Journal of Stochastic Analysis

Volume 4 | Number 3 Article 3

September 2023

Backward Stochastic Differential Equations in a Semi-Markov Chain Model

Robert J. Elliott
University of Calgary, Calgary, AB, T2N 1N4, Canada, relliott@ucalgary.ca

Zhe Yang
University of Calgary, Calgary, AB, T2N 1N4, Canada, yangzhezhe@gmail.com

Journal of Stochastic Analysis
Vol. 4, No. 3 (2023) Article 3 (12 pages)
DOI: 10.31390/josa.4.3.03

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS IN A SEMI-MARKOV CHAIN MODEL

ROBERT J. ELLIOTT* AND ZHE YANG

Abstract. In this paper we expand the results of Cohen and Elliott [2] to semi-Markov chains following the framework of Elliott and Malcolm [5]. We discuss backward stochastic differential equations in discrete time in a semi-Markov Chain Model. Existence and uniqueness of solutions to BSDEs in discrete time in a semi-Markov Chain Model are obtained using a Martingale Representation Theorem. Also the dimension of the solution $Z_.$ changes. Finally we establish comparison results for solutions of two one-dimensional BSDEs in a semi-Markov Chain framework.

1. Introduction

In 1990 Pardoux and Peng [4] considered general backward stochastic differential equations (BSDEs for short) of the following form:

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s,\ t \in [0, T].$$

Here B is a d-dimensional Brownian Motion, ξ is an $\mathcal{F}_T$-measurable, $\mathbb{R}^m$-valued, square integrable random variable so the above equation has a unique $(\mathbb{R}^m \times \mathbb{R}^{m\times d})$-valued solution $(Y_., Z_.)$. Since then, research into BSDEs has attracted extensive attention. In Cohen and Elliott [1], a Markov chain has a semimartingale representation involving a vector martingale $M_. = \{M_t \in \mathbb{R}^N,\ t \geq 0\}$. BSDEs in that framework were introduced by Cohen and Elliott [1] as

$$Y_t = \xi + \int_{]t,T]} F(s, Y_s, Z_s)ds - \int_{]t,T]} [G(s-, Y_{s-}) + Z_{s-}]dM_s,\ t \in [0, T],$$

where $M_.$ is a martingale and ξ is an $\mathcal{F}_T$-measurable, $\mathbb{R}^K$-valued, square integrable random variable. Then under suitable conditions this equation has a unique $(\mathbb{R}^K \times \mathbb{R}^{K\times N})$-valued solution $(Y_., Z_.)$.

Elliott and Malcolm [5] in 2021 derived semimartingale dynamics for a semi-Markov chain and give them in a new vector form which explicitly exhibits the times at which jump-events occur and the probabilities of state transitions. However, a useful result in [5] is the new vector lattice state-space representation for

Received 2023-8-21; Accepted 2023-9-25; Communicated by the editors.
2020 *Mathematics Subject Classification.* Primary 60K15; Secondary 60J10.
Key words and phrases. BSDEs, discrete-time, Markov chain models, semi-Markov chain models.
* Corresponding author.

a general finite-state, discrete-time semi-Markov chain. On this space the semi-Markov chain and its occupation times are a Markov process with dynamics described by transition matrices. Following the framework of Elliott and Malcolm [5], we discuss BSDEs in a semi-Markov Chain Model and we expand the results of Cohen and Elliott [2] to semi-Markov chains. Moreover, we find the solution Z_i takes values in spaces whose dimension is increasing as the time i increases.

The present paper is structured as follows: Section 2 presents the model and some preliminary results. In Section 3 we obtain the Martingale Representation Theorem. We provide existence and uniqueness of the solution to BSDEs in discrete time in a semi-Markov Chain Model. The final section gives the comparison results for solutions of two one-dimensional BSDEs in a semi-Markov Chain Model.

2. The Model and Some Preliminary Results

Consider a discrete time, finite state space semi-Markov chain $X_. = \{X_k, k = 0, 1, 2, ...\}$. If the state space has $N \in \mathbb{N}$ elements, they can be identified with the set of unit vectors $S = \{e_1, e_2 \cdots, e_N\}$, where $e_i = (0, \cdots 0, 1, 0, \cdots, 0)' \in \mathbb{R}^N$ with 1 in the i-th position. Suppose the initial state $X_0 \in S$ is given or its probability distribution $p_0 = (p_0^1, p_0^2, ..., p_0^N) \in [0, 1]^N$ is known. The chain will change state at random discrete times τ_N. State transitions at these times are of the type $e_i \to e_j$, with $i \neq j$. Set $\tau_0 := 0$. Then we know successive jump event times form a strictly increasing sequence, that is, $\tau_0 < \tau_1 < \tau_2 < ....$ Set $\mathcal{F}_m := \sigma\{X_k,\ k \leq m\}$.

Now we give the definition of a time-homogeneous semi-Markov chain coming from Elliott and Malcolm [5].

Definition 2.1. The stochastic process $X_.$ is a time-homogeneous semi-Markov chain if, independently of n,

$$P(X_{\tau_{n+1}} = e_j, \tau_{n+1} - \tau_n = m \mid \mathcal{F}_{\tau_n}) = P(X_{\tau_{n+1}} = e_j, \tau_{n+1} - \tau_n = m \mid X_{\tau_n} = e_i).$$

If $X_{\tau_n} = e_i$, we write this as $q(e_j, e_i, m)$.

Set

$$f_{j,i}(m) := P(\tau_{n+1} - \tau_n = m \mid X_{\tau_{n+1}} = e_j, X_{\tau_n} = e_i) \text{ and}$$
$$p_{j,i} := P(X_{\tau_{n+1}} = e_j | X_{\tau_n} = e_i).$$

Consequently

$$\begin{aligned} q(e_j, e_i, m) &= P(\tau_{n+1} - \tau_n = m \mid X_{\tau_{n+1}} = e_j, X_{\tau_n} = e_i)P(X_{\tau_{n+1}} = e_j | X_{\tau_n} = e_i) \\ &= f_{j,i}(m)p_{j,i}. \end{aligned}$$

We can also consider the factorization: Set

$$\pi_i(m) := P(\tau_{n+1} - \tau_n = m \mid X_{\tau_n} = e_i) \text{ and}$$
$$p_{j,i}(m) := P(X_{\tau_{n+1}} = e_j \mid \tau_{n+1} - \tau_n = m, X_{\tau_n} = e_i).$$

Thus,

$$\begin{aligned}&q(e_j, e_i, m)\\ =&P(\tau_{n+1}-\tau_n = m \mid X_{\tau_n}=e_i)P(X_{\tau_{n+1}}=e_j \mid \tau_{n+1}-\tau_n=m, X_{\tau_n}=e_i)\\ =&\pi_i(m)p_{j,i}(m).\end{aligned}$$

Write

$$G_i(m) := P(\tau_{n+1}-\tau_n \le m \mid X_{\tau_n}=e_i) = \sum_{l=1}^{m}\pi_i(l),$$

$$F_i(m) := P(\tau_{n+1}-\tau_n > m \mid X_{\tau_n}=e_i) = 1 - G_i(m).$$

Denote by $\Delta^i(m)$ the conditional probability for a state-transition to occur at the next discrete time, that is,

$$\Delta^i(m) := \frac{\pi_i(m)}{F_i(m-1)}.$$

Definition 2.2. For each index i, $1 \le i \le N$, we define the recursive process $h^i_k := \langle X_k, e_i\rangle + \langle X_k, e_i\rangle\langle X_k, X_{k-1}\rangle h^i_{k-1}$, with $h^i_0 := \langle X_0, e_i\rangle \in \{0,1\}$. Thus, the $h^i_\cdot$ processes are non-zero only at times when $X_\cdot = e_i$. The process $h^i_\cdot$ returns the cumulative time spent in state e_i. If $h_k = \sum_{i=1}^{N} h^i_k$, then $h_0 = 1$ and $h_k = 1 + \langle X_k, X_{k-1}\rangle h_{k-1}$. The process h_k measures the amount of time since the last transition event. This process is never zero.

For $m = 1, 2, ...$, denote $A(m)$ for the $N \times N$ matrix with entries $a_{i,i}(m) = 1 - \Delta^i(m)$ and $a_{j,i}(m) = p_{j,i}(m)\Delta^i(m)$.

The following lemma comes from Elliott and Malcolm [5].

Lemma 2.3. *The semi-Markov chain $X_\cdot$ has the following semi-martingale dynamics:*

$$X_{k+1} = A(h_k)X_k + M_{k+1} \in \mathbb{R}^N.$$

Here M_{k+1} is a martingale increment: $E[M_{k+1}|X_k, h_k] = 0 \in \mathbb{R}^N$.

3. Martingale Representation Theorem

Recall Section 3 in Elliott and Malcolm [5], which represents the process $(X_\cdot, h_\cdot)$ as a Markov chain.

The state of the chain $X_k \in \{e_1, e_2, ..., e_N\}$ and the number of time steps h_k describe the semi-Markov chain $X_\cdot$. That is, a state space $\bar{S}$ for the chain $\bar{X}_k := (X_k, h_k)$, $k \in \{1, 2, ...\}$ can be identified with countably many copies of S as follows: Elements of $\bar{S}$ can be thought of as infinite column vectors:

$$(e_{l,i}) \text{ corresponds to } (0,0,...,0|\cdots|\underbrace{0,...,1,...,0}_{h_k=i}|\cdots|0,0,...,0|\cdots)',$$

with $e_l = (0, ..., 1, ..., 0)'$ in the i^{th} block, where $l = 1, 2, ..., N$, and $i = 0, 1, ...$ As a basis of unit vectors for the process $\bar{X}_\cdot = (X_\cdot, h_\cdot)$, we take the unit vectors $e_{l,i}$

with $l = 1, 2, ..., N$, and $i = 0, 1, ...$ Here l denotes the state e_l in $\{e_1, e_2, ..., e_N\}$ and the i corresponds to the sojourn time in the state e_l since the last jump. Write $\bar{S} = \{e_{l,i};\ l = 1, 2, ..., N,$ and $i = 0, 1, ...\}$ with $e_{l,i}$ is e_l in the i^{th} block. For any $m \in \{1, 2, ...\}$, set

$$\Pi(m) = \begin{pmatrix} 0 & p_{1,2}(m)\Delta^2(m) & \cdots & p_{1,N}(m)\Delta^N(m) \\ p_{2,1}(m)\Delta^1(m) & 0 & \cdots & p_{2,N}(m)\Delta^N(m) \\ \cdots & & & \\ p_{N,1}(m)\Delta^1(m) & p_{N,2}(m)\Delta^2(m) & \cdots & 0 \end{pmatrix}$$

and $D(m) = \text{diag}\{1 - \Delta^1(m), 1 - \Delta^2(m), ..., 1 - \Delta^N(m)\}$. With 0 representing the $N \times N$ zero matrix write

$$C = \begin{pmatrix} \Pi(1) & \Pi(2) & \Pi(3) & \cdots & \Pi(T) & \cdots \\ D(1) & 0 & 0 & \cdots & 0 & \cdots \\ 0 & D(2) & 0 & \cdots & 0 & \cdots \\ \cdots & & & & & \\ 0 & 0 & 0 & \cdots & D(T) & \cdots \\ \cdots & & & & & \end{pmatrix}.$$

Write the enlarged vectors as $\bar{X}_i$. Then following Elliott and Malcolm [5] the semimartingale dynamics of Markov chain can be written

$$\bar{X}_{i+1} = C\bar{X}_i + \bar{M}_{i+1} \in \bar{S}$$

with $E[\bar{M}_{i+1}|\bar{X}_i] = E[\bar{X}_{i+1}|\bar{X}_i] - C\bar{X}_i = 0$. When the time horizon is $T < +\infty$, the matrix C is finite, and the state space of $\bar{X}_.$ at time i only has $(i+1)N$ elements.

Definition 3.1. For any integer K, and any $i \in \{0, 1, ..., T-1\}$, we shall denote by $\|\cdot\|_{\bar{M}_{i+1}}$ the seminorm on the space of $\sigma(\bar{X}_i)$-measurable and $\mathbb{R}^{K\times((i+1)N)}$-valued random variables Z_i, given by

$$\begin{aligned} \|Z_i\|^2_{\bar{M}_{i+1}} &:= E\text{Tr}[\sum_{0\le u\le i} Z_u \cdot E[\bar{M}_{u+1}\bar{M}'_{u+1}|\bar{X}_u] \cdot Z'_u] \\ &= \sum_{0\le u\le i} \text{Tr}E[(Z_u\bar{M}_{u+1})(Z_u\bar{M}_{u+1})']. \end{aligned}$$

Lemma 3.2. *For any $i \in \{0, 1, ..., T-1\}$, the following statements are equivalent:*

(i) $\|Z_i^{(1)} - Z_i^{(2)}\|^2_{\bar{M}_{i+1}} = 0.$

(ii) $E\,Tr[(Z_i^{(1)} - Z_i^{(2)})\bar{M}_{i+1}\bar{M}'_{i+1}(Z_i^{(1)} - Z_i^{(2)})'] = 0.$

(iii) $Z_i^{(1)}\bar{M}_{i+1} = Z_i^{(2)}\bar{M}_{i+1}$, $\mathbb{P}$*-a.s.*

(iv) $\sum_{0\le u\le i} Z_u^{(1)}\bar{M}_{u+1} = \sum_{0\le u\le i} Z_u^{(2)}\bar{M}_{u+1}$, $\mathbb{P}$*-a.s.*

In this case we shall write $Z_i^{(1)} \sim_{\bar{M}_{i+1}} Z_i^{(2)}$.

Proof. $(i) \Longleftrightarrow (ii)$: As for any $i \in \{0, 1, ..., T-1\}$, $\mathrm{Tr}E[(Z_i\bar{M}_{i+1})(Z_i\bar{M}_{i+1})'] \geq 0$, we have for any $i \in \{0, 1, ..., T-1\}$,

$$\|Z_i^{(1)} - Z_i^{(2)}\|^2_{\bar{M}_{i+1}} = 0$$

$$\Longleftrightarrow \mathrm{Tr}E[((Z_u^{(1)} - Z_u^{(2)})\bar{M}_{u+1})((Z_u^{(1)} - Z_u^{(2)})\bar{M}_{u+1})'] = 0, \ u \in \{0, 1, ..., i\}.$$

Then the following holds: for any $i \in \{0, 1, ..., T-1\}$,

$$\|Z_i^{(1)} - Z_i^{(2)}\|^2_{\bar{M}_{i+1}} = 0 \Longleftrightarrow \quad E\mathrm{Tr}[(Z_i^{(1)} - Z_i^{(2)})\bar{M}_{i+1}\bar{M}'_{i+1}(Z_i^{(1)} - Z_i^{(2)})'] = 0.$$

$(ii) \Longleftrightarrow (iii)$: As a sum of squares, (ii) is true if and only if each term is zero, that is, for any $l \in \{1, 2, ..., N\}, \ i \in \{0, 1, ..., T-1\}$,

$$E[(e'_{l,i}(Z_i^{(1)} - Z_i^{(2)})\bar{M}_{i+1})^2] = 0.$$

Considering all components at once, this is equivalent to $Z_i^{(1)}\bar{M}_{i+1} = Z_i^{(2)}\bar{M}'_{i+1}$, a.s., for any $i \in \{0, 1, ..., T-1\}$. Therefore, $(ii) \Longleftrightarrow (iii)$ holds.

$(iii) \Longleftrightarrow (iv)$ Taking a sum over the values $0 = u \leq i$ in statement (iii) gives statement (iv). Taking the differences of statement (iv) at times i and $i-1$ gives statement (iii). □

Definition 3.3. For any two $\sigma(\bar{X}_{i-1})$-measurable random variables $Z_{i-1}^{(1)}$ and $Z_{i-1}^{(2)}$, we shall write $Z_{i-1}^{(1)} \sim_{\bar{M}_i} Z_{i-1}^{(2)}$ if $Z_{i-1}^{(1)}\bar{M}_i = Z_{i-1}^{(2)}\bar{M}_i$, $\mathbb{P}$-a.s.

Recall the following Martingale Representation Theorem (from [6]).

Lemma 3.4. *For any* $i \in \{1, 2, ...\}$, *any* $\{\sigma(\bar{X}_i)\}$*-measurable,* $\mathbb{R}^K$*-valued martingale* $L_.$, *we know there exists a* $\sigma(\bar{X}_i)$*-measurable* $\mathbb{R}^{K\times(N(i+1))}$*-valued random variable* Z_i *such that*

$$L_i = L_0 + \sum_{0\leq u<i} Z_u\bar{M}_{u+1}.$$

Here, this Z_i *is unique up to equivalence* $\sim_{\bar{M}_{i+1}}$.

Proof. Because for any $i \in \{1, 2, ...\}$, L_i is $\sigma(\bar{X}_i)$-measurable, from the Doob-Dynkin Lemma (see [7], P174), we know there exists a $\sigma(\bar{X}_{i-1})$-measurable function $g_i : \mathbb{R}^{N(i+1)} \to \mathbb{R}^K$ satisfying

$$L_i = L_{i-1} + g_i(\bar{X}_i).$$

Since $\bar{X}_i$ takes $(i+1)N$ possible values, we can denote them by $\{\bar{e}_j \in \mathbb{R}^{N(i+1)};$ $j = 1, 2, ..., (i+1)N\}$. Then we create a $\sigma(\bar{X}_{i-1})$-measurable $\mathbb{R}^{K\times(N(i+1))}$ matrix Z_{i-1} with entries

$$Z_{i-1} = (g_i(\bar{e}_1)|g_i(\bar{e}_2)|\cdots|g_i(\bar{e}_{N(i+1)}))$$

which will satisfy

$$L_i = L_{i-1} + Z_{i-1}\bar{X}_i.$$

As $\bar{X}_i = \bar{M}_i + C\bar{X}_{i-1}$, we have $L_i = L_{i-1} + Z_{i-1}\bar{M}_i + Z_{i-1}C\bar{X}_{i-1}$. Since $L.$ is a martingale, we deduce $E[L_i|\bar{X}_{i-1}] = L_{i-1}$. Hence

$$L_{i-1} = L_{i-1} + Z_{i-1}E[\bar{M}_i|\bar{X}_{i-1}] + Z_{i-1}C\bar{X}_{i-1}.$$

Because $E[\bar{M}_i|\bar{X}_{i-1}] = 0$, we deduce $Z_{i-1}C\bar{X}_{i-1} = 0$. Thus

$$L_i = L_{i-1} + Z_{i-1}\bar{M}_i.$$

Similarly, we have a sequence of equations: $L_{i-1} = L_{i-2} + Z_{i-2}\bar{M}_{i-1}$, $L_{i-2} = L_{i-3} + Z_{i-3}\bar{M}_{i-2}$,..., $L_1 = L_0 + Z_0\bar{M}_1$. So for any $i \in \{1, 2, ...\}$, the following holds:

$$L_i = L_0 + \sum_{0 \le u < i} Z_u\bar{M}_{u+1}.$$

If for any $i \in \{1, 2, ...\}$, there are two random variables $Z_i^{(1)}$ and $Z_i^{(2)}$ such that

$$L_i = L_0 + \sum_{0 \le u < i} Z_u^{(1)}\bar{M}_{u+1} = L_0 + \sum_{0 \le u < i} Z_u^{(2)}\bar{M}_{u+1}.$$

By Lemma 3.2 (iv), we have $Z_i^{(1)} \sim_{\bar{M}_{i+1}} Z_i^{(2)}$. □

Lemma 3.5. *For any $\sigma(\bar{X}_i)$-measurable, $\mathbb{R}^K$-valued random variable ξ satisfying $E[\xi|\bar{X}_{i-1}] = 0$, there exists a $\sigma(\bar{X}_{i-1})$-measurable, $\mathbb{R}^{K\times(Ni)}$-valued random variable Z_{i-1} such that, $\mathbb{P}$-a.s.,*

$$\xi = Z_{i-1}\bar{M}_i.$$

Here, this random variable Z_{i-1} is unique up to equivalence $\sim_{\bar{M}_i}$.

Proof. By Lemma 3.4, we deduce for any $m \in \{1, 2, ..., i\}$, there exists a $\sigma(\bar{X}_m)$-measurable $\mathbb{R}^{K\times(N(m+1))}$-valued random variable Z_m such that

$$E[\xi|\bar{X}_m] = E[\xi] + \sum_{0 \le u < m} Z_u\bar{M}_{u+1}.$$

So we have

$$\xi = E[\xi|\bar{X}_i] = E[\xi] + \sum_{u=0}^{i-1} Z_u\bar{M}_{u+1} \tag{3.1}$$

and

$$0 = E[\xi|\bar{X}_{i-1}] = E[\xi] + \sum_{u=0}^{i-2} Z_u\bar{M}_{u+1}. \tag{3.2}$$

Using Equations (3.1) and (3.2), we conclude

$$\xi = Z_{i-1}\bar{M}_i.$$

□

4. Existence and Uniqueness of the Solution

Although the dimension of $\bar{M}_i$ is increasing with i, the duality between the solutions to linear BSDEs and linear SDEs, adapted to our one-dimensional time-discrete case on semi-Markov chain noise, still holds by Theorem 2 in Cohen and Elliott [3] and Lemma 2.5 in Yang, Ramarimbahoaka and Elliott [8]. That is,

Lemma 4.1. *(Duality) Set $K = 1$ and for $i \in \{0,1,...,T-1\}$, set $\Psi_i := diag(C\bar{X}_i) - diag(\bar{X}_i)C' - C\,diag(\bar{X}_i)$. Let $(u_\cdot, v_\cdot, g_\cdot)$ be a $du \times P$-a.s. bounded $(\mathbb{R}, \mathbb{R}^{1\times(N(i+1))}, \mathbb{R})$-valued adapted process for any $i \in \{0,1,...,T-1\}$, and ξ be an essentially bounded, valued in $\mathbb{R}$, $\sigma(\bar{X}_T)$ -measurable random variable. Then the linear BSDE given by*

$$\begin{cases} Y_i = \xi + \sum\limits_{k=i}^{T-1}(u_k Y_k + v_k Z_k' + g_k) - \sum\limits_{k=i}^{T-1} Z_k \bar{M}_{k+1}, \ i \in \{0,1,...,T-1\}; \\ Y_T = \xi \end{cases}$$

has an adapted solution $(Y_\cdot, Z_\cdot)$. Here for any $i \in \{0,1,...,T-1\}$, we have (Y_i, Z_i) is a $\sigma(\bar{X}_i)$-measurable and $\mathbb{R}\times\mathbb{R}^{1\times((i+1)N)}$-valued random variable, (up to appropriate sets of measure zero). Moreover, if for all $k \in \{i, i+1, ..., T-1\}$,

$$1 + v_k \Psi_k^\dagger (e_{l,i} - \bar{X}_k)$$

is non-zero (invertible for the multi-dimensional case) and non-negative for all l, i such that $e_{l,i}' C\bar{X}_k > 0$, except possibly on some evanescent set, where $(\cdot)^\dagger$ means the Moore-Penrose pseudoinverse of a matrix, then $Y_\cdot$ is given by the explicit formula

$$Y_i = E[\xi U_T + \sum_{k=i}^{T-1} g_k U_k | \bar{X}_i], \ i \in \{0,1,...,T-1\}$$

up to indistinguishability. Here $U_\cdot$ is the solution of the following one-dimensional SDE:

$$\begin{cases} U_i = 1 + \sum\limits_{k=t+1}^{i} U_k u_k + \sum\limits_{k=t+1}^{i} U_k v_k (\Psi_k^\dagger)' \bar{M}_k, \ i \in \{t+1, ..., T\}; \\ U_t = 1. \end{cases}$$

Theorem 4.2. *Suppose f is such that the following two assumptions hold:*
(i) For any $i \in \{0,1,...,T-1\}$ and $Y_\cdot$, if $Z_i^{(1)} \sim_{\bar{M}_{i+1}} Z_i^{(2)}$, then $f(i, Y_i, Z_i^{(1)}) = f(i, Y_i, Z_i^{(2)})$, $\mathbb{P}$-a.s.
(ii) For any $i \in \{0,1,...,T-1\}$ and $Z_\cdot$, the map $Y_i \mapsto Y_i - f(i, Y_i, Z_i)$ is $\mathbb{P}$-a.s. a bijection from $\mathbb{R}^K \to \mathbb{R}^K$, up to equality $\mathbb{P}$-a.s.
Then for any terminal condition Q essentially bounded, $\sigma(\bar{X}_T)$ -measurable, and with values in $\mathbb{R}^K$, BSDE

$$\begin{cases} Y_i = Q + \sum\limits_{u=i}^{T-1} f(u, Y_u, Z_u) - \sum\limits_{u=i}^{T-1} Z_u \bar{M}_{u+1}, \ i \in \{0,1,...,T-1\}; \\ Y_T = Q \end{cases} \tag{4.1}$$

has an adapted solution $(Y_\cdot, Z_\cdot)$. Here for any $i \in \{0,1,...,T-1\}$, we have (Y_i, Z_i) is a $\sigma(\bar{X}_i)$-measurable and $\mathbb{R}^K \times \mathbb{R}^{K\times((i+1)N)}$-valued random variable. Moreover, *this solution is unique up to indistinguishability for $Y_\cdot$ and equivalence $\sim_{\bar{M}_{i+1}}$ for Z_i.*

Proof. When $i = T-1$, we wish to find a solution (Y_{T-1}, Z_{T-1}). Then equation (4.1) can be simplified to

$$Y_{T-1} = Q + f(T-1, Y_{T-1}, Z_{T-1}) - Z_{T-1}\bar{M}_T.$$

Taking a conditional expectation gives

$$Y_{T-1} = E[Q|\bar{X}_{T-1}] + f(T-1, Y_{T-1}, Z_{T-1}).$$

Noting $Q - E[Q|\bar{X}_{T-1}]$ is a martingale difference term, by Lemma 3.5, there is a random variable Z_{T-1} such that $Q - E[Q|\bar{X}_{T-1}] = Z_{T-1}\bar{M}_T$ and this Z_{T-1} is unique up to equivalence $\sim_{\bar{M}_T}$. Using this Z_{T-1}, consider the equation

$$Y_{T-1} = E[Q|\bar{X}_{T-1}] + f(T-1, Y_{T-1}, Z_{T-1}).$$

By assumption (ii), the above equation has a unique solution Y_{T-1}, up to equality $\mathbb{P}$-a.s. The pair (Y_{T-1}, Z_{T-1}) is the solution of equation (4.1) at time $T-1$. Similarly we have (Y_i, Z_i) is the solution of equation (4.1) at time i, for any $i = T-2, T-3, ..., 1, 0$. The solution Y_i is unique up to equality $\mathbb{P}$-a.s., for all $i = 0, 1, ..., T-1$, and hence $Y_\cdot$ is unique up to indistinguishability (as we are working in discrete time), and the solution Z_i is unique up to equivalence $\sim_{\bar{M}_{i+1}}$, for any $i = 0, 1, ..., T-1$. □

5. Comparison Theorem

In this section, we only consider the case $K = 1$. Let $(Y_\cdot^{(1)}, Z_\cdot^{(1)})$ and $(Y_\cdot^{(2)}, Z_\cdot^{(2)})$ be the solutions of the following two BSDEs with semi-Markov chain noise, respectively:

$$\begin{cases} Y_i^{(1)} = Q_1 + \sum\limits_{u=i}^{T-1} f_1(u, Y_u^{(1)}, Z_u^{(1)}) - \sum\limits_{u=i}^{T-1} Z_u^{(1)}\bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T^{(1)} = Q_1 \end{cases}$$

and

$$\begin{cases} Y_i^{(2)} = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u^{(2)}, Z_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u^{(2)}\bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T^{(2)} = Q_2. \end{cases}$$

A comparison theorem for one-dimensional BSDEs with Markov chain noise in Yang, Ramarimbahoaka and Elliott [8] (Theorem 3.2) also extends to our one-dimensional time-discrete case with semi-Markov chain noise:

Theorem 5.1. *Assume Q_1, Q_2 are essentially bounded, $\sigma(\bar{X}_T)$ -measurable, and with values in $\mathbb{R}$, and f_1, f_2 satisfy conditions such that the above BSDEs have unique solutions. Suppose the following conditions hold:*

(I) $Q_1 \le Q_2$, $\mathbb{P}$-a.s.

(II) for any $i \in \{0, 1, ..., T-1\}$,

$$f_1(i, Y_i^{(2)}, Z_i^{(2)}) \le f_2(i, Y_i^{(2)}, Z_i^{(2)}), \ \mathbb{P}\text{-a.s.}$$

(III) there exist two constant $\omega_1, \omega_2 > 0$ the following holds: for each $i \in \{0, 1, ..., T-1\}$, $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{R}^{1\times((i+1)N)}$,

$$|f(i, y_1, z_1) - f(i, y_2, z_2)| \le \omega_1 |y_1 - y_2| + \omega_2((z_1 - z_2)'\Psi_i(z_1 - z_2))^{\frac{1}{2}}.$$

Moreover, ω_2 satisfies

$$6\omega_2^2(Tr(C'C))^{\frac{1}{2}} Tr((\Psi_i^{\dagger})'\Psi_i^{\dagger}) < 1.$$

Then $Y_i^{(1)} \le Y_i^{(2)}$, for any $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s.

However, we use another method to obtain the following comparison result:

Theorem 5.2. *Assume Q_1, Q_2 are essentially bounded, $\sigma(\bar{X}_T)$-measurable, and with values in $\mathbb{R}$, $f_2(i, \cdot, Z_i^{(2)})$ is essentially bounded for any $i \in \{0, 1, ..., T-1\}$, and f_1, f_2 satisfy $(i), (ii)$ in Theorem 4.2. Suppose the following conditions hold:*

(I) $Q_1 \le Q_2$, $\mathbb{P}$-a.s.

(II) for any $i \in \{0, 1, ..., T-1\}$, $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\times((i+1)N)}$,

$$f_1(i, y, z) \le f_2(i, y, Z_i^{(2)}),\ \mathbb{P}\text{-a.s.}$$

(III) if $y_1 \le y_2$, the following holds: for any $i \in \{0, 1, ..., T-1\}$, $z \in \mathbb{R}^{1\times((i+1)N)}$,

$$f_2(i, y_1, z) \le f_2(i, y_2, z),\ \mathbb{P}\text{-a.s.}$$

(IV) there exists a constant $\mu > 0$ such that for any $i \in \{0, 1, ..., T-1\}$, $y_1, y_2 \in \mathbb{R}$, $z \in \mathbb{R}^{1\times((i+1)N)}$,

$$|f_2(i, y_1, z) - f_2(i, y_2, z)| \le \mu |y_1 - y_2|.$$

Then $Y_i^{(1)} \le Y_i^{(2)}$, for any $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s.

Proof. From Theorem 4.2, we know the following equation

$$\begin{cases} Y_i^{(3)} = Q_2 + \sum_{u=i}^{T-1} f_2(u, Y_u^{(1)}, Z_u^{(2)}) - \sum_{u=i}^{T-1} Z_u^{(3)} \bar{M}_{u+1},\ i \in \{0, 1, ..., T-1\}; \\ Y_T^{(3)} = Q_2 \end{cases}$$

has a solution $(Y_\cdot^{(3)}, Z_\cdot^{(3)})$. Set $\epsilon := \inf\{q \ge 0 :\ |Q_2| \le q,\ \mathbb{P}\text{-a.s.}\}$ and for any $i \in \{0, 1, ..., T-1\}$, set

$$\rho_i := \inf\{p \ge 0 :\ |f_2(i, y, Z_i^{(2)})| \le p,\ \mathbb{P}\text{-a.s., for any } y \in \mathbb{R}\}.$$

Denote $\rho = \max\{\rho_i :\ i = 0, 1, ..., T-1\}$. Then we have $|Y_i^{(3)}| \le \epsilon + T\rho$, for any $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s. On the other hand, we know

$$\begin{aligned} Y_i^{(3)} &= E[Q_2 + \sum_{u=i}^{T-1} f_2(u, Y_u^{(1)}, Z_u^{(2)}) | \bar{X}_i] \\ &\ge E[Q_1 + \sum_{u=i}^{T-1} f_1(u, Y_u^{(1)}, Z_u^{(1)}) | \bar{X}_i] \\ &= Y_i^{(1)},\ i \in \{0, 1, ..., T-1\},\ \mathbb{P}\text{-a.s.}. \end{aligned}$$

Also from Theorem 4.2, we know equation

$$\begin{cases} Y_i^{(4)} = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u^{(3)}, Z_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u^{(4)} \bar{M}_{u+1},\ i \in \{0,1,...,T-1\}; \\ Y_T^{(4)} = Q_2 \end{cases}$$

has a solution $(Y_\cdot^{(4)}, Z_\cdot^{(4)})$. Moreover, $|Y_i^{(4)}| \leq \epsilon + T\rho$, for any $i \in \{0,1,...,T-1\}$, $\mathbb{P}$-a.s. By (III) we have $Y_i^{(3)} \leq Y_i^{(4)},\ i \in \{0,1,...,T-1\}$, $\mathbb{P}$-a.s. Similarly we know for any $n = 5,6,...$, the solution $(Y_\cdot^{(n)}, Z_\cdot^{(n)})$ of the equation

$$\begin{cases} Y_i^{(n)} = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u^{(n-1)}, Z_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u^{(n)} \bar{M}_{u+1},\ i \in \{0,1,...,T-1\}; \\ Y_T^{(n)} = Q_2 \end{cases}$$

satisfies

$$Y_i^{(1)} \leq Y_i^{(3)} \leq Y_i^{(4)} \leq Y_i^{(5)} \leq ... \leq Y_i^{(n)} \leq Y_i^{(n+1)} \leq ...,$$

for any $i \in \{0,1,...,T-1\}$, $\mathbb{P}$-a.s. Moreover, for any $n = 3,4,...$, $|Y_i^{(n)}| \leq \epsilon + T\rho$, for any $i = 0,1,...,T-1$, $\mathbb{P}$-a.s. This is because: For $n = 3,4,...$, we define the set $A_n := \{Y_i^{(n)} > \epsilon + T\rho$, for which there exists an $i = 0,1,...,T-1\}$. Hence,

$$P(\bigcup_{n=3}^{+\infty} A_n) \leq \sum_{n=3}^{+\infty} P(A_n) = 0.$$

Now define $\bar{Y}_i := \lim\limits_{n\to\infty} Y_i^{(n)},\ i \in \{0,1,...,T-1\}$. Thus, $Y_i^{(1)} \leq \bar{Y}_i$ and $|\bar{Y}_i| \leq \epsilon + T\rho$, for any $i \in \{0,1,...,T-1\}$, $\mathbb{P}$-a.s. By the Lebesgue Dominated Convergence Theorem, we get $\lim\limits_{n\to\infty} E[|Y_i^{(n)} - \bar{Y}_i|] = 0,\ i \in \{0,1,...,T-1\}$. Condition (IV) implies for any $i \in \{0,1,...,T-1\}$,

$$0 \leq E[|f_2(i, \bar{Y}_i, Z_i^{(2)}) - f_2(i, Y_i^{(n)}, Z_i^{(2)})|] \leq \mu E[|\bar{Y}_i - Y_i^{(n)}|] \to 0, \quad \text{when } n \to \infty.$$

Hence, we conclude

$$\begin{aligned} \bar{Y}_i &= \lim_{n\to\infty} Y_i^{(n)} \\ &= \lim_{n\to\infty} E[Q_2 + \sum_{u=i}^{T-1} f_2(u, Y_u^{(n-1)}, Z_u^{(2)}) | \bar{X}_i] \\ &= E[Q_2 + \sum_{u=i}^{T-1} f_2(u, \bar{Y}_u, Z_u^{(2)}) | \bar{X}_i],\ i \in \{0,1,...,T-1\},\ \mathbb{P}\text{-a.s.} \end{aligned}$$

Consider the case $i = T-1$. Thus, $\bar{Y}_{T-1} = E[Q_2|\bar{X}_{T-1}] + f_2(T-1, \bar{Y}_{T-1}, Z_{T-1}^{(2)})$. From Lemma 3.5, there is a $\bar{Z}_{T-1}$ such that $Q_2 - E[Q_2|\bar{X}_{T-1}] = \bar{Z}_{T-1}\bar{M}_T$ and this $\bar{Z}_{T-1}$ is unique up to equivalence $\sim_{\bar{M}_T}$. So

$$\bar{Y}_{T-1} = Q_2 + f_2(T-1, \bar{Y}_{T-1}, Z_{T-1}^{(2)}) - \bar{Z}_{T-1}\bar{M}_T.$$

Similarly we have $(\bar{Y}_i, \bar{Z}_i)$ is the solution of equation

$$\bar{Y}_i = \bar{Y}_{i+1} + f_2(i, \bar{Y}_i, Z_i^{(2)}) - \bar{Z}_i \bar{M}_{i+1}$$

at time i, for any $i = T-2, T-3, ..., 1, 0$. The solution $\bar{Y}_i$ is unique up to equality $\mathbb{P}$-a.s., for all $i = 0, 1, ..., T-1$, and hence $\bar{Y}_\cdot$ is unique up to indistinguishability (as we are working in discrete time), moreover, the solution $\bar{Z}_i$ is unique up to equivalence $\sim_{\bar{M}_{i+1}}$ for any $i = 0, 1, ..., T-1$. That means $(\bar{Y}_\cdot, \bar{Z}_\cdot)$ is the solution of the following equation

$$\begin{cases} Y_i = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u, Z_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u \bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T = Q_2. \end{cases}$$

Noting $(Y_\cdot^{(2)}, Z_\cdot^{(2)})$ also satisfies

$$\begin{cases} Y_i = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u, Z_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u \bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T = Q_2, \end{cases}$$

by the uniqueness of the solution (Theorem 4.2), we conclude $\bar{Y}_i = Y_i^{(2)}$, for $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s. So $Y_i^{(1)} \leq Y_i^{(2)}$, for $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s. □

Corollary 5.3. *Consider the following two BSDEs with semi-Markov chain noise, respectively:*

$$\begin{cases} Y_i^{(1)} = Q_1 + \sum\limits_{u=i}^{T-1} f_1(u, Y_u^{(1)}, Z_u^{(1)}) - \sum\limits_{u=i}^{T-1} Z_u^{(1)} \bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T^{(1)} = Q_1 \end{cases}$$

and

$$\begin{cases} Y_i^{(2)} = Q_2 + \sum\limits_{u=i}^{T-1} f_2(u, Y_u^{(2)}) - \sum\limits_{u=i}^{T-1} Z_u^{(2)} \bar{M}_{u+1}, \ i \in \{0, 1, ..., T-1\}; \\ Y_T^{(2)} = Q_2. \end{cases}$$

Assume Q_1, Q_2 are essentially bounded, $\sigma(\bar{X}_T)$ -measurable, and with values in $\mathbb{R}$, $f_2(i, \cdot)$ is essentially bounded for any $i \in \{0, 1, ..., T-1\}$, and f_1, f_2 satisfy $(i), (ii)$ in Theorem 4.2. Suppose the following conditions hold:

(I) $Q_1 \leq Q_2$, $\mathbb{P}$-a.s.

(II) for any $i \in \{0, 1, ..., T-1\}$, $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\times((i+1)N)}$,

$$f_1(i, y, z) \leq f_2(i, y), \ \mathbb{P}\text{-a.s.}$$

(III) if $y_1 \leq y_2$, the following holds: for any $i \in \{0, 1, ..., T-1\}$,

$$f_2(i, y_1) \leq f_2(i, y_2), \ \mathbb{P}\text{-a.s.}$$

(IV) there exists a constant $\mu > 0$ such that for any $i \in \{0, 1, ..., T-1\}$, $y_1, y_2 \in \mathbb{R}$,

$$|f_2(i, y_1) - f_2(i, y_2)| \leq \mu |y_1 - y_2|.$$

Then $Y_i^{(1)} \leq Y_i^{(2)}$, for any $i \in \{0, 1, ..., T-1\}$, $\mathbb{P}$-a.s.

Conclusion. Results for BSDEs with noise given by a finite state Markov chains are extended to semi-Markov chains. Existence and uniqueness results and a new comparison theorem are established.

Acknowledgment. The authors thank the Natural Sciences and Engineering Research Council of Canada for continuing support.

References

1. Cohen, S. N. and Elliott, R. J.: Solutions of Backward Stochastic Differential Equations in Markov Chains, *Communications on Stochastic Analysis* **2** (2008), no. 1, 251–262.
2. Cohen, S. N. and Elliott, R. J.: A general theory of finite state backward stochastic difference equations, *Stochastic Processes and their Applications* **120** (2010), 442–466.
3. Cohen, S. N. and Elliott, R. J.: Comparison Theorems for Finite State Backward Stochastic Differential Equations, *in Contemporary Quantitative Finance*, Springer, 2010.
4. Pardoux, E. and Peng, S.: Adapted solution of a backward differential equation, *Systems Controls Lett.* **14** (1990), 61–74.
5. Elliott, R. J. and Malcolm, W. P.: New representations for a semi-Markov chain and related filters, *Journal of Stochastic Analysis* **2** (March 2021), no. 1, 1–16.
6. Elliott, R. J. and Yang, H.: How to count and guess well: Discrete adaptive filters, *Applied Mathematics and Optimization* **30** (1994), no. 1, 51–78.
7. Shiryaev, A. N.: *Probability, 2nd ed.*, Springer, 2000.
8. Yang, Z., Ramarimbahoaka, D. and Elliott, R. J.: Comparison and converse comparison theorems for backward stochastic differential equations with Markov chain noise, *Electron. Commun. Probab.* **21** (2016), no. 25, 1–10.

Robert J. Elliott: Haskayne School of Business, University of Calgary, Calgary, AB, T2N 1N4, Canada; Business, University of South Australia, Adelaide, Australia 5001

Email address: relliott@ucalgary.ca

Zhe Yang: Haskayne School of Business, University of Calgary, Calgary, AB, T2N 1N4, Canada

Email address: yangzhezhe@gmail.com