Chapter 5
Numerical Methods for FBSDEs

Abstract We investigate numerical methods for forward-backward stochastic differential equations driven by a Brownian motion and a compensated Poisson random measure. We consider three approaches to solving FBSDEs. We apply discrete-time approximations and we derive recursive representations of the solution involving conditional expected values. In order to estimate the conditional expected values, we use Least Squares Monte Carlo which overcomes nested Monte Carlo simulations. In the case of a FBSDE driven by a Brownian motion and a compensated Poisson process we replace the original driving noises by discrete-space martingales. We also use the connection with partial integro-differential equations and we present an explicit-implicit finite difference method for solving a PIDE.

We continue the study of (decoupled) forward-backward stochastic differential equations driven by a Brownian motion and a compensated Poisson random measure, which we introduced in the previous chapter. In most cases we cannot derive the solution to a FBSDE in an explicit form and we have to apply a numerical method to solve a FBSDE. In this chapter we investigate three approaches to solving FBSDEs numerically.

5.1 Discrete-Time Approximation and Least Squares Monte Carlo

Let assumptions (B1)–(B5) from Chap. 4 hold. We deal with the forward-backward stochastic differential equation

$$\begin{aligned}\mathscr{X}(t) &= x + \int_0^t \mu\big(\mathscr{X}(s-)\big)ds + \int_0^t \sigma\big(\mathscr{X}(s-)\big)dW(s)\\ &\quad + \int_0^t\int_{\mathbb{R}} \gamma\big(\mathscr{X}(s-),z\big)\tilde{N}(ds,dz), \quad 0 \le t \le T,\\ Y(t) &= g\big(\mathscr{X}(T)\big) + \int_t^T f\left(s, \mathscr{X}(s-), Y(s-), Z(s), \int_{\mathbb{R}} U(s,z)\delta(z)\nu(dz)\right)ds\\ &\quad - \int_t^T Z(s)dW(s) - \int_t^T\int_{\mathbb{R}} U(s,z)\tilde{N}(ds,dz), \quad 0 \le t \le T.\end{aligned} \tag{5.1}$$

Ł. Delong, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, EAA Series, DOI 10.1007/978-1-4471-5331-3_5,

We denote

$$\Psi(t) = \int_{\mathbb{R}} U(t,z)\delta(z)\nu(dz), \quad 0 \le t \le T.$$

First, we consider the case of the random measure N generated by a compound Poisson process.

We aim to solve the FBSDE (5.1). An intuitive idea is to discretize the forward and the backward equation in the spirit of the Euler method. We choose a regular time grid $\pi = \{t_i = ih, i = 0, 1, \dots, n\}$ with step $h = \frac{T}{n}$. The solution to the forward equation (5.1) is approximated by

$$\begin{aligned} \mathscr{X}^n(0) &= x, \\ \mathscr{X}^n(t_{i+1}) &= \mathscr{X}^n(t_i) + \mu\big(\mathscr{X}^n(t_i)\big)h + \sigma\big(\mathscr{X}^n(t_i)\big)\Delta W(i+1) \\ &\quad + \int_{\mathbb{R}} \gamma\big(\mathscr{X}^n(t_i), z\big)\tilde{N}\big((t_i, t_{i+1}], dz\big), \quad i = 0, \dots, n-1. \end{aligned} \tag{5.2}$$

where $\Delta W(i+1) = W(t_{i+1}) - W(t_i)$ denotes the increment of the Brownian motion. Clearly, there exists a unique $\mathscr{F}$-adapted, square integrable solution $\mathscr{X}^n$ to (5.2). We set $\mathscr{X}^n(t) = \mathscr{X}^n(t_i)$, $t_i \le t < t_{i+1}$. If we apply the Euler-type discretization to the backward equation (5.1), we obtain

$$\begin{aligned} Y^n(T) &= g\big(\mathscr{X}^n(T)\big), \\ Y^n(t_i) &= Y^n(t_{i+1}) + f\left(t_i, \mathscr{X}^n(t_i), Y^n(t_i), Z^n(t_i), \int_{\mathbb{R}} U^n(t_i, z)\delta(z)\nu(dz)\right)h \\ &\quad - Z^n(t_i)\Delta W(i+1) - \int_{\mathbb{R}} U^n(t_i, z)\tilde{N}\big((t_i, t_{i+1}], dz\big), \quad i = n-1, \dots, 0. \end{aligned} \tag{5.3}$$

Unfortunately, the discrete-time equation (5.3) does not have a solution since the time-discretized Brownian motion and compound Poisson process do not have the predictable representation property, see Briand et al. (2002). However, we use the following backward recursion

$$\begin{aligned} Y^n(T) &= g\big(\mathscr{X}^n(T)\big), \\ Z^n(t_i) &= \frac{1}{h}\mathbb{E}\big[Y^n(t_{i+1})\Delta W(i+1)|\mathscr{F}_{t_i}\big], \quad i = n-1, \dots, 0, \\ \Psi^n(t_i) &= \frac{1}{h}\mathbb{E}\left[Y^n(t_{i+1})\int_{\mathbb{R}} \delta(z)\tilde{N}\big((t_i, t_{i+1}], dz\big)|\mathscr{F}_{t_i}\right], \quad i = n-1, \dots, 0, \\ Y^n(t_i) &= \mathbb{E}\big[Y^n(t_{i+1})|\mathscr{F}_{t_i}\big] \\ &\quad + f\big(t_i, \mathscr{X}^n(t_i), Y^n(t_i), Z^n(t_i), \Psi^n(t_i)\big)h, \quad i = n-1, \dots, 0. \end{aligned} \tag{5.4}$$

We set $Y^n(t) = Y^n(t_i)$, $Z^n(t) = Z^n(t_i)$, $\Psi^n(t) = \Psi^n(t_i)$, $t_i \le t < t_{i+1}$. The backward recursion (5.4) can be derived by a heuristic reasoning. Let us first recall that for square integrable martingales M_1 and M_2 we have

$$\mathbb{E}\big[M_1(T)M_2(T)\big] = \mathbb{E}\big[[M_1, M_2](T)\big], \tag{5.5}$$

see Corollary II.27.3 in Protter (2004). If we multiply (5.3) by $\Delta W(i+1)$ and $\int_{\mathbb{R}} \delta(z)\tilde{N}((t_i, t_{i+1}], dz)$, take the conditional expected value and use (5.5), then we obtain

$$\begin{aligned}
&\mathbb{E}\big[Y^n(t_{i+1})\Delta W(i+1)|\mathscr{F}_{t_i}\big] = \mathbb{E}\big[Z^n(t_i)|\Delta W(i+1)|^2|\mathscr{F}_{t_i}\big] = Z^n(t_i)h,\\
&\mathbb{E}\bigg[Y^n(t_{i+1})\int_{\mathbb{R}} \delta(z)\tilde{N}\big((t_i, t_{i+1}], dz\big)|\mathscr{F}_{t_i}\bigg]\\
&\quad = \mathbb{E}\bigg[\int_{\mathbb{R}} U^n(t_i, z)\tilde{N}\big((t_i, t_{i+1}], dz\big)\int_{\mathbb{R}} \delta(z)\tilde{N}\big((t_i, t_{i+1}], dz\big)|\mathscr{F}_{t_i}\bigg]\\
&\quad = \mathbb{E}\bigg[\bigg[\int_{t_i}^{\cdot}\int_{\mathbb{R}} U^n(t_i, z)\tilde{N}(dt, dz), \int_{t_i}^{\cdot}\int_{\mathbb{R}} \delta(z)\tilde{N}(dt, dz)\bigg](t_{i+1})|\mathscr{F}_{t_i}\bigg]\\
&\quad = \mathbb{E}\bigg[\int_{\mathbb{R}} U^n(t_i, z)\delta(z)N\big((t_i, t_{i+1}], dz\big)|\mathscr{F}_{t_i}\bigg]\\
&\quad = \int_{\mathbb{R}} U^n(t_i, z)\delta(z)\nu(dz)h = \Psi^n(t_i)h,
\end{aligned}$$

and the formulas for Z^n, Ψ^n can be established. If we take the condition expected value on both sides of (5.3), then the formula for Y^n can be established.

The next theorem justifies the approximations (5.2) and (5.4), see Theorem 2.1, Corollary 2.1 and Remark 2.7 in Bouchard and Elie (2008).

Theorem 5.1.1 *Consider the FBSDE* (5.1) *and the random measure N generated by a compound Poisson process. Assume that* (B2)–(B5) *from Sect.* 4.1 *hold and let the generator f be* 1/2*-Hölder continuous in t. We deal with the approximations* (5.2) *and* (5.4) *of the solution to the FBSDE* (5.1). *We have*

$$\begin{aligned}
&\max_{i=0,1,\dots,n-1}\mathbb{E}\Big[\sup_{t\in[t_i,t_{i+1}]}\big|\mathscr{X}(t) - \mathscr{X}^n(t_i)\big|^2\Big]\\
&\quad + \max_{i=0,1,\dots,n-1}\mathbb{E}\Big[\sup_{t\in[t_i,t_{i+1}]}\big|Y(t) - Y^n(t_i)\big|^2\Big]\\
&\quad + \mathbb{E}\bigg[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\big|Z(t) - Z^n(t_i)\big|^2 dt\bigg] \le K\frac{1}{n},\\
&\mathbb{E}\bigg[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\big|\Psi(t) - \Psi^n(t_i)\big|^2 dt\bigg] \le K\Big(\frac{1}{n}\Big)^{1-\varepsilon}, \quad \varepsilon > 0.
\end{aligned}$$

In addition, if for each $z \in \mathbb{R}$ the mapping $x \mapsto \gamma(x, z)$ is differentiable and

$$\big|\gamma_x(x, z) + 1\big| \ge K > 0, \quad (x, z) \in \mathbb{R}\times\mathbb{R},$$

then

$$
\begin{aligned}
&\max_{i=0,1,\dots,n-1}\mathbb{E}\Big[\sup_{t\in[t_i,t_{i+1}]}\big|\mathscr{X}(t)-\mathscr{X}^n(t_i)\big|^2\Big]\\
&\quad+\max_{i=0,1,\dots,n-1}\mathbb{E}\Big[\sup_{t\in[t_i,t_{i+1}]}\big|Y(t)-Y^n(t_i)\big|^2\Big]\\
&\quad+\mathbb{E}\Bigg[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\big|Z(t)-Z^n(t_i)\big|^2dt\Bigg]\\
&\quad+\mathbb{E}\Bigg[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\big|\Psi(t)-\Psi^n(t_i)\big|^2dt\Bigg]\le K\frac{1}{n}.
\end{aligned}
$$

We remark that deriving Y^n from (5.4) involves solving a fixed point equation. For a Lipschitz continuous generator the fixed point procedure of solving (5.4) has a convergence rate of $1/n$. Hence, numerical cost is small. To overcome the fixed point procedure, we can use the following scheme

$$
\begin{aligned}
Y^n(T)&=g\big(\mathscr{X}^n(T)\big),\\
Z^n(t_i)&=\frac{1}{h}\mathbb{E}\big[Y^n(t_{i+1})\Delta W(i+1)|\mathscr{F}_{t_i}\big],\quad i=n-1,\dots,0,\\
\Psi^n(t_i)&=\frac{1}{h}\mathbb{E}\Big[Y^n(t_{i+1})\int_{\mathbb{R}}\delta(z)\tilde{N}\big((t_i,t_{i+1}],dz\big)|\mathscr{F}_{t_i}\Big],\quad i=n-1,\dots,0,\\
Y^n(t_i)&=\mathbb{E}\big[Y^n(t_{i+1})\\
&\qquad+f\big(t_i,\mathscr{X}^n(t_i),Y^n(t_{i+1}),Z^n(t_i),\Psi^n(t_i)\big)h|\mathscr{F}_{t_i}\big],\quad i=n-1,\dots,0,
\end{aligned}
$$

but more complicated conditional expected values have to be calculated instead.

The algorithm (5.4) is still not an implementable scheme since the conditional expectations have to be estimated. Performing Monte Carlo simulations at each point t_i would lead to so-called nested simulations and an enormous numerical cost. Least Squares Monte Carlo can overcome nested simulations.

By the Markov property the conditional expected values in (5.4) can be represented as functions of the state process $\mathscr{X}$. The idea is to approximate the unknown functions by their projections on finite-dimensional function bases. At each point t_i we choose 3 function bases $(b_{l,i}(\cdot))_{l=0,1,2}$ and we approximate each conditional expected value in a vector space spanned by the basis. Each basis $b_{l,i}(\cdot)$ is a $d_{l,i}$-dimensional vector of scalar functions. The vector space spanned by $b_{l,i}$ is denoted by $\alpha b_{l,i}(\cdot)=\sum_{k=1}^{d_{l,i}}\alpha_k b_{l,i}^k(\cdot)$ where $\alpha\in\mathbb{R}^{d_{l,i}}$.

We can use the Least Squares Monte Carlo algorithm:

1. Simulate L independent copies of $(\Delta W^m(i+1)$, $i=0,1,\dots,n-1$, $m=1,\dots,L)$ and $(\tilde{N}^m((t_i,t_{i+1}],.)$, $i=0,1,\dots,n-1$, $m=1,\dots,L)$,
2. Simulate L independent copies of $(\mathscr{X}^{n,m}(t_i)$, $i=1,\dots,n$, $m=1,\dots,L)$,

3. Set $\hat{Y}(T, \mathscr{X}^{n,m}(T)) = g(\mathscr{X}^{n,m}(T))$ for $m = 1, \ldots, L$,
4. Choose function bases $(b_{l,i}(.))_{l=0,1,2,i=0,1,\ldots,n-1}$,
5. Going backwards, for $i = n-1, \ldots, 0$ solve the least squares regression problems

$$\alpha_{1,i} = \arg\inf_{\alpha}\left\{\frac{1}{L}\sum_{i=1}^{L}\left|\frac{1}{h}\hat{Y}\big(t_{i+1}, \mathscr{X}^{n,m}(t_{i+1})\big)\Delta W^m(i+1) - \alpha b_{1,i}\big(\mathscr{X}^{n,m}(t_i)\big)\right|^2\right\},$$

$$\alpha_{2,i} = \arg\inf_{\alpha}\left\{\frac{1}{L}\sum_{i=1}^{L}\left|\frac{1}{h}\hat{Y}\big(t_{i+1}, \mathscr{X}^{n,m}(t_{i+1})\big)\int_{\mathbb{R}}\delta(z)\tilde{N}^m\big((t_i, t_{i+1}], dz\big)\right.\right.$$
$$\left.\left. - \alpha b_{2,i}\big(\mathscr{X}^{n,m}(t_i)\big)\right|^2\right\},$$

6. Set $\hat{Z}(t_i, \mathscr{X}^{n,m}(t_i)) = \alpha_{1,i} b_{1,i}(\mathscr{X}^{n,m}(t_i))$ and $\hat{\Psi}(t_i, \mathscr{X}^{n,m}(t_i)) = \alpha_{2,i} \times b_{2,i}(\mathscr{X}^{n,m}(t_i))$,
7. Solve the least squares regression problem

$$\alpha_{0,i} = \arg\inf_{\alpha}\left\{\frac{1}{L}\sum_{i=1}^{L}\left|\hat{Y}\big(t_{i+1}, \mathscr{X}^{n,m}(t_{i+1})\big) + f\big(t_i, \mathscr{X}^{n,m}(t_i),\right.\right.$$
$$\hat{Y}\big(t_{i+1}, \mathscr{X}^{n,m}(t_{i+1})\big), \hat{Z}\big(t_i, \mathscr{X}^{n,m}(t_i)\big), \hat{\Psi}\big(t_i, \mathscr{X}^{n,m}(t_i)\big)\big)h$$
$$\left.\left. - \alpha b_{0,i}\big(\mathscr{X}^{n,m}(t_i)\big)\right|^2\right\},$$

8. Set $\hat{Y}(t_i, \mathscr{X}^{h,m}(t_i)) = \alpha_{0,i} b_{0,i}(\mathscr{X}^{n,m}(t_i))$,
9. Continue till $t_0 = 0$.

Polynomials, hypercubes and Voronoi partitions are usually used as basis functions, see Gobet et al. (2005). Notice that when we apply the Least Squares Monte Carlo algorithm we additionally face the error of approximating the conditional expectations by estimated regression functions. The total error of the Least Squares Monte Carlo algorithm depends on the number of time steps n, the number of simulations L and the number of basis functions d. The total error is studied in Bouchard and Touzi (2004), Gobet et al. (2005), Gobet et al. (2006), Gobet and Lemor (2006). We also point out that some truncation procedures can be useful in the final application of the algorithm, see Gobet et al. (2005).

We comment on one modification of the algorithm presented. It is shown by Bouchard and Touzi (2004), in the case of BSDEs driven by Brownian motions, that the error of approximating the conditional expectation by an estimator explodes when the mesh of time partition goes to zero, given the accuracy of the estimator. In order to control this approximation error one is forced to simulate more paths as the time partition becomes finer. This significantly increases computational cost. The idea of Bender and Denk (2007), who also investigate BSDEs driven by Brownian motions, is first to approximate the solution (Y, Z) by the Picard iterations

$$\begin{aligned} Y^I(t) = g\big(\mathscr{X}(T)\big) + \int_t^T f\big(s, \mathscr{X}(s), Y^{I-1}(s), Z^{I-1}(s)\big)ds \\ - \int_t^T Z^I(s)dW(s), \quad 0 \le t \le T, \end{aligned} \tag{5.6}$$

and to apply the algorithm (5.4) to derive (Y^I, Z^I). The Picard procedure clearly introduces an additional error, which converges to zero at geometric rate. The advantage of the scheme proposed by Bender and Denk (2007) is that the error of approximating of the conditional expectation by an estimator is reduced and this error does not explode when the mesh of time partition tends to zero and the number of the Picard iteration goes to infinity.

Let us now comment on the case when a FBSDE is driven by a general compensated Poisson random measure. It is known that we cannot simulate small jumps of a Lévy process with an infinite Lévy measure, see Chap. 6 in Cont and Tankov (2004). The usual procedure is to cut off small jumps of a Lévy process and approximate them by an independent Brownian motion. After cutting off small jumps, we can investigate a FBSDE driven by a compensated compound Poisson process and we can apply the algorithm presented in this chapter, see Aazizi (2011) for details.

Finally, we remark that in many applications we end up with a BSDE with zero generator or with a BSDE with generator independent of (Y, Z, U) for which we can derive representations of the solution Y and the control processes (Z, U) in the form of conditional expectations of the state process, see Proposition 4.1.2 and Chap. 8. In those cases the Monte Carlo algorithm is much simpler.

5.2 Discrete-Time and Discrete-Space Martingale Approximation

We deal with the forward-backward stochastic differential equation

$$\begin{aligned} \mathscr{X}(t) &= x + \int_0^t \mu\big(\mathscr{X}(s-)\big)ds + \int_0^t \sigma\big(\mathscr{X}(s-)\big)dW(s) \\ &\quad + \int_0^t \gamma\big(\mathscr{X}(s-)\big)\tilde{N}(ds), \quad 0 \le t \le T, \\ Y(t) &= g\big(\mathscr{X}(T)\big) + \int_t^T f\big(s, \mathscr{X}(s-), Y(s-), Z(s), U(s)\big)ds \\ &\quad - \int_t^T Z(s)dW(s) - \int_t^T U(s)\tilde{N}(ds), \quad 0 \le t \le T, \end{aligned} \tag{5.7}$$

where the measure N is the jump measure of a Poisson process with intensity λ. We consider a discrete-time approximation to (5.7) and we approximate the Brownian motion and the compensated random measure by two discrete-space martingales.

We choose a regular time grid $\pi = \{t_i = ih, i = 0, 1, \ldots, n\}$ with step $h = \frac{T}{n}$. We define two random walks $W^n := (W^n(k), k = 0, 1, \ldots, n)$ and $\tilde{N}^n := (\tilde{N}^n(k),\ k = 0, 1, \ldots, n)$ by

$$\begin{aligned} W^n(0) &= 0, \qquad W^n(k) = \sqrt{h}\sum_{i=1}^{k}\xi_i^n, \quad k = 1, 2, \ldots, n, \\ \tilde{N}^n(0) &= 0, \qquad \tilde{N}^n(k) = \sum_{i=1}^{k}\zeta_i^n, \quad k = 1, 2, \ldots, n, \end{aligned} \tag{5.8}$$

where $\xi_1^n, \ldots, \xi_n^n$ are independent Bernoulli random variables with probabilities

$$\mathbb{P}\big(\xi_k^n = 1\big) = \mathbb{P}\big(\xi_k^n = -1\big) = \frac{1}{2},$$

and $\zeta_1^n, \ldots, \zeta_n^n$ are independent Bernoulli random variables with probabilities

$$\mathbb{P}\big(\zeta_k^n = e^{-\lambda h} - 1\big) = 1 - \mathbb{P}\big(\zeta_k^n = e^{-\lambda h}\big) = e^{-\lambda h}.$$

We also introduce the filtration $\mathscr{F}_k^n = \sigma(\xi_1^n, \ldots, \xi_k^n, \zeta_1^n, \ldots, \zeta_k^n)$, $k = 1, \ldots, n$. The random walks W^n and $\tilde{N}^n$ are $\mathscr{F}^n$-discrete-time-space-martingales.

The first result shows that the random walks are good approximations of the Brownian motion and the compensated Poisson process, see Lemma 3 in Lejay et al. (2010).

Proposition 5.2.1 *The processes* $(W^n([\frac{t}{h}]), \tilde{N}^n([\frac{t}{h}]), 0 \le t \le T)$ *converge in the* J_1*-Skorokhod topology in probability to* $(W(t), \tilde{N}(t), 0 \le t \le T)$ *as* $n \to \infty$.

It is intuitive to approximate the solution to the forward equation (5.7) in the following way

$$\begin{aligned} \mathscr{X}^n(0) &= x, \\ \mathscr{X}^n(t_{i+1}) &= \mathscr{X}^n(t_i) + \mu\big(X^n(t_i)\big)h + \sqrt{h}\sigma\big(\mathscr{X}^n(t_i)\big)\xi_{i+1}^n \\ &\quad + \gamma\big(\mathscr{X}^n(t_i)\big)\zeta_{i+1}^n, \quad i = 0, 1, \ldots, n-1. \end{aligned} \tag{5.9}$$

Clearly, there exists a unique $\mathscr{F}^n$-adapted, square integrable solution $\mathscr{X}^n$ to (5.9). We set $\mathscr{X}^n(t) = \mathscr{X}^n(t_i)$, $t_i \le t < t_{i+1}$. In Sect. 5.1 we claim that the time-discretized Brownian motion and compound Poisson process do not have the predictable representation property. However, in some cases an orthogonal martingale term can be added to recover the predictable representation property, see Briand et al. (2002) and Lejay et al. (2010). We approximate the solution to the backward stochastic differential equation (5.7) by solving the backward stochastic difference equation

$$\begin{aligned} Y^n(T) &= g\big(\mathscr{X}^n(T)\big), \\ Y^n(t_i) &= Y^n(t_{i+1}) + f\big(t_i, Y^n(t_i), Z^n(t_i), U^n(t_i)\big)h \\ &\qquad - \sqrt{h} Z^n(t_i)\xi^n_{i+1} - U^n(t_i)\zeta^n_{i+1} - V^n(t_i)\varsigma^n_{i+1}, \quad i = 0, 1, \dots, n-1, \end{aligned} \tag{5.10}$$

where $(\varsigma^n_i, i = 1, \dots, n)$ denotes the increments of a third orthogonal discrete-time-space martingale. By the predictable representation property, for an $\mathscr{F}^n$-measurable $\mathscr{X}^n(T)$ there exists a unique $\mathscr{F}^n$-adapted solution (Y^n, Z^n, U^n, V^n) to the backward equation (5.10). We can also derive that solution. Multiplying both sides of (5.10) by ξ^n_{i+1} or ζ^n_{i+1} and taking the conditional expected values, we obtain the representations

$$\begin{aligned} Y^n(T) &= g\big(\mathscr{X}^n(T)\big), \\ Z^n(t_i) &= \frac{1}{\sqrt{h}} \mathbb{E}\big[Y^n(t_{i+1})\xi^n_{i+1} | \mathscr{F}^n_i\big], \quad i = n-1, \dots, 0, \\ U^n(t_i) &= \frac{1}{e^{-\lambda h}(1 - e^{-\lambda h})} \mathbb{E}\big[Y^n(t_{i+1})\zeta^n_{i+1} | \mathscr{F}^n_i\big], \quad i = n-1, \dots, 0, \\ Y^n(t_i) &= \mathbb{E}\big[Y^n(t_{i+1}) | \mathscr{F}^n_i\big] \\ &\qquad + f\big(t_i, \mathscr{X}^n(t_i), Y^n(t_i), Z^n(t_i), U^n(t_i)\big)h, \quad i = n-1, \dots, 0. \end{aligned} \tag{5.11}$$

We set $Y^n(t) = Y^n(t_i)$, $Z^n(t) = Z^n(t_i)$, $U^n(t) = U^n(t_i)$, $t_i \le t < t_{i+1}$. The process V^n can also be derived from (5.10) but it is not needed for the approximation of the solution to (5.7). Again, a fixed point procedure has to be applied to derive Y^n from (5.11).

We state the main result of this chapter, see Theorem 1 and Proposition 5 in Lejay et al. (2010).

Theorem 5.2.1 *Consider the FBSDE* (5.7) *and the random measure N generated by a Poisson process. Assume that* (B2)–(B5) *from Sect.* 4.1 *hold and let the generator f satisfy*

$$\begin{aligned} &\big| f(t, x, y, z, u) - f\big(t', x', y', z', u'\big)\big| \\ &\quad \le \varphi\big(t' - t\big) + K\big(\big|x - x'\big| + \big|y - y'\big| + \big|z - z'\big| + \big|u - u'\big|\big), \end{aligned}$$

for all $(t, x, y, z, u), (t', x', y', z', u') \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, where φ is a bounded, non-decreasing, continuous function such that $\varphi(0) = 0$. We deal with the approximations (5.9) *and* (5.11) *of the solution to the FBSDE* (5.7).

(a) *The process $\mathscr{X}^n$ converges in the J_1-Skorokhod topology in probability to $\mathscr{X}$ as $n \to \infty$.*

(b) *The processes* $(Y^n, \int_0^{\cdot} Z^n(s)ds, \int_0^{\cdot} U^n(s)ds)$ *converge in the* J_1*-Skorokhod topology in probability to* $(Y, \int_0^{\cdot} Z(s)ds, \int_0^{\cdot} U(s)ds)$ *as* $n \to \infty$.

The efficiency of the algorithm is studied numerically in Lejay et al. (2010).

In the discrete filtration $\mathscr{F}^n$ it is straightforward to calculate the conditional expected values in (5.11). This is the key advantage of the approximation by discrete-space martingales compared to the Least Squares Monte Carlo method. We can use the formula

$$\begin{aligned}
&\mathbb{E}\big[F\big(\xi_1^n, \dots, \xi_k^n, \xi_{k+1}^n, \zeta_1^n, \dots, \zeta_k^n, \zeta_{k+1}^n\big)|\mathscr{F}_k^n\big] \\
&\quad = F\big(\xi_1^n, \dots, \xi_k^n, 1, \zeta_1^n, \dots, \zeta_k^n, e^{-\lambda h} - 1\big)\frac{e^{-\lambda h}}{2} \\
&\qquad + F\big(\xi_1^n, \dots, \xi_k^n, -1, \zeta_1^n, \dots, \zeta_k^n, e^{-\lambda h} - 1\big)\frac{e^{-\lambda h}}{2} \\
&\qquad + F\big(\xi_1^n, \dots, \xi_k^n, 1, \zeta_1^n, \dots, \zeta_k^n, e^{-\lambda h}\big)\frac{1 - e^{-\lambda h}}{2} \\
&\qquad + F\big(\xi_1^n, \dots, \xi_k^n, -1, \zeta_1^n, \dots, \zeta_k^n, e^{-\lambda h}\big)\frac{1 - e^{-\lambda h}}{2}.
\end{aligned}$$

In a low dimension the random walk approximation can provide a numerically efficient alternative to the Monte Carlo simulation. However, complexity becomes very large in multidimensional problems.

5.3 Finite Difference Method

In Sect. 4.2 we establish the connection between the solution to a FBSDE and the solution to a PIDE. The results of that chapter show that in order to derive the solution to the BSDE (4.5) or (5.1) we can solve the PIDE

$$\begin{aligned}
&-u_t(t,x) - \mathscr{L}u(t,x) \\
&\quad - f\big(t, x, u(t,x), u_x(t,x)\sigma(x), \mathscr{J}u(t,x)\big) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}, \quad (5.12) \\
&u(T,x) = g(x), \quad x \in \mathbb{R}.
\end{aligned}$$

We can apply a finite difference method to solve (5.12).

Let the random measure N be generated by a compound Poisson process. In order to construct a finite difference scheme for (5.12), we have to consider the following steps:

- Localization: the PIDE (5.12) is given on the unbounded domain $\mathbb{R}$. We reduce the original domain to a bounded domain $[-A, A]$ and we impose boundary conditions. The domain of the integral in the operator $\mathscr{J}$ is localized to $[-B, B]$.

- Discretization in space and time: we choose discrete grids $t_k = \frac{T}{n}k, k = 0, 1, \ldots, n$, and $x_i = -A + \frac{2A}{m}i, i = 0, 1, \ldots, m$.
- Approximation of the derivatives: we use finite differences.
- Approximation of the integral in the operator $\mathscr{J}$: we use the trapezoidal quadrature rule.

If we deal with a Poisson random measure N with an infinite Lévy measure, then an additional step is needed to approximate small jumps of a Lévy process by an independent Brownian motion. Consequently, we end up with a random measure N generated by a compound Poisson process.

Using the results from Sect. 12.4 in Cont and Tankov (2004), we can state the following explicit-implicit scheme for solving the PIDE (5.12):

1. Choose n and m which define the spatial and time grid steps: $\Delta t = \frac{T}{n}$ and $\Delta x = \frac{2A}{m}$,
2. Set $u^{n,m}(t_n, x_i) = g(x_i)$ and extend the grid values to all $x \in [-A, A]$ by linear interpolation,
3. Going backward, for $k = n-1, \ldots, 0$ determine the grid values $u(t_k, x_i)$ by solving the difference equation

$$
\begin{aligned}
0 = {} & \frac{u^{n,m}(t_{k+1}, x_i) - u^{n,m}(t_k, x_i)}{\Delta t} + \Bigg(\mu(x_i) \\
& - \sum_{j=0}^{m} \gamma(x_i, z_j)\nu\big((z_j - 1/2, z_j + 1/2]\big)\Bigg) \frac{u^{n,m}(t_k, x_{i+1}) - u^{n,m}(t_k, x_i)}{\Delta x} \\
& + \frac{1}{2}\sigma^2(x_i) \frac{u^{n,m}(t_k, x_{i+1}) - 2u^{n,m}(t_k, x_i) + u^{n,m}(t_k, x_{i-1})}{|\Delta x|^2} \\
& + \sum_{j=0}^{m} \big(u^{n,m}\big(t_{k+1}, x_i + \gamma(x_i, z_j)\big) - u^{n,m}(t_{k+1}, x_i)\big)\nu\big((z_j - 1/2, z_j + 1/2]\big) \\
& + f\Bigg(t_{k+1}, x_i, u^{n,m}(t_{k+1}, x_i), \sigma(x_i)\frac{u^{n,m}(t_{k+1}, x_{i+1}) - u^{n,m}(t_{k+1}, x_i)}{\Delta x}, \\
& \sum_{j=0}^{m} \big(u^{n,m}\big(t_{k+1}, x_i + \gamma(x_i, z_j)\big) - u^{n,m}(t_{k+1}, x_i)\big) \\
& \cdot \delta(z_j)\nu\big((z_j - 1/2, z_j + 1/2]\big)\Bigg),
\end{aligned}
\tag{5.13}
$$

where $z_j = -B + \frac{2B}{m}j, j = 0, 1, \ldots, m$, and extend the grid values to all $x \in [-A, A]$ by linear interpolation.

The implicit scheme is used for the differential operator and the explicit scheme is used for the integral operator. Convergence of explicit-implicit schemes for PIDEs is discussed in Sect. 12.4 in Cont and Tankov (2004).

We point out that solving a PIDE by a finite difference method is efficient in low dimensions (when we deal with few state processes). Least Squares Monte Carlo algorithms perform much better than finite difference methods in high dimensions.

Since in actuarial and financial applications we deal with many risk factors and we consider multidimensional state processes, we can conclude that BSDEs and Monte Carlo methods are more efficient than PDEs (HJB equations) and finite difference methods in solving applied problems. It should be noticed that in many cases the solution to a problem does not involve all control processes of the BSDE and we do not have to estimate all expected values in the algorithms (5.4), (5.11), which simplifies numerical implementations of the algorithms.

Bibliographical Notes The Malliavin calculus plays an important role in proving convergence results for discrete-time approximations of BSDEs. Zhang (2004) was the first who applied the Malliavin calculus to prove path regularity of the solution and convergence of a discrete-time approximation under a deterministic regular time mesh. Bouchard and Elie (2008) followed the arguments from Zhang (2004) and showed path regularity and convergence for BSDEs with Poisson jumps. Various modifications of the Least Squares Monte-Carlo algorithm can be found in Gobet et al. (2006), Gobet and Lemor (2006). An alternative to the Least Square Monte Carlo is to apply Malliavin weights, see Bouchard et al. (2004). A comparison of the regression based approach, the Malliavin weights and the random walk approximation can be found in Bouchard and Warin (2010). Convergence results for discrete-time and martingale approximations to BSDEs driven by Brownian motions are investigated in Briand et al. (2002). In the case of a fully coupled BSDE driven by a Brownian motion, Douglas et al. (1996) provide a modification of a finite difference method from Sect. 5.3. Douglas et al. (1996) also prove convergence of an approximation of the derivative of the value function, which is needed to obtain the control process of a BSDE. Results on numerics for quadratic decoupled FBSDE can be found in Imkeller et al. (2010).