A regression Monte-Carlo method for Backward Doubly Stochastic Differential Equations

Omar Aboura

A REGRESSION MONTE-CARLO METHOD FOR BACKWARD DOUBLY STOCHASTIC DIFFERENTIAL EQUATIONS

OMAR ABOURA

Abstract. This paper extends the idea of E.Gobet, J.P.Lemor and X.Warin from the setting of Backward Stochastic Differential Equations to that of Backward Doubly Stochastic Differential equations. We propose some numerical approximation scheme of these equations introduced by E.Pardoux and S.Peng.

1. Introduction

Since the pioneering work of E. Pardoux and S. Peng [11], backward stochastic differential equations (BSDEs) have been intensively studied during the two last decades. Indeed, this notion has been a very useful tool to study problems in many areas, such as mathematical finance, stochastic control, partial differential equations; see e.g. [9] where many applications are described. Discretization schemes for BSDEs have been studied by several authors. The first papers on this topic are that of V.Bally [4] and D.Chevance [6]. In his thesis, J.Zhang made an interesting contribution which was the starting point of intense study among, which the works of B. Bouchard and N.Touzi [5], E.Gobet, J.P. Lemor and X. Warin[7],... The notion of BSDE has been generalized by E. Pardoux and S. Peng [12] to that of Backward Doubly Stochastic Differential Equation (BDSDE) as follows. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, T denote some fixed terminal time which will be used throughout the paper, $(W_t)_{0\leq t\leq T}$ and $(B_t)_{0\leq t\leq T}$ be two independent standard Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in $\mathbb{R}$. On this space we will deal with the following families of σ-algebras:

$$\mathcal{F}_t := \mathcal{F}^W_{0,t} \vee \mathcal{F}^B_{t,T} \vee \mathcal{N}, \quad \widehat{\mathcal{F}}_t := \mathcal{F}^W_{0,t} \vee \mathcal{F}^B_{0,T} \vee \mathcal{N}, \quad \mathcal{H}_t = \mathcal{F}^W_{0,T} \vee \mathcal{F}^B_{t,T} \vee \mathcal{N}, \tag{1.1}$$

where $\mathcal{F}^B_{t,T} := \sigma\left(B_r - B_t; t \leq r \leq T\right)$, $\mathcal{F}^W_{0,t} := \sigma\left(W_r; 0 \leq r \leq t\right)$ and $\mathcal{N}$ denotes the class of $\mathbb{P}$ null sets. We remark that $(\widehat{\mathcal{F}}_t)$ is a filtration, $(\mathcal{H}_t)$ is a decreasing family of σ-albegras, while $(\mathcal{F}_t)$ is neither increasing nor decreasing. Given an initial condition $x \in \mathbb{R}$, let (X_t) be the diffusion process defined by

$$X_t = x + \int_0^t b\left(X_s\right) ds + \int_0^t \sigma\left(X_s\right) dW_s. \tag{1.2}$$

Let $\xi \in L^2(\Omega)$ be an $\mathbb{R}$-valued, $\mathcal{F}_T$-measurable random variable, f and g be regular enough coefficients; consider the BDSDE defined as follows:

$$Y_t \;=\; \xi + \int_t^T f\left(s, X_s, Y_s, Z_s\right) ds + \int_t^T g\left(s, X_s, Y_s, Z_s\right) d\overleftarrow{B}_s - \int_t^T Z_s dW_s. \tag{1.3}$$

In this equation, dW is the forward stochastic integral and $d\overleftarrow{B}$ is the backward stochastic integral (we send the reader to [10] for more details on backward integration). A solution

Date: July 8, 2011.

to (1.3) is a pair of real-valued process (Y_t, Z_t), such that Y_t and Z_t are $\mathcal{F}_t$-measurable for every $t \in [0,T]$, such that (1.3) is satisfied and

$$\mathbb{E}\Big(\sup_{0\leq s\leq T}|Y_s|^2\Big)+\mathbb{E}\int_0^T|Z_s|^2ds<+\infty. \tag{1.4}$$

In [12] Pardoux and Peng have proved that under some Lipschitz property on f and g which will be stated later, (1.3) has a unique solution (Y,Z). They also proved that

$$Y_t=u\left(t,X_t,\left(\overleftarrow{\Delta}B_s\right)_{t\leq s\leq T}\right),\quad Z_t=v\left(t,X_t,\left(\overleftarrow{\Delta}B_s\right)_{t\leq s\leq T}\right),$$

for some Borel functions u and v.

The time discretization of BDSDEs has been addressed in [2] when the coefficient g does not depend on Z; see also [1] in the more general setting for g which may also depend on Z as in [12]. Both papers follow Zhang's approach and provide a theoretical approximation only using a constant time mesh.

In order to obtain a more tractable discretization which could be implemented, a natural idea is to see whether the methods introduced in [7] can be extended from the framework of BSDEs to that more involved of BDSDEs ; this is the aim of this paper.

We use three consecutive steps, and each time we give a precise estimate of the corresponding error. Thus, we start with a time discretization $(Y_{t_k}^N, Z_{t_k}^N)$ with a constant time mesh T/N. We can prove that

$$Y_{t_k}^N=u^N\left(t_k,X_{t_k}^N,\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_k\right),\quad Z_{t_k}^N=v^N\left(t_k,X_{t_k}^N,\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_k\right),$$

where for $k=1,\dots,N-1$, $t_k=kT/N$ and $\overleftarrow{\Delta}B_k=B_{t_{k+1}}-B_{t_k}$. Furthermore, if either $f=0$ or if the scheme is not implicit as in [1] then we have the more precise description:

$$Y_{t_k}^N=u_N^N\left(t_k,X_{t_k}^N\right)+\sum_{j=k}^{N-1}u_j^N\left(t_k,X_{t_k}^N,\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_{j+1}\right)\overleftarrow{\Delta}B_j,$$

$$Z_{t_k}^N=v_N^N\left(t_k,X_{t_k}^N\right)+\sum_{j=k}^{N-1}v_j^N\left(t_k,X_{t_k}^N,\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_{j+1}\right)\overleftarrow{\Delta}B_j,$$

with the convention that if $j+1>N-1$, $\left(\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_{j+1}\right)=\emptyset$. The main time discretization result in this direction is Theorem 3.4. In order to have a numerical scheme, we use this decomposition and the ideas of E.Gobet, J.P.Lemor and X.Warin [7]. Thus we introduce the following hypercubes, that is approximate random variables $u_j^N\left(t_k,X_{t_k}^N,\overleftarrow{\Delta}B_{N-1},\dots,\overleftarrow{\Delta}B_{j+1}\right)\overleftarrow{\Delta}B_j$ by their orthogonal projection on some finite vector space generated by some bases (u_j) and (v_j) defined below. For $k=1,\dots,N$ we have

$$\begin{aligned}Y_{t_k}^N\approx&\sum_{i_N}\mathbb{E}\left(Y_{t_k}^Nu_{i_N}\left(X_{t_k}^N\right)\right)u_{i_N}\left(X_{t_k}^N\right)\\&+\sum_{j=k}^{N-1}\sum_{i_N,i_{N-1},\dots,i_j}\mathbb{E}\left(Y_{t_k}^Nu_{i_N}\left(X_{t_k}^N\right)v_{i_{N-1}}\left(\overleftarrow{\Delta}B_{N-1}\right)\dots v_{i_{k+1}}\left(\overleftarrow{\Delta}B_{j+1}\right)\frac{\overleftarrow{\Delta}B_j}{\sqrt{h}}\right)\\&\qquad u_{i_N}\left(X_{t_k}^N\right)v_{i_{N-1}}\left(\overleftarrow{\Delta}B_{N-1}\right)\dots v_{i_{k+1}}\left(\overleftarrow{\Delta}B_{j+1}\right)\frac{\overleftarrow{\Delta}B_j}{\sqrt{h}}.\end{aligned}$$

We use a linear regression operator of the approximate solution. Thus, we at first use an orthogonal projection on a finite dimensional space $\mathcal{P}_k$. This space consists in linear

combinations of an orthonormal family of properly renormalized indicator functions of disjoint intervals composed either with the diffusion X or with increments of the Brownian motion B. As in [7], in order not to introduce error terms worse that those due to the time discretization, we furtherore have to use a Picard iteration scheme. The error due to this regression operator is estimated in Theorem 4.1.

Then the coefficients (α, β) of the decomposition of the projection of $(Y^N_{t_k}, Z^N_{t_k})$ are shown to solve a regression minimization problem and are expressed in terms of expected values. Note that a general regression approach has also been used by Bouchard and Touzi for BSDEs in [5]. Finally, the last step consists in replacing the minimization problem for the pair (α, β) in terms of expectations by similar expressions described in terms of an average over a sample of size M of the Brownian motions W and B. Then, a proper localization is needed to get an L^2 bound of the last error term. This requires another Picard iteration and the error term due to this Monte Carlo method is described in Theorem 5.8.

A motivation to study BSDEs is that these equations are widely used in financial models, so that having an efficient and fast numerical methods is important. As noted in [12], BDSDEs are connected with stochastic partial differential equations and the discretization of (2.2) is motivated by its link with the following SPDE:

$$\begin{aligned} u(t,x) &= \phi(x) + \int_t^T \Big(\mathcal{L}u(s,x) + f\left(s,x,u(s,x), \nabla u(s,x)\sigma(x)\right) \Big) ds \\ &\quad + \int_t^T g\left(s,x,u(s,x), \nabla u(s,x)\sigma(x)\right) d\overleftarrow{B}_s, \end{aligned} \tag{1.5}$$

Discretizations of SPDEs are mainly based on PDE techniques, such as finite differences or finite elements methods. Another approach for special equations is given by particle systems. We believe that this paper gives a third way to deal with this problem. As usual, the presence of the gradient in the diffusion coefficient is the most difficult part to handle when dealing with SPDEs. Only few results are obtained in the classical discretization framework when PDE methods are extended to the stochastic case.

Despite the fact that references [2] and [3] deal with a problem similar to that we address in section 3, we have kept the results and proofs of this section. Indeed, on one hand we study here an implicit scheme as in [7] and wanted the paper to be self contained. Furthermore, because of measurability properties of Y_0 and Y_0^π, the statements and proofs of Theorem 3.6 in [2] and Theorem 4.6 in [3] are unclear and there is a gap in the corresponding proofs because of similar measurability issues for (Y_t) and (Y_t^π).

The paper is organized as follows. Section 2 gives the main notations concerning the time discretization and the function basis. Section 3 describes the time discretization and results similar to those in [2] are proved in a more general framwork. The fourth section describes the projection error. Finally section 5 studies the regression technique and the corresponding Monte Carlo method. Note that the presence of increments of the Brownian motion B, which drives the backward stochastic integrals, requires some new arguments such as Lemma 5.16 which is a key ingredient of the last error estimates. As usual C denotes a constant which can change from line to line.

2. Notations

Let $(W_t, t \geq 0)$ and $(B_t, t \geq 0)$ be two mutually independent standard Brownian motions. For each $x \in \mathbb{R}$, let $(X_t, Y_t, Z_t, t \in [0,T])$ denote the solution of the following Backward Doubly Stochastic Differential Equation (BDSDE) introduced by E.Pardoux

and S.Peng in [12]:

$$X_t = x + \int_0^t b(X_s)\,ds + \int_0^t \sigma(X_s)\,dW_s, \tag{2.1}$$

$$Y_t = \Phi(X_T) + \int_t^T f(X_s, Y_s, Z_s)\,ds + \int_t^T g(X_s, Y_s)\,d\overleftarrow{B}_s - \int_t^T Z_s dW_s. \tag{2.2}$$

Assumption. *We suppose that the coefficients f and g satisfy the following:*

$$\Phi(X_T) \in L^2,$$

$$\left|f(x,y,z) - f(x',y',z')\right|^2 \leq L_f\left(|x-x'|^2 + |y-y'|^2 + |z-z'|^2\right), \tag{2.3}$$

$$\left|g(x,y) - g(x',y')\right|^2 \leq L_g\left(|x-x'|^2 + |y-y'|^2\right), \tag{2.4}$$

Note that (2.3) and (2.4) yield that f and g have linear growth in their arguments. We use two approximations. We at first discretize in time with a constant time mesh $h = T/N$, which yields the processes $\left(X^N, Y^N, Z^N\right)$. We then approximate the pair $\left(Y^N, Z^N\right)$ by some kind of Picard iteration scheme with I steps $\left(Y^{N,i,I}, Z^{N,I}\right)$ for $i = 1, \ldots, I$.

In order to be as clear as possible, we introduce below all the definitions used in the paper. Most of them are same as in [7].

(N0) *For $0 \leq t \leq t' \leq T$, set $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_{t,T}^B$ and*

$$\mathcal{F}_t^W = \sigma(W_s; 0 \leq s \leq t) \vee \mathcal{N}, \quad \mathcal{F}_{t,t'}^B = \sigma\left(B_s - B_{t'}; t \leq s \leq t'\right) \vee \mathcal{N}.$$

$\mathbb{E}_k$ is the conditionnal expectation with respect to $\mathcal{F}_{t_k}$.

(N1) *N is the number of steps of the time discretization, the integer I corresponds to the number of steps of the Picard iteration, $h := T/N$ is the size of the time mesh and for $k = 0, 1, \ldots, N$ we set $t_k := kh$ and $\overleftarrow{\Delta} B_k = B_{t_{k+1}} - B_{t_k}$, $\Delta W_{k+1} = W_{t_{k+1}} - W_{t_k}$. Let $\pi = t_0, t_1, \ldots, t_N = T$ denote the corresponding subdivision on $[0,T]$.*

(N2) *The function basis for $X_{t_k}^N$ is defined as follows: let $a_k < b_k$ be two reals and $(\mathcal{X}_i^k)_{i=1\ldots L}$ denote a partition of $[a_k, b_k]$; for $i = 1, \ldots, L$ set*

$$u_i\left(X_{t_k}^N\right) := 1_{\mathcal{X}_i^k}\left(X_{t_k}^N\right) / \sqrt{P\left(X_{t_k}^N \in \mathcal{X}_i^k\right)} \tag{2.5}$$

(N3) *The function basis for $N \sim \mathcal{N}(0,h)$ is defined as follows: let $a < b$ two reals and $(\mathcal{B}_i)_{i=1\ldots L}$ denote a partition of $[a,b]$. For $i = 1, \ldots, L$ set*

$$v_i(N) := 1_{\mathcal{B}_i}(N) / \sqrt{P(N \in \mathcal{B}_i)} \tag{2.6}$$

(N4) *For fixed $k = 1, \ldots, N$, let p_k denote the following vector whose components belong to $L^2(\Omega)$. It is defined blockwise as follows:*

$$\left(u_{i_N}\left(X_{t_k}^N\right)\right)_{i_N}, \left(u_{i_N}\left(X_{t_k}^N\right)\frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}}\right)_{i_N}, \left(u_{i_N}\left(X_{t_k}^N\right) v_{i_{N-1}}\left(\overleftarrow{\Delta} B_{N-1}\right)\frac{\overleftarrow{\Delta} B_{N-2}}{\sqrt{h}}\right)_{i_N, i_{N-1}},$$

$$\ldots$$

$$\left(u_{i_N}\left(X_{t_k}^N\right)\prod_{j=k+1}^{N-1} v_{i_j}\left(\overleftarrow{\Delta} B_j\right)\frac{\overleftarrow{\Delta} B_k}{\sqrt{h}}\right)_{i_N, i_{N-1}, \ldots, i_{k+1}}$$

where $i_N, \ldots, i_{k+1} \in \{1, \ldots, L\}$. Note that p_k is $\mathcal{F}_{t_k}$-measurable and $\mathbb{E} p_k p_k^* = Id$

3. Approximation result: step 1

We first consider a time discretization of equations (2.1) and (2.2). The forward equation (2.1) is approximated using the Euler scheme: $X_{t_0}^N = x$ and for $k = 0, \dots, N-1$,

$$X_{t_{k+1}}^N = X_{t_k}^N + hb(X_{t_k}^N) + \sigma(X_{t_k}^N)\Delta W_{k+1}. \tag{3.1}$$

The following result is well know: (see e.g. [8])

Theorem 3.1. *There exists a constant C such that for every N*

$$\max_{k=1,\dots,N} \sup_{t_{k-1}\le r\le t_k} \mathbb{E}\left|X_r - X_{t_{k-1}}^N\right|^2 \le Ch, \qquad \max_{k=0,\dots,N} \mathbb{E}\left|X_{t_k}^N\right|^2 = C < \infty.$$

The following time regularity is proved in [2] (see also Theorem 2.3 in [1]), it extends the original result of Zhang [13].

Lemma 3.2. *There exists a constant C such that for every integer $N \ge 1$, $s, t \in [0, T]$,*

$$\sum_{k=1}^N \mathbb{E}\int_{t_{k-1}}^{t_k} \left(\left|Z_r - Z_{t_{k-1}}\right|^2 + \left|Z_r - Z_{t_k}\right|^2\right) dr \le Ch, \quad \mathbb{E}\left|Y_t - Y_s\right|^2 \le C\left|t - s\right|.$$

The backward equation (2.2) is approximated by backward induction as follows:

$$Y_{t_N}^N := \Phi(X_{t_N}^N), \qquad Z_{t_N}^N := 0, \tag{3.2}$$

$$Z_{t_k}^N := \frac{1}{h}\mathbb{E}_k\left(Y_{t_{k+1}}^N \Delta W_{k+1}\right) + \frac{1}{h}\overleftarrow{\Delta} B_k \mathbb{E}_k\left(g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^N\right)\Delta W_{k+1}\right), \tag{3.3}$$

$$Y_{t_k}^N := \mathbb{E}_k Y_{t_{k+1}}^N + hf\left(X_{t_k}^N, Y_{t_k}^N, Z_{t_k}^N\right) + \overleftarrow{\Delta} B_k \mathbb{E}_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^N\right), \tag{3.4}$$

Note that as in [2], [3] and [7] we have introduced an implicit scheme, thus different from that in [1]. However, it differs from that in [2] and [3] since the conditional expectation we use is taken with respect to $\mathcal{F}_{t_k}$ which is different from $\sigma\left(X_{t_j}^N, j \le k\right) \vee \sigma\left(B_{t_j}, j \le k\right)$ used in [3].

Proposition 3.3 (Existence of the scheme)**.** *For sufficiently large N, the above scheme has a unique solution. Moreover, for all $k = 0, \dots, N$, we have $Y_{t_k}^N, Z_{t_k}^N \in L^2\left(\mathcal{F}_{t_k}\right)$.*

The following theorem is the main result of this section.

Theorem 3.4. *There exists a constant $C > 0$ such that for h small enough*

$$\max_{0\le k\le N} \mathbb{E}\left|Y_{t_k} - Y_{t_k}^N\right|^2 + \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left|Z_r - Z_{t_k}^N\right|^2 dr \le Ch + C\mathbb{E}\left|\phi\left(X_{t_N}^N\right) - \phi\left(X_T\right)\right|^2.$$

The rest of this section is devoted to the proof of this theorem; it requires several steps. First of all, we define a process $(Y_t^\pi, Z_t^\pi)_{t\in[0,T]}$ such that $Y_{t_k}^\pi$ and $Z_{t_k}^\pi$ are $\mathcal{F}_{t_k}$ measurable, and a family of $\mathcal{F}_{t_k}$ measurable random variables $Z_{t_k}^{\pi,1}$, $k = 0, \dots, N$ as follows. For $t = T$, set

$$Y_T^\pi := \Phi\left(X_{t_N}^N\right), \ Z_T^\pi := 0, \ Z_{t_N}^{\pi,1} := 0. \tag{3.5}$$

Suppose that the scheme (Y_t^π, Z_t^π) is defined for all $t \in [t_k, T]$ and that $Z_{t_j}^{\pi,1}$ has been defined for $j = N, \dots, k$. Then for h small enough the following equation

$$M_{t_{k-1}}^k := \mathbb{E}_{k-1}\left(Y_{t_k}^\pi + f\left(X_{t_{k-1}}^N, M_{t_{k-1}}^k, Z_{t_{k-1}}^N\right)\Delta t_{k-1} + g\left(X_{t_k}^N, Y_{t_k}^\pi\right)\overleftarrow{\Delta} B_{k-1}\right) \tag{3.6}$$

has a unique solution.
Using Proposition 3.3 and the linear growth of f, we deduce that the map F_ξ defined by

$$F_\xi(Y) = \mathbb{E}_{k-1}\left(\xi + hf\left(X^N_{t_{k-1}}, Y, Z^N_{t_{k-1}}\right)\right) \tag{3.7}$$

is such that $F_\xi\left(L^2\left(\mathcal{F}_{t_{k-1}}\right)\right) \subset L^2\left(\mathcal{F}_{t_{k-1}}\right)$. Futhermore, given $Y, Y' \in L^2\left(\mathcal{F}_{t_{k-1}}\right)$, the L^2 contraction property of $\mathbb{E}_{k-1}$ and the Lipschitz condition (2.3) imply $\mathbb{E}\left|F_\xi(Y) - F_\xi\left(Y'\right)\right|^2 \leq h^2 L_f \mathbb{E}\left|Y - Y'\right|^2$. Then F_ξ is a contraction for h small enough and the fixed point theorem concludes the proof.
We can extend $M^k_\cdot$ to the interval $t \in [t_{k-1}, t_k]$ letting

$$M^k_t := \mathbb{E}\left(Y^\pi_{t_k} + f\left(X^N_{t_{k-1}}, M^k_{t_{k-1}}, Z^N_{t_{k-1}}\right)\Delta t_{k-1} + g\left(X^N_{t_k}, Y^\pi_{t_k}\right)\overleftarrow{\Delta}B_{k-1}\middle|\, \mathcal{F}^W_t \vee \mathcal{F}^B_{t_{k-1},T}\right),$$

which is consistent at time t_{k-1}.
By an extension of the martingale representation theorem (see e.g. [12] p.212), there exists a $\left(\mathcal{F}^W_t \vee \mathcal{F}^B_{t_{k-1},T}\right)_{t_{k-1}\leq t\leq t_k}$-adapted and square integrable process $\left(N^k_t\right)_{t\in[t_{k-1},t_k]}$ such that for any $t \in [t_{k-1}, t_k]$, $M^k_t = M^k_{t_{k-1}} + \int_{t_{k-1}}^t N^k_s dW_s$ and hence $M^k_t = M^k_{t_k} - \int_t^{t_k} N^k_s dW_s$. Since,

$$M^k_{t_k} = Y^\pi_{t_k} + f\left(X^N_{t_{k-1}}, M^k_{t_{k-1}}, Z^N_{t_{k-1}}\right)\Delta t_{k-1} + g\left(X^N_{t_k}, Y^\pi_{t_k}\right)\overleftarrow{\Delta}B_{k-1},$$

we deduce that for $t \in [t_{k-1}, t_k]$

$$M^k_t = Y^\pi_{t_k} + f\left(X^N_{t_{k-1}}, M^k_{t_{k-1}}, Z^N_{t_{k-1}}\right)\Delta t_{k-1} + g\left(X^N_{t_k}, Y^\pi_{t_k}\right)\overleftarrow{\Delta}B_{k-1} - \int_t^{t_k} N^k_s dW_s. \tag{3.8}$$

For $t \in [t_{k-1}, t_k)$, we set

$$Y^\pi_t := M^k_t,\ Z^\pi_t := N^k_t,\ Z^{\pi,1}_{t_{k-1}} := \frac{1}{h}\mathbb{E}_{k-1}\left(\int_{t_{k-1}}^{t_k} Z^\pi_r dr\right). \tag{3.9}$$

Lemma 3.5. *For all $k = 0, \ldots, N$,*

$$Y^\pi_{t_k} = Y^N_{t_k},\ Z^{\pi,1}_{t_k} = Z^N_{t_k} \tag{3.10}$$

and hence for $k = 1, \ldots, N$

$$Y^\pi_{t_{k-1}} = Y^\pi_{t_k} + \int_{t_{k-1}}^{t_k} f\left(X^N_{t_{k-1}}, Y^\pi_{t_{k-1}}, Z^{\pi,1}_{t_{k-1}}\right)dr + \int_{t_{k-1}}^{t_k} g\left(X^N_{t_k}, Y^\pi_{t_k}\right)d\overleftarrow{B}_r - \int_{t_{k-1}}^{t_k} Z^\pi_r dW_r \tag{3.11}$$

Proof. We proceed by backward induction. For $k = N$, (3.10) is true by (3.2) and (3.5). Suppose that (3.10) holds for $l = N, N-1, \ldots, k$, so that $Y^\pi_{t_k} = Y^N_{t_k}$, $Z^{\pi,1}_{t_k} = Z^N_{t_k}$. Then (3.10) holds for $l = k-1$; indeed, for $\xi := Y^N_{t_k} + \overleftarrow{\Delta}B_{k-1}g\left(X^N_{t_k}, Y^N_{t_k}\right)$ we deduce from (3.4) and (3.6) that $M^k_{t_{k-1}} = F_\xi\left(M^k_{t_{k-1}}\right)$, $Y^N_{t_{k-1}} = F_\xi\left(Y^N_{t_{k-1}}\right)$ and $Y^\pi_{t_{k-1}} = M^k_{t_{k-1}} = F_\xi\left(M^k_{t_{k-1}}\right)$, where F_ξ is defined by (3.7). So using the uniqueness of the fixed point of the map F_ξ, we can conclude that $Y^\pi_{t_{k-1}} = Y^N_{t_{k-1}} (= M^k_{t_{k-1}})$. Therefore, (3.8) and (3.9) imply (3.11). Ito's formula yields

$$\Delta W_k \int_{t_{k-1}}^{t_k} Z^\pi_r dW_r = \int_{t_{k-1}}^{t_k} (W_r - W_{t_{k-1}}) Z^\pi_r dW_r + \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{r} Z^\pi_s dW_s dW_r + \int_{t_{k-1}}^{t_k} Z^\pi_r dr,$$

so that $\mathbb{E}_{k-1}\left(\Delta W_k \int_{t_{k-1}}^{t_k} Z^\pi_r dW_r\right) = \mathbb{E}_{k-1}\left(\int_{t_{k-1}}^{t_k} Z^\pi_r dr\right) = hZ^{\pi,1}_{t_{k-1}}$. Hence multiplying (3.11) by ΔW_k and taking conditional expectation with respect to $\mathcal{F}_{t_{k-1}} = \mathcal{F}^W_{t_{k-1}} \vee \mathcal{F}^B_{t_{k-1},T}$. We deduce

$$hZ^{\pi,1}_{t_{k-1}} = \mathbb{E}_{k-1}\left(Y^N_{t_k}\Delta W_k\right) + \overleftarrow{\Delta}B_{k-1}\mathbb{E}_{k-1}\left(g\left(X^N_{t_k}, Y^N_{t_k}\right)\Delta W_k\right)$$

Comparing this with (3.3) concludes the proof of (3.10) for $l = k-1$. □

Lemma 3.5 shows that for $r \in [t_k, t_{k+1}]$ one can upper estimate the L^2 norm of $Z_r - Z^N_{t_k}$ by that of $Z_r - Z^\pi_r$ and increments of Z. Indeed, using (3.10) we have for $k = 0, \dots, N-1$ and $r \in [t_k, t_{k+1}]$

$$\mathbb{E}\left|Z_r - Z^N_{t_k}\right|^2 = \mathbb{E}\left|Z_r - Z^{\pi,1}_{t_k}\right|^2 \leq 2\mathbb{E}\left|Z_r - Z_{t_k}\right|^2 + 2\mathbb{E}\left|Z_{t_k} - Z^{\pi,1}_{t_k}\right|^2$$

Furthermore, (3.9) and Cauchy-Schwarz's inequality yield for $k = 0, \dots, N-1$

$$\begin{aligned}\mathbb{E}\left|Z_{t_k} - Z^{\pi,1}_{t_k}\right|^2 \leq& \frac{1}{h}\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_{t_k} - Z^\pi_r|^2\, dr \\ \leq& \frac{2}{h}\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_{t_k} - Z_r|^2\, dr + \frac{2}{h}\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_r - Z^\pi_r|^2\, dr.\end{aligned}$$

Hence we deduce

$$\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left|Z_r - Z^N_{t_k}\right|^2 dr \leq 6\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left|Z_r - Z_{t_k}\right|^2 dr + 4\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \mathbb{E}\left|Z_r - Z^\pi_r\right|^2 dr. \tag{3.12}$$

Using Lemma 3.2 and (3.12) we see that Theorem 3.4 is a straightforward consequence of the following:

Theorem 3.6. *There exists a constant C such that for h small enough,*

$$\max_{0\leq k\leq N} \mathbb{E}\left|Y_{t_k} - Y^\pi_{t_k}\right|^2 + \int_0^T \mathbb{E}\left|Z_r - Z^\pi_r\right|^2 dr \leq Ch + C\mathbb{E}\left|\Phi\left(X^N_{t_N}\right) - \Phi\left(X_T\right)\right|^2.$$

Proof. For any $k = 1, \dots, N$ set

$$I_{k-1} := \mathbb{E}\left|Y_{t_{k-1}} - Y^\pi_{t_{k-1}}\right|^2 + \mathbb{E}\int_{t_{k-1}}^{t_k} |Z_r - Z^\pi_r|^2\, dr. \tag{3.13}$$

Since $Y_{t_{k-1}} - Y^\pi_{t_{k-1}}$ is $\mathcal{F}_{t_{k-1}}$-measurable while for $r \in [t_k, t_{k+1}]$ the random variable $Z_r - Z^\pi_r$ is $\mathcal{F}^W_r \vee \mathcal{F}^B_{t_{k-1},T}$-measurable, we deduce that $Y_{t_{k-1}} - Y^\pi_{t_{k-1}}$ is orthogonal to $\int_{t_{k-1}}^{t_k} (Z_r - Z^\pi_r)\, dW_r$. Therefore, the identities (2.2) and (3.11) imply that

$$\begin{aligned}I_{k-1} =& \mathbb{E}\left|Y_{t_{k-1}} - Y^\pi_{t_{k-1}} + \int_{t_{k-1}}^{t_k} (Z_r - Z^\pi_r)\, dW_r\right|^2 \\ =& \mathbb{E}\left|Y_{t_k} - Y^\pi_{t_k} + \int_{t_{k-1}}^{t_k} \left(f\left(X_r, Y_r, Z_r\right) - f\left(X^N_{t_{k-1}}, Y^\pi_{t_{k-1}}, Z^{\pi,1}_{t_{k-1}}\right)\right) dr\right. \\ &\left. + \int_{t_{k-1}}^{t_k} \left(g\left(X_r, Y_r\right) - g\left(X^N_{t_k}, Y^\pi_{t_k}\right)\right) d\overleftarrow{B}_r\right|^2.\end{aligned}$$

Notice that for $t_{k-1} \leq r \leq t_k$ the random variable $g\left(X_r, Y_r\right) - g\left(X^N_{t_k}, Y^\pi_{t_k}\right)$ is $\mathcal{F}^W_{t_k} \vee \mathcal{F}^B_{r,T}$-measurable. Hence $Y_{t_k} - Y^\pi_{t_k}$, which is $\mathcal{F}_{t_k}$-measurable, and $\int_{t_{k-1}}^{t_k} \left(g\left(X_r, Y_r\right) - g\left(X^N_{t_k}, Y^\pi_{t_k}\right)\right) d\overleftarrow{B}_r$ are orthogonal. The inequality $(a+b+c)^2 \leq \left(1+\frac{1}{\lambda}\right)(a^2+c^2) + (1+2\lambda)\, b^2 + 2ac$ valid for $\lambda > 0$, Cauchy-Schwarz's inequality and the isometry of backward stochastic integrals

yield for $\lambda := \frac{\epsilon}{h}$, $\epsilon > 0$:

$$\begin{aligned}
I_{k-1} \leq & \left(1+\frac{h}{\epsilon}\right)\left[\mathbb{E}\left|Y_{t_k}-Y^{\pi}_{t_k}\right|^2+\mathbb{E}\left|\int_{t_{k-1}}^{t_k}\left(g\left(X_r,Y_r\right)-g\left(X^N_{t_k},Y^{\pi}_{t_k}\right)\right)d\overleftarrow{B}_r\right|^2\right] \\
& +\left(1+2\frac{\epsilon}{h}\right)\mathbb{E}\left|\int_{t_{k-1}}^{t_k}\left(f\left(X_r,Y_r,Z_r\right)-f\left(X^N_{t_{k-1}},Y^{\pi}_{t_{k-1}},Z^{\pi,1}_{t_{k-1}}\right)\right)dr\right|^2 \\
\leq & \left(1+\frac{h}{\epsilon}\right)\left[\mathbb{E}\left|Y_k-Y^{\pi}_{t_k}\right|^2+\mathbb{E}\int_{t_{k-1}}^{t_k}\left|g\left(X_r,Y_r\right)-g\left(X^N_{t_k},Y^{\pi}_{t_k}\right)\right|^2dr\right] \\
& +\left(h+2\epsilon\right)\mathbb{E}\int_{t_{k-1}}^{t_k}\left|f\left(X_r,Y_r,Z_r\right)-f\left(X^N_{t_{k-1}},Y^{\pi}_{t_{k-1}},Z^{\pi,1}_{t_{k-1}}\right)\right|^2dr.
\end{aligned}$$

The Lipschitz properties (2.3) and (2.4) of f and g imply

$$\begin{aligned}
I_{k-1} \leq & \left(1+\frac{h}{\epsilon}\right)\left[\mathbb{E}\left|Y_{t_k}-Y^{\pi}_{t_k}\right|^2+L_g\mathbb{E}\int_{t_{k-1}}^{t_k}\left(\left|X_r-X^N_{t_k}\right|^2+\left|Y_r-Y^{\pi}_{t_k}\right|^2\right)dr\right] \\
& +\left(h+2\epsilon\right)L_f\mathbb{E}\int_{t_{k-1}}^{t_k}\left(\left|X_r-X^N_{t_{k-1}}\right|^2+\left|Y_r-Y^{\pi}_{t_{k-1}}\right|^2+\left|Z_r-Z^{\pi,1}_{t_{k-1}}\right|^2\right)dr. \quad (3.14)
\end{aligned}$$

Using the definition of $Z^{\pi,1}_{t_k}$ in (3.9), the L^2 contraction property of $\mathbb{E}_k$ and Cauchy-Schwarz's inequality, we have

$$h\mathbb{E}\left|Z_{t_k}-Z^{\pi,1}_{t_k}\right|^2 \leq \frac{1}{h}\mathbb{E}\left|\mathbb{E}_k\left(\int_{t_k}^{t_{k+1}}\left(Z_{t_k}-Z^{\pi}_r\right)dr\right)\right|^2 \leq \mathbb{E}\int_{t_k}^{t_{k+1}}\left|Z_{t_k}-Z^{\pi}_r\right|^2dr.$$

Thus, by Young's inequality, we deduce for $k=1,\ldots,N$

$$\begin{aligned}
\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z^{\pi,1}_{t_{k-1}}\right|^2dr \leq & 2\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2dr+2h\mathbb{E}\left|Z_{t_{k-1}}-Z^{\pi,1}_{t_{k-1}}\right|^2 \\
\leq & 2\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2dr+4\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z^{\pi}_r-Z_r\right|^2dr \\
& +4\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2dr \\
\leq & 6\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2dr+4\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z^{\pi}_r-Z_r\right|^2dr.
\end{aligned}$$

We now deal with increments of Y. Using Lemma 3.2, we have

$$\begin{aligned}
\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Y_r-Y^{\pi}_{t_{k-1}}\right|^2dr \leq & 2\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Y_r-Y_{t_{k-1}}\right|^2dr+2\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Y_{t_{k-1}}-Y^{\pi}_{t_{k-1}}\right|^2dr \\
\leq & Ch^2+2h\mathbb{E}\left|Y_{t_{k-1}}-Y^{\pi}_{t_{k-1}}\right|^2,
\end{aligned}$$

while a similar argument yields

$$\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Y_r-Y^{\pi}_{t_k}\right|^2dr \leq Ch^2+2h\mathbb{E}\left|Y_{t_k}-Y^{\pi}_{t_k}\right|^2.$$

Using Theorem 3.1 and the previous upper estimates in (3.14), we deduce

$$\begin{aligned} I_{k-1} \leq & \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|Y_{t_k}-Y_{t_k}^{\pi}\right|^2 + L_f\left(h+2\epsilon\right)\Big[Ch^2+2h\mathbb{E}\left|Y_{t_{k-1}}-Y_{t_{k-1}}^{\pi}\right|^2 \\ & + 6\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2 dr + 4\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r^{\pi}-Z_r\right|^2 dr\Big] \\ & + L_g\left(1+\frac{h}{\epsilon}\right)\left[Ch^2+2h\mathbb{E}\left|Y_{t_k}-Y_{t_k}^{\pi}\right|^2\right]. \end{aligned}$$

Thus, (3.13) implies that for any $\epsilon > 0$

$$\begin{aligned} & \left[1-2L_f\left(h+2\epsilon\right)h\right]\mathbb{E}\left|Y_{t_{k-1}}-Y_{t_{k-1}}^{\pi}\right|^2 + \left[1-4L_f\left(h+2\epsilon\right)\right]\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_r^{\pi}\right|^2 dr \\ \leq & \left(1+\frac{h}{\epsilon}+2L_g\left(1+\frac{h}{\epsilon}\right)h\right)\mathbb{E}\left|Y_{t_k}-Y_{t_k}^{\pi}\right|^2 + \left(L_f\left(h+2\epsilon\right)+L_g\left(1+\frac{h}{\epsilon}\right)\right)Ch^2 \\ & + 6L_f\left(h+2\epsilon\right)\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2 dr. \end{aligned}$$

Now we choose ϵ such that $8\epsilon L_f = \frac{1}{2}$. Then we have for $\widetilde{C} = 4L_f$, h small enough and some positive constant $\overline{C}$ depending on L_f and L_g:

$$\begin{aligned} & \left(1-\widetilde{C}h\right)\mathbb{E}\left|Y_{t_{k-1}}-Y_{t_{k-1}}^{\pi}\right|^2 + \left(\frac{1}{2}-\widetilde{C}h\right)\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_r^{\pi}\right|^2 dr \\ & \qquad \leq \left(1+\overline{C}h\right)\mathbb{E}\left|Y_{t_k}-Y_{t_k}^{\pi}\right|^2 + \overline{C}h^2 + \overline{C}\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2 dr. \end{aligned} \tag{3.15}$$

We need the following

Lemma 3.7. *Let $L > 0$; then for h^* small enough (more precisely $Lh^* < 1$) there exists $\Gamma := \frac{L}{1-Lh^*} > 0$ such that for all $h \in (0, h^*)$ we have $\frac{1}{1-Lh} < 1 + \Gamma h$*

Proof. Let $h \in (0, h^*)$; then we have $1 - Lh > 1 - Lh^* > 0$. Hence $\frac{L}{1-Lh} < \frac{L}{1-Lh^*} = \Gamma$, so that $Lh < \Gamma h(1 - Lh)$, which yields $1 + \Gamma h - Lh - \Gamma L h^2 = (1 + \Gamma h)(1 - Lh) > 1$. This concludes the proof. □

Lemma 3.7 and (3.15) imply the existence of a constant $C > 0$ such that for h small enough and $k = 1, 2, \ldots, N$ we have

$$\mathbb{E}\left|Y_{t_{k-1}}-Y_{t_{k-1}}^{\pi}\right|^2 \leq (1+Ch)\,\mathbb{E}\left|Y_{t_k}-Y_{t_k}^{\pi}\right|^2 + Ch^2 + C\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r-Z_{t_{k-1}}\right|^2 dr. \tag{3.16}$$

The final step relies on the following discrete version of Gronwall's lemma (see [7]).

Lemma 3.8 (Gronwall's Lemma)**.** *Let $(a_k), (b_k), (c_k)$ be nonnegative sequences such that for some $K > 0$ we have for all $k = 1, \ldots, N-1$, $a_{k-1} + c_{k-1} \leq (1 + Kh)a_k + b_{k-1}$. Then, for all $k = 0, \ldots, N-1$, $a_k + \sum_{i=k}^{N-1} c_i \leq e^{K(T-t_k)}\left(a_N + \sum_{i=k}^{N-1} b_i\right)$*

Use Lemma 3.8 with $c_k = 0$, $a_{k-1} = \mathbb{E}\left|Y_{t_{k-1}} - Y^{\pi}_{t_{k-1}}\right|^2$ and $b_k = C\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r - Z_{t_{k-1}}\right|^2 + Ch^2$; this yields

$$\begin{aligned}\sup_{0\le k\le N}\mathbb{E}\left|Y_{t_k} - Y^{\pi}_{t_k}\right|^2 \le & C\left(\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2 + \sum_{k=1}^{N}\mathbb{E}\int_{t_{k-1}}^{t_k}\left|Z_r - Z_{t_{k-1}}\right|^2 dr + Ch\right)\\ \le & C\left(\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2 + Ch\right),\end{aligned} \tag{3.17}$$

where the last upper estimate is deduced from Lemma 3.2. We sum (3.15) from $k = 1$ to $k = N$; using (3.17) we deduce that for some constant $\bar{C}$ depending on L_f and L_g we have

$$\begin{aligned}\left(\frac{1}{2} - \widetilde{C}h\right)\mathbb{E}\int_0^T |Z_r - Z^{\pi}_r|^2\, dr \le & \overline{C}h\sum_{k=1}^{N-1}\mathbb{E}\left|Y_{t_k} - Y^{\pi}_{t_k}\right|^2 + \overline{C}h + \overline{C}\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2\\ \le & \overline{C}h + \overline{C}\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2 + \overline{C}h\left(\overline{C} + N\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2\right)\\ \le & \overline{C}h + \overline{C}\mathbb{E}\left|Y_T - Y^{\pi}_{t_N}\right|^2.\end{aligned}$$

The definitions of Y_T and $Y^N_{t_N}$ from (2.2) and (3.2) conclude the proof of Theorem 3.6. □

4. Approximation results: step 2

In order to approximate $\left(Y^N_{t_k}, Z^N_{t_k}\right)_{k=0,\dots,N}$ we use the idea of E.Gobet, J.P. Lemor and X.Warin [7], that is a projection on the function basis and a Picard iteration scheme. In this section, N and I are fixed positive integers. We define the sequences $\left(Y^{N,i,I}_{t_k}\right)_{i=0,\dots,I\ k=0,\dots,N}$ and $\left(Z^{N,I}_{t_k}\right)_{k=0,\dots,N-1}$ using backward induction on k, and for fixed k forward induction on i for $Y^{N,i,I}_{t_k}$ as follows: For $k = N$, $Z^{N,I}_{t_N} = 0$ and for $i = 0,\dots,I$, set $Y^{N,i,I}_{t_N} := P_N\Phi\left(X^N_{t_N}\right)$. Assume that $Y^{N,I,I}_{t_{k+1}}$ has been defined and set

$$Z^{N,I}_{t_k} := \frac{1}{h}P_k\left[Y^{N,I,I}_{t_{k+1}}\Delta W_{k+1}\right] + \frac{1}{h}P_k\left[\overleftarrow{\Delta}B_k g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right)\Delta W_{k+1}\right]. \tag{4.1}$$

Let $Y^{N,0,I}_{t_k} := 0$ and for $i = 1,\dots,I$ define inductively by the following Picard iteration scheme:

$$Y^{N,i,I}_{t_k} := P_k Y^{N,I,I}_{t_{k+1}} + hP_k\left[f\left(X^N_{t_k}, Y^{N,i-1,I}_{t_k}, Z^{N,I}_{t_k}\right)\right] + P_k\left[\overleftarrow{\Delta}B_k g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right)\right], \tag{4.2}$$

where P_k is the orthogonal projection on the Hilbert space $\mathcal{P}_k \subset L^2\left(\mathcal{F}_{t_k}\right)$ generated by the function p_k defined by (N4). Set $R_k := I - P_k$. Note that P_k is a contraction of $L^2\left(\mathcal{F}_{t_k}\right)$. Furthermore, given $Y \in L^2(\Omega)$,

$$\mathbb{E}_k P_k Y = P_k\mathbb{E}_k Y = P_k Y. \tag{4.3}$$

Indeed, since $\mathcal{P}_k \subset L^2\left(\mathcal{F}_{t_k}\right)$, $\mathbb{E}_k P_k Y = P_k Y$. Let $Y \in L^2$; for every, $U_k \in \mathcal{P}_k$, since U_k is $\mathcal{F}_{t_k}$-measurable, we have $\mathbb{E}\left(U_k R_k Y\right) = 0 = \mathbb{E}\left(U_k\mathbb{E}_k R_k Y\right)$; so that, $P_k\mathbb{E}_k R_k(Y) = 0$. Futhermore $Y = P_k Y + R_k Y$ implies $P_k\mathbb{E}_k Y = P_k P_k Y + P_k\mathbb{E}_k R_k Y = P_k Y$ which yields (4.3). Now we state the main result of this section.

Theorem 4.1. *For h small enough, we have*

$$\max_{0\leq k\leq N}\mathbb{E}\left|Y_{t_k}^{N,I,I}-Y_{t_k}^{N}\right|^2+h\sum_{k=0}^{N-1}\mathbb{E}\left|Z_{t_k}^{N,I}-Z_{t_k}^{N}\right|^2\leq Ch^{2I-2}+C\sum_{k=0}^{N-1}\mathbb{E}\left|R_kY_{t_k}^{N}\right|^2$$
$$+C\mathbb{E}\left|\Phi\left(X_{t_N}^{N}\right)-P_N\Phi\left(X_{t_N}^{N}\right)\right|^2+Ch\sum_{k=0}^{N-1}\mathbb{E}\left|R_kZ_{t_k}^{N,I}\right|^2.$$

Proof of Theorem 4.1. The proof will be deduced from severals lemmas. The first result gives integrability properties of the scheme defined by (4.1) and (4.2).

Lemma 4.2. *For every $k=0,\dots,N$ and $i=0,\dots,I$ we have $Y_{t_k}^{N,i,I},Z_{t_k}^{N,I}\in L^2\left(\mathcal{F}_{t_k}\right)$.*

Proof. We prove this by backward induction on k, and for fixed k by forward induction on i. By definition $Y_{t_N}^{N,i,I}=P_N\Phi(X_{t_N}^N)$ and $Z_{t_N}^{N,I}=0$. Suppose that $Z_{t_j}^{N,I}$ and $Y_{t_j}^{N,l,I}$ belong to $L^2\left(\mathcal{F}_{t_j}\right)$ for $j=N,N-1,\dots,k+1$ and any l, and for $j=k$ and $l=0,\dots,i-1$; we will show that $Y_{t_k}^{N,i,I},Z_{t_k}^{N,I}\in L^2\left(\mathcal{F}_{t_k}\right)$.
The measurability is obvious since $\mathcal{P}_k\subset L^2\left(\mathcal{F}_k\right)$. We at first prove the square integrability of $Z_{t_k}^{N,I}$. Using (4.3), the conditional Cauchy-Schwarz inequality and the independence of ΔW_{k+1} and $\mathcal{F}_{t_k}$, we deduce

$$\begin{aligned}\mathbb{E}\left|P_k\left(Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right)\right|^2=&\mathbb{E}\left|P_k\mathbb{E}_k\left(Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right)\right|^2\leq\mathbb{E}\left|\mathbb{E}_k\left(Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right)\right|^2\\&\leq\mathbb{E}\left(\mathbb{E}_k\left|\Delta W_{k+1}\right|^2\mathbb{E}_k\left|Y_{t_{k+1}}^{N,I,I}\right|^2\right)\leq h\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2.\end{aligned}$$

A similar computation using the independence of ΔW_{k+1} and $\mathcal{F}_{t_k}$, and of $\overleftarrow{\Delta}B_k$ and $\mathcal{F}_{t_{k+1}}$ as well as the growth condition deduced from (2.4) yields

$$\begin{aligned}&\mathbb{E}\left|P_k\left(\overleftarrow{\Delta}B_k\Delta W_{k+1}g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2=\mathbb{E}\left|P_k\mathbb{E}_k\left(\overleftarrow{\Delta}B_k\Delta W_{k+1}g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2\\&\leq\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta}B_k\Delta W_{k+1}g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2\leq h\mathbb{E}\mathbb{E}_{k+1}\left|\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right|^2\\&\leq h^2\mathbb{E}\left|g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right|^2\leq 2h^2\left|g(0,0)\right|^2+2h^2L_g\left(\mathbb{E}\left|X_{t_{k+1}}^N\right|^2+\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2\right).\end{aligned}$$

The two previous upper estimates and the induction hypothesis proves that $Z_{t_k}^{N,I}\in L^2\left(\mathcal{F}_{t_k}\right)$. A similar easier proof shows that $Y_{t_k}^{N,i,I}\in L^2\left(\mathcal{F}_{t_k}\right)$. □

The following lemma gives L^2 bounds for multiplication by ΔW_{k+1}

Lemma 4.3. *For every $Y\in L^2$ we have $\mathbb{E}\left|\mathbb{E}_k\left(Y\Delta W_{k+1}\right)\right|^2\leq h\left(\mathbb{E}|Y|^2-\mathbb{E}\left|\mathbb{E}_kY\right|^2\right)$*

Proof. Using the fact that $\mathbb{E}_k\left(\Delta W_{k+1}\mathbb{E}_kY\right)=0$ we have

$$\mathbb{E}\left|\mathbb{E}_k\left(Y\Delta W_{k+1}\right)\right|^2=\mathbb{E}\left|\mathbb{E}_k\left(\left(Y-\mathbb{E}_kY\right)\Delta W_{k+1}\right)\right|^2$$

Using the conditional Cauchy-Schwarz inequality and the independence of ΔW_{k+1} and $\mathcal{F}_{t_k}$, we deduce $\mathbb{E}\left|\mathbb{E}_k\left(Y\Delta W_{k+1}\right)\right|^2\leq h\mathbb{E}\left|Y-\mathbb{E}_kY\right|^2\leq h\left(\mathbb{E}|Y|^2-\mathbb{E}\left|\mathbb{E}_kY\right|^2\right)$; this concludes the proof. □

The following result gives orthogonality properties of several projections.

Lemma 4.4. *Let $k=0,\dots,N-1$, and $M_{t_{k+1}},N_{t_{k+1}}\in L^2\left(\mathcal{F}_{t_{k+1}}\right)$. Then*

$$\mathbb{E}\left(P_kM_{t_{k+1}}P_k\left(\overleftarrow{\Delta}B_kN_{t_{k+1}}\right)\right)=0.$$

Proof. Let $M_{t_{k+1}} \in L^2\left(\mathcal{F}_{t_{k+1}}\right)$; the definition of P_k yields

$$P_k M_{t_{k+1}} = \sum_{1\le i_N\le L} \alpha(i_N) u_{i_N}\left(X^N_{t_k}\right) + \sum_{1\le i_N\le L} \alpha(N-1,i_N) u_{i_N}\left(X^N_{t_k}\right)\frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}} \tag{4.4}$$
$$+ \sum_{k\le l\le N-1}\ \sum_{1\le i_N,\dots,i_{l+1}\le L} \alpha\left(l,i_N,\dots,i_{l+1}\right) u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}},$$

where $\alpha(i_N) = \mathbb{E}\left[M_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\right]$, $\alpha\left(N-1,i_N\right) = \mathbb{E}\left[M_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}}\right]$, and $\alpha\left(l,i_N,\dots,i_{l+1}\right) = \mathbb{E}\left[M_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}}\right]$. Taking conditional expectation with respect to $\mathcal{F}_{t_{k+1}}$, we deduce that for any $i_N,\dots,i_k \in \{1,\dots,L\}$

$$\alpha\left(k,i_N,\dots,i_{k+1}\right) = \mathbb{E}\left[M_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\prod_{r=k+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\mathbb{E}_{k+1}\frac{\overleftarrow{\Delta} B_k}{\sqrt{h}}\right] = 0.$$

A similar decomposition of $P_k\left(\overleftarrow{\Delta} B_k N_{t_{k+1}}\right)$ yields

$$P_k\left(\overleftarrow{\Delta} B_k N_{t_{k+1}}\right) = \sum_{1\le i_N\le L} \beta\left(i_N\right) u_{i_N}\left(X^N_{t_k}\right) + \sum_{1\le i_N\le L} \beta\left(N-1,i_N\right) u_{i_N}\left(X^N_{t_k}\right)\frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}}$$
$$+ \sum_{k\le l\le N-1}\ \sum_{1\le i_N,\dots,i_{l+1}\le L} \beta\left(l,i_N,\dots,i_{l+1}\right) u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}} \tag{4.5}$$

where $\beta\left(i_N\right) = \mathbb{E}\left[\overleftarrow{\Delta} B_k N_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\right]$, $\beta\left(N-1,i_N\right) = \mathbb{E}\left[\overleftarrow{\Delta} B_k N_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}}\right]$ and $\beta\left(l,i_N,\dots,i_{l+1}\right) = \mathbb{E}\left[\overleftarrow{\Delta} B_k N_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}}\right]$. In the above sum, all terms except those corresponding to $l = k$ are equal to 0. Indeed, let $l \in \{k+1,\dots,N-1\}$; then using again the conditional expectation with respect to $\mathcal{F}_{t_{k+1}}$ we obtain

$$\beta\left(l,i_N,\dots,i_{l+1}\right) = \mathbb{E}\left[N_{t_{k+1}} u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}}\mathbb{E}_{k+1}\overleftarrow{\Delta} B_k\right] = 0$$

The two first terms in the decomposition of $P_k\left(\overleftarrow{\Delta} B_k N_{t_{k+1}}\right)$ are dealt with by a similar argument. Notice that for any $l \in \{k+1,\dots,N-1\}$ and any $i_N,\dots,i_l,j_N,\dots,j_{k+1} \in \{1,\dots,L\}$ we have, (conditioning with respect to $\mathcal{F}_{t_k}$):

$$\mathbb{E}\left[u_{i_N}\left(X^N_{t_k}\right)\prod_{r=l+1}^{N-1} v_{i_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_l}{\sqrt{h}} u_{j_N}\left(X^N_{t_k}\right)\prod_{r=k+1}^{N-1} v_{j_r}\left(\overleftarrow{\Delta} B_r\right)\frac{\overleftarrow{\Delta} B_k}{\sqrt{h}}\right] = 0.$$

A similar computation proves that for any $i_N,j_N,\dots,j_{k+1} \in \{1,\dots,L\}$, $\xi \in \left\{1, \frac{\overleftarrow{\Delta} B_{N-1}}{\sqrt{h}}\right\}$

$$\mathbb{E}\left[u_{i_N}\left(X^N_{t_k}\right)\xi u_{j_N}\left(X^N_{t_k}\right)\prod_{r=k+1}^{N-1} v_r\left(\overleftarrow{\Delta} B_r\right)\right] = 0.$$

The decompositions (4.4) and (4.5) conclude the proof. □

The next lemma provides upper bounds of the L^2-norm of $Z^{N,I}_{t_k}$ and $Z^{N,I}_{t_k} - Z^N_{t_k}$.

Lemma 4.5. *For small h enough and for $k = 0, \dots, N-1$, we have the following L^2 bounds*

$$\begin{aligned}\mathbb{E}\left|Z_{t_k}^{N,I}\right|^2 \leq & \frac{1}{h}\left(\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2 - \mathbb{E}\left|\mathbb{E}_k Y_{t_{k+1}}^{N,I,I}\right|^2\right) \\ & + \frac{1}{h}\left(\mathbb{E}\left|\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right|^2 - \mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2\right),\end{aligned} \tag{4.6}$$

$$\begin{aligned}\mathbb{E}\left|Z_{t_k}^{N,I} - Z_{t_k}^N\right|^2 \leq & \mathbb{E}\left|R_k Z_{t_k}^N\right|^2 + \frac{1}{h}\left(\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N\right|^2 - \mathbb{E}\left|\mathbb{E}_k\left(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N\right)\right|^2\right) \\ & + \frac{1}{h}\left(\mathbb{E}\left|\overleftarrow{\Delta} B_k\left[g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right) - g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^N\right)\right]\right|^2\right. \\ & \left. - \mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta} B_k\left[g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right) - g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^N\right)\right]\right)\right|^2\right).\end{aligned} \tag{4.7}$$

Proof. Lemma 4.4 implies that both terms in the right hand side of (4.1) are orthogonal. Hence squaring both sides of equation (4.1), using (4.3) and Lemma 4.3, we deduce

$$\begin{aligned}\mathbb{E}\left|Z_{t_k}^{N,I}\right|^2 = & \frac{1}{h^2}\mathbb{E}\left|P_k\left(Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right)\right|^2 + \frac{1}{h^2}\mathbb{E}\left|P_k\left(\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\Delta W_{k+1}\right)\right|^2 \\ = & \frac{1}{h^2}\mathbb{E}\left|P_k\mathbb{E}_k\left[Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right]\right|^2 + \frac{1}{h^2}\mathbb{E}\left|P_k\mathbb{E}_k\left(\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\Delta W_{k+1}\right)\right|^2 \\ \leq & \frac{1}{h^2}\mathbb{E}\left|\mathbb{E}_k\left[Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}\right]\right|^2 + \frac{1}{h^2}\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\Delta W_{k+1}\right)\right|^2 \\ \leq & \frac{1}{h}\left(\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2 - \mathbb{E}\left|\mathbb{E}_k Y_{t_{k+1}}^{N,I,I}\right|^2\right) \\ & + \frac{1}{h}\left(\mathbb{E}\left|\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right|^2 - \mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2\right);\end{aligned}$$

this proves (4.6).
Using the orthogonal decomposition $Z_{t_k}^N = P_k Z_{t_k}^N + R_k Z_{t_k}^N$, since $Z_{t_k}^{N,I} \in \mathcal{P}_k$ we have $\mathbb{E}\left|Z_{t_k}^{N,I} - Z_{t_k}^N\right|^2 = \mathbb{E}\left|Z_{t_k}^{N,I} - P_k Z_{t_k}^N\right|^2 + \mathbb{E}\left|R_k Z_{t_k}^N\right|^2$. Futhermore (3.3), (4.1) and (4.3) yield

$$\begin{aligned}Z_{t_k}^{N,I} - P_k Z_{t_k}^N = & \frac{1}{h}P_k\left[\left(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N\right)\Delta W_{k+1}\right] \\ & + \frac{1}{h}P_k\left[\overleftarrow{\Delta} B_k\left(g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right) - g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^N\right)\right)\Delta W_{k+1}\right].\end{aligned}$$

Lemma 4.4 shows that the above decomposition is orthogonal; thus using (4.3), the contraction property of P_k and Lemma 4.3, we deduce

$$\begin{aligned}
\mathbb{E}\left|Z_{t_k}^{N,I}-P_kZ_{t_k}^N\right|^2 =&\frac{1}{h^2}\mathbb{E}\left|P_k\mathbb{E}_k\left[\left(Y_{t_{k+1}}^{N,I,I}-Y_{t_{k+1}}^N\right)\Delta W_{k+1}\right]\right|^2\\
&+\frac{1}{h^2}\mathbb{E}\left|P_k\mathbb{E}_k\left[\overleftarrow{\Delta}B_k\Delta W_{k+1}\left(g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)-g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^N\right)\right)\right]\right|^2\\
\leq&\frac{1}{h^2}\mathbb{E}\left|\mathbb{E}_k\left[\left(Y_{t_{k+1}}^{N,I,I}-Y_{t_{k+1}}^N\right)\Delta W_{k+1}\right]\right|^2\\
&+\frac{1}{h^2}\mathbb{E}\left|\mathbb{E}_k\left[\overleftarrow{\Delta}B_k\Delta W_{k+1}\left(g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)-g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^N\right)\right)\right]\right|^2\\
\leq&\frac{1}{h}\left(\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}-Y_{t_{k+1}}^N\right|^2-\mathbb{E}\left|\mathbb{E}_k\left(Y_{t_{k+1}}^{N,I,I}-Y_{t_{k+1}}^N\right)\right|^2\right)\\
&+\frac{1}{h}\left(\mathbb{E}\left|\overleftarrow{\Delta}B_k\left[g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)-g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^N\right)\right]\right|^2\right.\\
&\left.-\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta}B_k\left[g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)-g\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^N\right)\right]\right)\right|^2\right).
\end{aligned}$$

This concludes the proof of (4.7). □

For $Y\in L^2\left(\mathcal{F}_{t_k}\right)$, let $\chi_k^{N,I}(Y)$ be defined by:

$$\chi_k^{N,I}(Y):=P_k\left(Y_{t_{k+1}}^{N,I,I}+hf\left(X_{t_k}^N,Y,Z_{t_k}^{N,I}\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right).$$

The growth conditions of f and g deduced from (2.3), (2.4) and the orthogonality of $\overleftarrow{\Delta}B_k$ and $\mathcal{F}_{t_{k+1}}$ imply that $\chi_k^{N,I}\left(L^2\left(\mathcal{F}_{t_k}\right)\right)\subset\mathcal{P}_k\subset L^2\left(\mathcal{F}_{t_k}\right)$. Futhermore, (2.3) implies that for $Y_1,Y_2\in L^2\left(\mathcal{F}_{t_k}\right)$

$$\mathbb{E}\left|\chi_k^{N,I}(Y_2)-\chi_k^{N,I}(Y_1)\right|^2\leq L_fh^2\mathbb{E}\left|Y_2-Y_1\right|^2, \tag{4.8}$$

and (4.2) shows that $Y_{t_k}^{N,i,I}=\chi_k^{N,I}\left(Y_{t_k}^{N,i-1,I}\right)$ for $i=1,\ldots,I$.

Lemma 4.6. *For small h (i.e., $h^2L_f<1$) and for $k=0,\ldots,N-1$, there exists a unique $Y_{t_k}^{N,\infty,I}\in L^2\left(\mathcal{F}_{t_k}\right)$ such that*

$$Y_{t_k}^{N,\infty,I}=P_k\left[Y_{t_{k+1}}^{N,I,I}+hf\left(X_{t_k}^N,Y_{t_k}^{N,\infty,I},Z_{t_k}^{N,I}\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right], \tag{4.9}$$

$$\mathbb{E}\left|Y_{t_k}^{N,\infty,I}-Y_{t_k}^{N,i,I}\right|^2\leq L_f^ih^{2i}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2, \tag{4.10}$$

and there exists some constant $K>0$ such that for every N,k,I,

$$\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2\leq Kh+(1+Kh)\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2. \tag{4.11}$$

Proof. The fixed point theorem applied to the map $\chi_k^{N,I}$, which is a contration for $h^2L_f<1$, proves (4.9) ; (4.10) is straightforward consequence from (4.2) by induction on i. Lemma 4.4 shows that $P_kY_{t_{k+1}}^{N,I,I}$ and $P_k\left(\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)$ are orthogonal. Hence for any $\epsilon>0$, using Young's inequality, (4.3), the L^2 contracting property of P_k, the growth

condition of g deduced from (2.4) we obtain

$$\begin{aligned}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2 \leq& \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|P_kY_{t_{k+1}}^{N,I,I}\right|^2 + (h^2+2\epsilon h)\mathbb{E}\left|P_k\left[f\left(X_{t_k}^N, Y_{t_k}^{N,\infty,I}, Z_{t_k}^{N,I}\right)\right]\right|^2\\ &+\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|P_k\left[\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right]\right|^2\\ \leq&\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\mathbb{E}_kY_{t_{k+1}}^{N,I,I}\right|^2 + 2(h^2+2\epsilon h)\left|f(0,0,0)\right|^2\\ &+2L_f(h^2+2\epsilon h)\left(\mathbb{E}\left|X_{t_k}^N\right|^2+\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2+\mathbb{E}\left|Z_{t_k}^{N,I}\right|^2\right)\\ &+\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\mathbb{E}_k\left[\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right]\right|^2.\end{aligned}$$

Using the upper estimate (4.6) in Lemma 4.5, we obtain

$$\begin{aligned}&[1-2L_f(h^2+2\epsilon h)]\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2\\ &\leq\left(1+\frac{h}{\epsilon}-2L_f\left(h+2\epsilon\right)\right)\mathbb{E}\left|\mathbb{E}_kY_{t_{k+1}}^{N,I,I}\right|^2+2\left(h^2+\epsilon h\right)\left(|f(0,0,0)|^2+L_f\mathbb{E}\left|X_{t_k}^N\right|^2\right)\\ &\quad+2L_f\left(h+2\epsilon\right)\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2+2L_f\left(h+2\epsilon\right)\mathbb{E}\left|\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right|^2\\ &\quad+\left(1+\frac{h}{\epsilon}-2L_f\left(h+2\epsilon\right)\right)\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right)\right|^2.\end{aligned}$$

Choose ϵ such that $4L_f\epsilon=1$. Then $\left(1+\frac{h}{\epsilon}\right)-2L_f\left(h+2\epsilon\right)=2L_fh$ and $2L_f(h+2\epsilon)=2L_fh+1$. Using Theorem 3.1 we deduce the exitence of $C>0$ such that,

$$\begin{aligned}&\left[1-2L_f(h^2+2\epsilon h)\right]\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2\\ &\qquad\leq Ch+(1+4L_fh)\left[\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2+\mathbb{E}\left|\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right|^2\right].\end{aligned}$$

Then for $h^*\in(0,1]$ small enough (ie $(2L_f+1)h^*<1$), using Lemma 3.7, we deduce that for $\Gamma:=\frac{2L_f+1}{1-(2L_f+1)h^*}$ and $h\in(0,h^*)$, we have $(1-(2L_f+1)h)^{-1}\leq 1+\Gamma h$. Thus using the independence of $\overleftarrow{\Delta}B_k$ and $\mathcal{F}_{t_{k+1}}$, the growth condition (2.4) and Lemma 3.1, we deduce the existence of a constant $C>0$, such that for $h\in(0,h^*)$,

$$\begin{aligned}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2\leq&Ch+(1+Ch)\,\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2+Ch\mathbb{E}\left|g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right)\right|^2\\ \leq&Ch+(1+Ch)\,\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}\right|^2.\end{aligned}$$

This concludes the proof of (4.11). □

Let $\eta_k^{N,I}:=\mathbb{E}\left|Y_{t_k}^{N,I,I}-Y_{t_k}^N\right|^2$ for $k=0,\ldots,N$; the following lemma gives an upper bound of the L^2-norm of $Y_{t_k}^{N,\infty,I}-P_kY_{t_k}^N$ in terms of $\eta_{k+1}^{N,I}$.

Lemma 4.7. *For small h and for $k=0,\ldots,N-1$ we have:*

$$\mathbb{E}\left|Y_{t_k}^{N,\infty,I}-P_kY_{t_k}^N\right|^2\leq(1+Kh)\eta_{k+1}^{N,I}+Kh\left[\mathbb{E}\left|R_kY_{t_k}^N\right|^2+\mathbb{E}\left|R_kZ_{t_k}^N\right|^2\right].$$

Proof. The argument, which is similar to that in the proof of Lemmas 4.5 and 4.6 is more briefly sketched. Applying the operator P_k to both sides of equation (3.4) and using (4.3), we obtain

$$P_kY^N_{t_k} = P_kY^N_{t_{k+1}} + hP_k\left[f\left(X^N_{t_k}, Y^N_{t_k}, Z^N_{t_k}\right)\right] + P_k\left[\overleftarrow{\Delta}B_k g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right)\right].$$

Hence Lemma 4.6 implies that

$$\begin{aligned}Y^{N,\infty,I}_{t_k} - P_kY^N_{t_k} =& P_k\left[Y^{N,I,I}_{t_{k+1}} - Y^N_{t_{k+1}}\right] + hP_k\left[f\left(X^N_{t_k}, Y^{N,\infty,I}_{t_k}, Z^{N,I}_{t_k}\right) - f\left(X^N_{t_k}, Y^N_{t_k}, Z^N_{t_k}\right)\right]\\ &+ P_k\left(\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right).\end{aligned}$$

Lemma 4.4 proves the orthogonality of the first and third term of the above decomposition. Squaring this equation, using Young's inequality and (4.3), the L^2-contraction property of P_k and the Lipschitz property of g given in (2.4), computations similar to that made in the proof of Lemma 4.6 yield

$$\begin{aligned}\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - P_kY^N_{t_k}\right|^2 =& \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|P_k\left[Y^{N,I,I}_{t_{k+1}} - Y^N_{t_{k+1}}\right]\right|^2\\ &+ h^2\left(1+2\frac{\epsilon}{h}\right)\mathbb{E}\left|P_k\left[f\left(X^N_{t_k}, Y^{N,\infty,I}_{t_k}, Z^{N,I}_{t_k}\right) - f\left(X^N_{t_k}, Y^N_{t_k}, Z^N_{t_k}\right)\right]\right|^2\\ &+ \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|P_k\left(\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right)\right|^2\\ \le& \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\mathbb{E}_k\left[Y^{N,I,I}_{t_{k+1}} - Y^N_{t_{k+1}}\right]\right|^2 + L_f\left(h+2\epsilon\right)h\left(\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - Y^N_{t_k}\right|^2 + \mathbb{E}\left|Z^{N,I}_{t_k} - Z^N_{t_k}\right|^2\right)\\ &+ \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right)\right|^2. \qquad (4.12)\end{aligned}$$

By construction $Y^{N,\infty,I}_{t_k} \in \mathcal{P}_k$. Hence

$$\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - Y^N_{t_k}\right|^2 = \mathbb{E}\left|Y^{N,\infty,I}_{t_k} - P_kY^N_{t_k}\right|^2 + \mathbb{E}\left|R_kY^N_{t_k}\right|^2. \qquad (4.13)$$

Using Lemma 4.5 we deduce that for any $\epsilon > 0$

$$\begin{aligned}&\left(1 - L_f\left(h^2+2\epsilon h\right)\right)\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - P_kY^N_{t_k}\right|^2\\ \le& L_f\left(h+2\epsilon\right)\eta^{N,I}_{k+1} + hL_f\left(h+2\epsilon\right)\left[\mathbb{E}\left|R_kY^N_{t_k}\right|^2 + \mathbb{E}\left|R_kZ^N_{t_k}\right|^2\right]\\ &+ \left(\left(1+\frac{h}{\epsilon}\right) - L_f\left(h+2\epsilon\right)\right)\mathbb{E}\left|\mathbb{E}_k\left(Y^{N,I,I}_{t_{k+1}} - Y^N_{t_{k+1}}\right)\right|^2\\ &+ L_f\left(h+2\epsilon\right)\mathbb{E}\left|\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right|^2\\ &+ \left(\left(1+\frac{h}{\epsilon}\right) - L_f\left(h+2\epsilon\right)\right)\mathbb{E}\left|\mathbb{E}_k\left(\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right)\right|^2.\end{aligned}$$

Let $\epsilon > 0$ satisfy $2L_f\epsilon = 1$; then $\left(1+\frac{h}{\epsilon}\right) - L_f\left(h+2\epsilon\right) = L_fh$ and $L_f\left(h+2\epsilon\right) = L_fh + 1$. Thus, since $\mathbb{E}_k$ contracts the L^2-norm, we deduce

$$\begin{aligned}&\left(1 - L_f\left(h^2+2\epsilon h\right)\right)\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - P_kY^N_{t_k}\right|^2\\ \le& \left(1+2L_fh\right)\eta^{N,I}_{k+1} + h\left(1+L_fh\right)\left[\mathbb{E}\left|R_kY^N_{t_k}\right|^2 + \mathbb{E}\left|R_kZ^N_{t_k}\right|^2\right]\\ &+ \left(1+2L_fh\right)\mathbb{E}\left|\overleftarrow{\Delta}B_k\left[g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right]\right|^2.\end{aligned}$$

Let $h^* \in (0, \frac{1}{L_f+1})$ and set $\Gamma = \frac{L_f+1}{1-(L_f+1)h^*}$. Lemma 3.7 shows that for $h \in (0, h^*)$ we have $\left(1 - L_f\left(h^2 + 2\epsilon h\right)\right)^{-1} \leq 1 + \Gamma h$. The previous inequality, the independence of $\overleftarrow{\Delta} B_k$ and $\mathcal{F}_{t_{k+1}}$ and the Lipschitz property (2.4) imply that for some constant K which can change for one line to the next

$$\begin{aligned}
\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - P_k Y^N_{t_k}\right|^2 \leq{}& (1+Kh)\eta^{N,I}_{k+1} + Kh\left[\mathbb{E}\left|R_k Y^N_{t_k}\right|^2 + \mathbb{E}\left|R_k Z^N_{t_k}\right|^2\right] \\
&+ Kh\mathbb{E}\left|g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right) - g\left(X^N_{t_{k+1}}, Y^N_{t_{k+1}}\right)\right|^2 \\
\leq{}& (1+Kh)\eta^{N,I}_{k+1} + Kh\left[\mathbb{E}\left|R_k Y^N_{t_k}\right|^2 + \mathbb{E}\left|R_k Z^N_{t_k}\right|^2\right].
\end{aligned}$$

This concludes the proof of Lemma 4.7 □

The following Lemma provides L^2-bounds of $Y^{N,I,I}_{t_k}$, $Y^{N,\infty,I}_{t_k}$ and $Z^{N,I}_{t_k}$ independent of N and I.

Lemma 4.8. *There exists a constant K such that for large N and for every $I \geq 1$,*

$$\max_{0\leq k\leq N}\mathbb{E}\left|Y^{N,I,I}_{t_k}\right|^2 + \max_{0\leq k\leq N-1}\mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2 + \max_{0\leq k\leq N} h\mathbb{E}\left|Z^{N,I}_{t_k}\right|^2 \leq K.$$

Proof. Using inequality (4.10) and Young's inequality, we have the following bound, for $i = 1, \ldots, I$, $h < 1$ and some constant K depending on L_f:

$$\begin{aligned}
\mathbb{E}\left|Y^{N,i,I}_{t_k}\right|^2 &\leq \left(1 + \frac{1}{h}\right)\mathbb{E}\left|Y^{N,\infty,I}_{t_k} - Y^{N,i,I}_{t_k}\right|^2 + (1+h)\mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2 \\
&\leq \left(1 + \frac{1}{h}\right)L^i_f h^{2i}\mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2 + (1+h)\mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2 \leq (1+Kh)\mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2.
\end{aligned} \tag{4.14}$$

Choosing $i = I$ and using (4.11) we deduce that for some constant K which can change from line to line, $\mathbb{E}\left|Y^{N,I,I}_{t_k}\right|^2 \leq Kh + (1+Kh)\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2$. Hence Lemma 3.8 yields $\max_k \mathbb{E}\left|Y^{N,I,I}_{t_k}\right|^2 \leq K$. Plugging this relation into inequality (4.11) proves that

$$\max_k \mathbb{E}\left|Y^{N,I,I}_{t_k}\right|^2 + \max_k \mathbb{E}\left|Y^{N,\infty,I}_{t_k}\right|^2 \leq K < \infty.$$

Using (4.6) and the independence of $\overleftarrow{\Delta} B_k$ and $\mathcal{F}_{t_{k+1}}$, we deduce

$$\begin{aligned}
h\mathbb{E}\left|Z^{N,I}_{t_k}\right|^2 \leq{}& \mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2 + \mathbb{E}\left|\overleftarrow{\Delta} B_k g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right)\right|^2 \\
\leq{}& \mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2 + h\mathbb{E}\left|g\left(X^N_{t_{k+1}}, Y^{N,I,I}_{t_{k+1}}\right)\right|^2
\end{aligned}$$

Finally, the Lipschitz property (2.4) yields

$$\begin{aligned}
h\mathbb{E}\left|Z^{N,I}_{t_k}\right|^2 \leq{}& \mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2 + 2h\left|g\left(0,0\right)\right|^2 + 2hL_g\left(\mathbb{E}\left|X^N_{t_{k+1}}\right|^2 + \mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2\right) \\
\leq{}& \left(1 + 2hL_g\right)\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}\right|^2 + 2h\left|g\left(0,0\right)\right|^2 + 2hL_g\mathbb{E}\left|X^N_{t_{k+1}}\right|^2.
\end{aligned}$$

Theorem 3.1 and the L^2-upper estimates of $Y^{N,I,I}_{t_{k+1}}$ conclude the proof. □

The following lemma provides a backward recursive upper estimate of $\eta^{N,I}_\cdot$ Recall that $\eta^{N,I}_k = \mathbb{E}\left|Y^{N,I,I}_{t_k} - Y^N_{t_k}\right|^2$

Lemma 4.9. *For $0 \leq k < N$, we have:*

$$\eta_k^{N,I} \leq (1+Kh)\eta_{k+1}^{N,I} + Ch^{2I-1} + K\mathbb{E}\left|R_k Y_{t_k}^N\right|^2 + Kh\mathbb{E}\left|R_k Z_{t_k}^N\right|^2.$$

Proof. For $k = N$, $Y_{t_N}^N = \Phi\left(X_{t_N}^N\right)$ and $Y_{t_N}^{N,I,I} = P_N\Phi\left(X_{t_N}^N\right)$ so that $\eta_N^{N,I} = \mathbb{E}\left|\Phi\left(X_{t_N}^N\right) - P_N\Phi\left(X_{t_N}^N\right)\right|^2$. Let $k \in \{0,\dots,N-1\}$; using inequality (4.10) and Young's inequality, we obtain

$$\begin{aligned}
\eta_k^{N,I} =& \mathbb{E}\left|Y_{t_k}^{N,I,I} - Y_{t_k}^N\right|^2 \\
\leq& \left(1+\frac{1}{h}\right)\mathbb{E}\left|Y_{t_k}^{N,I,I} - Y_{t_k}^{N,\infty,I}\right|^2 + (1+h)\mathbb{E}\left|Y_{t_k}^{N,\infty,I} - Y_{t_k}^N\right|^2 \\
\leq& \left(1+\frac{1}{h}\right)L_f^I h^{2I}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2 + (1+h)\mathbb{E}\left|Y_{t_k}^{N,\infty,I} - P_k Y_{t_k}^N\right|^2 + (1+h)\mathbb{E}\left|R_k Y_{t_k}^N\right|^2.
\end{aligned}$$

Finally, Lemmas 4.8 and 4.7 imply that for some constant K we have for every N any $k = 1,\dots,N$:

$$\begin{aligned}
\eta_k^{N,I} \leq& Kh^{2I-1} + (1+h)\mathbb{E}\left|Y_{t_k}^{N,\infty,I} - P_k Y_{t_k}^N\right|^2 + (1+h)\mathbb{E}\left|R_k Y_{t_k}^N\right|^2 \qquad (4.15)\\
\leq& (1+Kh)\eta_{k+1}^{N,I} + Kh^{2I-1} + K\mathbb{E}\left|R_k Y_{t_k}^N\right|^2 + Kh\mathbb{E}\left|R_k Z_{t_k}^N\right|^2;
\end{aligned}$$

this concludes the proof. □

Gronwall's Lemma 3.8 and Lemma 4.9 prove the existence of C such that for h small enough

$$\begin{aligned}
\max_{0\leq k\leq N}\mathbb{E}\left|Y_{t_k}^{N,I,I} - Y_{t_k}^N\right|^2 \leq& Ch^{2I-2} + C\sum_{k=0}^{N-1}\mathbb{E}\left|R_k Y_{t_k}^N\right|^2 + Ch\sum_{k=0}^{N-1}\mathbb{E}\left|R_k Z_{t_k}^{N,I}\right|^2 \\
&+ C\mathbb{E}\left|\Phi\left(X_{t_N}^N\right) - P_N\Phi\left(X_{t_N}^N\right)\right|^2 \qquad (4.16)
\end{aligned}$$

which is part of Theorem 4.1. Let $\zeta^N := h\sum_{k=0}^{N-1}\mathbb{E}\left|Z_{t_k}^{N,I} - Z_{t_k}^N\right|^2$. In order to conclude the proof Theorem 4.1, we need to upper estimate ζ^N, which is done in the next lemma.

Lemma 4.10. *There exits a constant C such that for h small enough and every $I \geq 1$*

$$\zeta^N \leq Ch^{2I-2} + Ch\sum_{k=0}^{N-1}\mathbb{E}\left|R_k Z_k^N\right|^2 + C\sum_{k=0}^{N-1}\mathbb{E}\left|R_k Y_k^N\right|^2 + C\max_{0\leq k\leq N-1}\eta_k^{N,I}.$$

Proof. Multiply inequality (4.7) by h, use the independence of $\overleftarrow{\Delta}B_k$ and $\mathcal{F}_{t_{k+1}}$ and the Lipschitz property (2.4); this yields

$$\zeta^N \leq h\sum_{k=0}^{N-1}\mathbb{E}\left|R_k Z_{t_k}^N\right|^2 + \sum_{k=0}^{N-1}\left((1+L_g h)\,\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N\right|^2 - \mathbb{E}\left|\mathbb{E}_k\left(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N\right)\right|^2\right). \qquad (4.17)$$

Multiply inequality(4.12) by $(1+L_gh)(1+h)$, use the independence of $\overleftarrow{\Delta}B_k$ and $\mathcal{F}_{t_{k+1}}$ and the Lipschitz property (2.4); this yields for $\epsilon>0$:

$$\begin{aligned}(1+L_gh)(1+h)&\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-P_kY^N_{t_k}\right|^2\\
&\leq\left(1+\frac{h}{\epsilon}\right)(1+L_gh)(1+h)\mathbb{E}\left|\mathbb{E}_k\left[Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right]\right|^2\\
&\quad+L_f\left(h+2\epsilon\right)h(1+L_gh)(1+h)\left(\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-Y^N_{t_k}\right|^2+\mathbb{E}\left|Z^{N,I}_{t_k}-Z^N_{t_k}\right|^2\right)\\
&\quad+\left(1+\frac{h}{\epsilon}\right)(1+L_gh)(1+h)L_gh\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right|^2. \qquad (4.18)\end{aligned}$$

Multiply inequality (4.15) by $(1+L_gh)$ and use (4.18); this yields for some constants K, C, $\bar{C}$ and $h\in(0,1]$, $\epsilon>0$:

$$\begin{aligned}\Delta_{k+1}:=&(1+L_gh)\,\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right|^2-\mathbb{E}\left|\mathbb{E}_k\left(Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right)\right|^2\\
\leq&Kh^{2I-1}+K\mathbb{E}\left|R_kY^N_{t_k}\right|^2+\left(\left(1+\frac{h}{\epsilon}\right)(1+L_gh)(1+h)-1\right)\mathbb{E}\left|\mathbb{E}_k\left[Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right]\right|^2\\
&+C\left(h+2\epsilon\right)h\left(\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-Y^N_{t_k}\right|^2+\mathbb{E}\left|Z^{N,I}_{t_k}-Z^N_{t_k}\right|^2\right)\\
&+\left(1+\frac{h}{\epsilon}\right)Ch\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right|^2.\end{aligned}$$

Now we choose ϵ such that $2C\epsilon=\frac{1}{4}$; then we have for some constant K and $h\in(0,1]$:

$$\begin{aligned}\Delta_{k+1}\leq&Kh^{2I-1}+K\mathbb{E}\left|R_kY^N_{t_k}\right|^2+Kh\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right|^2\\
&+\left(Ch+\frac{1}{4}\right)h\left(\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-Y^N_{t_k}\right|^2+\mathbb{E}\left|Z^{N,I}_{t_k}-Z^N_{t_k}\right|^2\right).\end{aligned}$$

Thus, for h small enough (so that $Ch\leq\frac{1}{4}$), summing over k we obtain

$$\begin{aligned}\sum_{k=0}^{N-1}&\left((1+L_gh)\mathbb{E}\left|Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right|^2-\mathbb{E}\left|\mathbb{E}_k\left(Y^{N,I,I}_{t_{k+1}}-Y^N_{t_{k+1}}\right)\right|^2\right)\\
&\leq Kh^{2I-2}+K\sum_{k=0}^{N-1}\mathbb{E}\left|R_kY^N_{t_k}\right|^2+K\max_k\eta^{N,I}_k\\
&\quad+\frac{1}{2}h\sum_{k=0}^{N-1}\left(\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-Y^N_{t_k}\right|^2+\mathbb{E}\left|Z^{N,I}_{t_k}-Z^N_{t_k}\right|^2\right).\end{aligned}$$

Plugging this inequality in (4.17) yields

$$\begin{aligned}\frac{1}{2}\zeta^N\leq&Kh^{2I-2}+h\sum_{k=0}^{N-1}\mathbb{E}\left|R_kZ^N_{t_k}\right|^2+K\sum_{k=0}^{N-1}\mathbb{E}\left|R_kY^N_{t_k}\right|^2+K\max_k\eta^{N,I}_{t_k}\\
&+\frac{1}{2}h\sum_{k=0}^{N-1}\mathbb{E}\left|Y^{N,\infty,I}_{t_k}-Y^N_{t_k}\right|^2.\end{aligned}$$

Using (4.13) and Lemma 4.7, we obtain for some constant K and every $h \in (0,1]$

$$h\sum_{k=0}^{N-1}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}-Y_{t_k}^{N}\right|^2 \leq (1+Kh)h\sum_{k=0}^{N-1}\eta_{k+1}^{N,I}+Kh^2\sum_{k=0}^{N-1}\left[\mathbb{E}\left|R_kY_{t_k}^{N}\right|^2+\mathbb{E}\left|R_kZ_{t_k}^{N}\right|^2\right]$$
$$+h\sum_{k=0}^{N-1}\mathbb{E}\left|R_kY_{t_k}^{N}\right|^2 \leq K\max_k\eta_k^{N,I}+K\sum_{k=0}^{N-1}\mathbb{E}\left|R_kY_{t_k}^{N}\right|^2+Kh\sum_{k=0}^{N-1}\mathbb{E}\left|R_kZ_{t_k}^{N}\right|^2.$$

This concludes the proof of Lemma 4.10. □

Theorem 4.1 is a straightforward consequence of inequality (4.16) and Lemma 4.10.

5. Approximation step 3

In this section we will use regression approximations and introduce some minimization problem for a M-sample of (B,W) denoted by $(B^m,W^m,m=1,\dots,M)$. This provides a Monte Carlo approximation of Y^N,I,I and Z^N,I on the time grid.

5.1. Some more notations for the projection.

We at first introduce some notations

(N5) *For fixed $k=1,\dots,N$ and $m=1,\dots,M$, let p_k^m denote the orthonormal family of $L^2(\Omega)$ similar to p_k in (N4) replacing X^N by $X^{N,m}$ and B by B^m.*

(N6) *For a real $n\times n$ symmetric matrix A, $\|A\|$ is the maximum of the absolute value of its eigenvalues and $\|A\|_F=\left(\sum_{i,j}A_{i,j}^2\right)^{\frac{1}{2}}$ its Frobenius norm. If $A:\mathbb{R}^n\to\mathbb{R}^n$ also denotes the linear operator whose matrix in the canonical basis is A, then $\|A\|$ is the operator-norm of A when $\mathbb{R}^n$ is endowed with the Euclidian norm. Note that $\|A\|\leq\|A\|_F$ follows from Schwarz's inequality.*

(N7) *For $k=0,\dots,N-1$ and $m=1,\dots,M$ let v_k^m and v_k be column vectors whose entries are the components in the canonical base of the vectors*

$$\left(p_k^m,p_k^m\frac{\Delta W_{k+1}^m}{\sqrt{h}}\right),\ \text{and}\ \left(p_k,p_k\frac{\Delta W_{k+1}}{\sqrt{h}}\right) \tag{5.1}$$

respectively. Note that $\mathbb{E}v_kv_k^*=Id$, since the entries of p_k are an orthonormal family of $L^2(\mathcal{F}_k)$ and $\frac{\Delta W_{k+1}}{h}$ is a normed vector in L^2 independent of p_k.

(N8) *For $k=0,\dots,N-1$ let V_k^M,P_k^M be symmetric matrices defined by*

$$V_k^M:=\frac{1}{M}\sum_{m=1}^{M}v_k^m\left[v_k^m\right]^*,P_k^M:=\frac{1}{M}\sum_{m=1}^{M}p_k^m(p_k^m)^*. \tag{5.2}$$

(N9) *We denote by $\mathcal{N}$ the σ-algebra of measurable sets A with $\mathbb{P}(A)=0$ and set:*

$$\mathcal{F}_t^{W,m}=\sigma\left(W_s^m;0\leq s\leq t\right)\vee\mathcal{N},\quad \mathcal{F}_{t,t'}^{B,m}=\sigma\left(B_s^m-B_{t'}^m;t\leq s\leq t'\right)\vee\mathcal{N},$$
$$\mathcal{F}_t^{W,M}=\mathcal{F}_t^W\vee\bigvee_{m=1}^{M}\mathcal{F}_t^{W,m},\quad \mathcal{F}_{t,T}^{B,M}=\mathcal{F}_{t,T}^B\vee\bigvee_{m=1}^{M}\mathcal{F}_{t,T}^{B,m},\quad \mathcal{F}_t=\mathcal{F}_t^W\vee\mathcal{F}_{t,T}^B.$$

Note that $(\mathcal{F}_t)_t$ and $\left(\mathcal{F}_{t,T}^B\right)_t$ are not filtrations.

(N10) *In the sequel we will need to localize some processes using the following events*

$$\mathfrak{A}_j:=\left\{\|V_j^M-Id\|\leq h,\|P_j^M-Id\|\leq h\right\}\in\mathcal{F}_{t_{j+1}}^{W,M}\vee\mathcal{F}_{t_j,T}^{B,M}, \tag{5.3}$$

$$A_k^M:=\bigcap_{j=k}^{N-1}\mathfrak{A}_j\in\mathcal{F}_{t_N}^{W,M}\vee\mathcal{F}_{t_k,T}^{B,M}. \tag{5.4}$$

(N11) *For $x = (x_1, \dots, x_m) \in \mathbb{R}^M$, we denote $|x|_M^2 := \frac{1}{M}\sum_{m=1}^M |x_m|^2$.*

5.2. **Another look at the previous results.** *We introduce the following random variables*

$$\zeta_k^N := \rho_k^N := \left(|p_k|\sqrt{C_0}\right) \vee 1,$$

where C_0 is constant in the Lemma 4.8. Since $Y_{t_k}^{N,i,I}$ and $Z_{t_k}^{N,I}$ are in $\mathcal{P}_k$ (see (4.1) and (4.2)), we can rewrite these random variables as follows:

$$Y_{t_k}^{N,i,I} = \alpha_k^{i,I}.p_k = \left(\alpha_k^{i,I}\right)^* p_k, \quad Z_{t_k}^{N,I} = \beta_k^I.p_k = \left(\beta_k^I\right)^* pk, \tag{5.5}$$

where $\alpha_k^{i,I}$ (resp. β_k^I) is the vector of the coefficient in the basis p_k of the random variable $Y_{t_k}^{N,i,I}$ (resp. $Z_{t_k}^{N,I}$), identified with the column matrix of the coefficients in the canonical basis.

Remark 5.1. Note that the vectors $\alpha_k^{i,I}$ and β_k^I are deterministic.

The following Proposition gives a priori estimates of $Y_{t_k}^{N,i,I}$ and $Z_{t_k}^{N,I}$.

Proposition 5.2. *For $i \in \{1, \dots, I\} \cup \{\infty\}$ and for $k = 0, \dots, N$, we have $\left|Y_{t_k}^{N,i,I}\right| \leq \rho_k^N$, $\sqrt{h}\left|Z_{t_k}^{N,I}\right| \leq \zeta_k^N$. Moreover, for every I and $i = 0, \dots, I$:*

$$\left|\alpha_k^{i,I}\right|^2 \leq \mathbb{E}\left|\rho_k^N\right|^2, \quad \left|\beta_k^I\right|^2 \leq \frac{1}{h}\mathbb{E}\left|\zeta_k^N\right|^2. \tag{5.6}$$

Proof. Let $i \in \{1, \dots, I\} \cup \{\infty\}$ and $k = 0, \dots, N$. Squaring $Y_{t_k}^{N,i,I}$, taking expectation and using the previous remark, we obtain

$$\mathbb{E}\left|Y_{t_k}^{N,i,I}\right|^2 = \left(\alpha_k^{i,I}\right)^* \mathbb{E}\left(p_k p_k^*\right)\alpha_k^{i,I} \geq \left(\alpha_k^{i,I}\right)^* \alpha_k^{i,I} = \left|\alpha_k^{i,I}\right|^2$$

Using Lemma 4.8, we deduce that $\left|\alpha_k^{i,I}\right|^2 \leq C_0$. The Cauchy-Schwarz inequality implies

$$\left|Y_{t_k}^{N,i,I}\right| \leq \left|\alpha_k^{i,I}\right||p_k| \leq |p_k|\sqrt{C_0} \leq \left(|p_k|\sqrt{C_0}\right) \vee 1.$$

A similar computation based on Lemma 4.8 proves that $\sqrt{h}\left|Z_{t_k}^{N,I}\right| \leq \zeta_k^N$. The upper estimates of $\left|\alpha_k^{i,I}\right|^2$ and $\left|\beta_k^I\right|^2$ are straightforward consequences of the previous ones. □

We now prove that $\left(\alpha_k^{i,I}, \beta_k^I\right)$ solves a minimization problem.

Proposition 5.3. *The vector $\left(\alpha_k^{i,I}, \beta_k^I\right)$ solves the following minimization problem: for $k = 0, \dots, N-1$ and for every $i = 1, \dots, I$, we have:*

$$\begin{aligned}\left(\alpha_k^{i,I}, \beta_k^I\right) = \arg\min_{(\alpha,\beta)} \mathbb{E}\Big|Y_{k+1}^{N,I,I} - \alpha.p_k + hf\left(X_{t_k}^N, \alpha_k^{i-1,I}.p_k, Z_{t_k}^{N,I}\right) \\ + \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, Y_{t_{k+1}}^{N,I,I}\right) - \beta.p_k \Delta W_{k+1}\Big|^2.\end{aligned} \tag{5.7}$$

Proof. Let $(Y,Z)\in\mathcal{P}_k\times\mathcal{P}_k$; then since $\mathcal{P}_k\subset L^2\left(\mathcal{F}_{t_k}\right)$ and ΔW_{k+1} is independent of $\mathcal{F}_{t_k}$, we have

$$\begin{aligned}
&\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}-Y+hf\left(X_{t_k}^N,Y_{t_k}^{N,i-1,I},Z_{t_k}^{N,I}\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)-Z\Delta W_{k+1}\right|^2\\
=&\mathbb{E}\left|Y_{t_{k+1}}^{N,I,I}-Y+hf\left(X_{t_k}^N,Y_{t_k}^{N,i-1,I},Z_{t_k}^{N,I}\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right|^2\\
&+h\mathbb{E}\left|Z-\frac{1}{h}\left(Y_{t_{k+1}}^{N,I,I}+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)\Delta W_{k+1}\right|^2\\
&-\frac{1}{h}\mathbb{E}\left|\left(Y_{t_{k+1}}^{N,I,I}+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right)\Delta W_{k+1}\right|^2.
\end{aligned}$$

The minimun on pairs of elements of $\mathcal{P}_k$ is given by the orthogonal projections, that is by the random variables $Y=Y_{t_k}^{N,i,I}$ and $Z=Z_{t_k}^{N,I}$ defined by (4.2) and (4.1) respectively. This concludes the proof using the notations introduced in (5.5). □

For $i\in\{1,\dots,I\}\cup\{\infty\}$, we define $\theta_k^{i,I}:=\left(\alpha_k^{i,I},\sqrt{h}\beta_k^I\right)$. The following lemma gives some properties on $\theta_k^{i,I}$.

Lemma 5.4. *For all $i\in\{1,\dots,I\}\cup\{\infty\}$, we have for $k=0,\dots,N$ (resp. for $k=0,\dots,N-1$)*

$$\left|\theta_k^{i,I}\right|^2\le\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2,\quad resp.\quad \left|\theta_k^{\infty,I}-\theta_k^{i,I}\right|^2\le L_f^ih^{2i}\mathbb{E}\left|\rho_k^N\right|^2.$$

Furthermore, we have the following explicit expression of $\theta_k^{\infty,I}$ for v_k defined by (5.1):

$$\theta_k^{\infty,I}=\mathbb{E}\left[v_k\left(\alpha_{k+1}^{I,I}.p_{k+1}+hf\left(X_k^N,\alpha_k^{\infty,I}.p_k,\beta_k^I.p_k\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,\alpha_{k+1}^{I,I}.p_{k+1}\right)\right)\right].\tag{5.8}$$

Proof. Proposition 5.2 implies that $\left|\theta_k^{i,I}\right|^2=\left|\alpha_k^{i,I}\right|^2+h\left|\beta_k^I\right|^2\le\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2$. Using inequality (4.10) and Proposition 5.2, since $\mathbb{E}\left|p_k\right|^2=1$ we obtain

$$\left|\theta_k^{\infty,I}-\theta_k^{i,I}\right|^2=\mathbb{E}\left|Y_{t_k}^{N,\infty,I}-Y_{t_k}^{N,i,I}\right|^2\le L_f^ih^{2i}\mathbb{E}\left|Y_{t_k}^{N,\infty,I}\right|^2\le L_f^ih^{2i}\mathbb{E}\left|\rho_k^N\right|^2.$$

Using equation (4.9) and the fact that the components of p_k are an orthonormal family of L^2, we have

$$\begin{aligned}
\alpha_k^{\infty,I}=&\mathbb{E}\left[p_kY_k^{N,\infty,I}\right]\\
=&\mathbb{E}\left(p_kP_k\left[Y_{t_{k+1}}^{N,I,I}+hf\left(X_{t_k}^N,Y_{t_k}^{N,\infty,I},Z_{t_k}^{N,I}\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\right]\right)\\
=&\mathbb{E}\left[p_k\left(\alpha_{k+1}^{I,I}.p_{k+1}+hf\left(X_k^N,\alpha_k^{\infty,I}.p_k,\beta_k^I.p_k\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,\alpha_{k+1}^{I,I}.p_{k+1}\right)\right)\right].
\end{aligned}$$

A similar computation based on equation (4.1) and on the independence of $\mathcal{F}_{t_k}$ and ΔW_{k+1} yields

$$\begin{aligned}
\sqrt{h}\beta_k^I=&\mathbb{E}\left[\sqrt{h}p_kZ_{t_k}^{N,I}\right]\\
=&\mathbb{E}\left[\frac{1}{\sqrt{h}}p_kP_k\left(Y_{t_{k+1}}^{N,I,I}\Delta W_{k+1}+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,Y_{t_{k+1}}^{N,I,I}\right)\Delta W_{k+1}\right)\right]\\
=&\mathbb{E}\left[p_k\left(\alpha_{k+1}^{I,I}.p_{k+1}\frac{\Delta W_{k+1}}{\sqrt{h}}+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,\alpha_{k+1}^{I,I}.p_{k+1}\right)\frac{\Delta W_{k+1}}{\sqrt{h}}\right)\right]\\
=&\mathbb{E}\left[p_k\frac{\Delta W_{k+1}}{\sqrt{h}}\left(\alpha_{k+1}^{I,I}.p_{k+1}+hf\left(X_k^N,\alpha_k^{\infty,I}.p_k,\beta_k^I.p_k\right)+\overleftarrow{\Delta}B_kg\left(X_{t_{k+1}}^N,\alpha_{k+1}^{I,I}.p_{k+1}\right)\right)\right].
\end{aligned}$$

Finally, we recall by (5.1) that $v_k := \left(p_k, p_k \frac{\Delta W_{k+1}}{\sqrt{h}}\right)$; this concludes the proof. □

5.3. **The numerical scheme.** Let $\xi : \mathbb{R} \to \mathbb{R}$ be a C_b^2 function, such that $\xi(x) = x$ for $|x| \leq 3/2, |\xi|_\infty \leq 2$ and $|\xi'|_\infty \leq 1$. We define the random truncation functions

$$\widehat{\rho}_k^N(x) := \rho_k^N \xi\left(\frac{x}{\rho_k^N}\right), \quad \widehat{\zeta}_k^N(x) := \zeta_k^N \xi\left(\frac{x}{\zeta_k^N}\right). \tag{5.9}$$

The following lemma states some properties of these functions.

Lemma 5.5. *Let $\widehat{\rho}_k^N$ and $\widehat{\zeta}_k^N$ be defined by (5.9), then*

(1) *$\widehat{\rho}_k^N$ (resp. $\widehat{\zeta}_k^N$) leaves $Y_{t_k}^{N,I,I}$ (resp. $\sqrt{h} Z_{t_k}^{N,I}$) invariant, that is:*

$$\widehat{\rho}_k^N\left(\alpha_k^{I,I}.p_k\right) = \alpha_k^{I,I}.p_k, \quad \widehat{\zeta}_k^N\left(\sqrt{h}\beta_k^I.p_k\right) = \sqrt{h}\beta_k^I.p_k.$$

(2) *$\widehat{\rho}_k^N, \widehat{\zeta}_k^N$ are 1-Lipschitz and $\left|\widehat{\rho}_k^N(x)\right| \leq |x|$ for every $x \in \mathbb{R}$.*

(3) *$\widehat{\rho}_k^N$ (resp. $\widehat{\zeta}_k^N$) is bounded by $2\left|\rho_k^N\right|$ (resp. by $2\left|\zeta_k^N\right|$).*

Proof. In part (1)-(3) we only give the proof for $\widehat{\rho}_k^N$, since that for $\widehat{\zeta}_k^N$ is similar.

1. By Proposition 5.2, $\left|\frac{\alpha_k^{I,I}.p_k}{\rho_k^N}\right| \leq 1$. Hence, $\xi\left(\frac{\alpha_k^{I,I}.p_k}{\rho_k^N}\right) = \frac{\alpha_k^{I,I}.p_k}{\rho_k^N}$.

2. Let $y, y' \in \mathbb{R}$; since $|\xi'|_\infty \leq 1$,

$$\left|\widehat{\rho}_k^N(y) - \widehat{\rho}_k^N(y')\right| = \left|\rho_k^N\right| \left|\xi\left(\frac{y}{\rho_k^N}\right) - \xi\left(\frac{y'}{\rho_k^N}\right)\right| \leq |y - y'|.$$

Since $\widehat{\rho}_k^N(0) = 0$, we deduce $\left|\widehat{\rho}_k^N(x)\right| \leq |x|$.

3. This upper estimate is a straightforward consequence of $|\xi|_\infty \leq 2$; this concludes the proof. □

Let $\left(X_\cdot^{N,m}\right)_{1\leq m\leq M}$, $\left(\Delta W_\cdot^m\right)_{1\leq m\leq M}$ and $\left(\overleftarrow{\Delta} B_\cdot^m\right)_{1\leq m\leq M}$ be independent realizations of X^N, ΔW and $\overleftarrow{\Delta} B$ respectively. In a similar way, we introduce the following random variables and random functions:

$$\begin{aligned} \zeta_k^{N,m} :=& \rho_k^{N,m} := \left|p_k^m\right| \sqrt{C_0} \vee 1, \\ \widehat{\zeta}_k^{N,m}(x) :=& \zeta_k^{N,m} \xi\left(\frac{x}{\zeta_k^{N,m}}\right), \quad \widehat{\rho}_k^{N,m}(x) := \rho_k^{N,m} \xi\left(\frac{x}{\rho_k^{N,m}}\right), \ x \in \mathbb{R}. \end{aligned} \tag{5.10}$$

An argument similar to that used to prove Lemma 5.5 yields the following:

Lemma 5.6. *The random functions $\widehat{\rho}_k^{N,m}(.)$ defined above satisfy the following properties:*

(1) *$\widehat{\rho}_k^{N,m}$ is bounded by $2\left|\rho_k^{N,m}\right|$ and is 1-Lipschitz.*

(2) *$\rho_k^{N,m}$ and ρ_k^N have the same distribution.*

We now describe the numerical scheme

Definition 5.7. *Initialization.* At time $t = t_N$, set $Y_{t_N}^{N,i,I,M} := \alpha_N^{i,I,M}.p_N := P_N \Phi\left(X_{t_N}^N\right)$ and $\beta_N^{i,I,M} = 0$ for all $i \in \{1, \ldots, I\}$.

Induction Assume that an approximation $Y_{t_l}^{N,i,I,M}$ is built for $l = k+1, \ldots, N$ and set $Y_{t_{k+1}}^{N,I,I,M,m} := \widehat{\rho}_k^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)$ its realization along the mth simulation.

We use backward induction in time and forward induction on i. For $i = 0$, let $\alpha_k^{0,I,M} =$

$\beta_k^{0,I,M}=0$. For $i=1,\dots,I$, the vector $\theta_k^{i,I,M}:=\left(\alpha_k^{i,I,M},\sqrt{h}\beta_k^{i,I,M}\right)$ is defined by (forward) induction as the arg min in (α,β) of the quantity:

$$\frac{1}{M}\sum_{m=1}^{M}\Big|Y_{t_{k+1}}^{N,I,I,M,m}-\alpha.p_k^m+hf\left(X_{t_k}^{N,m},\alpha_k^{i-1,I,M}.p_k^m,\beta_k^{i-1,I,M}.p_k^m\right)$$
$$+\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},Y_{t_{k+1}}^{N,I,I,M,m}\right)-\beta.p_k^m\Delta W_{k+1}^m\Big|^2. \tag{5.11}$$

This minimization problem is similar to (5.7) replacing the expected value by an average over M independent realizations. Note that $\theta_k^{i,I,M}=\left(\alpha_k^{i,I,M},\sqrt{h}\beta_k^{i,I,M}\right)$ is a random vector. We finally set:

$$Y_{t_k}^{N,I,I,M}:=\widehat{\rho}_k^N\left(\alpha_k^{I,I,M}.p_k\right),\sqrt{h}Z_{t_k}^{N,I,I,M}:=\widehat{\zeta}_k^N\left(\sqrt{h}\beta_k^{i,I,M}.p_k\right), \tag{5.12}$$

The following theorem gives an upper estimate of the L^2 error beetween $\left(Y_\cdot^{N,I,I},Z_\cdot^{N,I}\right)$ and $\left(Y_\cdot^{N,I,I,M},Z_\cdot^{N,I,I,M}\right)$ in terms of $\left|\zeta_\cdot^N\right|$ and $\left|\rho_\cdot^N\right|$; it is the main result of this section. We recall that by (5.4) $A_k^M=\bigcap_{j=k}^{N-1}\left\{\left\|V_j^M-Id\right\|\le h,\|P_j^M-Id\|\le h\right\}\in\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_k,T}^{B,M}$. For $k=1,\dots,N-1$ set

$$\begin{aligned}\epsilon_k:=&\mathbb{E}\|v_kv_k^*-Id\|_F^2\left(\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right)+\mathbb{E}\left[|v_k|^2\left|p_{k+1}\right|^2\right]\mathbb{E}\left|\rho_{k+1}^N\right|^2\\&+h^2\mathbb{E}\left[|v_k|^2\left(1+\left|X_k^N\right|^2+|p_k|^2\mathbb{E}\left|\rho_k^N\right|^2+\frac{1}{h}|p_k|^2\mathbb{E}\left|\zeta_k^N\right|^2\right)\right]\\&+h\mathbb{E}\left[\left(|v_k|^2+\left|w_k^p\right|^2\right)\left(1+\left|X_{t_{k+1}}^N\right|^2+|p_{k+1}|^2\mathbb{E}\left|\rho_{k+1}^N\right|^2\right)\right].\end{aligned} \tag{5.13}$$

Choosing N and then M large enough, the following result gives the speed of convergence of the Monte Carlo approximation scheme of $Y^{N,I,I}$ and $Z^{N,I}$.

Theorem 5.8. *There exists a constant $C>0$ such that for h small enough, for any $k=0,\dots,N-1$ and $M\ge1$:*

$$\begin{aligned}\mathfrak{E}_M:=&\mathbb{E}\left|Y_{t_k}^{N,I,I}-Y_{t_k}^{N,I,I,M}\right|^2+h\sum_{j=k}^{N-1}\mathbb{E}\left|Z_{t_j}^{N,I}-Z_{t_j}^{N,I,I,M}\right|^2\\\le&16\sum_{j=k}^{N-1}\mathbb{E}\left[\left(\left|\zeta_j^N\right|^2+\left|\rho_j^N\right|^2\right)1_{\left[A_k^M\right]^c}\right]+Ch^{I-1}\sum_{j=k}^{N-1}\left(h^2+h\mathbb{E}\left|\rho_{j+1}^N\right|^2+\mathbb{E}\left|\rho_j^N\right|^2+\mathbb{E}\left|\zeta_j^N\right|^2\right)\\&+\frac{C}{hM}\sum_{j=k}^{N-1}\epsilon_j.\end{aligned}$$

5.4. **Proof of Theorem 5.8.** Before we start the proof, let us recall some results on regression (i.e. orthogonal projections). Let $v=(v^m)_{1\le m\le M}$ be a sequence of vectors in $\mathbb{R}^n$. Let use define the $n\times n$ matrix $V^M:=\frac{1}{M}\sum_{m=1}^M v^mv^{m*}$, suppose that V^M is invertible and denote by $\lambda_{\min}\left(V^M\right)$ its smallest eigenvalue.

Lemma 5.9. *Under the above hypotheses, we have the following results: Let $(x^m,m=1,\dots,M)$ be a vector in $\mathbb{R}^M$.*

1. *There exists a unique $\mathbb{R}^n$ valued vector θ_x satisfying $\theta_x=\underset{\theta\in\mathbb{R}^n}{\arg\inf}|x-\theta.v|_M^2$ where $\theta.v$ denotes the vector $(\sum_{i=1}^n\theta(i)v^m(i),m=1\dots,M)$.*
2. *Moreover, we have $\theta_x=\frac{1}{M}\left(V^M\right)^{-1}\sum_{m=1}^M x^mv^m\in\mathbb{R}^n$*

(3) *The map* $x \mapsto \theta_x$ *is linear from* $\mathbb{R}^M$ *to* $\mathbb{R}^n$ *and* $\lambda_{\min}(V^M)|\theta_x|^2 \leq |\theta_x.v|_M^2 \leq |x|_M^2$.

The following lemma gives a first upper estimate of $\mathfrak{E}_M$.

Lemma 5.10. *For every* M *and* $k = 0, \dots, N-1$, *we have the following upper estimate*

$$\begin{aligned}\mathfrak{E}_M \leq & \mathbb{E}\left[\left|\alpha_k^{I,I} - \alpha_k^{I,I,M}\right|^2 1_{A_k^M}\right] + h\sum_{j=k}^{N-1}\mathbb{E}\left[\left|\beta_j^I - \beta_j^{I,I,M}\right|^2 1_{A_j^M}\right] \\ & + 16\mathbb{E}\left[\left|\rho_k^N\right|^2 1_{[A_k^M]^c}\right] + 16\sum_{j=k}^{N-1}\mathbb{E}\left[\left|\zeta_j^N\right|^2 1_{[A_j^M]^c}\right].\end{aligned}$$

This lemma should be compared with inequality (31) in [7].

Proof. Using the decomposition of $Y^{N,I,I}$, $Y^{N,I,I,M}$, $Z^{N,I}$ and $Z^{N,I,I,M}$, Lemma 5.5 (1) , we deduce

$$\begin{aligned}\mathfrak{E}_M = & \mathbb{E}\left[\left|\widehat{\rho}_k^N\left(\alpha_k^{I,I}.p_k\right) - \widehat{\rho}_k^N\left(\alpha_k^{I,I,M}.p_k\right)\right|^2\right] \\ & + h\sum_{j=k}^{N-1}\mathbb{E}\left[\left|\frac{1}{\sqrt{h}}\widehat{\zeta}_j^N\left(\sqrt{h}\beta_j^I.p_j\right) - \frac{1}{\sqrt{h}}\widehat{\zeta}_j^N\left(\sqrt{h}\beta_j^{I,I,M}.p_j\right)\right|^2\right].\end{aligned}$$

Using hte partition A_k^M, $\left(A_k^M\right)^c$ where A_k^M is defined by (5.4), Cauchy-Schwarz's inequality, Lemma 5.5 and the independence of $\left(\alpha_k^{I,I,M}, \beta_j^{I,I,M}, 1_{A_k^M}\right)$ and p_k we deduce:

$$\begin{aligned}\mathfrak{E}_M \leq & \mathbb{E}\left[\left|\alpha_k^{I,I}.p_k - \alpha_k^{I,I,M}.p_k\right|^2 1_{A_k^M}\right] + h\sum_{j=k}^{N-1}\mathbb{E}\left[\left|\beta_j^I.p_j - \beta_j^{I,I,M}.p_j\right|^2 1_{A_j^M}\right] \\ & + 2\mathbb{E}\left[\left(\left|\widehat{\rho}_k^N\left(\alpha_k^{I,I}.p_k\right)\right|^2 + \left|\widehat{\rho}_k^N\left(\alpha_k^{I,I,M}.p_k\right)\right|^2\right)1_{[A_k^M]^c}\right] \\ & + 2\sum_{j=k}^{N-1}\mathbb{E}\left[\left(\left|\widehat{\zeta}_j^N\left(\sqrt{h}\beta_j^I.p_j\right)\right|^2 + \left|\widehat{\zeta}_j^N\left(\sqrt{h}\beta_j^{I,I,M}.p_j\right)\right|^2\right)1_{[A_j^M]^c}\right] \\ \leq & \mathbb{E}\left[\left(\alpha_k^{I,I} - \alpha_k^{I,I,M}\right)^* p_k p_k^* \left(\alpha_k^{I,I} - \alpha_k^{I,I,M}\right) 1_{A_k^M}\right] \\ & + h\sum_{j=k}^{N-1}\mathbb{E}\left[\left(\beta_j^I - \beta_j^{I,I,M}\right)^* p_j p_j^* \left(\beta_j^I - \beta_j^{I,I,M}\right) 1_{A_j^M}\right] \\ & + 2\mathbb{E}\left[8\left|\rho_k^N\right|^2 1_{[A_k^M]^c}\right] + 2\sum_{j=k}^{N-1}\mathbb{E}\left[8\left|\zeta_j^N\right|^2 1_{[A_j^M]^c}\right] \\ \leq & \mathbb{E}p_k p_k^* \mathbb{E}\left[\left|\alpha_k^{I,I} - \alpha_k^{I,I,M}\right|^2 1_{A_k^M}\right] + h\sum_{j=k}^{N-1}\mathbb{E}p_j p_j^* \mathbb{E}\left[\left|\beta_j^I - \beta_j^{I,I,M}\right|^2 1_{A_j^M}\right] \\ & + 16\mathbb{E}\left[\left|\rho_k^N\right|^2 1_{[A_k^M]^c}\right] + 16\sum_{j=k}^{N-1}\mathbb{E}\left[\left|\zeta_j^N\right|^2 1_{[A_j^M]^c}\right].\end{aligned}$$

This concludes the proof. □

We now upper estimate $\left|\theta_k^{I,I,M} - \theta_k^{I,I}\right|^2$ on the event A_k^M. This will be done in severals lemmas below. By definition $\|V_k^M - I\| \leq h$ on A_k^M for any $k = 1, \dots, N$. Hence for

$h \in (0,1)$

$$1 - h \leq \lambda_{\min}\left(V_k^M(\omega)\right) \quad \text{on } A_k^M. \tag{5.14}$$

Lemma 5.11. *For every $\alpha \in \mathbb{R}^n$ and $k = 1, \ldots, N$, we have $\frac{1}{M}\sum_{m=1}^M |\alpha.p_k^m|^2 \leq |\alpha|^2 \left\|P_k^M\right\|$.*

Proof. The definition of the Euclidian norm and of P_k^M imply

$$\frac{1}{M}\sum_{m=1}^M |\alpha.p_k^m|^2 = \alpha^* \frac{1}{M}\sum_{m=1}^M p_k^m \left(p_k^m\right)^* \alpha = \alpha^* P_k^M \alpha \leq \left\|P_k^M\right\| |\alpha|^2;$$

this concludes the proof. □

For $i = 0, \ldots, I$, we introduce the vector $x_k^{i,I,M} := \left(x_k^{i,I,m,M}\right)_{m=1,\ldots,M}$ defined for $m = 1, \ldots, M$ by:

$$\begin{aligned} x_k^{i,I,m,M} :=& \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right) + hf\left(X_k^{N,m}, \alpha_k^{i,I,M}.p_k^m, \beta_k^{i,I,M}.p_k^m\right) \\ &+ \overleftarrow{\Delta} B_k^m g\left(X_{k+1}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right). \end{aligned} \tag{5.15}$$

Using Lemma 5.9, we can rewrite equation (5.11) as follows:

$$\theta_k^{i,I,M} = \arg\inf_\theta \left|x_k^{i-1,I,M} - \theta.v_k^m\right|_M^2 = \frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^M x_k^{i-1,I,m,M} v_k^m. \tag{5.16}$$

We will need the following

Lemma 5.12. *For all $k = 0, \ldots, N-1$ and every I, the random variables $\alpha_k^{I,I,M}$ are $\mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$ measurable.*

Proof. The proof uses backward indution on k and forward induction on i.
Initialization. Let $k = N-1$. By definition $\alpha_{N-1}^{0,I,M} = 0$. Let $i \geq 1$ and suppose $\alpha_{N-1}^{i-1,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{N-1},T}^{B,M}$. Using (5.1) (resp. (5.2)), we deduce that $v_{N-1}^m \in \mathcal{F}_T^{W,m} \vee \mathcal{F}_{t_{N-1},T}^{B,m}$ (resp. $V_{N-1}^M \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{N-1},T}^{B,M}$).
Futhermore (5.15) shows that $x_{N-1}^{i-1,I,m,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{N-1},T}^{B,M}$ and hence (5.16) implies that $\alpha_{N-1}^{i,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{N-1},T}^{B,M}$.
Induction. Suppose that $\alpha_{k+1}^{I,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{k+1},T}^{B,M}$; we will prove by forward induction on i that $\alpha_k^{i,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$ for $i = 0, \ldots, I$.
By definition $\alpha_k^{0,I,M} = 0$. Suppose $\alpha_k^{i-1,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$; we prove that $\alpha_k^{i,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$ by similar arguments. Indeed, (5.1) (resp. (5.2)) implies that $v_k^m \in \mathcal{F}_T^{W,m} \vee \mathcal{F}_{t_k,T}^{B,m}$ (resp. $V_k^M \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$), while (5.15) (resp. (5.16)) yields $x_k^{i-1,I,m,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$ (resp. $\alpha_k^{i,I,M} \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_k,T}^{B,M}$). This concludes the proof. □

The following Lemma gives an inductive upper estimate of $\left|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}\right|^2$.

Lemma 5.13. *There exists $\widetilde{C} > 0$ such that for small h, for $k = 0, \ldots, N-1$ and for $i = 1, ..., I-1$ $\left|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}\right|^2 \leq \widetilde{C}h\left|\theta_k^{i,I,M} - \theta_k^{i-1,I,M}\right|^2$ on A_k^M.*

Proof. Using (5.14) and Lemma 5.9 (4), we obtain on A_k^M

$$(1-h)\left|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}\right|^2 \leq \lambda_{\min}\left(V_k^M\right)\left|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}\right|^2 \leq \left|x_k^{i,I,M} - x_k^{i-1,I,M}\right|_M^2.$$

Plugging equation (5.15) and using the Lipschitz property (2.3) of f, we deduce

$$(1-h)\left|\theta_k^{i+1,I,M}-\theta_k^{i,I,M}\right|^2 \leq \frac{h^2L_f}{M}\sum_{m=1}^{M}\left(\left|\left(\alpha_k^{i,I,M}-\alpha_k^{i-1,I,M}\right).p_k^m\right|^2 + \left|\left(\beta_k^{i,I,M}-\beta_k^{i-1,I,M}\right).p_k^m\right|^2\right).$$

Lemma 5.11 and the inequality $\|P_k^M\| \leq 2$, yield

$$\begin{aligned}(1-h)\left|\theta_k^{i+1,I,M}-\theta_k^{i,I,M}\right|^2 \leq& \left(\left|\alpha_k^{i,I,M}-\alpha_k^{i-1,I,M}\right|^2+\left|\beta_k^{i,I,M}-\beta_k^{i-1,I,M}\right|^2\right)h^2L_f\left\|P_k^M\right\| \\ \leq& 2hL_f\left|\theta_k^{i,I,M}-\theta_k^{i-1,I,M}\right|^2.\end{aligned}$$

This concludes the proof. □

For $\theta=\left(\alpha,\sqrt{h}\beta\right)$ set $F_k(\theta):=\arg\inf_{\theta^*}\left|x_k^{I,M}(\theta)-\theta^*.v_k\right|^2$ where
$x_k^{I,M}(\theta):=\rho_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)+hf\left(X_{t_k}^{N,m},\alpha.p_k^m,\beta.p_k^m\right)+\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},\hat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)$.

Lemma 5.14. *On A_k^M, the map F_k is Lipschitz with a Lipschitz constant $2hL_f(1-h)^{-1}$.*

Proof. Using (5.14) and Lemma 5.9 (3), we obtain on A_k^M

$$(1-h)\left|F_k(\theta_1)-F_k(\theta_2)\right|^2 \leq \lambda_{\min}\left(V_k^M\right)\left|F_k(\theta_1)-F_k(\theta_2)\right|^2 \leq \left|x_k^{I,M}(\theta_1)-x_k^{I,M}(\theta_2)\right|^2.$$

Using the Lipschitz property (2.3) of f, Lemma 5.11 and the inequality $\|P_k^M\| \leq 2$, we deduce that on A_k^M:

$$\begin{aligned}(1-h)\left|F_k(\theta_1)-F_k(\theta_2)\right|^2 \leq& \frac{h^2L_f}{M}\sum_{m=1}^{M}\left(|\alpha_1.p_k^m-\alpha_2.p_k^m|^2+|\beta_1.p_k^m-\beta_2.p_k^m|^2\right). \\ \leq& |\alpha_1-\alpha_1|^2h^2L_f\left\|P_k^M\right\|+|\beta_1-\beta_2|^2h^2L_f\left\|P_k^M\right\| \\ \leq& 2hL_f|\theta_1-\theta_2|^2;\end{aligned}$$

this concludes the proof. □

The Lipschitz property of F_k yields the following:

Corollary 5.15. *(i) For h small enough, on A_k^M, there exists a unique random vector $\theta_k^{\infty,I,M}:=\left(\alpha_k^{\infty,I,M},\sqrt{h}\beta_k^{\infty,I,M}\right)$ such that*

$$\theta_k^{\infty,I,M}=\frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^{M}x_k^{\infty,I,m,M}v_k^m=\arg\inf_{\theta}\left|x_k^{I,M}\left(\theta_k^{\infty,I,M}\right)-\theta.v_k\right|_M^2, \tag{5.17}$$

where for $\theta=\left(\alpha,\sqrt{h}\beta\right)$, $x_k^{I,M}(\theta):=\left(x_k^{I,m,M}(\theta)\right)_{m=1,\ldots,M}$ denotes the vector with components

$$\begin{aligned}x_k^{I,m,M}(\theta):=&\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)+hf\left(X_{t_k}^{N,m},\alpha.p_k^m,\beta.p_k^m\right) \\ &+\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right).\end{aligned}$$

Let $x_k^{\infty,I,M} = \left(x_k^{\infty,I,m,M}\right)_{m=1,\dots,M} = \left(x^{I,m,M}\left(\theta_k^{\infty,I,M}\right)\right)_{m=1,\dots,M}$.
(ii) Moreover there exits a constant $C>0$ *such that for small* h *and any* $k=0,\dots,N-1$

$$\left|\theta_k^{\infty,I,M}-\theta_k^{I,I,M}\right|^2 \le Ch^I\left|\theta_k^{\infty,I,M}\right|^2.$$

Proof. (i) This is a consequence of Lemma 5.14 since $2hL_f(1-h)^{-1}<1$ for small h.
(ii) An argument similar to that used to prove Lemma 5.14 implies that for $i=1,\dots,I$

$$(1-h)\left|\theta_k^{\infty,I,M}-\theta_k^{I,I,M}\right|^2 \le 2hL_f\left|\theta_k^{\infty,I,M}-\theta_k^{I-1,I,M}\right|^2$$

Since $\theta_k^{0,I,M}=0$, we conclude the proof. □

The following result, similar to Lemma 4.4, will be crucial in subsequent estimates. It requires some additional argument compared with similar estimates in [7].

Lemma 5.16. *Let* U_{k+1}^m *be a* $\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_{k+1},T}^{B,M}$ *measurable random variable. Then we have*

$$\mathbb{E}\left[1_{A_k^M}U_{k+1}^m\overleftarrow{\Delta}B_k^m\right]=0.$$

Proof. Using (5.3) and (5.4) we deduce

$$\mathbb{E}\left(1_{A_k^M}U_{k+1}^m\overleftarrow{\Delta}B_k^m\right)=\mathbb{E}\left(1_{A_{k+1}^M}U_{k+1}^m\mathbb{E}\left(\overleftarrow{\Delta}B_k^m1_{\mathfrak{A}_k}\middle|\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_{k+1},T}^{B,M}\right)\right)$$

Recall that $\mathfrak{A}_k=\left\{\|V_k^M-Id\|\le h,\|P_k^M-Id\|\le h\right\}$. We will prove that

$$1_{\mathfrak{A}_k}=f\left(\overleftarrow{\Delta}B_k^1,\dots,\overleftarrow{\Delta}B_k^M\right) \tag{5.18}$$

with a symmetric function f, that is $f(\beta_1,\dots,\beta_M)=f(-\beta_1,\dots,-\beta_M)$ for any $\beta\in\mathbb{R}^M$. Suppose at first that (5.18) is true. Since the distribution of the vectors $\left(\overleftarrow{\Delta}B_k^1,\dots,\overleftarrow{\Delta}B_k^M\right)$ and $\left(-\overleftarrow{\Delta}B_k^1,\dots,-\overleftarrow{\Delta}B_k^M\right)$ are the same, the independence of $\left(\overleftarrow{\Delta}B_k^l,l=1,\dots,M\right)$ and $\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_{k+1},T}^{B,M}$ yields

$$\begin{aligned}\mathbb{E}\left(\overleftarrow{\Delta}B_k^m1_{\mathfrak{A}_k}\middle|\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_{k+1}}^{B,M}\right)=&\mathbb{E}\left(\overleftarrow{\Delta}B_k^mf\left(\overleftarrow{\Delta}B_k^1,\dots,\overleftarrow{\Delta}B_k^M\right)\right)\\=&\mathbb{E}\left(-\overleftarrow{\Delta}B_k^mf\left(-\overleftarrow{\Delta}B_k^1,\dots,-\overleftarrow{\Delta}B_k^M\right)\right)\\=&-\mathbb{E}\left(\overleftarrow{\Delta}B_k^mf\left(\overleftarrow{\Delta}B_k^1,\dots,\overleftarrow{\Delta}B_k^M\right)\right).\end{aligned}$$

Which concludes the proof.
Let us now prove (5.18). Clearly, it is enough to prove to prove that each norm involved in the definition of $\mathfrak{A}_k$ is of this form. Let A be one of the matrices V_k^M or P_k^M. Now we will compute the characteristic polynomial χ of the matrix $A-Id$ and prove that its coefficients are symmetric.

Let q^m be p_k^m or v_k^m. We reorganize q^m as $q^m=\left(q_1^m,q_2^m\overleftarrow{\Delta}B_k^m\right)^*$, where q_1^m are the elements of q^m independent of $\overleftarrow{\Delta}B_k^m$, and q_2^m is independent of $\overleftarrow{\Delta}B_k^m$. So we have

$$q^m\left(q^m\right)^*=\left(\begin{array}{c|c}q_1^m\left(q_1^m\right)^* & q_1^m\left(q_2^m\right)^*\overleftarrow{\Delta}B_k^m\\ \hline q_2^m\left(q_1^m\right)^*\overleftarrow{\Delta}B_k^m & q_2^m\left(q_2^m\right)^*\overleftarrow{\Delta}B_k^m\end{array}\right)$$

Let $A=\frac{1}{M}\sum_{m=1}^M q^m\left(q^m\right)^*$; then the characteristic polynomial of the matrix $A-Id$ is given by

$$\chi(A-Id)(X)=\det\left(\begin{array}{c|c}B-(X+1)Id & C\\ \hline C^* & D-(X+1)Id\end{array}\right)$$

where

$$B := \frac{1}{M}\sum_{m=1}^{M} q_1^m q_1^{m,*} \in M_{I_1\times I_1}(\mathbb{R}), \quad C := \frac{1}{M}\sum_{m=1}^{M} q_1^m q_2^{m,*} \overleftarrow{\Delta} B_k^m \in M_{I_1\times I_2}(\mathbb{R}),$$
$$D := \frac{1}{M}\sum_{m=1}^{M} q_2^m q_2^{m,*} \left|\overleftarrow{\Delta} B_k^m\right|^2 \in M_{I_2\times I_2}(\mathbb{R}).$$

Set $J_1 = \{1,\dots,I_1\}$ and $J_2 = \{I_1+1,\dots,I_1+I_2\}$, and for $\sigma \in \mathfrak{S}_{I_1+I_2}$ the following sets $\mathcal{H}(\alpha,\sigma,\beta) = \{i \in J_\alpha, \sigma(i) \in J_\beta\}$, for $\alpha,\beta \in \{1,2\}$. Using the definition of the determinant, we have

$$\chi(A-Id)(X) = \sum_{\sigma\in\mathfrak{S}_{I_1+I_2}} \epsilon(\sigma) \prod_{i\in\mathcal{H}(1,\sigma,1)} \left[B(i,\sigma(i)) - (X+1)\delta_{i,\sigma(i)}\right] \prod_{i\in\mathcal{H}(1,\sigma,1)} C(i,\sigma(i)) \prod_{i\in\mathcal{H}(2,\sigma,1)} C(\sigma(i),i) \prod_{i\in\mathcal{H}(2,\sigma,2)} \left[D(i,\sigma(i)) - (X+1)\delta_{i,\sigma(i)}\right]$$

Since we have the relation $|\mathcal{H}(1,\sigma,1)|+|\mathcal{H}(1,\sigma,2)| = |J_1| = I_1$ and $|\mathcal{H}(1,\sigma,1)|+|\mathcal{H}(2,\sigma,1)| = |J_1| = I_1$, we deduce that $|\mathcal{H}(1,\sigma,1)| + |\mathcal{H}(2,\sigma,1)|$ is even. Therefore, the power of $\overleftarrow{\Delta} B_k^m$ in $\chi(A-Id)(X)$ is even, which concludes the proof. □

As a corollary, we deduce the following identities

Corollary 5.17. *For $k = 0,\dots,N-1$, we have*

$$\mathbb{E}\left[1_{A_k^M} \sum_{m=1}^{M} \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right) \overleftarrow{\Delta} B_k^m g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right] = 0, \tag{5.19}$$

$$\mathbb{E}\Big[1_{A_k^M}\Big(\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) - \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\Big) \overleftarrow{\Delta} B_k^m\left(g\left(X_{t_{k+1}}^{N,m}, \alpha_{k+1}^{I,I}.p_{k+1}^m\right) - g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right)\Big] = 0 \tag{5.20}$$

Proof. Indeed, $X_{t_{k+1}}^{N,m} \in \mathcal{F}_T^{W,M}$. Futhermore, (5.10), Lemma 5.12 and the definition of p_{k+1}^m imply that $\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right), \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) \in \mathcal{F}_T^{W,M} \vee \mathcal{F}_{t_{k+1},T}^{B,M}$. Thus Lemma 5.16 concludes the proof. □

The following result provides an L^2 bound of $\theta_k^{\infty,I,M}$ in terms of ρ_{k+1}^N.

Lemma 5.18. *There exists a constant C such that, for every N and $k = 0,\dots,N-1$, $\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2\right] \le C\mathbb{E}\left|\rho_{k+1}^N\right|^2 + Ch$.*

Proof. Using (5.14), Lemma 5.9 (3) and Corollary 5.15 (i) we have on A_k^M

$$(1-h)\left|\theta_k^{\infty,I,M}\right|^2 \le \lambda_{\min}(V_k^M)\left|\theta_k^{\infty,I,M}\right|^2 \le \left|x_k^{\infty,I,M}\right|_M^2.$$

Using (N11), taking expectation, using Young's inequality and (5.19), we deduce for any $\epsilon > 0$, $k = 0,\dots,N-1$,

$$(1-h)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2\right] \le \sum_{j=1}^{3} T_k^{I,M}(j),$$

where

$$T_k^{I,M}(1) := \frac{1}{M}\left(1+\frac{h}{\epsilon}\right)\sum_{m=1}^{M}\mathbb{E}\left[1_{A_k^M}\left|\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right|^2\right],$$

$$T_k^{I,M}(2) := \frac{h^2}{M}\left(1+2\frac{\epsilon}{h}\right)\sum_{m=1}^{M}\mathbb{E}\left[1_{A_k^M}\left|f\left(X_{t_k}^{N,m},\alpha_k^{\infty,I,M}.p_k^m,\beta_k^{\infty,I,M}.p_k^m\right)\right|^2\right],$$

$$T_k^{I,M}(3) := \frac{1}{M}\left(1+\frac{h}{\epsilon}\right)\sum_{m=1}^{M}\mathbb{E}\left[1_{A_k^M}\left|\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right|^2\right].$$

Lemma 5.6 yields

$$T_k^{I,M}(1) \leq 4\frac{1}{M}\left(1+\frac{h}{\epsilon}\right)\sum_{m=1}^{M}\mathbb{E}\left|\rho_{k+1}^{N,m}\right|^2 \leq 4\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\rho_{k+1}^{N}\right|^2. \tag{5.21}$$

The Lipschitz condition (2.3) of f, Lemma 5.11 and the inequalities $\left\|P_k^M\right\| \leq 2$ valid on A_k^M imply

$$\begin{aligned}
T_k^{I,M}(2) \leq & 2L_f h(h+2\epsilon)\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{\infty,I,M}.p_k^m\right|^2+1_{A_k^M}\left|\beta_k^{\infty,I,M}.p_k^m\right|^2\right]\\
&+2h(h+2\epsilon)\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left|f\left(X_{t_k}^{N,m},0,0\right)\right|^2\\
\leq & 2L_f h(h+2\epsilon)\mathbb{E}\left\{1_{A_k^M}\left(\left|\alpha_k^{\infty,I,M}\right|^2+\left|\beta_k^{\infty,I,M}\right|^2\right)\left\|P_k^M\right\|\right\}+2h(h+2\epsilon)\mathbb{E}\left|f\left(X_{t_k}^{N},0,0\right)\right|^2\\
\leq & 4L_f h(h+2\epsilon)\mathbb{E}\left[1_{A_k^M}\left(\left|\alpha_k^{\infty,I,M}\right|^2+\left|\beta_k^{\infty,I,M}\right|^2\right)\right]+2h(h+2\epsilon)\mathbb{E}\left|f\left(X_{t_k}^{N},0,0\right)\right|^2.
\end{aligned} \tag{5.22}$$

Finally, since $\overleftarrow{\Delta}B_k^m$ is independent of $\mathcal{F}_T^{W,M}\vee\mathcal{F}_{t_{k+1},T}^{B,M}$ for every $m=1,\ldots,M$, the Lipschitz property (2.4) of g and Lemma 5.6 (1) yield for $m=1,\ldots,M$

$$\begin{aligned}
&\mathbb{E}\left[1_{A_k^M}\left|\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right|^2\right]\\
&\quad=\mathbb{E}\left[1_{A_{k+1}^M}1_{\mathfrak{A}_k}\left|\overleftarrow{\Delta}B_k^m g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right|^2\right]\\
&\quad=\mathbb{E}\left[1_{A_{k+1}^M}\left|g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right|^2\mathbb{E}\left(1_{\mathfrak{A}_k}\left|\overleftarrow{\Delta}B_k^m\right|^2\left|\mathcal{F}_{t_N}^{W,M}\vee\mathcal{F}_{t_{k+1},T}^{B,M}\right.\right)\right]\\
&\quad\leq h\mathbb{E}\left[1_{A_{k+1}^M}\left|g\left(X_{t_{k+1}}^{N,m},\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right|^2\right]\\
&\quad\leq 8L_g h\mathbb{E}\left[1_{A_{k+1}^M}\left|\rho_{k+1}^{N,m}\right|^2\right]+2h\mathbb{E}\left[1_{A_{k+1}^M}\left|g\left(X_{t_{k+1}}^{N,m},0\right)\right|^2\right].
\end{aligned}$$

Therefore,

$$T_k^{I,M}(3) \leq 8L_g h\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|\rho_{k+1}^{N}\right|^2+2h\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|g\left(X_{t_{k+1}}^{N},0\right)\right|^2. \tag{5.23}$$

The inequalities (5.21)-(5.23) imply that for any $\epsilon > 0$ and $h \in (0,1]$,

$$\begin{aligned}(1-h)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2\right] \leq & \left\{4\left(1+\frac{h}{\epsilon}\right)+8L_g h\left(1+\frac{h}{\epsilon}\right)\right\}\mathbb{E}\left|\rho_{k+1}^N\right|^2 \\ & + 4L_f h(h+2\epsilon)\mathbb{E}\left[1_{A_k^M}\left(\left|\alpha_k^{\infty,I,M}\right|^2+\left|\beta_k^{\infty,I,M}\right|^2\right)\right] \\ & + 2h(h+2\epsilon)\mathbb{E}\left|f\left(X_{t_k}^N,0,0\right)\right|^2 + 2h\left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left|g\left(X_{t_{k+1}}^N,0\right)\right|^2.\end{aligned}$$

Choose ϵ such that $8L_f\epsilon = \frac{1}{4}$ so that $4L_f(h+2\epsilon) = \frac{1}{4} + 4L_f h$. For h small enough (that is $h \leq \frac{1}{4(4L_f+\frac{1}{2})}$), we have $4L_f(h+2\epsilon) \leq \frac{1}{2}(1-h)$. Hence, we deduce $\frac{1}{2}(1-h)\mathbb{E}1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2 \leq C\mathbb{E}\left|\rho_{k+1}^N\right|^2 + Ch$, which concludes the proof. □

The next result yields an upper estimate of the L^2-norm of $\theta_k^{I,I,M} - \theta_k^{I,I}$ in terms of $\theta_k^{\infty,I,M} - \theta_k^{\infty,I}$.

Lemma 5.19. *There is a constant C such that for every N large enough and all $k = 0,\ldots,N-1$,*

$$\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{I,I}\right|^2\right] \leq (1+Ch)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right] + Ch^{I-1}\left(\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right).$$

Proof. We decompose $\theta_k^{I,I,M} - \theta_k^{I,I}$ as follows:

$$\theta_k^{I,I,M} - \theta_k^{I,I} = \left(\theta_k^{\infty,I,M} - \theta_k^{\infty,I}\right) + \left(\theta_k^{I,I,M} - \theta_k^{\infty,I,M}\right) - \left(\theta_k^{I,I} - \theta_k^{\infty,I}\right).$$

Young's inequality implies

$$\left|\theta_k^{I,I,M}-\theta_k^{I,I}\right|^2 = (1+h)\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2 + 2\left(1+\frac{1}{h}\right)\left(\left|\theta_k^{I,I}-\theta_k^{\infty,I}\right|^2+\left|\theta_k^{I,I,M}-\theta_k^{\infty,I,M}\right|^2\right).$$

Taking expectation over the set A_k^M, using Lemma 5.4 and the fact that $\alpha_k^{i,I}$ and β_k^I are deterministic, we deduce

$$\begin{aligned}\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{I,I}\right|^2\right] \leq & (1+h)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right] + 2\left(1+\frac{1}{h}\right)L_f^I h^{2I}\mathbb{E}\left|\rho_k^N\right|^2 \\ & + 2\left(1+\frac{1}{h}\right)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{\infty,I,M}\right|^2\right].\end{aligned}$$

Since $\theta_k^{\infty,I}$ is deterministic, Corollary 5.15 (ii) and again Lemma 5.4 yield

$$\begin{aligned}\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{\infty,I,M}\right|^2\right] \leq & Ch^I\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2\right] \\ \leq & Ch^I\left(\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right) + Ch^I\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right].\end{aligned}$$

Therefore, we deduce

$$\begin{aligned}\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{I,I}\right|^2\right] \leq & (1+h)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right] + 2\left(1+\frac{1}{h}\right)Ch^I\left(\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right) \\ & + 2\left(1+\frac{1}{h}\right)Ch^I\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right] + 2\left(1+\frac{1}{h}\right)L_f^I h^{2I}\mathbb{E}\left|\rho_k^N\right|^2 \\ \leq & (1+Ch)\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right] + Ch^{I-1}\left(\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right),\end{aligned}$$

which concludes the proof. □

The rest of this section is devoted to upper estimate $\theta_k^{\infty,I,M} - \theta_k^{\infty,I}$ on A_k^M. We at first decompose $\theta_k^{\infty,I} - \theta_k^{\infty,I,M}$ as follows:

$$\theta_k^{\infty,I} - \theta_k^{\infty,I,M} = \sum_{i=1}^{5} \mathfrak{B}_i, \tag{5.24}$$

where $\mathfrak{B}_2$, $\mathfrak{B}_3$ and $\mathfrak{B}_5$ introduce a Monte-Carlo approximation of some expected value by an average over the M-realization: for $k = 0, \dots, N-1$,

$$\begin{aligned}
\mathfrak{B}_1 &:= \left(Id - \left(V_k^M\right)^{-1}\right) \theta_k^{\infty,I},\\
\mathfrak{B}_2 &:= \left(V_k^M\right)^{-1} \left[\mathbb{E}\left(v_k \widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right) - \frac{1}{M}\sum_{m=1}^{M} v_k^m \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right],\\
\mathfrak{B}_3 &:= \left(V_k^M\right)^{-1} h\left[\mathbb{E}\left(v_k f\left(X_k^N, \alpha_k^{\infty,I}.p_k, \beta_k^I.p_k\right)\right) - \frac{1}{M}\sum_{m=1}^{M} v_k^m f\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I}.p_k^m, \beta_k^I.p_k^m\right)\right],\\
\mathfrak{B}_4 :=& \frac{1}{M}\left(V_k^M\right)^{-1} \sum_{m=1}^{M} v_k^m \left[\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) - \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right.\\
&\quad + hf\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I}.p_k^m, \beta_k^I.p_k^m\right) - hf\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I,M}.p_k^m, \beta_k^{\infty,I,M}.p_k^m\right)\\
&\quad \left. + \overleftarrow{\Delta} B_k^m\left[g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right) - g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right]\right],\\
\mathfrak{B}_5 &:= \left(V_k^M\right)^{-1}\left[\mathbb{E}\left(v_k \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right) - \frac{1}{M}\sum_{m=1}^{M} v_k^m \overleftarrow{\Delta} B_k^m g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right)\right].
\end{aligned}$$

Note that compared to the similar decomposition in [7], $\mathfrak{B}_4$ is slightly different and $\mathfrak{B}_5$ is new. Indeed, using equation (5.8) and (5.17) and Lemma 5.5 (1), we obtain:

$$\begin{aligned}
\theta_k^{\infty,I} - \theta_k^{\infty,I,M} &= \left(Id - \left(V_k^M\right)^{-1}\right)\theta_k^{\infty,I} + \left(V_k^M\right)^{-1}\theta_k^{\infty,I} - \theta_k^{\infty,I,M}\\
&= \mathfrak{B}_1 + \left(V_k^M\right)^{-1}\mathbb{E}\left[v_k\left(\alpha_{k+1}^{I,I}.p_{k+1} + hf\left(X_k^N, \alpha_k^{\infty,I}.p_k, \beta_k^I.p_k\right) + \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right)\right]\\
&\quad - \frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^{M} v_k^m\left[\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right) + hf\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I,M}.p_k^m, \beta_k^{\infty,I,M}.p_k^m\right)\right.\\
&\qquad \left. + \overleftarrow{\Delta} B_k^m g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right]
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j\in\{1,2,3,5\}} \mathfrak{B}_j + \frac{1}{M}\left(V_k^M\right)^{-1}\left[\sum_{m=1}^{M} v_k^m \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) - \sum_{m=1}^{M} v_k^m \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right]\\
&\quad + \frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^{M} v_k^m\left[hf\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I}.p_k^m, \beta_k^I.p_k^m\right) - hf\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I,M}.p_k^m, \beta_k^{\infty,I,M}.p_k^m\right)\right]\\
&\quad + \frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^{M} v_k^m \overleftarrow{\Delta} B_k^m\left[g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right) - g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\right],
\end{aligned}$$

which concludes the proof of (5.24).

The following lemmas provide upper bounds of the error terms $\mathfrak{B}_i$. Recall that if F is a matrix such that $\|Id - F\| < 1$, then F is inversible, $F^{-1} - Id = \sum_{k\geq 1}(Id-F)^k$ and

$$\|Id - F^{-1}\| \leq \frac{\|Id - F\|}{1 - \|Id - F\|} \tag{5.25}$$

Indeed, $F^{-1} = (Id - (Id-F))^{-1} = \sum_{k\geq 0}(Id-F)^k$ and $\|Id - F^{-1}\| \leq \sum_{k\geq 1}\|(Id-F)^k\|$.

Lemma 5.20. *(i) Let $(U_1, ..., U_M)$ be a sequence of iid centered random variables. Then we have $\mathbb{E}\left|\sum_{m=1}^M U_j\right|^2 = M\mathbb{E}\left|U_1\right|^2$.*
(ii) We have $\mathbb{E}\left\|\sum_{m=1}^M \left(v_k^m \left(v_k^m\right)^* - Id\right)\right\|_F^2 = M\mathbb{E}\|v_k v_k^* - Id\|_F^2$.

Proof. (i) The proof is straightforward.
(ii) Using (i) (N6) and (N7), we deduce

$$\begin{aligned}\mathbb{E}\left\|\sum_{m=1}^M \left[v_k^m \left(v_k^m\right)^* - Id\right]\right\|_F^2 =& \sum_{i,j}\mathbb{E}\left|\sum_{m=1}^M \left[v_k^m \left(v_k^m\right)^* - Id\right](i,j)\right|^2 \\ =& M\sum_{i,j}\mathbb{E}\left|\left[v_k \left(v_k\right)^* - Id\right](i,j)\right|^2 = M\mathbb{E}\|v_k v_k^* - Id\|_F^2;\end{aligned}$$

this concludes the proof of the Lemma. □

The following lemma provides a L^2 upper bound of $\mathfrak{B}_1$. Recall that A_k^M is defined by (5.4).

Lemma 5.21 (Upper estimate of $\mathfrak{B}_1$)**.** *There exist a constant C such that for small h and every $M \geq 1$,*

$$\mathbb{E}\left[\left|\mathfrak{B}_1\right|^2 1_{A_k^M}\right] \leq \frac{C}{M}\mathbb{E}\|v_k v_k^* - Id\|_F^2 \left(\mathbb{E}\left|\rho_k^N\right|^2 + \mathbb{E}\left|\zeta_k^N\right|^2\right).$$

Proof. On A_k^M we have $\|Id - V_k^M\| \leq h < 1$; and hence (5.25) implies $\|Id - (V_k^M)^{-1}\| \leq \frac{\|Id - V_k^M\|}{1 - \|Id - V_k^M\|} \leq \frac{\|Id - V_k^M\|}{1-h}$. Using the inequality $\|.\| \leq \|.\|_F$ we deduce

$$\mathbb{E}\left[\|Id - (V_k^M)^{-1}\|^2 1_{A_k^M}\right] \leq \frac{1}{(1-h)^2}\mathbb{E}\left[1_{A_k^M}\left\|Id - V_k^M\right\|_F^2\right].$$

By definition $V_k^M = \frac{1}{M}\sum_{m=1}^M v_k^m \left(v_k^m\right)^*$; so using Lemma 5.20 we obtain $\mathbb{E}\left[1_{A_k^M}\left\|Id - V_k^M\right\|_F^2\right] \leq \frac{1}{M}\mathbb{E}\|v_k v_k^* - Id\|_F^2$. Therefore, since $\theta_k^{\infty,I}$ is deterministic, Lemma 5.4 yields

$$\begin{aligned}\mathbb{E}\left[\left|\mathfrak{B}_1\right|^2 1_{A_k^M}\right] \leq& \left|\theta_k^{\infty,I}\right|^2 \mathbb{E}\left[\|Id - (V_k^M)^{-1}\|^2 1_{A_k^M}\right] \\ \leq& \frac{C}{M}\left(\mathbb{E}\left|\rho_k^N\right|^2 + \mathbb{E}\left|\zeta_k^N\right|^2\right)\mathbb{E}\|v_k v_k^* - Id\|_F^2;\end{aligned}$$

this concludes the proof. □

The next lemma gives an upper bound of $\left\|\left(V_k^M\right)^{-1}\right\|$ on A_k^M.

Lemma 5.22. *For $h \in (0, \frac{1}{2})$, we have $\|(V_k^M)^{-1}\| \leq 2$ on A_k^M.*

Proof. Using the triangular inequality and inequality (5.25), we obtain on A_k^M

$$\|(V_k^M)^{-1}\| \leq \|Id\| + \|Id - (V_k^M)^{-1}\| \leq 1 + \frac{\|Id - V_k^M\|}{1 - \|Id - V_k^M\|} \leq 1 + \frac{h}{1-h} = \frac{1}{1-h}.$$

Since $h < \frac{1}{2}$, the proof is complete. □

The following result provides an upper bound of $\mathfrak{B}_2$. This estimate should be compared with that given in [7] page 2192.

Lemma 5.23 (Upper estimate of $\mathfrak{B}_2$)**.** *There exists a constant $C>0$ such that for large N and $k=0,\dots,N-1$, $\mathbb{E}\left[\left|\mathfrak{B}_2\right|^2 1_{A_k^M}\right] \leq \frac{C}{M}\mathbb{E}\left[\left|v_k\right|^2\left|p_{k+1}\right|^2\right]\mathbb{E}\left|\rho_{k+1}^N\right|^2$.*

Proof. We can rewrite $\mathfrak{B}_2$ as follows:

$$\mathfrak{B}_2 = -\frac{(V_k^M)^{-1}}{M}\sum_{m=1}^{M}\left(v_k^m\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)-\mathbb{E}\left[v_k\widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right]\right).$$

Using Lemmas 5.22 and 5.20 (i), we obtain for small h

$$\mathbb{E}\left[\left|\mathfrak{B}_2\right|^2 1_{A_k^M}\right] \leq \frac{4}{M^2}\mathbb{E}\left[1_{A_k^M}\left|\sum_{m=1}^{M}\left(v_k^m\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)-\mathbb{E}\left[v_k\widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right]\right)\right|^2\right]$$
$$\leq \frac{4}{M}\mathbb{E}\left|v_k\widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)-\mathbb{E}\left[v_k\widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right]\right|^2 \leq \frac{4}{M}\mathbb{E}\left|v_k\widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2.$$

Using Lemma 5.5 (2), Cauchy-Schwarz's inequality and Proposition 5.2, since $\alpha_{k+1}^{I,I}$ is deterministic we deduce

$$\mathbb{E}\left[\left|\mathfrak{B}_2\right|^2 1_{A_k^M}\right] \leq \frac{4}{M}\mathbb{E}\left[\left|v_k\right|^2\left|\alpha_{k+1}^{I,I}.p_{k+1}\right|^2\right] \leq \frac{4}{M}\mathbb{E}\left[\left|v_k\right|^2\left|p_{k+1}\right|^2\right]\mathbb{E}\left|\rho_{k+1}^N\right|^2,$$

which concludes the proof. □

The next lemma gives an upper estimate of the L^2-norm of $\mathfrak{B}_3$

Lemma 5.24 (Upper estimate of $\mathfrak{B}_3$)**.** *There exists a constant C such that for large N and $k=0,\dots,N-1$,*

$$\mathbb{E}\left[1_{A_k^M}\left|\mathfrak{B}_3\right|^2\right] \leq C\frac{h^2}{M}\mathbb{E}\left[\left|v_k\right|^2\left(1+\left|X_k^N\right|^2+\left|p_k\right|^2\mathbb{E}\left|\rho_k^N\right|^2+\frac{1}{h}\left|p_k\right|^2\mathbb{E}\left|\zeta_k^N\right|^2\right)\right]$$

Proof. We take expectation on A_k^M, use Lemmas 5.22 and 5.20 (i); this yields for small h

$$\mathbb{E}\left[1_{A_k^M}\left|\mathfrak{B}_3\right|^2\right] \leq 4\frac{h^2}{M}\mathbb{E}\left(\left|v_k\right|^2\left|f\left(X_k^N,\alpha_k^{\infty,I}.p_k,\beta_k^I.p_k\right)\right|^2\right) \tag{5.26}$$

The Lipschitz condition (2.3), Cauchy-Schwarz's inequality and Proposition 5.2 imply

$$\left|f\left(X_k^N,\alpha_k^{\infty,I}.p_k,\beta_k^I.p_k\right)\right|^2 \leq 2L_f\left(\left|X_k^N\right|^2+\left|\alpha_k^{\infty,I}.p_k\right|^2+\left|\beta_k^I.p_k\right|^2\right)+2\left|f(0,0,0)\right|^2$$
$$\leq 2L_f\left(\left|X_k^N\right|^2+\left|p_k\right|^2\mathbb{E}\left|\rho_k^N\right|^2+\frac{1}{h}\left|p_k\right|^2\mathbb{E}\left|\zeta_k^N\right|^2\right)+2\left|f(0,0,0)\right|^2,$$

which together with (5.26) concludes the proof. □

The next result gives an upper estimate of $\mathfrak{B}_4$ in L^2.

Lemma 5.25 (Upper estimate of $\mathfrak{B}_4$)**.** *Fix $\epsilon>0$; there exist constants C and $C(\epsilon)$ such that for N large and $k=0,\dots,N-2$,*

$$(1-h)\mathbb{E}\left[1_{A_k^M}\left|\mathfrak{B}_4\right|^2\right] \leq (1+C(\epsilon)h)\,\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]$$
$$+C\left(h+2\epsilon\right)h\left(\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{\infty,I}-\alpha_k^{\infty,I,M}\right|^2\right]+\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{\infty,I,M}\right|^2\right]\right).$$

Proof. By definition, we have $\mathfrak{B}_4 = \frac{1}{M}\left(V_k^M\right)^{-1}\sum_{m=1}^M v_k^m x_4^m$. Let $x_4 := (x_4^m, m = 1, \dots, M)$; then Lemma 5.9 and inequality (5.14) imply that on A_k^M, $(1-h)\left|\mathfrak{B}_4\right|^2 \leq \lambda_{\min}(V_k^M)\left|\mathfrak{B}_4\right|^2 \leq |x_4|_M^2$. Taking expectation, using Young's inequality and (5.20) in Corollary 5.17, we obtain for $\epsilon > 0$: $(1-h)\mathbb{E}\left[1_{A_k^M}\left|\mathfrak{B}_4\right|^2\right] \leq \sum_{i=1}^3 T_i$, where:

$$
\begin{aligned}
T_1 &:= \left(1+\frac{h}{\epsilon}\right)\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left[1_{A_k^M}\left|\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) - \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right|^2\right],\\
T_2 &:= \left(1+\frac{2\epsilon}{h}\right)h^2\frac{1}{M}\sum_{m=1}^M \mathbb{E}\Big[1_{A_k^M}\Big|f\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I}.p_k^m, \beta_k^I.p_k^m\right)\\
&\qquad\qquad -f\left(X_{t_k}^{N,m}, \alpha_k^{\infty,I,M}.p_k^m, \beta_k^{\infty,I,M}.p_k^m\right)\Big|^2\Big],\\
T_3 &:= \left(1+\frac{h}{\epsilon}\right)\frac{1}{M}\sum_{m=1}^M \mathbb{E}\Big[1_{A_k^M}\Big|\overleftarrow{\Delta}B_k^m\Big[g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right)\\
&\qquad\qquad -g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\Big]\Big|^2\Big].
\end{aligned}
$$

Lemma 5.6 (1) and Lemma 5.11 yield

$$
\begin{aligned}
T_1 &\leq \left(1+\frac{h}{\epsilon}\right)\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left[1_{A_k^M}\left|\alpha_{k+1}^{I,I}.p_{k+1}^m - \alpha_{k+1}^{I,I,M}.p_{k+1}^m\right|^2\right]\\
&\leq \left(1+\frac{h}{\epsilon}\right)\mathbb{E}\left[1_{A_k^M}\left|\alpha_{k+1}^{I,I} - \alpha_{k+1}^{I,I,M}\right|^2\left\|P_{k+1}^M\right\|\right].
\end{aligned}
$$

Since $A_k^M \subset A_{k+1}^M$ and $\|P_{k+1}^M\| \leq 1+h$ on A_k^M, we deduce

$$
T_1 \leq \left(1+\frac{h}{\epsilon}\right)(1+h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I} - \alpha_{k+1}^{I,I,M}\right|^2\right]. \tag{5.27}
$$

Using property (2.3), Lemma 5.11 and a similar argument, we obtain for $0 < h \leq 1$:

$$
\begin{aligned}
T_2 &\leq L_f h\,(h+2\epsilon)\,\mathbb{E}\left[1_{A_k^M}\left(\left|\alpha_k^{\infty,I} - \alpha_k^{\infty,I,M}\right|^2 + \left|\beta_k^I - \beta_k^{\infty,I,M}\right|^2\right)\left\|P_k^M\right\|\right]\\
&\leq 2L_f h\,(h+2\epsilon)\,\mathbb{E}\left[1_{A_k^M}\left(\left|\alpha_k^{\infty,I} - \alpha_k^{\infty,I,M}\right|^2 + \left|\beta_k^I - \beta_k^{\infty,I,M}\right|^2\right)\right].
\end{aligned} \tag{5.28}
$$

Finally, since $A_k^M = A_{k+1}^M \cap \mathfrak{A}_k$ and $\overleftarrow{\Delta}B_k^m$ is independent of $\mathcal{F}_{t_k}^W \vee \mathcal{F}_{t_{k+1},T}^B$, we have using the Lipschitz property (2.4):

$$
\begin{aligned}
T_3 \leq \left(1+\frac{h}{\epsilon}\right)\frac{1}{M}\sum_{m=1}^M \mathbb{E}\Big[1_{A_{k+1}^M}\Big|g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right)\right) - g\left(X_{t_{k+1}}^{N,m}, \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right)\Big|^2\\
\mathbb{E}\left(1_{\mathfrak{A}_k}\left|\overleftarrow{\Delta}B_k^m\right|^2\left|\mathcal{F}_{t_N}^W \vee \mathcal{F}_{t_{k+1},T}^B\right.\right)\Big]\\
\leq L_g h\left(1+\frac{h}{\epsilon}\right)\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left[1_{A_{k+1}^M}\left|\widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I}.p_{k+1}^m\right) - \widehat{\rho}_{k+1}^{N,m}\left(\alpha_{k+1}^{I,I,M}.p_{k+1}^m\right)\right|^2\right].
\end{aligned}
$$

So using Lemma again 5.6 (1) and Lemma 5.11, we deduce

$$
T_3 \leq L_g h\left(1+\frac{h}{\epsilon}\right)(1+h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I} - \alpha_{k+1}^{I,I,M}\right|^2\right]. \tag{5.29}
$$

The inequalities (5.27)-(5.29) conclude the proof. □

We decompose v_k as $v_k = (v_k^o, v_k^p)$ where v_k^o contains all the elements in the basis which are independent to $\overleftarrow{\Delta} B_k$ and $v_k^p = \frac{\overleftarrow{\Delta} B_k}{\sqrt{h}} w_k^p$. with w_k^p independent to $\overleftarrow{\Delta} B_k$. The next lemma gives an upper estimate of the L^2 norm of $\mathfrak{B}_5$ on A_k^M.

Lemma 5.26 (Upper estimate of $\mathfrak{B}_5$)**.** *There exists constant C such that for small h and $k = 0, \dots, N-1$,*

$$\mathbb{E}\left[1_{A_k^M} |\mathfrak{B}_5|^2\right] \leq \frac{Ch}{M} \mathbb{E}\left[\left(|v_k|^2 + \left|w_k^p\right|^2\right)\left(1 + \left|X_{k+1}^N\right|^2 + |p_{k+1}|^2 \mathbb{E}\left|\rho_{k+1}^N\right|^2\right)\right].$$

Proof. The proof is similar to that of Lemma 5.24 which deals with $\mathfrak{B}_3$. Lemmas 5.22, 5.20 and 5.5 (1) yield for small h

$$\begin{aligned}\mathbb{E}\left[1_{A_k^M} |\mathfrak{B}_5|^2\right] \leq & \frac{4}{M} \mathbb{E}\left|v_k \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right) - \mathbb{E}\left[v_k \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \widehat{\rho}_{k+1}^N\left(\alpha_{k+1}^{I,I}.p_{k+1}\right)\right)\right]\right|^2 \\ \leq & \frac{4}{M} \mathbb{E}\left|v_k \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2.\end{aligned}$$

Then the decompostion of v_k yields

$$\mathbb{E}\left[1_{A_k^M} |\mathfrak{B}_5|^2\right] \leq \frac{4}{M} \mathbb{E}\left|v_k^o \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 + \frac{4}{M} \mathbb{E}\left|v_k^p \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2.$$

Since $\overleftarrow{\Delta} B_k$ is independent of $\mathcal{F}_T^W \vee \mathcal{F}_{t_{k+1},T}^B$, we deduce

$$\begin{aligned}\mathbb{E}\left[1_{A_k^M} |\mathfrak{B}_5|^2\right] \leq & \frac{4}{M} \mathbb{E}\left|v_k^o \overleftarrow{\Delta} B_k g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 + \frac{4}{M} \mathbb{E}\left|w_k^p \frac{\left|\overleftarrow{\Delta} B_k\right|^2}{\sqrt{h}} g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 \\ \leq & \frac{Ch}{M} \mathbb{E}\left|v_k^o g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 + \frac{Ch}{M} \mathbb{E}\left|w_k^p g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 \\ \leq & \frac{Ch}{M} \mathbb{E}\left[\left(|v_k|^2 + \left|w_k^p\right|^2\right)\left|g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2\right].\end{aligned}$$

The Lipschitz condition (2.4), Cauchy-Schwarz's and Young's inequalities together with Proposition 5.2 yield

$$\begin{aligned}\left|g\left(X_{t_{k+1}}^N, \alpha_{k+1}^{I,I}.p_{k+1}\right)\right|^2 \leq & 2L_g\left(\left|X_{t_{k+1}}^N\right|^2 + \left|\alpha_{k+1}^{I,I}\right|^2 |p_{k+1}|^2\right) + 2|g(0,0)|^2 \\ \leq & 2L_g\left(\left|X_{t_{k+1}}^N\right|^2 + |p_{k+1}|^2 \mathbb{E}\left|\rho_{k+1}^N\right|^2\right) + 2|g(0,0)|^2.\end{aligned}$$

This concludes the proof. □

Final step of the proof of Theorem 5.8. Young's inequality implies that for $h \in (0,1]$, $(b_1+b_2+b_3+b_4+b_5)^2 \leq \frac{8}{h}(b_1^2+b_2^2+b_3^2+b_5^2) + (1+h)b_4^2$. Recall that ϵ_k has been defined in (5.13). Then the decomposition (5.24) and Lemmas 5.21 and 5.23-5.26 yield for $\epsilon > 0$, small h and ϵ_k defined by (5.13):

$$\begin{aligned}&\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}\right|^2\right] \leq \frac{8}{h}\mathbb{E}\left[1_{A_k^M} \sum_{j\in\{1,2,3,5\}} |\mathfrak{B}_j|^2\right] + (1+h)\mathbb{E}\left[1_{A_k^M} |\mathfrak{B}_4|^2\right] \\ &\leq \frac{C}{Mh}\epsilon_k + (1+C(\epsilon)h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I} - \alpha_{k+1}^{I,I,M}\right|^2\right] \\ &\quad + (1+Ch)C(h+2\epsilon)h\left(\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{\infty,I} - \alpha_k^{\infty,I,M}\right|^2\right] + \mathbb{E}\left[1_{A_k^M}\left|\beta_k^I - \beta_k^{\infty,I,M}\right|^2\right]\right),\end{aligned}$$

where in the last inequality, we have used Lemma 3.7. The definition of $\theta_k^{\infty,I,M}$ and $\theta_k^{\infty,I}$, yield for h small enough:

$$[1-(1+Ch)C(h+2\epsilon)h]\,\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{\infty,I}-\alpha_k^{\infty,I,M}\right|^2\right]+h\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{\infty,I,M}\right|^2\right]$$
$$\leq\frac{C}{Mh}\epsilon_k+(1+C(\epsilon)h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]+(1+Ch)C\,(h+2\epsilon)\,h\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{\infty,I,M}\right|^2\right].$$

Using again Lemma 3.7, we obtain for some constant C and h small enough

$$\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}-\theta_k^{\infty,I}\right|^2\right]\leq\frac{C}{Mh}\epsilon_k+(1+C(\epsilon)h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]$$
$$+(1+Ch)C\,(h+2\epsilon)\,h\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{\infty,I,M}\right|^2\right]. \qquad (5.30)$$

Using Corollary 5.15 (ii) and Lemma 5.18 we deduce

$$\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{\infty,I,M}\right|^2\right]\leq 2\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{I,I,M}\right|^2\right]+\frac{2}{h}\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{\infty,I,M}\right|^2\right]$$
$$\leq 2\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{I,I,M}\right|^2\right]+Ch^{I-1}\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{\infty,I,M}\right|^2\right]$$
$$\leq 2\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{I,I,M}\right|^2\right]+Ch^{I-1}\mathbb{E}\left|\rho_{k+1}^N\right|^2+Ch^I. \qquad (5.31)$$

Plugging (5.30) and (5.31) in Lemma 5.19, we obtain for some constant C and h small enough

$$\mathbb{E}\left[1_{A_k^M}\left|\theta_k^{I,I,M}-\theta_k^{I,I}\right|^2\right]\leq\frac{C}{Mh}\epsilon_k+(1+C(\epsilon)h)\mathbb{E}\left[1_{A_{k+1}^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]$$
$$+Ch^{I-1}\left(h^2+h\mathbb{E}\left|\rho_{k+1}^N\right|^2+\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right)$$
$$+(1+Ch)C\,(h+2\epsilon)\,h\mathbb{E}\left[1_{A_k^M}\left|\beta_k^I-\beta_k^{I,I,M}\right|^2\right]$$

But $(1+Ch)C\,(h+2\epsilon)=2\epsilon C+h(C+C^2h+2\epsilon C^2)$ and we may choose ϵ such that $2\epsilon C=\frac{1}{2}$, so that $1-(1+Ch)C\,(h+2\epsilon)=\frac{1}{2}-(C+C^2h+\frac{C}{2})h$. Using again Lemma 3.7 we obtain for some constant C and h small enough:

$$\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{I,I}-\alpha_k^{I,I,M}\right|^2\right]+h\frac{1}{2}\,(1-Ch)\,\mathbb{E}\left[1_{A_k^M}\left|\beta_k^{I,I,M}-\beta_k^I\right|^2\right]$$
$$\leq(1+Ch)\mathbb{E}\left[1_{A_k^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]+C\frac{\epsilon_k}{hM}+Ch^{I-1}\left(h^2+h\mathbb{E}\left|\rho_{k+1}^N\right|^2+\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right)$$

So for small h,

$$(1-Ch)\left\{\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{I,I}-\alpha_k^{I,I,M}\right|^2\right]+h\frac{1}{2}\mathbb{E}\left[1_{A_k^M}\left|\beta_k^{I,I,M}-\beta_k^I\right|^2\right]\right\}$$
$$\leq(1+Ch)\mathbb{E}\left[1_{A_k^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]+C\frac{\epsilon_k}{hM}+Ch^{I-1}\left(h^2+h\mathbb{E}\left|\rho_{k+1}^N\right|^2+\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right)$$

Using the Lemma 3.7, we obtain

$$\mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{I,I}-\alpha_k^{I,I,M}\right|^2\right]+h\frac{1}{2}\mathbb{E}\left[1_{A_k^M}\left|\beta_k^{I,I,M}-\beta_k^I\right|^2\right]$$
$$\leq(1+Ch)\mathbb{E}\left[1_{A_k^M}\left|\alpha_{k+1}^{I,I}-\alpha_{k+1}^{I,I,M}\right|^2\right]+C\frac{\epsilon_k}{hM}+Ch^{I-1}\left(h^2+h\mathbb{E}\left|\rho_{k+1}^N\right|^2+\mathbb{E}\left|\rho_k^N\right|^2+\mathbb{E}\left|\zeta_k^N\right|^2\right)$$

The Gronwall Lemma 3.8 applied with $a_k = \mathbb{E}\left[1_{A_k^M}\left|\alpha_k^{I,I} - \alpha_k^{I,I,M}\right|^2\right]$ and $c_k = h\frac{1}{2}\mathbb{E}\left[1_{A_k^M}\left|\beta_k^{I,I,M} - \beta_k^{I}\right|^2\right]$ and the fact that $\alpha_N^{I,I,M} = \alpha_N^{I,I}$ concludes the proof.

Acknowledgments: The author wishes to thank Annie Millet for many helpful comments.

References

[1] Aboura O., On the discretization of backward doubly stochastic differential equations *Arxiv 0907.1406*
[2] Aman A., Numerical scheme for backward doubly stochastic differential equations *Arxiv 0907.2035*
[3] Aman A., Numerical scheme for backward doubly stochastic differential equations *Arxiv 1011.6170*
[4] Bally V., Approximation scheme for solutions of BSDE *Backward stochastic differential equations (Paris, 1995–1996), 177–191, Pitman Res. Notes Math. Ser.*, 364, Longman, Harlow, 1997.
[5] Bouchard B., Touzi N., Discrete time approximation and Monte Carlo simulation of backward stochastic differential equations *Stochastic process and applications* **111** (2004) 175-206
[6] Chevance D., Numerical methods for backward stochastic differential equations *Numerical methods in finance, 232–244, Publ. Newton Inst., Cambridge Univ. Press, Cambridge,* 1997.
[7] Gobet E., Lemor J.P., Warin X., A regression-based Monte Carlo method to solve Backward stochastic differential equations *The Annals of Applied Probability* 2005, Vol. 15, No. 3
[8] Kloeden P. E., Platen E., *Numerical Solution of Stochastic Differential Equations* Springer
[9] Ma J., Yong J., *Forward-Backward Stochastic Differential Equations and their Applications* Lecture Note in Math. 1702 Springer 1999
[10] Nualart D., Pardoux E., Stochastic calculus with anticipatng integrands *Probability theory related fields* 78, 535-581 (1988)
[11] Pardoux E., Peng S., Backward stochastic differential equation and quasilinear parabolic partial differential equations. In: *B. L.Rozovski, R. B. Sowers (eds). Stochastic partial equations and their applications. Lect. Notes control Inf. Sci.* **176**, 200 − 217, Springer, Berlin, (1992).
[12] Pardoux E., Peng S., Backward doubly stochastic differential equations and systems of quasilinear SPDEs *Probability Theory and Related Fields* (1994) 209-227
[13] Zhang J., A numerical scheme for BSDEs *The Annals of Applied Probability* Vol. 14 No. 1 (2004) 459-488

SAMM (EA 4543), Université Paris 1 Panthéon Sorbonne, 90 Rue de Tolbiac, 75634 Paris Cedex France
E-mail address: omar.aboura@malix.univ-paris1.fr