{\eta_i > 0\}$; it is easy to verify that

$$\begin{aligned}\mathbb{E}[\max_{1\le i\le n} \eta_i] &= \mathbb{E}[\left(\max_{1\le i\le n} \eta_i I_A + \max_{1\le i\le n} \eta_i I_{A^c}\right] \\ &\ge \mathbb{E}[|\eta_1|] + \mathbb{E}[\max_{1\le i\le n} \eta_i^+].\end{aligned} \tag{31}$$

Using Fubini theorem, we can show that, for fix $\mu > 0$,

$$\mathbb{E}[\max_{1\le i\le n} \eta_i^+] > \mu\mathbb{P}\{\max_{1\le i\le n} \eta_i^+ > \mu\}. \tag{32}$$

It is well-known that for some $t > t_0$, since η_i is a Gaussian variable,

$$\frac{C}{t}\exp\left(-t^2/2\right) \le \mathbb{P}\{\eta_i > t\}. \tag{33}$$

Combining (30)–(33), we obtain

$$\mathbb{E}[\max_{1\le i\le n} \eta_i] \ge \mathbb{E}[|\eta_1|] + \mu\left(1 - \exp\left(-Cne^{-\mu^2/2}/\mu\right)\right) \tag{34}$$

Now, we choose $\mu = \sqrt{2\log n - \log\log n}$ and compute the value of the second part on the right side of (34):

$$ne^{-\mu^2/2}/\mu = n \times \frac{\log n}{2n} \times \frac{1}{\sqrt{2\log n - \log\log n}} \sim \sqrt{\log n} \to +\infty,$$

which implies

$$\left(1 - \exp\left(-Cne^{-\mu^2/2}/\mu\right)\right) \to 1.$$

Therefore, $\mathbb{E}[\max_{1\le i\le n}\eta_i] \ge \mathbb{E}[|\eta_1|] + \mu\left(1-\exp\left(-Cne^{-\mu^2/2}/\mu\right)\right) \sim C\sqrt{\log n}$, so, for some constants C_1, C_2 and a large arbitrary n, we have

$$C_1\sqrt{\log n} \le \mathbb{E}[\max_{1\le i\le n}\eta_i] \le C_2\sqrt{\log n}, \tag{35}$$

which completes the first proof.

We shall start to prove the last result. Since we have $\max_{1\le i\le n}\eta_i \le \max_{1\le i\le n}|\eta_i|$ and (35), it is obvious that we need only to verify

$$\mathbb{E}[\max_{1\le i\le n}|\eta_i|] \le C_3\sqrt{\log n} + C_4.$$

The same step is used as the proof in the beginning, noting that

$$\begin{aligned} E[\exp(\beta|\eta_i|)] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(\beta|x|)\exp(-\frac{x^2}{2})dx \\ &\le \int_{-\infty}^{\infty}\exp(\beta x)\exp(-\frac{x^2}{2})dx + \int_{-\infty}^{\infty}\exp(-\beta x)\exp(-\frac{x^2}{2})dx \\ &= E[\exp(\beta\eta_i)] + E[\exp(-\beta\eta_i)] \\ &= 2\exp(\frac{\beta^2}{2}), \end{aligned}$$

we get

$$\beta\mathbb{E}[\max_{1\le i\le n}\eta_i] \le \log 2n + \beta^2/2.$$

Taking $\beta = \sqrt{\log 2n}$, because of $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$, we can obtain

$$\mathbb{E}[\max_{1\le i\le n}\eta_i] \le \frac{3}{2}\sqrt{\log 2 + \log n} \le \frac{3}{2}(\sqrt{\log 2} + \sqrt{\log n}).$$

We, at length, complete the proof. □

The following Lemma is a simplified version of a result of [12], Theorem 1.

Lemma 6. *Let $\{\mathcal{F}_t^n\}_{0\le t\le T}$ be a sequence of filtrations and $\{\mathcal{F}_t\}_{0\le t\le T}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$, such that $\mathcal{F}^n \xrightarrow{w} \mathcal{F}$. Let X be a $\mathcal{F}$ adapted continuous process, such that $\sup_{0\le t\le T}|X_t|$ is integrable. Then, $\mathbb{E}[X_\cdot|\mathcal{F}_\cdot^n] \xrightarrow{u.c.p.} X_\cdot$.*

We also list Proposition 1.5 (b) and Corollary 1.9 in [15] for readers' convenience.

Lemma 7. *If a sequence of continuous local martingales M^n satisfies*

$$\sup_n \mathbb{E}\left[\sup_{0\le t\le T}\Delta|M_t^n|\right] < \infty,$$

then it has the UT condition. Moreover, if $M^n \xrightarrow{u.c.p.} M^\infty$, then the quadratic variation process $[M^n, M^n] \xrightarrow{u.c.p.} [M, M]$.

Author Contributions: Writing—original draft, Y.G. and N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bismut, J. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* **1973**, *44*, 384–404. [CrossRef]
2. Pardoux, É.; Peng, S. Adapted solution of a backward stochastic differential equation. *Syst. Control Lett.* **1990**, *14*, 55–61. [CrossRef]
3. El Karoui, N.; Peng, S.; Quenez, M. Backward stochastic differential equations in finance. *Math. Financ.* **1997**, *7*, 1–71. [CrossRef]
4. Ma, J.; Protter, P.; Yong, J. Solving forward-backward stochastic differential equations explicitly—A four step scheme. *Probab. Theory Relat. Fields* **1994**, *98*, 339–359. [CrossRef]
5. Douglas, J.; Ma, J.; Protter, P. Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.* **1996**, *6*, 940–968. [CrossRef]
6. Zhang, J. A numerical scheme for BSDEs. *Ann. Appl. Probab.* **2004**, *14*, 459–488. [CrossRef]
7. Chevance, D. Numerical methods for backward stochastic differential equations. In *Numerical Methods in Finance*; Publications of the Newton Institute; Cambridge University Press: Cambridge, UK, 1997; Volume 13, pp. 232–244.
8. Coquet, F.; Mackevičius, V.; Mémin, J. Stability in D of martingales and backward equations under discretization of filtration. *Stoch. Process. Appl.* **1998**, *75*, 235–248. [CrossRef]
9. Briand, P.; Delyon, B.; Mémin, J. Donsker-type theorem for BSDEs. *Electron. Commun. Probab.* **2001**, 6, 1–14. [CrossRef]
10. Donsker, M. An invariance principle for certain probability limit theorems. *Mem. Am. Math. Soc.* **1951**, *6*, 1–12.
11. Jacod, J.; Shiryaev, A. *Limit Theorems for Stochastic Processes*, 2nd ed.; Grundlehren der Mathematischen Wissenschaften; Springer: Berlin/Heidelberg, Germany, 2003; Volume 288.
12. Coquet, F.; Mémin, J.; Słominski, L. On weak convergence of filtrations. In *Séminaire de Probabiliés XXXV*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 2001; Volume 1755, pp. 306–328.
13. Aldous, D. Stopping times and tightness. *Ann. Probab.* **1978**, *6*, 335–340. [CrossRef]
14. Zhang, J. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*; Springer: New York, NY, USA, 2017.
15. Mémin, J.; Słominski, L. Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. In *Séminaire de Probabilités XXV*; Springer: Berlin/Heidelberg, Geramny, 2006; pp. 162–177.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.