The Annals of Applied Probability
2002, Vol. 12, No. 1, 302–316

NUMERICAL METHOD FOR BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

BY JIN MA,[1] PHILIP PROTTER,[2] JAIME SAN MARTÍN AND SOLEDAD TORRES[3]

Purdue University, Cornell University, Universidad de Chile and Universidad de Valparaíso

We propose a method for numerical approximation of backward stochastic differential equations. Our method allows the final condition of the equation to be quite general and simple to implement. It relies on an approximation of Brownian motion by simple random walk.

1. Introduction. In this paper we propose a new method of approximating solutions of backward stochastic differential equations (BSDEs). Our method allows the final condition of the equation to be quite general and it is simple to implement. It relies on an approximation of Brownian motion by simple random walk.

This type of equation appears in numerous problems in finance, in contingent claim valuation when there are constraints on the hedging portfolios (see [8]).

Some numerical methods for approximating solutions of BSDEs have already been developed. A four step algorithm developed by Ma, Protter and Yong to solve a class of more general equations called forward–backward SDEs has been proposed in [13]. A numerical scheme was developed based on this method in [7]. Bally [2] presents a random time scheme to approximate BSDEs. The convergence result only needs regularity assumptions. However, his scheme requires a further approximation to give an implementation. On the other hand Chevance [5] gives a numerical method for solving BSDEs associated with a forward stochastic differential equation (FSDE). His method requires strong regularity assumptions for its implementation. Finally, a new result of Bally and Pagès allows for the numerical treatment of BSDEs and reflected BSDEs [3].

We should note that there is another type of approximating solution to BSDEs, via the discretization of filtration (see, e.g., [1, 6]). Although our method uses ideas similar to those used in these results, in that we also approximate the Brownian motion by discrete processes, the main feature of our result is that we *do not* assume that the discretized filtrations "converge" to the original Brownian filtration

Received August 2000; revised May 2001.
[1]Supported in part by NSF Grant DMS-99-71720.
[2]Supported in part by NSF Grant DMS-99-71720 and NSA Grant 904-00-1-0035.
[3]Supported in part by FONDECYT Grant 1000270.
AMS 2000 subject classifications. 60H20, 65C99, 65L99, 60H99, 60G35.
Key words and phrases. Backward stochastic differential equations, numerical methods.

in order to guarantee the convergence of the solutions. Such a relaxation obviously reduces the complexity in constructing the approximating solutions.

Let $\Omega = C([0,1], \mathbb{R}^d)$ and consider the canonical Wiener space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$, in which $B_t(\omega) = \omega(t)$ is a standard d-dimensional Brownian motion. We consider the following BSDE:

$$Y_t = \xi + \int_t^1 f(s, Y_s)\, ds - \int_t^1 Z_s\, dB_s, \tag{1}$$

where ξ is an $\mathcal{F}_1$-measurable square integrable random variable and f is Lipschitz continuous in the space variable with Lipschitz constant L. The solution of (1) is a pair of adapted processes (Y, Z) which satisfies the equation. Existence and uniqueness for the solutions of such equations are proved in [14, 15], and further results for the extension of uniqueness under more general assumptions for the generator have been developed, for example, in [12] in the one-dimensional case. We note that for technical reasons in this paper we only treat the BSDEs with generator f independent of the component Z. Nevertheless, our result for f independent of Z, combined with a simple Girsanov argument, yields a result that can include a Z term in the drift in a linear way. We explain this in the remark at the end of the article. We hope to be able to address the general case in our future publications.

2. A numerical scheme for BSDEs. The numerical scheme we propose in this paper is based upon a discretization of (1) and replacing B by a simple random walk. To be more precise, let $t_i^{(n)} = i/n,\ i = 0, \dots, n$. For the sake of simplicity we shall write t_i instead of $t_i^{(n)}$ when dependence on n is possible; also, we shall consider only the one-dimensional case, although the generalization to the d-dimensional case is rather clear. We let $M^{(n)}$ denote the approximating binomial random walk, whose increments are $1/\sqrt{n}$ and $-1/\sqrt{n}$ with probability $1/2$. Further, we assume that the sequence $\{M^{(n)}\}$ is i.i.d. We note that, while most of the results presented in this paper also hold for other approximations of B, we work with $M^{(n)}$ mainly for simplicity.

In what follows $(\mathcal{F}^n)$ denotes the natural filtration of $M^{(n)}$. In some of our computations we shall use the linear interpolation associated with the discrete process $M^{(n)}$, which will then become a continuous process and will still be denoted as $M^{(n)}$ itself. Consequently, if F is a functional defined on Ω, then by a slight abuse of notation we shall identify $F(M_0^{(n)}, \dots, M_1^{(n)})$ and $F(M^{(n)})$ as the same, although the latter means F is evaluated over the linear interpolation of $M^{(n)}$, while in the former F is considered as a function on $\mathbb{R}^{d(m+1)}$.

Let us now consider the discrete version of the BSDE (1):

$$\tilde{Y}_{t_i}^{(n)} = F(M^{(n)}) + \frac{1}{n}\sum_{j=i}^{n} f\big(t_j, \tilde{Y}_{t_j}^{(n)}\big) - \sum_{j=i}^{n-1} \tilde{Z}_{t_j}^{(n)} \Delta M_{t_{j+1}}^{(n)}. \tag{2}$$

This equation has a unique solution $(\tilde{Y}^{(n)}, \tilde{Z}^{(n)})$ since the martingale $M^{(n)}$ has the predictable representation property (see [4]). It can be checked that solving this equation is equivalent to finding a solution to the following implicit iteration problem:

$$\tilde{Y}_{t_i}^{(n)} = \mathbb{E}\left\{\tilde{Y}_{t_{i+1}}^{(n)} + \frac{1}{n} f\big(t_i, \tilde{Y}_{t_i}^{(n)}\big) \mid \mathcal{F}_{t_i}^{(n)}\right\},$$

which, due to the adaptedness condition, is equivalent to

$$\tilde{Y}_{t_i}^{(n)} - \frac{1}{n} f\big(t_i, \tilde{Y}_{t_i}^{(n)}\big) = \mathbb{E}\{\tilde{Y}_{t_{i+1}}^{(n)} \mid \mathcal{F}_{t_i}^{(n)}\}. \tag{3}$$

We point out, as we shall prove in Lemma 3.1, that one can in fact assume without loss of generality that the *generator* f is bounded, and henceforth we denote its bound by R. Furthermore, once $\tilde{Y}_{t_{i+1}}^{(n)}$ is determined, $\tilde{Y}_{t_i}^{(n)}$ is solved via (3) by a fixed point technique:

$$X^0 = \mathbb{E}\{\tilde{Y}_{t_{i+1}}^{(n)} \mid \mathcal{F}_{t_i}^{(n)}\},$$

$$X^{k+1} = X_0 + \frac{1}{n} f(t_i, X^k).$$

It is standard to show that, if f is uniformly Lipschitz in the spatial variable x with Lipschitz constant L, then the iterations of this procedure will converge to the true solution of (3) at a geometric rate L/n. Therefore, in the case when n is large enough, one iteration would already give us the error estimate: $|\tilde{Y}_{t_i}^{(n)} - X^1| \le LR/n^2$, producing a good approximate solution of (3). Consequently, we propose the following explicit numerical scheme for the BSDE (1):

$$\hat{Y}_1^{(n)} = F(M^{(n)}), \qquad \hat{Z}_1^{(n)} = 0,$$

$$\hat{X}_{t_i} = \mathbb{E}\{\hat{Y}_{t_{i+1}}^{(n)} \mid \mathcal{F}_{t_i}^{(n)}\},$$

$$\hat{Y}_{t_i}^{(n)} = \hat{X}_{t_i} + \frac{1}{n} f(t_i, \hat{X}_{t_i}),$$

$$\hat{Z}_{t_i}^{(n)} = \mathbb{E}\left\{\left[\hat{Y}_{t_{i+1}}^{(n)} + \frac{1}{n} f\big(t_i, \hat{Y}_{t_i}^{(n)}\big) - \hat{Y}_{t_i}^{(n)}\right]\big(\Delta M_{t_{i+1}}^{(n)}\big)^{-1} \mid \mathcal{F}_{t_i}^{(n)}\right\}.$$

Let us now analyze the error of this scheme. Clearly, the error produced by this method is bounded by

$$\begin{aligned}|\tilde{Y}_{t_i}^{(n)} - \hat{Y}_{t_i}^{(n)}| &\le |\mathbb{E}\{\tilde{Y}_{t_{i+1}}^{(n)} - \hat{Y}_{t_{i+1}}^{(n)} \mid \mathcal{F}_{t_i}^{(n)}\}| \\ &\quad + \frac{1}{n}\big|f\big(t_i, \tilde{Y}_{t_i}^{(n)}\big) - f\big(t_i, \hat{Y}_{t_i}^{(n)}\big)\big| + \frac{1}{n}\big|f\big(t_i, \hat{Y}_{t_i}^{(n)}\big) - f(t_i, \hat{X}_{t_i})\big|,\end{aligned}$$

which yields

$$\sup_{\omega} |\tilde{Y}_{t_i}^{(n)} - \hat{Y}_{t_i}^{(n)}| \le \gamma \left(\sup_{\omega} |\tilde{Y}_{t_{i+1}}^{(n)} - \hat{Y}_{t_{i+1}}^{(n)}| + \frac{LR}{n^2} \right) \quad \text{a.s.},$$

where $\gamma = (1 - L/n)^{-1}$. Iterating this inequality one obtains

$$\sup_{\omega} |\tilde{Y}_{t_i}^{(n)} - \hat{Y}_{t_i}^{(n)}| \le \gamma^{n-i} \sup_{\omega} |\tilde{Y}_1^{(n)} - \hat{Y}_1^{(n)}| + \frac{LR\gamma}{n^2} \frac{\gamma^{n-i} - 1}{\gamma - 1}.$$

For large n ($2L \le n$ works) one obtains that $(\frac{1}{1-L/n})^{n-i} \le (1 + \frac{2L}{n})^{n-i} \le e^{2L}$. Therefore we get

$$\sup_{\omega} |\tilde{Y}_{t_i}^{(n)} - \hat{Y}_{t_i}^{(n)}| \le \frac{R(e^{2L} - 1)}{n}.$$

Using this bound we obtain the corresponding bound for $\tilde{Z}^{(n)} - \hat{Z}^{(n)}$. In fact, we have

$$\begin{aligned} \tilde{Z}_{t_i}^{(n)} - \hat{Z}_{t_i}^{(n)} = {} & \left[\tilde{Y}_{t_{i+1}}^{(n)} + \frac{1}{n} f\big(t_i, \tilde{Y}_{t_i}^{(n)}\big) - \tilde{Y}_{t_i}^{(n)} \right] \big(\Delta M_{t_{i+1}}^{(n)}\big)^{-1} \\ & - \mathbb{E}\left\{ \left[\hat{Y}_{t_{i+1}}^{(n)} + \frac{1}{n} f\big(t_i, \hat{Y}_{t_i}^{(n)}\big) - \hat{Y}_{t_i}^{(n)} \right] \big(\Delta M_{t_{i+1}}^{(n)}\big)^{-1} \mid \mathcal{F}_{t_i}^{(n)} \right\}, \end{aligned}$$

which by the adaptedness of $\tilde{Z}^{(n)}$ yields

$$\begin{aligned} \tilde{Z}_{t_i}^{(n)} - \hat{Z}_{t_i}^{(n)} = \mathbb{E}\Bigg\{ & \left[\tilde{Y}_{t_{i+1}}^{(n)} - \hat{Y}_{t_{i+1}}^{(n)} + \frac{1}{n}\big(f\big(t_i, \tilde{Y}_{t_i}^{(n)}\big) - f\big(t_i, \hat{Y}_{t_i}^{(n)}\big)\big) + \tilde{Y}_{t_i}^{(n)} - \hat{Y}_{t_i}^{(n)} \right] \\ & \times \big(\Delta M_{t_{i+1}}^{(n)}\big)^{-1} \mid \mathcal{F}_{t_i}^{(n)} \Bigg\}. \end{aligned}$$

Finally, one obtains

$$\sup_{\omega} |\tilde{Z}_{t_i}^{(n)} - \hat{Z}_{t_i}^{(n)}| \le \frac{R(e^{2L} - 1)(2 + L/n)}{\sqrt{n}} \quad \text{a.s.}$$

This means that for the convergence of the numerical method we just need to concentrate on the solution of (2).

We remark that the conditional expectations with respect to the discrete filtration $(\mathcal{F}^{(n)})$ can be computed explicitly as follows. We assume Γ is an $\mathcal{F}_{t_{k+1}}^{(n)}$-measurable random variable, and we take the 2^k atoms corresponding to the trajectories of the martingale $M^{(n)}$ in $\mathcal{F}_{t_k}^{(n)}$. Each atom in $\mathcal{F}_{t_k}^{n}$ splits into two atoms of $\mathcal{F}_{t_{k+1}}^{n}$. Then we have

$$\mathbb{E}\big(\Gamma \,|\, \mathcal{F}_{t_k}^{(n)}\big)(\omega) = \tfrac{1}{2}(a + b),$$

where a, b are the values of Γ in the two atoms of $\mathcal{F}_{t_{k+1}}^{(n)}$ coming from the corresponding atom in $\mathcal{F}_{t_k}^{(n)}$ containing ω.

Our main result is the following:

THEOREM 1. *Assume that in the BSDE* (1) *the following conditions hold*:

(i) $\xi = F(B)$, *where* $F\colon \Omega \to \mathbb{R}^d$ *is a bounded Lipschitz function with respect to the uniform topology on* Ω; *that is*, *there exists a constant* κ *such that for all* $\omega,\ \omega' \in \Omega$ *it holds that*

$$|F(\omega) - F(\omega')| \le \kappa \sup_{0\le t\le 1} |\omega(t) - \omega'(t)|;$$

(ii) $f\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}^d$ *is a continuous function and it is uniformly Lipschitz*; *that is*, *there exists a constant* L *such that*, *for all* $x, y \in \mathbb{R}^d$,

$$\sup_{0\le t\le 1} |f(t,x) - f(t,y)| \le L|x-y|.$$

Let $(\tilde{Y}^{(n)}, \tilde{Z}^{(n)})$ *be the solution of* (2), *and let*

$$U_{t_i}^{(n)} = \sum_{j=0}^{i-1} \tilde{Z}_{t_j}^{(n)} \Delta M_{t_{j+1}}^{(n)} = \tilde{Y}^{(n)} - F\big(M^{(n)}\big) - \frac{1}{n}\sum_{j=1}^{i} f\big(t_j, \tilde{Y}_{t_j}^{(n)}\big).$$

Then the sequence $(\tilde{Y}^{(n)}, U^{(n)})$ *converges weakly in the Skorohod topology to* $(Y, \int Z\,dB)$, *where* (Y, Z) *is the unique solution of* (1).

3. Proof of the main result. We first show that under the assumptions of Theorem 1 we can in fact assume without loss of generality that the function f is bounded. To this end, notice that f is continuous, it is bounded on any compact set $[0,1]\times[-A, A]$, $A > 0$. Thus it suffices to show that there exists a constant $A > 0$ such that the solution of (1) satisfies $|Y_t| \le A$, $\forall t$, a.s. Denoting $R_0 = \sup_{s\in[0,1]} |f(s,0)|$, we have the following lemma.

LEMMA 3.1. *Under the assumptions on* F *and* f *in Theorem* 1, *the solution* Y *of* (1) *is bounded by* $(D + R_0)e^L$, *where* D *is a bound for* F.

PROOF. In the one-dimensional case this can be proved by the comparison theorem. However, we shall provide an argument that does not involve any comparison results; therefore it works for the general case.

For any $A > 0$ we consider the following BSDE:

$$X_t^A = F(B) + \int_t^1 f\big(s, \phi_A(X_s^A)\big)\,ds - \int_t^1 \Lambda_s\,dB_s, \tag{4}$$

where ϕ_A is any bounded Lipschitz function with constant 1, such that $\phi_A(x) = x$ for $|x| \le A$, and we shall determine the constant A later. Since f is continuous, it

is bounded on $[0,1]\times[-2A,2A]$ and we let C_A denote this bound. Note that X^A also solves the problem

$$X_t^A = \mathbb{E}\Big\{F(B) + \int_t^1 f\big(s,\phi_A(X_s^A)\big)\,ds \mid \mathcal{F}_t\Big\},$$

we see that X^A is bounded by $D + C_A$ and the following estimate holds:

$$|X_t^A| \le D + \mathbb{E}\Big\{\int_t^1 \big|f\big(s,\phi_A(X_s^A)\big)\big|\,ds \mid \mathcal{F}_t\Big\} \le D + R_0 + L\mathbb{E}\Big\{\int_t^1 |X_s^A|\,ds \mid \mathcal{F}_t\Big\}.$$

Let $h^A(t) := |X_t|_\infty$ denote the $L^\infty(\Omega)$-norm of the random variable X_t. Then h^A is a deterministic, measurable function and satisfies the inequality

$$h^A(t) \le D + R_0 + L\int_t^1 h^A(s)\,ds. \tag{5}$$

Since $h^A(t) \le D + C_A$ for all t, iterating (5) we derive that, for any $N > 0$,

$$h^A(t) \le (D+R_0)\sum_{k=0}^{N}\frac{L^k(1-t)^k}{k!} + (D+C_A)\frac{L^{N+1}(1-t)^{N+1}}{(N+1)!}.$$

Letting $N\to\infty$ we obtain a *generalized* Gronwall inequality:

$$|X_t^A| \le h^A(t) \le (D+R_0)e^{L(1-t)} \le (D+R_0)e^L.$$

Note that the right-hand side of this inequality does not depend on A; therefore if we choose $A = (D+R_0)e^L$ in (4), then X^A actually solves the equation

$$X_t^A = F(B) + \int_t^1 f\big(s, X_s^A\big)\,ds - \int_t^1 \Lambda_s\,dB_s.$$

The uniqueness of the solution of the BSDE (1) then implies that $(X^A,\Lambda)\equiv(Y,Z)$, proving the lemma. $\square$

The next two lemmas give some fine properties of the discretized solution $\{\tilde Y^{(n)}\}$.

LEMMA 3.2. *Let $\tilde Y^{(n)}$ be the solution of* (2). *Then the jumps of $\tilde Y^{(n)}$ converge uniformly to zero. Moreover*

$$\sup_\omega |\tilde Y_{t_{i+1}}^{(n)} - \tilde Y_{t_i}^{(n)}| \le \frac{\kappa e^{2L}}{\sqrt{n}} + \frac{R}{n}.$$

PROOF. We use induction. Let us start with $\tilde Y_1^{(n)} - \tilde Y_{t_{n-1}}^{(n)}$, which is given by

$$\tilde Y_1^{(n)} - \tilde Y_{t_{n-1}}^{(n)} = F\big(M^{(n)}\big) - \mathbb{E}\big(F\big(M^{(n)}\big) \mid \mathcal{F}_{t_{n-1}}^n\big) - \frac{1}{n} f\big(t_{n-1}, \tilde Y_{t_{n-1}}^{(n)}\big).$$

Computing the conditional expectation and recalling our convention on the process $M^{(n)}$ and its linear interpolation (see Section 2) we see that

$$\tilde{Y}_1^{(n)} - \tilde{Y}_{t_{n-1}}^{(n)} = F\big(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)}, M_1^{(n)}\big) - \frac{1}{2}F\bigg(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)}, M_{t_{n-1}}^{(n)} + \frac{1}{\sqrt{n}}\bigg)$$
$$- \frac{1}{2}F\bigg(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)}, M_{t_{n-1}}^{(n)} - \frac{1}{\sqrt{n}}\bigg) - \frac{1}{n}f\big(t_{n-1}, \tilde{Y}_{t_{n-1}}^{(n)}\big),$$

from which we deduce, using that F is Lipschitz, the upper bound

$$|\tilde{Y}_1^{(n)} - \tilde{Y}_{t_{n-1}}^{(n)}| \le \frac{\kappa}{\sqrt{n}} + \frac{R}{n}.$$

On the other hand, since

$$\tilde{Y}_{t_{n-1}}^{(n)} - \frac{1}{n}f\big(t_{n-1}, \tilde{Y}_{t_{n-1}}^{(n)}\big) = \mathbb{E}\big\{F\big(M^{(n)}\big) \mid \mathcal{F}_{t_{n-1}}^n\big\},$$

we can let $G_{n-1}(\cdot)$ denote the inverse function of the mapping $y \mapsto y - \frac{1}{n}f(t_{n-1}, y)$, which exists if $L < n$. (Recall here that L is the Lipschitz constant for f in the spatial variable y.) Observe that G_{n-1} is a Lipschitz function with Lipschitz constant $\gamma = (1 - L/n)^{-1}$. Furthermore, it holds that

$$\tilde{Y}_{t_{n-1}}^{(n)} = G_{n-1}\bigg(\frac{1}{2}F\bigg(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)}, M_{t_{n-1}}^{(n)} + \frac{1}{\sqrt{n}}\bigg)$$
$$+ \frac{1}{2}F\bigg(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)}, M_{t_{n-1}}^{(n)} - \frac{1}{\sqrt{n}}\bigg)\bigg).$$

Clearly, the right-hand side above is a Lipschitz function of $(M_0^{(n)}, \dots, M_{t_{n-1}}^{(n)})$ with Lipschitz constant $\kappa\gamma$. If we take $\tilde{Y}_{t_{n-1}}^{(n)}$ as the terminal value for a discrete BSDE with generator f on $\{0, \dots, t_{n-1}\}$ we can apply the same argument as above to deduce that

$$|\tilde{Y}_{t_{n-1}}^{(n)} - \tilde{Y}_{t_{n-2}}^{(n)}| \le \frac{\kappa\gamma}{\sqrt{n}} + \frac{R}{n},$$

and by an inductive argument

$$|\tilde{Y}_{t_i}^{(n)} - \tilde{Y}_{t_{i-1}}^{(n)}| \le \frac{\kappa\gamma^{n-i}}{\sqrt{n}} + \frac{R}{n} \le \frac{\kappa e^{2L}}{\sqrt{n}} + \frac{R}{n},$$

completing the proof. □

LEMMA 3.3. *The sequence $\tilde{Y}^{(n)}$ is tight in the Skorohod topology.*

PROOF. We use the criterium given in [11], Theorem 2.3, for locally square integrable semimartingales. First notice that $\tilde{Y}^{(n)}$ has the following decomposition:

$\tilde{Y}^{(n)} = U^{(n)} + A^{(n)}$, where $A^{(n)}_{t_i} = \tilde{Y}^{(n)}_0 + \frac{1}{n}\sum_{j=0}^{i-1} f(t_j, \tilde{Y}^{(n)}_{t_j})$ is a predictable process with finite variation, and $U^{(n)}_{t_i} = \sum_{j=0}^{i-1} \tilde{Z}^{(n)}_{t_j} \Delta M^{(n)}_{t_{j+1}}$ is a locally square integrable martingale. Define $G^{(n)} = [U^{(n)}, U^{(n)}] + V(A^{(n)})$, where $V(A^{(n)})$ is the total variation of $A^{(n)}$. We have that $G^{(n)}$ is bounded by an increasing function g that only depends on t. In fact $V(A^{(n)}) \le C$ where C is a constant not depending on n. On the other hand $[U^{(n)}, U^{(n)}] = \sum(\Delta U^{(n)})^2 = \sum(\tilde{Z}^{(n)}_{t_i})^2(M^{(n)}_{t_{i+1}} - M^{(n)}_{t_i})^2$. Since $\tilde{Y}^{(n)}$ satisfies (2) we have

$$\tilde{Y}^{(n)}_{t_{i+1}} - \tilde{Y}^{(n)}_{t_i} = -\frac{1}{n} f\big(t_i, \tilde{Y}^{(n)}_{t_i}\big) + \tilde{Z}^{(n)}_{t_i} \Delta M^{(n)}_{t_{i+1}}.$$

Using Lemma 3.2 and the fact $|\Delta M^{(n)}| = \frac{1}{\sqrt{n}}$, we obtain

$$|\tilde{Z}^{(n)}_{t_i}| \le \kappa e^{2L} + \frac{2R}{\sqrt{n}}$$

and

$$[U^{(n)}, U^{(n)}] \le \left(\kappa e^{2L} + \frac{2R}{\sqrt{n}}\right)^2.$$

By choosing g equal to a constant, we have that $G^{(n)}$ satisfies conditions C1 and C2 of Theorem 2.3 given in [11] and the conclusions of this theorem imply that $(\tilde{Y}^{(n)})$ is relatively compact under the Skorohod topology. □

We point out that the sequence of predictable finite variation processes $\{\tilde{Y}^{(n)}_0 + \frac{1}{n}\sum f(t_j, \tilde{Y}^{(n)}_{t_j})\}$ is bounded in total variation, and therefore is relatively compact under the Skorohod topology. From the previous proof we also know that the sequence $\{U^{(n)}\}$ is relatively compact.

The following lemma is a standard result in the theory of BSDEs and will give us the basic estimates for the proof of our main result. To simplify presentation let us introduce some notation. For a process Y we denote the $L^{2,\infty}$-norm of Y as

$$\|Y\|_{L^{2,\infty}} := \Big\{\mathbb{E} \sup_{0\le t\le 1} |Y_t|^2\Big\}^{1/2}.$$

Further, we call the pair of functions F and f in (1) the *generator* of this BSDE; and we denote the adapted solution of (1) by $Y(F, f)$ [or $(Y(F, f), Z(F, f))$ if necessary], when the generators are to be specified. We have the following lemma.

LEMMA 3.4. *Let (G, g) and (F, f) be continuous and bounded generators. If $Y(G, g)$ and $Y(F, f)$ are the corresponding solutions for the BSDEs. Then there exists a constant $C < \infty$ which depends only on the Lipschitz constant L of f such that*

$$\|Y(G, g) - Y(F, f)\|_{L^{2,\infty}} \le C\big(|G(B) - F(B)|_{L^2(\Omega)} + \|f - g\|_\infty\big).$$

In particular if we have a sequence (G_k, g_k) *converging to* (F, f) *in* $L^2(\Omega) \times L^\infty([0,1] \times \mathbb{R}^d)$, *then the corresponding solutions converge in the* $L^{2,\infty}$*-norm.*

PROOF. For notational convenience we write $Y = Y(F, f)$ and $\bar{Y} = Y(G, g)$. Since $Y - \bar{Y}$ satisfies the equation

$$Y_t - \bar{Y}_t = \mathbb{E}\Big(F(B) - G(B) + \int_t^1 f(Y_s) - g(\bar{Y}_s)\,ds \,\big|\, \mathcal{F}_t\Big), \tag{6}$$

we deduce that $|Y_t - \bar{Y}_t| \le N_t = \mathbb{E}(|F(B) - G(B)| + \int_0^1 |f(Y_s) - g(\bar{Y}_s)|\,ds\,|\mathcal{F}_t)$, and therefore from Doob's maximal inequality

$$\begin{aligned}\mathbb{E}&\Big(\sup_{0\le t\le 1} |Y_t - \bar{Y}_t|^2\Big)\\ &\le 12\Big(\mathbb{E}\Big(|F(B) - G(B)|^2 + L^2 \int_0^1 |Y_s - \bar{Y}_s|^2\,ds\Big) + \|f - g\|_\infty^2\Big).\end{aligned}$$

On the other hand using (6) we obtain

$$\mathbb{E}(|Y_t - \bar{Y}_t|^2) \le 3\Big(\mathbb{E}\Big(|F(B) - G(B)|^2 + L^2 \int_t^1 |Y_s - \bar{Y}_s|^2\,ds\Big) + \|f - g\|_\infty^2\Big),$$

which by Gronwall's inequality gives

$$\mathbb{E}(|Y_t - \bar{Y}_t|^2) \le C_1\big(\mathbb{E}\big(|F(B) - G(B)|^2\big) + \|f - g\|_\infty^2\big),$$

for some finite constant C_1 which depends only on L, and the result follows. □

We note that if in the previous result the sequence $Y(G_k, g_k)$ is uniformly bounded by A, then we can replace $\|f - g\|_\infty$ by $\sup_{|x|\le A} |f(x) - g(x)|$. Using that $M^{(n)}$ converges weakly to a Brownian motion we obtain the following result, whose proof is in the same spirit as the previous lemma.

LEMMA 3.5. *Let* (G, g) *and* (F, f) *be continuous and bounded generators. Consider* $\tilde{Y}^{(n)}(G, g)$ *the solution of the discrete BSDE* (2) *associated with* (G, g) *and similarly for* $\tilde{Y}^{(n)}(F, f)$. *Then there exists a constant* $C < \infty$ *which depends only on the Lipschitz constant* L *of* f *such that*

$$\limsup_{n\to\infty} \|\tilde{Y}^{(n)}(G, g) - \tilde{Y}^{(n)}(F, f)\|_{L^{2,\infty}} \le C\big(|G(B) - F(B)|_{L^2(\Omega)} + \|f - g\|_\infty\big).$$

In particular if we have a sequence (G_k, g_k) *converging to* (F, f), *then*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|\tilde{Y}^{(n)}(G_k, g_k) - \tilde{Y}^{(n)}(F, f)\|_{L^{2,\infty}} = 0.$$

Our next step is to prove that our numerical method converges when the terminal functional F is of discrete type; that is, it depends only on a finite number of points of a continuous path. We also assume that (F, f) are in C_0^∞. For the sake of simplicity we assume that F depends only on two variables, leaving the obvious generalization for the reader.

We assume now the process Y satisfies the following backward stochastic differential equation:

$$Y_t = F(B_{\tau_0}, B_1) + \int_t^1 f(Y_s)\,ds + \int_t^1 Z_s\,dB_s, \tag{7}$$

where $\tau_0 \in (0, 1)$. Using Itô's formula the process Y is obtained by solving the following system of partial differential equations:

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x, y) + \frac{1}{2}\frac{\partial^2 u}{\partial y^2}(t, x, y) &= -f\big(u(t, x, y)\big), \qquad t \in [\tau_0, 1], \\ u(1, x, y) &= F(x, y), \\ \frac{\partial v}{\partial t}(t, z) + \frac{1}{2}\frac{\partial^2 v}{\partial z^2}(t, z) &= -f\big(v(t, z)\big), \qquad t \in [0, \tau_0], \\ v(\tau_0, z) &= u(\tau_0, z, z). \end{aligned} \tag{8}$$

In fact the solution is given by

$$Y_t = \begin{cases} u(t, B_{\tau_0}, B_t), & t \in [\tau_0, 1], \\ v(t, B_t), & t \in [0, \tau_0]. \end{cases} \tag{9}$$

In the previous system we can assume that u is defined on the whole interval $[0, 1]$.

To compute a discretization of u, we consider (x_i, y_j) such that $x_i = i\delta$, $y_j = j\delta$, where $\delta = \delta(n) = 1/\sqrt{n}$. We also write $\Delta t = 1/n$. We fix x_i and consider the following difference equation associated with the first part of (8) which comes from a Taylor expansion of $u(t + \Delta t, x_i, y_j + \delta) + u(t + \Delta t, x_i, y_j - \delta)$ [the idea is that $U^{(n)}(k, i, j) \approx u(k\,\Delta t, i\delta, j\delta)$]. We write $k_0 = [\tau_0 n]$,

$$\begin{aligned} &\tfrac{1}{2}\big(U^{(n)}(k+1, i, j+1) + U^{(n)}(k+1, i, j-1)\big) \\ &\qquad = U^{(n)}(k, i, j) - f\big(U^{(n)}(k, i, j)\big)\,\Delta t, \\ &U^{(n)}(n, i, j) = F(x_i, y_j). \end{aligned} \tag{10}$$

Using the Lipschitz condition on f one proves that the previous difference equation has a unique solution, for $k = k_0, \ldots, n-1$, $i, j \in \mathbb{Z}$. With this solution we approximate the second part of (8) by solving

$$\begin{aligned} &\tfrac{1}{2}\big(V^{(n)}(k+1, j+1) + V^{(n)}(k+1, j-1)\big) \\ &\qquad = V^{(n)}(k, j) - f\big(V^{(n)}(k, j)\big)\,\Delta t, \\ &V^{(n)}(k_0, j) = U^{(n)}(k_0, j, j), \end{aligned} \tag{11}$$

where $k = 0, \ldots, k_0 - 1$ and $j \in \mathbb{Z}$.

Using the results of [10], Chapter V, pages 353–381, we know that if (F, f) is C_0^4 then the solution of (8) is regular in (t, x, y) and it has bounded derivatives of order 2 in t and order 4 in (x, y). Let $C > 0$ be the common bound for these derivatives. Using this fact and a Taylor expansion for u on the (t, y) variables one obtains an upper bound for $\theta^{(n)}(k) = \sup_{i,j} |u(k\,\Delta t, x_i, y_j) - U^{(n)}(k, i, j)|$ for $k = k_0, \dots, n$. In fact if $\gamma = 1/(1 - L\,\Delta t)$, we get

$$\theta^{(n)}(k) \le \gamma\left[\theta^{(n)}(k+1) + \frac{4C}{n^{3/2}}\right],$$

which yields the inequality

$$\theta^{(n)}(k) \le \gamma^{n-k}\theta^{(n)}(n) + \frac{4C}{n^{3/2}}\sum_{p=1}^{n-k}\gamma^p = \gamma^{n-k}\theta^{(n)}(n) + \frac{4C}{n^{3/2}}\gamma\frac{\gamma^{n-k}-1}{\gamma-1}.$$

From $\Delta t = 1/n$ one obtains, as before, for large n that $(\frac{1}{1-L/n})^{n-k} \le (1 + 2L/n)^{n-k} \le e^{2L}$ and therefore

$$\max_{k_0 \le k \le n} \theta^{(n)}(k) \le e^{2L}\theta^{(n)}(n) + \frac{4C(e^{2L}-1)}{L\sqrt{n}}.$$

Since $\theta^{(n)}(n) = 0$ we get $\max_{k_0 \le k \le n} \theta^{(n)}(k) \le \frac{4C(e^{2L}-1)}{L\sqrt{n}}$.

For the other terms we proceed in the same way. Consider

$$\phi^{(n)}(k) = \sup_{i,j} |v(k\Delta t, y_j) - V^{(n)}(k, j)|,$$

for $0 \le k \le k_0$. Then we have $\phi^{(n)}(k_0) \le C(\tau_0 - k_0/n) + \theta^{(n)}(k_0) \le C(1 + 4(e^{2L} - 1)/L)/\sqrt{n}$ and, for $0 \le k \le k_0 - 1$,

$$\phi^{(n)}(k) \le \gamma\left[\phi^{(n)}(k+1) + \frac{4C}{n^{3/2}}\right],$$

which yields the upper bound

$$\begin{aligned}\max_{0 \le k \le k_0} \phi^{(n)}(k) &\le e^{2L}\phi^{(n)}(k_0) + \frac{4C(e^{2L}-1)}{L\sqrt{n}} \\ &\le \left(e^{2L} + 4e^{2L}\frac{e^{2L}-1}{L} + 4\frac{e^{2L}-1}{L}\right)\frac{C}{\sqrt{n}}.\end{aligned}$$

Therefore, using a suitable constant A, we obtain the following estimate:

$$\begin{aligned}\sup_{i,j\in\mathbb{Z}}\Big\{&\sup_{k_0 \le k \le n} |u(k\,\Delta t, x_i, y_j) - U^{(n)}(k, i, j)|, \\ &\sup_{0 \le k \le k_0} |v(k\,\Delta t, y_j) - V^{(n)}(k, j)|\Big\} \le \frac{A}{\sqrt{n}}.\end{aligned} \tag{12}$$

With this estimate in hand we can prove that $\tilde{Y}^{(n)}$ converges to Y. Using Skorohod's embedding theorem we can assume that $M^{(n)}, B$ are defined in the same space and $M^{(n)}$ converges a.s. uniformly on $[0, 1]$ to B. On the other hand it is not hard to see that the unique solution of

$$\tilde{Y}^{(n)}_{t_k} = F\big(M^{(n)}_{t_{k_0}}, M^{(n)}_1\big) + \sum_{p=k}^{n-1} f\big(\tilde{Y}^{(n)}_{t_p}\big)\Delta t + \sum_{p=k}^{n} Z^{(n)}_{t_p}\big(M^{(n)}_{t_{p+1}} - M^{(n)}_{t_p}\big) \tag{13}$$

is given by

$$\tilde{Y}^{(n)}_{t_k} = \begin{cases} U^{(n)}\big(t_k, M^{(n)}_{t_{k_0}}, M^{(n)}_{t_k}\big), & t_k \in [t_{k_0}, 1], \\ V^{(n)}\big(t_k, M^{(n)}_{t_k}\big), & t_k \in [0, t_{k_0}]. \end{cases}$$

Finally, from (12) one obtains that $\tilde{Y}^{(n)}$ converges a.s. uniformly on compact sets to the continuous process (9), which is exactly the solution of (7). This means that Theorem 1 is proven when (F, f) is C_b^∞, and F depends on a finite number of coordinates.

Now, we are ready to complete the proof of Theorem 1 in the general case. First, we take a convergent subsequence of $\{\tilde{Y}^{(n)}\}$, which by simplicity we denote by $\{\tilde{Y}^{(n)}\}$ itself. By Skorohod's embedding theorem we can assume that all are defined on the same space and moreover they converge a.s. uniformly on $[0, 1]$ to a continuous process X. We consider the approximation F_m which corresponds to the linear interpolation at $t_0, \dots, t_m = 1$. The processes $\tilde{Y}^{(n,m)} = \tilde{Y}^{(n)}(F_m, f)$, $\tilde{Y}^{(n)} = \tilde{Y}^{(n)}(F, f)$, $Y(F_m, f)$ are uniformly bounded by some constant A and therefore we can approximate (F_m, f) with a sequence of C_0^∞-functions $((F_{m,p}, f_p))_p$, such that

$$\lim_{p\to\infty} \mathbb{E}\big(|F_{m,p}(B_{t_0}, \dots, B_{t_m}) - F_m(B_{t_0}, \dots, B_{t_m})|^2\big) = 0$$

and

$$\lim_{p\to\infty} \|f_p - f\|_\infty = 0.$$

Actually we just need $\lim_{p\to\infty} \sup_{|x|\le A} |f_p - f| = 0$, for an appropriate constant A. We write $\tilde{Y}^{(n,m,p)} = \tilde{Y}^{(n)}(F_{m,p}, f_p)$ and $Y^{(m,p)} = Y(F_{m,p}, f_p)$. Using the triangle inequality we obtain

$$\begin{aligned} \|Y - X\| \le{}& \|Y - Y^{(m)}\| + \|Y^{(m)} - Y^{(m,p)}\| + \|Y^{(m,p)} - \tilde{Y}^{(n,m,p)}\| \\ &+ \|\tilde{Y}^{(n,m,p)} - \tilde{Y}^{(n,m)}\| + \|\tilde{Y}^{(n,m)} - \tilde{Y}^{(n)}\| + \|\tilde{Y}^{(n)} - X\|. \end{aligned} \tag{14}$$

Taking limits first in n, then in p and finally in m we see that the right-hand side of (14) converges to 0. In fact, the first and the second terms converge to zero due to Lemma 3.4. The fourth and fifth ones converge to zero by Lemma 3.5. The last one tends to zero thanks to the a.s. uniform convergence of $\tilde{Y}^{(n)}$ to X

plus the dominated convergence theorem. Finally, the third term converges to 0 because $(F_{m,p}, f_p)$ is a smooth function, and $F_{m,p}$ depends on a finite number of coordinates. Therefore the only limit point for the sequence $\tilde{Y}^{(n)}$ is exactly Y, the solution of (1), proving the result.

REMARK. We note that our results in this paper can easily be used to treat a slightly more general case, where there is a term Z in the drift of the BSDE. Indeed, suppose that $(H_s)_{s>0}$ is a predictable process in L^2 for $dP\,dt$. Also suppose it is known to us. Then we can numerically solve the BSDE

$$Y_t = \xi + \int_t^1 \big(f(s, Y_s) + H_s Z_s\big)\,ds - \int_t^1 Z_s\,dB_s \tag{15}$$

by first solving

$$Y_t = \xi + \int_t^1 f(s, Y_s)\,ds - \int_t^1 Z_s\,dB_s. \tag{16}$$

We do this as follows. Given equation (15) on $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t, P)$, we form the equation

$$M_t = 1 + \int_0^1 M_s H_s\,dB_s \tag{17}$$

and then set $\beta = B_t - \int_0^t H_s\,ds$, which is a Brownian motion under the new probability measure Q defined by $dQ = M_T\,dP$. We then solve (16) under Q with β replacing B. The solution Y of (16) is then also a solution of (15) under P and using the original Brownian motion B. Last, we note that equations of the type (15) are of interest in financial asset pricing theory; see, for example, [9], Section 1.

Summary. We summarize here the proposed algorithm to solve the BSDE

$$Y_t = F(B) + \int_t^1 f(s, Y_s)\,ds - \int_t^1 Z_s\,dB_s,$$

by means of a random walk appoximation $M^{(n)}$, in the following scheme:

$$\begin{aligned}
&\hat{Y}_1^{(n)} = F(M^{(n)}), \qquad \hat{Z}_1^{(n)} = 0,\\
&\hat{X}_{t_i} = \mathbb{E}\{\hat{Y}_{t_{i+1}}^{(n)} \mid \mathcal{F}_{t_i}^{(n)}\},\\
&\hat{Y}_{t_i}^{(n)} = \hat{X}_{t_i} + \frac{1}{n} f(t_i, \hat{X}_{t_i}),\\
&\hat{Z}_{t_i}^{(n)} = \mathbb{E}\left\{\left[\hat{Y}_{t_{i+1}}^{(n)} + \frac{1}{n} f\big(t_i, \hat{Y}_{t_i}^{(n)}\big) - \hat{Y}_{t_i}^{(n)}\right]\big(\Delta M_{t_{i+1}}^{(n)}\big)^{-1} \mid \mathcal{F}_{t_i}^{(n)}\right\},
\end{aligned}$$

where the conditional expectations with respect to the discrete σ-field $\mathcal{F}_{t_i}^{(n)}$ are

computed using a tree structure. For example we obtain

$$\hat{X}_{t_{n-1}} = \frac{1}{2}\left\{F\left(M_{t_{n-1}}^{(n)} + \frac{1}{\sqrt{n}}\right) + F\left(M_{t_{n-1}}^{(n)} - \frac{1}{\sqrt{n}}\right)\right\}.$$

REFERENCES

[1] ANTONELLI, F. and KOHATSU-HIDA, A. (2000). Filtration stability of backward SDE's. *Stochastic Anal. Appl.* **18** 11–37.

[2] BALLY, V. (1997). Approximation scheme for solutions of BSDE. *Backward Stochastic Differential Equations* (N. El Karoui and L. Mazliak, eds.) 177–191. Pitman, London.

[3] BALLY, V., PAGÈS, G. and PRINTEMS, J. (2001). A stochastic quantization method for nonlinear problems. *Monte Carlo Methods Appl.* **7** 21–34.

[4] BUCKDAHN, R. (2001). Backward stochastic differential equations driven by a martingale. Preprint 93-05, Univ. Provence.

[5] CHEVANCE, D. (1997). Discrétisation des équations différentieles stochastiques rétrogrades. In *Numerical Methods in Finance* (L. C. G. Rogers and D. Talay, eds.) 232–244. Cambridge Univ. Press.

[6] COQUET, F., MACKEVIČIUS, V. and MÉMIN, J. (1998). Stability in D of martingales and backward equations under discretization of filtration. *Stochastic Process. Appl.* **75** 235–248.

[7] DOUGLAS, J., MA, J. and PROTTER, P. (1996). Numerical methods for forward–backward stochastic differential equations. *Ann. Appl. Probab.* **6** 940–968.

[8] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71.

[9] EL KAROUI, N. and QUENEZ, M. C. (1997). Nonlinear pricing theory and backward stochastic differential equations. *Financial Mathematics*. *Lecture Notes in Math.* **1656** 191–246. Springer, New York.

[10] FREIDLIN, M. (1985). *Functional Integration and Partial Differential Equations*. Princeton Univ. Press.

[11] JACOD, J., MÉMIN, J. and METIVIER, M. (1983). On tightness and stopping times. *Stochastic Process. Appl.* **14** 109–146.

[12] LEPELTIER, J. P. and SAN MARTIN, J. (1997). Backward stochastic differential equations with continuous generator. *Statist. Probab. Lett.* **32** 425–430.

[13] MA, J., PROTTER, P. and YONG, J. (1994). Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Relat/Fields* **98** 339–359.

[14] PARDOUX, E. and PENG, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61.

[15] PARDOUX, E. and PENG, S. (1992). Backward stochastic differential equation and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control and Inform. Sci.* **176** 200–217. Springer, New York.

[16] PROTTER, P. (1990). *Stochastic Integration and Differential Equations*. Springer, New York.

J. MA
DEPARTMENT OF MATHEMATICS AND STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1395

P. PROTTER
ORIE
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-3801
E-MAIL: protter@orie.cornell.edu

J. SAN MARTÍN
FACULTAD DE CIENCIAS FÍSICAS
Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA
UNIVERSIDAD DE CHILE
AND
CENTRO DE MODELAMIENTO MATEMÁTICO
UMO 2071-CNRS
CASILLA 170-3 CORREO 3, SANTIAGO
CHILE

S. TORRES
DEPARTAMENTO DE ESTADÍSTICA
IMAFI
FACULTAD DE CIENCIAS
UNIVERSIDAD DE VALPARAÍSO
CASILLA 5030, VALPARAÍSO
CHILE
AND
CENTRO DE MODELAMIENTO MATEMÁTICO
UMO 2071-CNRS
CASILLA 170-3 CORREO 3, SANTIAGO
CHILE