MAXIMUM PRINCIPLE FOR DISCRETE-TIME STOCHASTIC OPTIMAL CONTROL PROBLEM AND STOCHASTIC GAME

ZHEN WU
School of Mathematics
Shandong University
Jinan 250100, Shandong Province, China

FENG ZHANG*
School of Mathematics and Quantitative Economics
Shandong University of Finance and Economics
Jinan 250014, Shandong Province, China

(Communicated by Qi Lü)

Abstract. This paper is first concerned with one kind of discrete-time stochastic optimal control problem with convex control domains, for which necessary condition in the form of Pontryagin’s maximum principle and sufficient condition of optimality are derived. The results are then extended to two kinds of discrete-time stochastic games. Two illustrative examples are studied, for which the explicit optimal strategies are given. This paper establishes a rigorous version of discrete-time stochastic maximum principle in a clear and concise way and paves a road for further related topics.

1. Introduction. As a milestone of optimal control theory, the maximum principle (MP) method is to derive necessary conditions that must be satisfied by any optimal solution. Deterministic MP was established by Pontryagin’s group. The general stochastic MP was established in Peng [22]. For other results on stochastic MP, see e.g. [7], [10], [17], [27], [28] and [30]. Game theory has been an active research area. The study of differential games could be traced to Isaacs [12]. The differential game theory was extended to the stochastic case in 1960s. MP for stochastic differential games can be seen in [1], [3], [11] and [25].

These mentioned results are mainly concentrated on the continuous-time problems. It’s worth pointing out that the discrete-time case is as important as the continuous-time one. In many situations it is sufficient or natural to describe a system by a discrete-time model. For example, the signal values are sometimes only available for measurement or manipulation at certain times, and discretization of the dynamics of continuous-time problems is needed when one wants to simulate or compute the continuous-time one after obtaining its dynamical characteristics. The discrete-time systems are usually described by difference equations. Difference equations are in their own right important mathematical models.

2020 Mathematics Subject Classification. Primary: 93E20, 93C55; Secondary: 91A15.
Key words and phrases. Maximum principle, stochastic optimal control, stochastic game, discrete-time system, optimal investment/consumption choice.

* Corresponding author: Feng Zhang.
MP for discrete-time deterministic optimal control problems could be found in [9] and [21]. Discrete-time stochastic linear quadratic (LQ) problems were studied in terms of Riccati-type difference equations (see e.g. [23], [2], [8], [18] and [19]). For other discrete-time stochastic control problems, see e.g. [6], [14], [15] and [29]. The study of discrete-time LQ games can be seen in [20] and [24].

From the authors’ viewpoint, some difficulties may be encountered in the study of discrete-time stochastic systems. The differentiation operation is not workable for discrete-time system, and therefore, Itô’s formula, which plays a key role in stochastic calculus as well as in the study of continuous-time stochastic MP, no longer works now. Besides, different from the classical Itô’s integral, the product terms involving the noises could not naturally admit good properties such as zero-mean and isometry, for example $X_kW_k$ and $u_kW_k$. This will also affect the integrability of the solution to the adjoint equation. To our best knowledge, only a few literatures have discussed discrete-time stochastic MP; see e.g. [13], [16] and [26]. The paper [16] derived the necessary condition and sufficient condition for the stochastic optimal controls. One kind of backward stochastic difference equation was introduced there for the first time to work as the adjoint equation. However, there are some deficiencies in [16]. In a recent paper [13], maximum principle for stochastic optimal control problem of forward-backward stochastic difference systems was studied. However, the results there cannot cover those in our present paper.

Based on the above concerns, the main purpose of this paper is to construct a strict mathematical framework for one kind of discrete-time stochastic optimal control problem and obtain a rigorous MP in a clear and concise way. We consider the convex control domains case. We overcome a difficulty arising from the lack of necessary integrability of the solution to the adjoint equation, by supposing higher yet relatively necessary integrability assumptions on the noises and the admissible controls. The results are also extended to two kinds of discrete-time stochastic games: a nonzero-sum game and a zero-sum one. Necessary as well as sufficient conditions are obtained for the equilibrium point of the nonzero-sum game and the saddle point of the zero-sum one. Two examples are presented to show the applications of the theoretical results. In a discrete-time investment/consumption choice problem, we derive the optimal investment and consumption strategy. For a discrete-time LQ nonzero-sum stochastic game, the equilibrium point is obtained by using the MP method and without introducing the Riccati equation. Our results for the discrete-time stochastic games, to the authors’ best knowledge, are new in the literature.

The rest of this paper is organized as follows. In Section 2, we study the MP for the discrete-time stochastic optimal control problem. Section 3 studies the MP for two kinds of discrete-time stochastic games. Two examples are studied in Section 4. In the last section, we give some comparisons with two closely related papers and show some concluding remarks.

2. Maximum principle for one kind of discrete-time stochastic optimal control problem.

2.1. Preliminaries and formulation of the problem. For $n \geq 1$, denote by $\mathbb{R}^n$ the $n$-dimensional Euclidean space with the usual norm $|\cdot|$ and inner product $\langle \cdot, \cdot \rangle$. With a slight abuse of notations, $|\cdot|$ is also used for the norm of a matrix. Use $A^\top$ for the transpose of any matrix $A$. Denote by $C$ a constant which can differ from line to line and use $C(\theta)$ to show its dependence on some parameter $\theta$.\n
}\n\n\}
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and \(E\) the expectation. Let \(N\) be a sufficiently large natural number, and denote \(\mathcal{T} = \{0, 1, \cdots, N - 1\}, \mathcal{T} = \{1, 2, \cdots, N\}\) and \(\mathcal{F} = \{0, 1, \cdots, N\}\). Assume that \(\{W_k, k \in \mathcal{T}\}\) is a sequence of \(\mathbb{R}^d\)-valued random variables (r.v.), with \(W_k = (W_k^1, \cdots, W_k^d)^\top\) for each \(k \in \mathcal{T}\). \(\{W_k, k \in \mathcal{T}\}\) is regarded as a sequence of noises. Define \(\mathcal{F}_0, \mathcal{F}_1, \cdots, \mathcal{F}_N \subset \mathcal{F}\) by \(\mathcal{F}_0 = \{\emptyset, \Omega\}\) and \(\mathcal{F}_k = \sigma\{W_0, W_1, \cdots, W_{k-1}\}\) for \(k \in \mathcal{T}\). Let us assume
\[
\mathbb{E}[|W_k|^\alpha |\mathcal{F}_k] \leq C(\alpha) < \infty, \quad \forall k \in \mathcal{T},
\]
where \(\alpha\) is a constant such that \(\alpha > 2\) and \(C(\alpha)\) is a positive constant depending only on \(\alpha\). Note that it is not assumed that \(W_k\) is independent of \(\mathcal{F}_k\). By \(\xi \in L^1(\Omega, \mathcal{F}_k, \mathbb{R}^n)\) with \(\lambda > 0\) and \(k \in \mathcal{T}\) we mean that \(\xi\) is an \(\mathcal{F}_k\)-measurable \(\mathbb{R}^n\)-valued r.v. satisfying \(\mathbb{E}[|\xi|^\lambda] < \infty\). When \(n = 1\), we use \(L^1(\Omega, \mathcal{F}_k)\) or even \(L^\lambda\) if no confusion occurs.

Given \(U_0, U_1, \cdots, U_{N-1}\) a sequence of nonempty convex subsets of \(\mathbb{R}^m\). We call \(u = \{u_k, k \in \mathcal{T}\}\) an admissible control, if for each \(k \in \mathcal{T}\), \(u_k\) is an \(\mathcal{F}_k\)-measurable \(U_k\)-valued r.v. such that
\[
\mathbb{E}[|u_k|^\beta] < \infty
\]
holds for some \(\beta > 2\). Denoted by \(\mathcal{U}\) the set of admissible controls. For later use, set \(\gamma = \min\{\alpha, \beta\}\). Then \(\gamma > 2\). The discrete-time stochastic optimal control problem is to minimize
\[
J(u) = \mathbb{E}\left\{\sum_{k \in \mathcal{T}} l(k, X_k, u_k) + h(X_N) \right\}, \quad u \in \mathcal{U},
\]
where \(\{X_k, k \in \mathcal{T}\}\) is described by
\[
\begin{cases}
X_{k+1} = b(k, X_k, u_k) + \sigma(k, X_k, u_k)W_k, & k \in \mathcal{T}, \\
X_0 = x_0 \in \mathbb{R}^n,
\end{cases}
\]
with \(b(k, \cdot, \cdot, \cdot): \mathbb{R}^n \times U_k \times \Omega \to \mathbb{R}^n, \sigma(k, \cdot, \cdot, \cdot): \mathbb{R}^n \times U_k \times \Omega \to \mathbb{R}^{n \times d}, l(k, \cdot, \cdot, \cdot): \mathbb{R}^n \times U_k \times \Omega \to \mathbb{R}\) and \(h(\cdot, \cdot): \mathbb{R}^n \times \Omega \to \mathbb{R}\) for \(k \in \mathcal{T}\). That is, we hope to find \(u^* = \{u_k^*, k \in \mathcal{T}\} \in \mathcal{U}\) which satisfies
\[
J(u^*) = \inf_{u \in \mathcal{U}} J(u).
\]
Such \(u^*\) is called an optimal control, and the corresponding system state is denoted by \(X^* = \{X_k^*, k \in \mathcal{T}\}\). This problem is called Problem \((OCP)\). For later use, define \(l(N, x, v) = h(x)\) for any \((x, v)\).

The following assumptions will be in force.

\textbf{(H1)} For each \(k\), \(b(k, x, v)\) and \(\sigma(k, x, v)\) are \(\mathcal{F}_k\)-measurable for any \((x, v)\), and they are continuously differentiable in \((x, v)\) with uniformly bounded partial derivatives. In addition, there exists \(C > 0\) such that \(|b(k, 0, 0)| + |\sigma(k, 0, 0)| \leq C\).

\textbf{(H2)} For each \(k\), \(l(k, x, v)\) is \(\mathcal{F}_k\)-measurable and \(h(x)\) is \(\mathcal{F}_N\)-measurable for any \((x, v)\), and they are continuously differentiable in \((x, v)\) with the partial derivatives being bounded by \(C(1 + |x| + |v|)\). Besides, \(\mathbb{E}\left[\sum_{k \in \mathcal{T}} |l(k, 0, 0)| + |h(0)|\right] < \infty\).

\textbf{Lemma 2.1.} Assume \((\text{H1})\). Then for any \(u \in \mathcal{U}\), the equation \((3)\) admits a unique solution \(X = \{X_k, k \in \mathcal{T}\}\) such that \(X_k \in L^\gamma(\Omega, \mathcal{F}_k, \mathbb{R}^n)\) for each \(k \in \mathcal{T}\). Moreover, there exists \(C = C(\alpha, \gamma, N) > 0\) satisfying
\[
\mathbb{E}\sum_{k \in \mathcal{T}} |X_k|^\gamma \leq CE\left\{1 + |x_0|^\gamma + \sum_{k \in \mathcal{T}} |u_k|^\gamma \right\}.
\]
Proof. We only need to check (4). Using Hölder’s inequality to (1) yields
\[ E[|W_k|\gamma |\mathcal{F}_k] \leq C(\alpha, \gamma). \]
Since \( X_k \in \mathcal{F}_k \), we have
\[ E[|X_k|^\gamma |W_k|^\gamma] = E \{ |X_k|^\gamma E[|W_k|^\gamma |\mathcal{F}_k] \} \leq C(\alpha, \gamma)E|X_k|^\gamma. \]
Similarly, \( E[|u_k|^\gamma |W_k|^\gamma] \leq C(\alpha, \gamma)E|u_k|^\gamma \). Then, by (H1) we get
\[ E|X_{k+1}|^\gamma \leq C(\alpha, \gamma)E \{ 1 + |X_k|^\gamma + |u_k|^\gamma \}, \quad k \in T. \]
Finally, the conclusion is obtained by induction. \( \square \)

The previous assumptions are not sufficient to get \( E \sum_{k \in T} |X_k|^r < \infty \) for \( r > \gamma \). Under (H1) and (H2), \( J \) is well defined on \( \mathcal{U} \).

2.2. The variational equation and variational inequality. For any \( \mu = \{ \mu_k, k \in T \} \in \mathcal{U} \), define \( \nu = \{ \nu_k, k \in T \} \) where \( \nu_k = \mu_k - u_k^* \) for each \( k \in T \), and \( u^\varepsilon = \{ u_k^\varepsilon, k \in T \} \) for \( \varepsilon \in [0, 1] \) where \( u_k^\varepsilon = u_k^* + \varepsilon \nu_k \) for each \( k \in T \). Denote by \( X^\varepsilon = \{ X_k^\varepsilon, k \in T \} \) the state corresponding to \( u^\varepsilon \). Since \( U_0, \ldots, U_{N-1} \) are convex sets, we have \( u^\varepsilon = \varepsilon \mu + (1 - \varepsilon)u^* \in \mathcal{U} \). In what follows, we use \( \rho(k) \) to denote \( \rho(k, X_k^\varepsilon, u_k^\varepsilon) \) for simplicity, where \( \rho \) represents \( b, \sigma, \lambda \) and their partial derivatives with respect to \( (x, v) \). Denote by \( \mathcal{D} \) the index set \( \{1, 2, \ldots, d\} \).

Let us introduce
\[
\left\{
\begin{array}{l}
Z_{k+1} = \left[ (b_x(k))^T Z_k + (b_v(k))^T \nu_k \right] \\
\quad + \sum_{i \in \mathcal{D}} \left[ (\sigma_x^i(k))^T Z_k + (\sigma_v^i(k))^T \nu_k \right] W_k^i, \quad k \in T, \\
Z_0 = 0,
\end{array}
\right.
\]
where \( \sigma = (\sigma^1, \sigma^2, \ldots, \sigma^d) \). It is called the variational equation. Note that (5) admits a unique solution \( Z = \{ Z_k, k \in T \} \) such that \( Z_k \in L^2(\Omega, \mathcal{F}_k, \mathbb{R}^n) \) for each \( k \). Moreover, we can get the explicit solution of (5). In fact, by defining
\[
M_k = (b_x(k))^T + \sum_{i \in \mathcal{D}} (\sigma_x^i(k))^T W_k^i, \quad R_k = (b_v(k))^T \nu_k + \sum_{i \in \mathcal{D}} (\sigma_v^i(k))^T W_k^i \nu_k,
\]
we have \( Z_{k+1} = M_k Z_k + R_k \) for each \( k \in T \). Since \( Z_0 = 0 \), using induction yields
\[
Z_k = \sum_{i=0}^{k-1} M_i^{k-1} R_i,
\]
where \( M_i^j \) is the \( j \)th minor of the \( M_i \) when \( i < j \), and \( M_i^i = I_n \) when \( i = j \), with \( I_n \) being the \( (n \times n) \)-identity matrix.

Let us define
\[
\mathcal{L} J(u^*) = E \left\{ \sum_{k \in T} \left[ c_x(k, Z_k) + c_v(k, \nu_k) \right] + h_z(X_{k+1}, Z_{k+1}) \right\}.
\]
Then we have the following variational inequality.

Lemma 2.2. Under (H1) and (H2), it holds that \( \mathcal{L} J(u^*) \geq 0 \).

Proof. By standard arguments it is easy to show that
\[
\mathcal{L} J(u^*) = \lim_{\varepsilon \to 0} \varepsilon^{-1} [J(u^\varepsilon) - J(u^*)].
\]
Then the conclusion follows from the optimality of \( u^* \). \( \square \)
2.3. The adjoint equation. Let us define

$$H(k, x, v, p, q) = \langle b(k, x, v), p \rangle + \sum_{i \in D} \langle \sigma^i(k, x, v), q^i \rangle + l(k, x, v),$$

where $q = (q^1, \cdots, q^d)$. Note that $\sum_{i \in D} \langle \sigma^i(k, x, v), q^i \rangle = \text{tr} \left[ \sigma(k, x, v)^\top q \right]$. $H$ is called the Hamiltonian function. Introduce the following adjoint equation:

$$\begin{align*}
P_k &= \mathbb{E} \left\{ H_x(k + 1, X_{k+1}^*, u_{k+1}^*, P_{k+1}, Q_{k+1}) \mid \mathcal{F}_k \right\}, \\
Q_k &= \mathbb{E} \left\{ H_x(k + 1, X_{k+1}^*, u_{k+1}^*, P_{k+1}, Q_{k+1}) W_k^\top \mid \mathcal{F}_k \right\}, \\
k &= 0, 1, \cdots, N - 2, \\
P_{N-1} &= \mathbb{E} \left[ h_x(X_N^*) \mid \mathcal{F}_{N-1} \right], \\
Q_{N-1} &= \mathbb{E} \left[ h_x(X_N^*) W_{N-1}^\top \mid \mathcal{F}_{N-1} \right],
\end{align*}$$

(8)

that is,

$$\begin{align*}
P_k &= \mathbb{E} \left\{ b_x(k + 1) P_{k+1} + \sum_{i \in D} \sigma^i_x(k + 1) Q_{k+1}^i + l_x(k + 1) \right\} \mid \mathcal{F}_k, \\
Q_k &= \mathbb{E} \left\{ b_x(k + 1) P_{k+1} + \sum_{i \in D} \sigma^i_x(k + 1) Q_{k+1}^i + l_x(k + 1) \right\} W_k^\top \mid \mathcal{F}_k, \\
k &= 0, 1, \cdots, N - 2, \\
P_{N-1} &= \mathbb{E} \left[ h_x(X_N^*) \mid \mathcal{F}_{N-1} \right], \\
Q_{N-1} &= \mathbb{E} \left[ h_x(X_N^*) W_{N-1}^\top \mid \mathcal{F}_{N-1} \right],
\end{align*}$$

(9)

where $Q_k = (Q^1_k, \cdots, Q^d_k)$ for $k \in T$.

**Remark 1.** (i). The form of (8) was first introduced in [16] as the adjoint equation. As one kind of backward stochastic difference equation, (8) is quite different from the adjoint equation in the continuous-time counterpart.

(ii). Note that [4] and [5] also studied backward stochastic difference equations and obtained the solvability results. However, (8) is different from the equations studied in [4] and [5]. On the one hand, they have different forms. On the other hand, (8) is introduced naturally as the dual equation of the variational equation (5).

(iii). It’s well known that the adjoint equation in the continuous-time case admits a square-integrable solution under the classical assumptions. However, the solution $\{(P_k, Q_k), k \in T\}$ of (8) has a problem of integrability, which was not taken into consideration in [16]. Let us show next where the problems are. On the one hand, if $\{W_k\}$ are only square-integrable ($\alpha = 2$), then $X_k$, $1 \leq k \leq N$, are at most square-integrable in general. In this case, $(P_k, Q_k)$, $1 \leq k \leq N - 2$, are not square-integrable, so $\mathbb{E}[P_k|x_k]$ and $\mathbb{E}[Q_k|x_k]$, $1 \leq k \leq N - 2$, are not well defined, which implies that the proof of the sufficient condition (Theorem 2.7) fails. On the other hand, if $\{u_k\}$ are only square-integrable ($\beta = 2$), then in general we can still at most get the square-integrability of the states $X_k$, $1 \leq k \leq N$, and $(P_k, Q_k)$, $1 \leq k \leq N - 2$, are still not square-integrable, which also means that the proof of the sufficient condition does not work. Besides, $\mathbb{E}[|P_k| u_k]$ and $\mathbb{E}[Q_k| u_k]$, $1 \leq k \leq N - 2$, are not well defined for all admissible controls $u = \{u_k\}$, which implies that the proof of the necessary condition (Theorem 2.6) no longer works.

Based on the previous discussions, let us assume moreover the following

(H3) $\alpha \geq N + 1$ and $\beta = 2\alpha(\alpha - N + 1)^{-1}$.

**Lemma 2.3.** Assume (H1)-(H3). Then it holds that

$$\mathbb{E}[|P_k| + |Q_k|] < \infty, \quad k \in T,$$

(10)
\[ \mathbb{E}[|P_{k+1}|W_k] + |Q_{k+1}|W_k| < \infty, \quad k = 0, 1, \cdots, N - 2, \]  
(11)  
\[ \mathbb{E}[|P_k||X_k| + |Q_k||X_k| + |P_k||u_k| + |Q_k||u_k|] < \infty, \quad k \in \mathcal{T}, \]  
(12)  

where \( \{X_k\} \) is the state corresponding to any admissible control \( \{u_k\} \) and \( \{(P_k, Q_k)\} \) is the solution of the adjoint equation (8).

**Proof.** The following result will be used for several times, which is a consequence of Hölder’s inequality. For \( \xi \in L^\mu \) and \( \eta \in L^{\nu} \) with \( \mu, \nu > 0 \), it holds that \( \xi \eta \in L^{\frac{\mu \nu}{\mu + \nu}} \), and \( \xi \eta \in L^1 \) if \( \frac{\mu \nu}{\mu + \nu} \geq 1 \).

Denote \( \theta_k = \frac{\alpha \gamma}{\alpha + \gamma(N-1-k)} \) for \( k \in \mathcal{T} \). Note that \( \theta_k \) is increasing in \( k \) and \( \theta_k \leq \gamma \).

Let us show that
\[ |P_k| \in L^{\theta_k}, \quad |Q_k| \in L^{\theta_k-1}, \quad k = 1, \cdots, N - 1. \]  
(13)

Since \( |h_x(X^*_N)| \leq C(1+|X^*_N|) \) and \( |X^*_N| \in L^\gamma \), so \( |P_{N-1}| \in L^\gamma \). Since \( |h_x(X^*_N)| \in L^\gamma \) and \( |W_{N-1}| \in L^{\alpha} \), we have \( |h_x(X^*_N)|W_{N-1}^{1/\gamma} \in L^{\frac{\alpha \gamma}{\alpha + \gamma}} \) and thus \( |Q_{N-1}| \in L^{\frac{\alpha \gamma}{\alpha + \gamma}} \). Thus (13) holds for \( k = N - 1 \). Let us use induction to proceed. Assume that (13) holds for \( k = j + 1 \). Note that \( |X^*_{j+1}| \in L^\gamma, |u^*_{j+1}| \in L^\beta \), and \( \theta_j < \theta_{j+1} \leq \gamma \leq \beta \). Considering the expression of \( P_j \) in (9), by (H1) and (H2) we get \( P_j \in L^{\theta_j} \). Similarly, since \( |W_j| \in L^\alpha, |X^*_j||W_j| \in L^{\frac{\alpha \gamma}{\alpha + \gamma}}, |u^*_{j+1}||W_j| \in L^{\frac{\alpha \gamma}{\alpha + \gamma}}, |P^*_j||W_j| \in L^{\theta_j}, |Q^*_{j+1}||W_j| \in L^{\theta_j-1}, \) and \( \theta_{j+1} < \theta_j \leq \frac{\alpha \gamma}{\alpha + \gamma} \leq \frac{\alpha \gamma}{\alpha + \beta} < \alpha \), from (9) we get \( |Q_j| \in L^{\theta_j-1} \). That is, (13) holds for \( k = j \). Thus, (13) is proved.

With the above preparations, (10)-(12) hold if \( |Q_1||W_0|, |Q_1||X_1|, |Q_2||X_2| \in L^1 \), and thus if \( \gamma(\alpha - N + 1) \geq 2\alpha \). Since \( \gamma = \min\{\alpha, \beta\} \), solving this inequality yields \( \alpha \geq N + 1 \) and \( \beta \geq 2\alpha(\alpha - N + 1)^{-1} \). Thus, it’s appropriate and sufficient to choose \( \alpha \geq N + 1 \) and \( \beta \geq 2\alpha(\alpha - N + 1)^{-1} \), which is just (H3).

**Remark 2.** (i). The proof of Lemma 2.3 shows why we need and how we worked out the assumption (H3). Lemma 2.3 guarantees that the adjoint equation (8) is well defined with a unique solution \( \{(P_k, Q_k), k \in \mathcal{T}\} \), and all the expectations involving \( \{(P_k, Q_k), k \in \mathcal{T}\} \) in this paper are well defined.

(ii). In the general case, (H3) is nearly the minimum condition that ensures (10)-(12). (H3) could be relaxed in some special cases such as when \( h_x(x) \) is uniformly bounded.

Next result gives the explicit solution of the adjoint equation (8).

**Lemma 2.4.** The solution \( \{(P_k, Q_k), k \in \mathcal{T}\} \) of the adjoint equation (8) is given by
\[
\left\{
\begin{array}{l}
P_k = \sum_{i=k+1}^{N} \mathbb{E}\left[ (M_i^{-1})^\top l_x(i) | \mathcal{F}_k \right], \\
Q_k = \sum_{i=k+1}^{N} \mathbb{E}\left[ (M_i^{-1})^\top l_x(i) W_k^\top | \mathcal{F}_k \right].
\end{array}
\right.
\]  
(14)

**Proof.** Recall the measurability of the related random variables, and the definitions of \( \{M_i\} \) and \( \{M_i^e\} \). By the properties of conditional expectations, we can use backward induction to show that \( \{(P_k, Q_k)\} \) given by (14) satisfies (8). \( \square \)

### 2.4. The maximum principle
The following result gives the duality relation between the variational equation (5) and the adjoint equation (8).
Lemma 2.5. Under (H1)-(H3), it holds that
\[
\mathbb{E} \sum_{k \in T} \left( b_v(k) P_k + \sum_{j \in D} \sigma_j^T(k) Q_k^j \right) \nu_k = \mathbb{E} \sum_{k \in T} \langle l_x(k), Z_k \rangle. \tag{15}
\]

Proof. Since \( b_v(k) \) and \( \nu_k \) are \( \mathcal{F}_k \)-measurable, in view of (14) we use the law of total expectation and then interchange the sums to get
\[
\mathbb{E} \sum_{k \in T} P_k^T (b_v(k))^\top \nu_k = \mathbb{E} \sum_{i \in T} \sum_{k=0}^{i-1} (l_x(i))^\top M_k^{i-1} (b_v(k))^\top \nu_k.
\]
In the same way,
\[
\mathbb{E} \sum_{k \in T} \sum_{j \in D} (Q_j^k)^T (\sigma_j^T(k))^\top \nu_k = \mathbb{E} \sum_{i \in T} \sum_{k=0}^{i-1} (l_x(i))^\top M_k^{i-1} \sum_{j \in D} (\sigma_j^T(k))^\top W_j^k \nu_k.
\]
Recalling the definition of \( \{ R_k \} \) and \( \{ Z_i \} \), we can add the previous two equations to get
\[
\mathbb{E} \sum_{k \in T} \left[ P_k^T (b_v(k))^\top + \sum_{j \in D} (Q_j^k)^T (\sigma_j^T(k))^\top \right] \nu_k
\]
\[
= \mathbb{E} \sum_{i \in T} (l_x(i))^\top \sum_{k=0}^{i-1} M_k^{i-1} R_k = \mathbb{E} \sum_{i \in T} (l_x(i))^\top Z_i,
\]
which is just (15). \( \square \)

Now we are in a position to establish the necessary MP for Problem (OCP).

Theorem 2.6. Assume (H1)-(H3). If \( u^* = \{ u_k^*, k \in T \} \) is an optimal control of Problem (OCP) with \( X^* = \{ X_k^*, k \in T \} \) being the corresponding system state and \( \{(P_k, Q_k), k \in T\} \) the solution of the adjoint equation (8), then it satisfies
\[
\langle H_v(k, X_k^*, u_k^*, P_k, Q_k), v - u_k^* \rangle \geq 0, \quad \forall v \in U_k, k \in T. \tag{16}
\]

Proof. Combining Lemma 2.2 and (15) yields \( \mathbb{E} \sum_{i \in T} \langle H_v(i, X_i^*, u_i^*, P_i, Q_i), \nu_i \rangle \geq 0. \) Then by the definition of \( \nu \), we have
\[
\mathbb{E} \sum_{i \in T} \langle H_v(i, X_i^*, u_i^*, P_i, Q_i), \mu_i - u_i^* \rangle \geq 0, \quad \forall \mu = \{ \mu_i, i \in T \} \in U. \tag{17}
\]

For any \( k \in T \), \( v \in U_k \) and \( A \in \mathcal{F}_k \), let us define an admissible control sequence \( \mu = \{ \mu_i \} \in U \) by \( \mu_k = v \chi_A + u_k^* \chi_{A^c} \), and \( \mu_i = u_i^* \) if \( i \neq k \). Applying this \( \mu \) to (17) gives \( \mathbb{E} [\Delta_k \chi_A] \geq 0 \), where \( \Delta_k = \langle H_v(k, X_k^*, u_k^*, P_k, Q_k), v - u_k^* \rangle \). Since \( \Delta_k \) is \( \mathcal{F}_k \)-measurable, and \( A \) is arbitrarily chosen from \( \mathcal{F}_k \), it follows that \( \Delta_k \geq 0 \). Thus we get (16), since \( k \in T \) and \( v \in U_k \) are also arbitrarily chosen. \( \square \)

Next let us give the sufficient MP for Problem (OCP).

Theorem 2.7. Assume (H1)-(H3). Let \( u^* = \{ u_k^*, k \in T \} \) be an admissible control, with \( X^* = \{ X_k^*, k \in T \} \) and \( \{(P_k, Q_k), k \in T\} \) being respectively the corresponding solutions of (3) and (8). Assume that \( h(\cdot) \) and \( H(\cdot, \cdot, \cdot, P_k, Q_k) \) are convex functions for each \( k \in T \). Then \( u^* \) is an optimal control of Problem (OCP) if it satisfies (16).
Proof. Let \( u = \{u_k, k \in \mathcal{T}\} \) be any admissible control and \( X = \{X_k, k \in \mathcal{T}\} \) the corresponding state. Set \( \dot{X}_k = X_k - X_k^* \), \( b(k) = b(k, X_k, u_k) - b(k, X_k^*, u_k^*) \) and \( \dot{\sigma}(k) = \sigma(k, X_k, u_k) - \sigma(k, X_k^*, u_k^*) \). By the definition of \( J \) and \( H \), if \( h(\cdot) \) and \( H(k, \cdot, \cdot, P_k, Q_k) \) are convex for each \( k \in \mathcal{T} \) and (16) holds true, then
\[
J(u) - J(u^*) \geq \mathbb{E} \left\{ \sum_{k \in \mathcal{T}} \langle l_x(k), \dot{X}_k \rangle + \langle h_x(X_k^*), \dot{X}_k \rangle - \sum_{k \in \mathcal{T}} \langle \alpha_k, P_k \rangle - \sum_{j \in D} \sum_{k \in \mathcal{T}} \langle \beta_j^k, Q_{j,k} \rangle \right\},
\]
where \( \alpha_k = \dot{b}(k) - (b_x(k))^T \dot{X}_k \) and \( \beta_j^k = \dot{\sigma}^j(k) - (\sigma_j^k(k))^T \dot{X}_k \) for \( k \in \mathcal{T} \) and \( j = 1, 2, \cdots, d \). Recall that \( l(N, x, v) = h(x) \). To conclude, it suffices to prove
\[
\mathbb{E} \sum_{k \in \mathcal{T}} \langle l_x(k), \dot{X}_k \rangle = \mathbb{E} \sum_{k \in \mathcal{T}} \left( \langle \alpha_k, P_k \rangle + \sum_{j \in D} \langle \beta_j^k, Q_{j,k} \rangle \right).
\]
In fact, \( \{\dot{X}_k\} \) satisfies \( \dot{X}_{k+1} = M_k \dot{X}_k + \delta_k \) for \( k \in \mathcal{T} \), with \( \dot{X}_0 = 0 \) and \( \delta_k = \alpha_k + \sum_{j \in D} \beta_j^k W_{k,j}^j \) for \( k \in \mathcal{T} \). Similar to (7), we can use induction to get
\[
\dot{X}_k = \sum_{i=0}^{k-1} M_i^{k-1} \delta_i, \quad k \in \mathcal{T}.
\]
With this expression, we can interchange two sums to get
\[
\mathbb{E} \sum_{k \in \mathcal{T}} \langle l_x(k), \dot{X}_k \rangle = \mathbb{E} \sum_{k \in \mathcal{T}} \sum_{i=k+1}^{N} (l_x(i))^T M_i^{i-1} \delta_i.
\]
Finally, we can use the law of total expectation and recall (14) to get (18). \( \square \)

3. Maximum principle for two kinds of discrete-time stochastic games.

3.1. Maximum principle for one kind of discrete-time nonzero-sum stochastic game. Let us consider a system described by
\[
\begin{align*}
X_{k+1} &= b(k, X_k, u_{1,k}, u_{2,k}) + \sigma(k, X_k, u_{1,k}, u_{2,k}) W_k, \quad k \in \mathcal{T}, \\
X_0 &= x_0 \in \mathbb{R}^n,
\end{align*}
\]
where \( u_1 = \{u_{1,k}, k \in \mathcal{T}\} \) and \( u_2 = \{u_{2,k}, k \in \mathcal{T}\} \) are respectively the controls of player 1 and player 2. \( W_k, u_{1,k} \) and \( u_{2,k} \) satisfy (1), (2) and (H3). Let \( U_{i,0}, U_{i,1}, \cdots, U_{i,N-1} \) be a sequence of nonempty convex subsets of \( \mathbb{R}^n \), \( i = 1, 2 \). For \( i = 1, 2 \), we denote by \( \mathcal{U}_i \) the class of \( u_i = \{u_{i,k}, k \in \mathcal{T}\} \) which share the same assumptions as the admissible control \( u = \{u_k, k \in \mathcal{T}\} \) in Section 2. Each element of \( \mathcal{U}_i \) is called an admissible control for player \( i, i = 1, 2 \). Two players intervene on the dynamics of the system and each one acts to save his own interest. The problem is to find \( (u_1^*, u_2^*) \in \mathcal{U}_1 \times \mathcal{U}_2 \) which satisfies
\[
J_1(u_1^*, u_2^*) = \inf_{u_1 \in \mathcal{U}_1} J_1(u_1, u_2), \quad J_2(u_1^*, u_2^*) = \inf_{u_2 \in \mathcal{U}_2} J_2(u_1, u_2),
\]
where
\[
J_i(u_1, u_2) = \mathbb{E}\left\{ \sum_{k \in \mathcal{T}} l_i(k, X_k, u_{1,k}, u_{2,k}) + h_i(X_N) \right\}, \quad i = 1, 2.
\]
We call such \( (u_1^*, u_2^*) \) an equilibrium point, and the corresponding state is denoted by \( X^* = \{X_k^*, k \in \mathcal{T}\} \). This problem is a kind of discrete-time nonzero-sum stochastic game. We call it Problem (NZG). Assume that the coefficients \( b, \sigma, l_i, h_1, h_2 \) satisfy the assumptions (H1) and (H2) with \( v \) there replaced by \( (u_1, u_2) \).
For $i = 1, 2$, let us define
\[ H_i(k, x, u_1, u_2, p, q) = \langle b(k, x, u_1, u_2), p \rangle + tr[\sigma(k, x, u_1, u_2)^\top q] + l_i(k, x, u_1, u_2), \]
and introduce the following adjoint equation:
\[
\begin{align*}
P_{i,k} &= \mathbb{E}[H_{ix}(k + 1, X_{k+1}^i, u_{1,k+1}^i, u_{2,k+1}^i, P_{i,k+1}, Q_{i,k+1}) | \mathcal{F}_k], \\
Q_{i,k} &= \mathbb{E}[H_{ix}(k + 1, X_{k+1}^i, u_{1,k+1}^i, u_{2,k+1}^i, P_{i,k+1}, Q_{i,k+1}) W_k^i | \mathcal{F}_k], \\
k &\in \{0, 1, \cdots, N-2\}, \\
P_{i,N-1} &= \mathbb{E}[h_{ix}(X_N^i) | \mathcal{F}_{N-1}], \\
Q_{i,N-1} &= \mathbb{E}[h_{ix}(X_N^i) W_{N-1}^i | \mathcal{F}_{N-1}],
\end{align*}
\]  
(21)
where $H_{ix}$ and $h_{ix}$ are respectively the partial derivatives of $H_i$ and $h_i$ with respect to $x$.

Note that Problem (NZG) consists of two discrete-time stochastic optimal control problems of Problem (OCP)'s type. Following Section 2, we have the following two results for the equilibrium point of Problem (NZG). Since they are similar to those in Section 2, we omit the proof.

**Theorem 3.1.** Suppose $(u_1^*, u_2^*)$ is an equilibrium point of Problem (NZG), $X^* = \{X_k^*, k \in T\}$ is the corresponding state, and $\{(P_i,k,Q_i,k), k \in T\}$ is the solution of adjoint equation (21), $i = 1, 2$. Then
\[
(H_{iu}^i(k, X_k^i, u_{1,k}^i, u_{2,k}^i, P_i,k,Q_i,k), v_i,k - u_{1,k}^* | v_i,k - u_{1,k}^*) \geq 0, \\
\forall v_i,k \in U_{i,k}, k \in T, i = 1, 2,
\]  
(22)
where $H_{iu}$ denotes the partial derivative of $H_i$ with respect to $u_i$, $i = 1, 2$.

**Theorem 3.2.** Let $(u_1^*, u_2^*) \in U_1 \times U_2$ be a pair of admissible controls for Problem (NZG), $X^* = \{X_k^*, k \in T\}$ the corresponding state and $\{(P_i,k,Q_i,k), k \in T\}$ the solution of adjoint equation (21), $i = 1, 2$. Suppose that $H_1(k, \cdot, \cdot, u_{1,k}^*, P_i,k,Q_i,k), H_2(k, \cdot, \cdot, u_{2,k}^*, P_i,k,Q_i,k)$, $h_1(\cdot)$ and $h_2(\cdot)$ are convex for each $k \in T$. Then $(u_1^*, u_2^*)$ is an equilibrium point of Problem (NZG) if it satisfies (22).

### 3.2. Maximum principle for one kind of discrete-time zero-sum stochastic game

We still consider the system (19). The problem here is to find $(u_1^*, u_2^*) \in U_1 \times U_2$ satisfying
\[
J(u_1^*, u_2^*) \leq J(u_1^*, u_2^*) \leq J(u_1, u_2^*), \quad \forall (u_1, u_2) \in U_1 \times U_2,
\]  
(23)
where
\[
J(u_1, u_2) = \mathbb{E} \left\{ \sum_{k \in T} l(k, X_k, u_{1,k}, u_{2,k}) + h(X_N) \right\}.
\]
This problem is a kind of discrete-time zero-sum stochastic game. We call it Problem (ZSG). The pair of admissible controls $(u_1^*, u_2^*)$ satisfying (23) is called a saddle point, and the corresponding state is denoted by $X^* = \{X_k^*, k \in T\}$. We still assume that $b,\sigma,l,h$ satisfy the assumptions (H1) and (H2) with $v$ there replaced by $(u_1, u_2)$. Assume also (H3).

Let us define $(l_1, h_1, J_1) = (l, h, J)$ and $(l_2, h_2, J_2) = (-l, -h, -J)$. Then the saddle point $(u_1^*, u_2^*)$ of Problem (ZSG) satisfies
\[
J_1(u_1^*, u_2^*) = \inf_{u_1 \in U_1} J_1(u_1, u_2^*), \quad J_2(u_1^*, u_2^*) = \inf_{u_2 \in U_2} J_2(u_1^*, u_2),
\]
where
\[
J_i(u_1, u_2) = \mathbb{E} \left\{ \sum_{k \in T} l_i(k, X_k, u_{1,k}, u_{2,k}) + h_i(X_N) \right\}, \quad i = 1, 2.
\]
Note that this new problem with system (19) and cost functions $J_1$ and $J_2$ is a nonzero-sum game of Problem (NZG)’s type. We denote it Problem (NZG1). That is, Problem (ZSG) can be turned to Problem (NZG1), and the saddle point $(u_1^*, u_2^*)$ of Problem (ZSG) is an equilibrium point of Problem (NZG1).

Define
\[
H(k, x, u_1, u_2, p, q) = \langle b(k, x, u_1, u_2), p \rangle + \text{tr} \left[ \sigma(k, x, u_1, u_2)^T q \right] + l(k, x, u_1, u_2),
\]
\[
H_i(k, x, u_1, u_2, p, q) = \langle b(k, x, u_1, u_2), p \rangle + \text{tr} \left[ \sigma(k, x, u_1, u_2)^T q \right] + l_i(k, x, u_1, u_2),
\]
$i = 1, 2$. Let us give the necessary condition of the saddle point.

**Theorem 3.3.** Suppose $(u_1^*, u_2^*)$ is a saddle point of Problem (ZSG) and $X^* = \{X_k^*, k \in \mathcal{T}\}$ is the corresponding state. Then for each $k \in \mathcal{T}$, it holds that
\[
\begin{cases}
\langle H_u\(k, X_k^*, u_1^*, u_2^*, P_k, Q_k\), v_1, k - u_1^* \rangle \geq 0, & \forall v_1, k \in U_1, k, \\
\langle H_u\(k, X_k^*, u_1^*, u_2^*, P_k, Q_k\), v_2, k - u_2^* \rangle \leq 0, & \forall v_2, k \in U_2, k,
\end{cases}
\]
where $H_u$ denotes the partial derivative of $H$ with respect to $u_i$, $i = 1, 2$, and $\{P_k, Q_k\}$ is the solution of the following adjoint equation:
\[
\begin{align*}
P_k &= E[H_x(k + 1, X_{k+1}^*, u_1^*_{k+1}, u_2^*_{k+1}, P_{k+1}, Q_{k+1})|F_k], \\
Q_k &= E[H_x(k + 1, X_{k+1}^*, u_1^*_{k+1}, u_2^*_{k+1}, P_{k+1}, Q_{k+1})W_k|F_k], \\
P_{N-1} &= E[h_x(X_N)|F_{N-1}], \\
Q_{N-1} &= E[h_x(X_N)W_N^T|F_{N-1}].
\end{align*}
\]
\[
(25)
\]

**Proof.** Recall that Problem (ZSG) is equivalent to Problem (NZG1). By Theorem 3.1, we have for each $k \in \mathcal{T}$,
\[
\begin{cases}
\langle H_{u1}\(k, X_k^*, u_1^*, u_2^*, P_1, Q_1\), v_1, k - u_1^* \rangle \geq 0, & \forall v_1, k \in U_1, k, \\
\langle H_{u2}\(k, X_k^*, u_1^*, u_2^*, P_2, Q_2\), v_2, k - u_2^* \rangle \geq 0, & \forall v_2, k \in U_2, k,
\end{cases}
\]
where $H_{ui}$ denotes the partial derivative of $H_i$ with respect to $u_i$, $i = 1, 2$, and $\{P_{1, k}, Q_{1, k}\}$, $i = 1, 2$, is the solution of the following adjoint equation
\[
\begin{align*}
P_{1, k} &= E[H_{ix}(k + 1, X_{k+1}^*, u_1^*_{k+1}, u_2^*_{k+1}, P_{1, k+1}, Q_{1, k+1})|F_k], \\
Q_{1, k} &= E[H_{ix}(k + 1, X_{k+1}^*, u_1^*_{k+1}, u_2^*_{k+1}, P_{1, k+1}, Q_{1, k+1})W_k|F_k], \\
P_{1, N-1} &= E[h_{ix}(X_N)|F_{N-1}], \\
Q_{1, N-1} &= E[h_{ix}(X_N)W_N^T|F_{N-1}].
\end{align*}
\]
By the definition of $H_i$, it’s easy to check that $\{P_{1, k}, Q_{1, k}\}$ satisfies
\[
\begin{align*}
P_{1, k} &= E\{b_x(k + 1)P_{1, k+1} + \sigma_x(k + 1)Q_{1, k+1} + l_x(k + 1)|F_k\}, \\
Q_{1, k} &= E\{b_x(k + 1)P_{1, k+1} + \sigma_x(k + 1)Q_{1, k+1} + l_x(k + 1)|W_k|F_k\}, \\
P_{1, N-1} &= E[h_{ix}(X_N)|F_{N-1}], \\
Q_{1, N-1} &= E[h_{ix}(X_N)W_N^T|F_{N-1}].
\end{align*}
\]
and $\{P_{2, k}, Q_{2, k}\}$ satisfies
\[
\begin{align*}
P_{2, k} &= E\{b_x(k + 1)P_{2, k+1} + \sigma_x(k + 1)Q_{2, k+1} - l_x(k + 1)|F_k\}, \\
Q_{2, k} &= E\{b_x(k + 1)P_{2, k+1} + \sigma_x(k + 1)Q_{2, k+1} - l_x(k + 1)|W_k|F_k\}, \\
P_{2, N-1} &= -E[h_{ix}(X_N)|F_{N-1}], \\
Q_{2, N-1} &= -E[h_{ix}(X_N)W_N^T|F_{N-1}].
\end{align*}
\]
Comparing these two equations yields \((P_{1,k}, Q_{1,k}) = (-P_{2,k}, -Q_{2,k})\) for each \(k \in T\). Define \((P_{k}, Q_{k}) = (P_{1,k}, Q_{1,k}), k \in T\). Then by the definition of \(H, \{(P_{k}, Q_{k})\}\) satisfies (25). Finally, by the definitions of \(H_i\) and \(H\) again, it’s easy to check that (26) is just (24).

Next result is the sufficient condition of the saddle point.

**Theorem 3.4.** For \((u_{1}^{*}, u_{2}^{*}) \in U_{1} \times U_{2}\), let \(X^{*} = \{X_{k}^{*}, k \in \mathcal{T}\}\) be the corresponding state and \(\{(P_{k}, Q_{k}), k \in \mathcal{T}\}\) the solution of adjoint equation (25). Assume that for each \(k\), \(H(k, \cdot, \cdot, u_{2}^{*}, P_{k}, Q_{k})\) is convex. By Theorem 2.7, we have the inequality in (26) holds. Moreover, it holds that \(H(k, \cdot, \cdot, u_{2}^{*}, P_{k}, Q_{k})\) is concave and \(h(x) = Rx + \xi\) with \(R\) being a bounded \(\mathcal{F}_{T}\)-measurable r.v. and \(\xi\) an \(\mathcal{F}_{T}\)-measurable integrable r.v. Then \((u_{1}^{*}, u_{2}^{*})\) is a saddle point of Problem (ZSG) if (24) holds.

**Proof.** The proof is separated into three parts.

**Part 1.** Assume that \(H(k, \cdot, \cdot, u_{2}^{*}, P_{k}, Q_{k})\) and \(h(\cdot)\) are convex, and the first inequality in (24) holds. Since \((h_{1}, l_{1}, J_{1}) = (h, l, J)\) and \((P_{1,k}, Q_{1,k}) = (P_{k}, Q_{k})\), it follows that \(H(k, x, u_{1}, u_{2}, P_{k}, Q_{k}) = H_{1}(k, x, u_{1}, u_{2}, P_{1,k}, Q_{1,k})\), and thus the first inequality in (26) holds. Moreover, it holds that \(H_{1}(k, \cdot, \cdot, u_{2}^{*}, P_{1,k}, Q_{1,k})\) and \(h_{1}(\cdot)\) are convex. By Theorem 2.7, we have \(J_{1}(u_{1}^{*}, u_{2}^{*}) = \inf_{u_{1} \in U_{1}}J_{1}(u_{1}, u_{2}^{*})\). That is, \(J(u_{1}^{*}, u_{2}^{*}) = \inf_{u_{1} \in U_{1}}J(u_{1}, u_{2}^{*})\).

**Part 2.** Assume that \(H(k, \cdot, \cdot, u_{2}^{*}, P_{k}, Q_{k})\) and \(h(\cdot)\) are concave, and the second inequality of (24) holds. Since \((h_{2}, l_{2}, J_{2}) = (-h, -l, -J)\) and \((-P_{2,k}, -Q_{2,k}) = (P_{k}, Q_{k})\), it follows that \(H(k, x, u_{1}, u_{2}, P_{k}, Q_{k}) = -H_{2}(k, x, u_{1}, u_{2}, P_{2,k}, Q_{2,k})\), and thus the second inequality in (26) holds. Moreover, \(H_{2}(k, \cdot, \cdot, u_{1}^{*}, P_{2,k}, Q_{2,k})\) and \(h_{2}(\cdot)\) are convex. By Theorem 2.7 again, we have \(J_{2}(u_{1}^{*}, u_{2}^{*}) = \inf_{u_{2} \in U_{2}}J_{2}(u_{1}^{*}, u_{2})\). That is, \(J(u_{1}^{*}, u_{2}^{*}) = \sup_{u_{2} \in U_{2}}\inf_{u_{1} \in U_{1}}J(u_{1}, u_{2}^{*})\).

**Part 3.** Since \(x \rightarrow h(x)\) is assumed to be linear, the conclusion is a combination of Parts 1 and 2. □

4. **Applications.** In this section, two illustrative examples are presented.

4.1. **Example 1.** Let us consider one kind of investment/consumption choice problem. Suppose that \(m + 1\) assets are traded discontinuously in a market. One asset is bond, whose price \(\{B_{k}, k \in \mathcal{T}\}\) is subject to

\[
B_{k+1} - B_{k} = r_{k}B_{k}, \quad k \in \mathcal{T}, \quad B_{0} = b_{0}.
\]

The other \(m\) assets are stocks, whose prices \(\{S_{i,k}, i \in \mathcal{M} = \{1, 2, \ldots, m\}, k \in \mathcal{T}\}\) satisfy

\[
S_{i,k+1} - S_{i,k} = S_{i,k}(\mu_{i,k} + \sigma_{i,k}W_{k}), \quad S_{i,0} = s_{i}.
\]

We assume that the coefficients are bounded random variables. The total wealth of the investor at time \(k\) is denoted by \(X_{k}\). Denote by \(u_{i,k} \in \mathcal{M}\) the amount of wealth in the \(i\)-th stock, and call \(u_{k} = (u_{1,k}, u_{2,k}, \ldots, u_{m,k})^{\top}\) a portfolio of the investor. Then

\[
X_{k+1} = \left(\frac{r_{k} + 1}{X_{k} + \sum_{i \in \mathcal{M}}(\mu_{i,k} - r_{k})u_{i,k}}\right) + \sum_{i \in \mathcal{M}}\sigma_{i,k}u_{i,k}W_{k}.
\]

Let us assume that consumption may occur at each time \(k\), and denote by \(\{v_{k}, k \in \mathcal{T}\}\) the consumption sequence. Thus, the dynamics of the wealth satisfies

\[
\begin{cases}
X_{k+1} = \left(\frac{r_{k} + 1}{X_{k} + b_{k}^{\top}u_{k} - v_{k}}\right) + \sigma_{k}^{\top}u_{k}W_{k}, \\
X_{0} = x_{0},
\end{cases}
\]
where \( b_k = (\mu_{1,k} - r_k, \mu_{2,k} - r_k, \cdots, \mu_{m,k} - r_k)^T \) and \( \sigma_k = (\sigma_{1,k}, \sigma_{2,k}, \cdots, \sigma_{m,k})^T \).

Let \( \mathcal{U} \) be the set of \( u = \{u_k, k \in \mathcal{T} \} \), where each \( u_k \) is an \( \mathcal{F}_k \)-measurable \( \mathbb{R}^m \)-valued r.v. and (2) holds true. Denote by \( \mathcal{V} \) the collection of all \( v = \{v_k, k \in \mathcal{T} \} \) satisfying (2), where each \( v_k \) is a real-valued \( \mathcal{F}_k \)-measurable r.v. such that \( v_k \geq a \) with some constant \( a > 0 \). Assume also (H3).

The objective of the investor is to find \((u^*, v^*) \in \mathcal{U} \times \mathcal{V}\) such that
\[
J(u^*, v^*) = \inf_{(u,v) \in \mathcal{U} \times \mathcal{V}} J(u,v)
\]
with
\[
J(u,v) = \mathbb{E}\left\{ \sum_{k \in \mathcal{T}} \left[ \frac{1}{2} L_k(u_k - \eta_k)^2 - G_k \frac{(v_k)^{1-\delta}}{1-\delta} \right] - RX \right\},
\]
where \( L_k, G_k \) and \( R \) are bounded scalar-valued random variables representing the weight factors, \( \eta_k \) is a bounded \( \mathbb{R}^m \)-valued r.v. representing a benchmark, and \( \delta \) is a deterministic constant representing the Arrow-Pratt index of the risk aversion.

Let us assume that \( 0 < \delta < 1, G_k > 0, L_k > c \) and \( R > c \) for some constant \( c > 0 \). Note that the objective of the investor consists of three parts. The first one is a minimization of the deviation between the portfolio and a given benchmark, the second is a maximization of a consumption utility and the last is a maximization of the terminal wealth.

For this example, the assumptions (H1) and (H2) hold true. The Hamiltonian takes the following form:
\[
H(k, x, u, v, p, q) = [(r_k + 1)x + b_k^T u - v] p + \sigma_k^T u q + \frac{L_k}{2} (u - \eta_k)^2 - G_k \frac{(v_k)^{1-\delta}}{1-\delta}.
\]
Note that \( H \) and \( h(x) = -Rx \) are convex functions of \((x, u, v)\). Applying Theorem 2.6 shows that the optimal control \((u^*, v^*)\) satisfies
\[
\langle b_k P_k + \sigma_k Q_k + L_k(u_k^* - \eta_k), u - u_k^* \rangle \geq 0, \quad \forall u \in \mathbb{R}^m, \quad (27)
\]
\[
[ - P_k - G_k^{(v_k^*)^{-\delta}} ] (v - v_k^*) \geq 0, \quad \forall v \geq a, \quad (28)
\]
for each \( k \), where \( \{ (P_k, Q_k), k \in \mathcal{T} \} \) uniquely solves
\[
\left\{ \begin{array}{l}
P_k = \mathbb{E} [(r_{k+1} + 1) P_{k+1} | \mathcal{F}_k ], \quad Q_k = \mathbb{E} [(r_{k+1} + 1) P_{k+1} W_k | \mathcal{F}_k ], \\
k = 0, 1, \cdots, N-2,
\end{array} \right.
\]
\[
P_{N-1} = - \mathbb{E} [R | \mathcal{F}_{N-1} ], \quad Q_{N-1} = - \mathbb{E} [RW_{N-1} | \mathcal{F}_{N-1} ]. \quad (29)
\]
By Theorem 2.7, an admissible control \((u^*, v^*)\) satisfying (27) and (28) is indeed an optimal control. From (27) we derive
\[
u_k^* = \eta_k - L_k^{-1} (b_k P_k + \sigma_k Q_k). \quad (30)
\]
Using induction we know that \( P_k \) is bounded and \( P_k < -c \) for each \( k \in \mathcal{T} \). Note that (28) implies that the convex function \( \mathcal{H}(v) := -P_k v - G_k \frac{v^{1-\delta}}{1-\delta} \), \( v \geq a \) takes its minimum at \( v_k^* \). Thus, it’s easy to get
\[
u_k^* = \begin{cases} \hat{v}_k, & \text{if } \hat{v}_k \geq a, \\ a, & \text{if } \hat{v}_k < a, \end{cases} \quad (31)
\]
with \( \hat{v}_k = \left( - \frac{G_k}{P_k} \right)^{-\frac{1}{\delta}} \). Next, let us show that \( u^* = \{u_k^*\} \) and \( v^* = \{v_k^*\} \) are admissible. Recall that \( G_k \) is positive and bounded, and \( P_k < -c \). Thus, each \( v_k^* \) defined by (31) is bounded and so \( v^* \) is admissible. Besides, since \( R \) and each \( P_k \) are bounded, it holds that \( Q_k \in L^\alpha \), and so \( Q_k \in L^\beta \) since \( \beta \leq \alpha \). This implies that \( u^* = \{u_k^*\} \) defined by (30) is indeed admissible. Thus we get the following result.
**Proposition 1.** The control \((u^*, v^*) = \{(u_k^*, v_k^*)\}\), where \(u_k^*\) and \(v_k^*\) are defined respectively by (30) and (31), is the optimal strategy of this investment/consumption choice problem.

**Remark 3.** In some special cases, we can solve (29) to get its solution in an explicit form. If \(R, r_0, \cdots, r_{N-1}\) are deterministic constants and \(\mathbb{E}[W_k | \mathcal{F}_k] = 0\) holds for each \(k \in \mathcal{T}\), then (29) is uniquely solved to get

\[
P_k = \begin{cases} 
  -R(r_{N-1} + 1)(r_{N-2} + 1) \cdots (r_{k+1} + 1), & \text{if } 0 \leq k \leq N-2, \\
  -R, & \text{if } k = N-1,
\end{cases}
\]

and \(Q_k = 0\) for each \(k \in \mathcal{T}\).

**4.2. Example 2.** Let us study one kind of discrete-time LQ nonzero-sum stochastic game. Take \(U_{i,k} = \mathbb{R}^m\) for all \(i = 1, 2\) and \(k \in \mathcal{T}\). Assume \(d = 1\). The problem is to find \((u_1^*, u_2^*) \in U_1 \times U_2\) such that (20) holds, where

\[
J_i(u_1, u_2) = \frac{1}{2} \mathbb{E} \left[ \sum_{k \in \mathcal{T}} \left( (G_{i,k} X_k, X_k) + (L_{i,k} u_{1,k}, u_{1,k}) + (\ell_{i,k} u_{2,k}, u_{2,k}) \right) + (R_i X_N, X_N) \right],
\]

and

\[
\begin{align*}
X_{k+1} &= (A_k X_k + B_{1,k} u_{1,k} + B_{2,k} u_{2,k} + C_k) \\
&\quad + (D_k X_k + E_{1,k} u_{1,k} + E_{2,k} u_{2,k} + F_k) W_k, \quad k \in \mathcal{T},
\end{align*}
\]

\[
X_0 = x_0 \in \mathbb{R}^n.
\]

The coefficients are deterministic and bounded matrices of appropriate dimensions. Moreover, \(G_{i,k}, R_i \geq 0, L_{i,k}, \ell_{i,k} \geq 0\) and \(L_{1,k}, \ell_{2,k} > 0\) for \(k \in \mathcal{T}, i = 1, 2\). We call it Problem (LQNZG).

For this example, the Hamiltonian \(H\) is defined by

\[
H_i(x, u_1, u_2, p, q) = (A_k x + B_{1,k} u_1 + B_{2,k} u_2 + C_k) \\
+ (D_k x + E_{1,k} u_{1,k} + E_{2,k} u_{2,k} + F_k) q \\
+ \frac{1}{2} \left( (G_{i,k} x, x) + (L_{i,k} u_{1,k}, u_{1,k}) + (\ell_{i,k} u_{2,k}, u_{2,k}) \right), \quad i = 1, 2,
\]

and the adjoint equation takes the following form:

\[
\begin{aligned}
P_{i,k} &= \mathbb{E} \left\{ [A_{k+1}^\top P_{i,k+1} + D_{i,k+1}^\top Q_{i,k+1} + G_{i,k+1} X_{k+1}] | \mathcal{F}_k \right\}, \\
Q_{i,k} &= \mathbb{E} \left\{ [A_{k+1}^\top P_{i,k+1} + D_{i,k+1}^\top Q_{i,k+1} + G_{i,k+1} X_{k+1}] W_k | \mathcal{F}_k \right\}, \\
P_{i,N-1} &= \mathbb{E}[R_i X_N^\top | \mathcal{F}_{N-1}], \\
Q_{i,N-1} &= \mathbb{E}[R_i X_N W_{N-1} | \mathcal{F}_{N-1}], \\
i &= 1, 2, \quad k = 0, 1, \cdots, N-2.
\end{aligned}
\]

The convexity conditions hold true. The MP condition turns out to be

\[
H_{iu_i}(k, X_k^*, u_{1,k}^*, u_{2,k}^*, P_{i,k}, Q_{i,k}) = 0, \quad k \in \mathcal{T}, i = 1, 2.
\]

Since \(L_{1,k}, \ell_{2,k} > 0\), it is equivalent to

\[
\begin{aligned}
u_{1,k}^* &= -L_{1,k}^{-1} (B_{1,k}^\top P_{i,k} + E_{1,k}^\top Q_{i,k}), \\
u_{2,k}^* &= -\ell_{2,k}^{-1} (B_{2,k}^\top P_{i,k} + E_{2,k}^\top Q_{i,k}).
\end{aligned}
\]

In the remaining part of this example, we consider the simple case when \(\{D_k\}\), \(\{E_{1,k}\}\) and \(\{E_{2,k}\}\) vanish. We call this special case Problem (NZG2). For this problem, a state-depending form of the equilibrium point \((u_1^*, u_2^*)\) is given. The
we can express
\[ u_{1,k}^* = -L_{1,k}^{-1} B_{1,k}^\top P_{1,k}, \quad u_{2,k}^* = -\ell_{2,k}^{-1} B_{2,k}^\top P_{2,k}, \] (33)
where
\[
\begin{align*}
P_{i,k} &= \mathbb{E} \left\{ [A_{i,k+1}^\top P_{i,k+1} + G_{i,k+1} X_{k+1}^\ast] | \mathcal{F}_k \right\}, \quad k = 0, 1, \ldots, N - 2, \\
P_{i,N-1} &= \mathbb{E} \left[ R_i X_N^\ast | \mathcal{F}_{N-1} \right].
\end{align*}
\]

In the same way, we can use induction to get
\[
E \left[ X_{k+1}^\ast | \mathcal{F}_k \right] = \Phi_k \sum_{i=1}^{k-1} E \left[ X_{i+1}^\ast | \mathcal{F}_i \right] + \psi_k, \quad k \in \mathcal{T},
\]
where
\[
\begin{align*}
\Phi_k &= I_n + B_{1,k} L_{1,k}^{-1} B_{1,k}^\top \alpha_{1,k} + B_{2,k} \ell_{2,k}^{-1} B_{2,k}^\top \alpha_{2,k}, \\
\psi_k &= C_k - B_{1,k} L_{1,k}^{-1} B_{1,k}^\top \beta_{1,k} - B_{2,k} \ell_{2,k}^{-1} B_{2,k}^\top \beta_{2,k}.
\end{align*}
\] (37)

Let us assume
\[
(A) \quad (\Phi_k)^{-1}
\]
exists and is bounded for each \( k \in \mathcal{T}. \)

Then
\[
E \left[ X_{k+1}^\ast | \mathcal{F}_k \right] = \phi_k X_k^\ast + \psi_k, \quad k \in \mathcal{T},
\]
with
\[
\phi_k = \Phi_k^{-1} A_k, \quad \psi_k = \Phi_k^{-1} \psi_k.
\] (39)

Consequently, by (36) and (38) we get
\[
\begin{align*}
\Phi_k \sum_{i=1}^{k-1} E \left[ X_{i+1}^\ast | \mathcal{F}_i \right] + \psi_k = \phi_k X_k^\ast + \psi_k.
\end{align*}
\] (40)

On the other hand, by (38), applying the tower property of conditional expectations yields
\[
E \left[ X_i^\ast | \mathcal{F}_k \right] = \phi_{i-1} E \left[ X_{i-1}^\ast | \mathcal{F}_k \right] + \phi_{i-1} \psi_{i-1}, \quad i = k + 2, k + 3, \ldots, N.
\]

In the same way, we can use induction to get
\[
E \left[ X_i^\ast | \mathcal{F}_k \right] = \phi_{i-1}^j E \left[ X_{j+1}^\ast | \mathcal{F}_k \right] + \sum_{j=k+1}^{i-1} \phi_{i-1}^j \psi_j, \quad i = k + 2, k + 3, \ldots, N,
\] (41)

where \( \phi_i^j \) is defined by: \( \phi_i^j = \phi_j \phi_{j+1} \cdots \phi_i \) if \( i < j \), and \( \phi_i^i = I_n \) if \( i = j \). That is, we can express \( E \left[ X_i^\ast | \mathcal{F}_k \right], i = k + 2, k + 3, \ldots, N, \) in terms of \( E \left[ X_{k+1}^\ast | \mathcal{F}_k \right] \) only.
Taking \( i = N \) and \( k = N - 2 \) in (41) gives
\[
\mathbb{E}[X_N^s | \mathcal{F}_{N-2}] = \phi_{N-1} \mathbb{E}[X_{N-1}^s | \mathcal{F}_{N-2}] + \psi_{N-1}.
\] (42)
Since \( P_{i,N-2} = A_i^{N-1} R_i \mathbb{E}[X_N^s | \mathcal{F}_{N-2}] + G_{i,N-1} \mathbb{E}[X_{N-1}^s | \mathcal{F}_{N-2}] \), by (42) we have
\[
P_{i,N-2} = A_i^{N-1} R_i \phi_{N-1} + G_{i,N-1} \mathbb{E}[X_{N-1}^s | \mathcal{F}_{N-2}] + A_i^{N-1} R_i \psi_{N-1}.
\]
Recall from (34) that \( P_{i,k} \) satisfies that
\[
\begin{align*}
P_{i,k} &= (A_i^{N-1})^\top R_i \mathbb{E}[X_N^s | \mathcal{F}_k] + \sum_{j=k+2}^{N-1} (A_j^{N-1})^\top G_{i,j} \mathbb{E}[X_j^s | \mathcal{F}_k] + G_{i,k+1} \mathbb{E}[X_{k+1}^s | \mathcal{F}_k], \\
\{ & k = 0, 1, \ldots, N - 3, \\
P_{i,N-2} &= [(A_{N-1})^\top R_i \phi_{N-1} + G_{i,N-1} \mathbb{E}[X_{N-1}^s | \mathcal{F}_{N-2}] + (A_{N-1})^\top R_i \psi_{N-1}], \\
P_{i,N-1} &= R_i \mathbb{E}[X_N^s | \mathcal{F}_{N-1}].
\end{align*}
\]
In view of (35) and (41), it follows that
\[
\begin{align*}
\alpha_{i,N-1} \mathbb{E}[X_N^s | \mathcal{F}_{N-1}] + \beta_{i,N-1} &= R_i \mathbb{E}[X_N^s | \mathcal{F}_{N-1}], \\
\alpha_{i,N-2} \mathbb{E}[X_N^s | \mathcal{F}_{N-2}] + \beta_{i,N-2} &= [(A_{N-1})^\top R_i \phi_{N-1} + G_{i,N-1} \mathbb{E}[X_{N-1}^s | \mathcal{F}_{N-2}] + (A_{N-1})^\top R_i \psi_{N-1}], \\
\alpha_{i,k} \mathbb{E}[X_{k+1}^s | \mathcal{F}_k] + \beta_{i,k} &= G_{i,k+1} \mathbb{E}[X_{k+1}^s | \mathcal{F}_k] + \sum_{j=k+2}^{N-1} (A_j^{N-1})^\top G_{i,j} [\phi_j^{N-1} \mathbb{E}[X_j^s | \mathcal{F}_k] + \sum_{s=k+1}^{j-1} \phi_s^{N-1} \psi_s] \\
&\quad + (A_k^{N-1})^\top R_i [\phi_k^{N-1} \mathbb{E}[X_{k+1}^s | \mathcal{F}_k] + \sum_{j=k+1}^{N-1} \phi_j^{N-1} \psi_j], \quad k = 0, 1, \ldots, N - 3.
\end{align*}
\]
Comparing the coefficients leads to
\[
\begin{align*}
\alpha_{i,N-1} &= R_i, \quad \beta_{i,N-1} = 0, \\
\alpha_{i,N-2} &= G_{i,N-1} + A_i^{N-1} R_i \phi_{N-1}, \quad \beta_{i,N-2} = A_i^{N-1} R_i \psi_{N-1}, \\
\alpha_{i,k} &= G_{i,k+1} + \sum_{j=k+2}^{N-1} (A_j^{N-1})^\top G_{i,j} \phi_k^{j-1} + (A_k^{N-1})^\top R_i \phi_k^{N-1}, \\
\beta_{i,k} &= \sum_{k+2}^{N-1} (A_j^{N-1})^\top G_{i,j} \sum_{s=k+1}^{j-1} \phi_s^{j-1} \psi_s + (A_k^{N-1})^\top R_i \sum_{j=k+1}^{N-1} \phi_j^{N-1} \psi_j, \\
&\quad k = 0, 1, \ldots, N - 3, \quad i = 1, 2.
\end{align*}
\] (43)

**Proposition 2.** Let \( \{\Phi_k, \Psi_k\} \), \( \{\phi_k, \psi_k\} \) and \( \{\alpha_{i,k}, \beta_{i,k}\} \), \( i = 1, 2 \) be defined by (37), (39) and (43), respectively, and assume that the assumption (A) holds. Then \( \{u_1^*, u_2^*\} \) defined by (40) is the equilibrium point of Problem (NZG2).

**Remark 4.** Note that no Riccati equation is introduced in this example. A trick is used to get the equilibrium point (40). In fact, we connect \( P_{i,k} \) with \( \mathbb{E}[X_{k+1}^s | \mathcal{F}_k] \) through (35) in the first step. However, if (35) is changed to \( P_{i,k} = \alpha_{i,k} X_k^s + \beta_{i,k} \) which connects \( P_{i,k} \) with \( X_k^s \) directly, then the same method will lead to a mess.

To conclude this example, let us give some numerical results. We only consider one-dimensional case. Assume that the coefficients of the system and the cost functional are time-invariant. Take \( A = B_1 = B_2 = C = F = G_1 = L_1 = l_2 = R_2 = 1, l_1 = R_1 = G_2 = L_2 = 0 \), and \( N = 200 \). Figures 1-4 show respectively the sequences \( \{\alpha_{1,k}\} \) and \( \{\beta_{1,k}\} \), \( \{\alpha_{2,k}\} \) and \( \{\beta_{2,k}\} \), \( \{\Psi_k\} \) and \( \{\Phi_k\} \), and \( \{\psi_k\} \) and \( \{\phi_k\} \). With these preparations, the equilibrium point \( \{u_1^*, u_2^*\} \) could be obtained through (40).
Figure 1. The sequences \( \{\alpha_{1,k}\} \) and \( \{\beta_{1,k}\} \)

Figure 2. The sequences \( \{\alpha_{2,k}\} \) and \( \{\beta_{2,k}\} \)

Figure 3. The sequences \( \{\Psi_k\} \) and \( \{\Phi_k\} \)
5. Comparisons and concluding remarks. As stated before, [13] and [16] are two closely related papers.

There are some distinct differences between [13] and our present paper. In [13], the noise is assumed to be a discrete-time martingale with independent increments. Based on this assumption, the solution of the adjoint difference equation, which is a triple of processes, is derived by the Galtchouk-Kunita-Watanabe decomposition. While in our paper, the noises are random variables without martingale property and the adjoint equation, whose solution is a pair of processes, is derived in another way. In addition, [13] assumes that the functions involving in the cost functional have bounded partial derivatives, which excludes quadratic functions.

Compared with [16], our present paper has some advantages. The stochastic optimal control problem studied in our paper is more general. Firstly, the control domains can be general convex sets, and the control domains of different times, i.e. $U_0, U_1, \ldots, U_{N-1}$, can be different sets. Secondly, the coefficients of the state equation and the cost functional can be time-varying and random, and neither the twice differentiability of the coefficients nor the boundedness of $h_x(x)$ is needed in our paper. Thirdly, under (H1) and (H2) in our paper, the square-integrability assumption of the noises and admissible controls is not sufficient in general since the solution of the adjoint equation may not admit necessary integrability. Fourthly, the necessary condition in our paper (Theorem 2.6) is more general, and the sufficient condition (Theorem 2.7) is more clear and simple. Besides, the method is also extended to two kinds of discrete-time stochastic games in our paper.

This paper studies the MP for discrete-time stochastic optimal control problems and stochastic games. There are some features worthy of being highlighted. Firstly, a rigorous discrete-time stochastic MP is obtained in a clear and concise way. Secondly, a difficulty arising from the absence of necessary integrability of the solution to the adjoint equation is overcome by supposing higher yet in general necessary integrability assumptions on the noises and the admissible controls. Thirdly, two kinds of discrete-time stochastic games are studied with the MP method, which, to our best knowledge, are new in the literature. Fourthly, one kind of investment/consumption choice problem is studied, which illustrates the extensive applicability of our theoretical results in the financial market. Last but not least, one kind of discrete-time LQ nonzero-sum stochastic game is studied, for which the
equilibrium point is obtained by using the MP method and without introducing the Riccati equation.

Acknowledgments. The authors would like to thank the editor and two anonymous referees for their constructive and insightful comments for improving the quality of this work. The first author is supported by the Natural Science Foundation of China (11831010,61961160732), and Shandong Provincial Natural Science Foundation (ZR2019ZD42). The second author is supported by the Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions (1716009), the Special Funds of Taishan Scholar Project (tsqn20161041), and the Colleges and Universities Youth Innovation Technology Program of Shandong Province (2019KJI011).

REFERENCES

[1] T. T. K. An and B. Øksendal, Maximum principle for stochastic differential games with partial information, *J. Optim. Theory Appl.*, 139 (2008), 463–483.
[2] A. Beghi and D. D’Alessandro, Discrete-time optimal control with control-dependent noise and generalized Riccati difference equations, *Automatica*, 34 (1998), 1031–1034.
[3] L. Chen and Z. Y. Yu, Maximum principle for nonzero-sum stochastic differential game with delays, *IEEE Trans. Automat. Control*, 60 (2015), 1422–1426.
[4] S. N. Cohen and R. J. Elliott, A general theory of finite state backward stochastic difference equations, *Stoch. Proc. Appl.*, 120 (2010), 442–466.
[5] S. N. Cohen and R. J. Elliott, Backward stochastic difference equations and nearly time-consistent nonlinear expectations, *SIAM J. Control Optim.*, 49 (2011), 125–139.
[6] O. L. V. Costa and A. de Oliveira, Optimal mean-variance control for discrete-time linear systems with Markovian jumps and multiplicative noises, *Automatica*, 48 (2012), 304–315.
[7] K. Du and Q. X. Meng, A maximum principle for optimal control of stochastic evolution equations, *SIAM J. Control Optim.*, 51 (2013), 4343–4362.
[8] R. J. Elliott, X. Li and Y. H. Ni, Discrete time mean-field stochastic linear-quadratic optimal control problems, *Automatica*, 49 (2013), 3222–3233.
[9] H. Halkin, A maximum principle of the Pontryagin type for systems described by nonlinear difference equations, *SIAM J. Control Optim.*, 4 (1966), 90–111.
[10] Y. C. Han, S. G. Peng and Z. Wu, Maximum principle for backward doubly stochastic control systems with applications, *SIAM J. Control Optim.*, 48 (2010), 4224–4241.
[11] E. C. M. Hui and H. Xiao, Maximum principle for differential games of forward-backward stochastic systems with applications, *J. Math. Anal. Appl.*, 386 (2012), 412–427.
[12] R. Isaacs, *Differential Games*, John Wiley and Sons, New York, 1965.
[13] S. L. Ji and H. D. Liu, Maximum principle for stochastic optimal control problem of forward-backward stochastic difference systems, *Int. J. Control.*, (2021).
[14] X. S. Jiang, S. P. Tian, T. L. Zhang and W. H. Zhang, Stability and stabilization of nonlinear discrete-time stochastic systems, *Int. J. Robust Nonlin.*, 29 (2019), 6419–6437.
[15] D. Li and C. W. Schmidt, Cost smoothing in discrete-time linear-quadratic control, *Automatica*, 33 (1997), 447–452.
[16] X. Y. Lin and W. H. Zhang, A maximum principle for optimal control of discrete-time stochastic systems with multiplicative noise, *IEEE Trans. Automat. Control*, 60 (2015), 1121–1126.
[17] Q. Liu and X. Zhang, *General Pontryagin-type Stochastic Maximum Principle and Backward Stochastic Evolution Equations in Infinite Dimensions*, SpringerBriefs in Mathematics, Springer, Cham, 2014.
[18] J. B. Moore, X. Y. Zhou and A. E. B. Lim, Discrete time LQG controls with control dependent noise, *Syst. Control Lett.*, 36 (1999), 199–206.
[19] Y. H. Ni, R. J. Elliott and X. Li, Discrete-time mean-field stochastic linear-quadratic optimal control problems, II: Infinite horizon case, *Automatica*, 57 (2015), 65–77.
[20] M. Pachter and K. D. Pham, Discrete-time linear-quadratic dynamic games, *J. Optim. Theory Appl.*, 146 (2010), 151–179.
[21] P. Paruchuri and D. Chatterjee, Discrete time pontryagin maximum principle under state-action-frequency constraints, *IEEE Trans. Automat. Control*, 64 (2019), 4202–4208.
[22] S. G. Peng, A general stochastic maximum principle for optimal control problems, *SIAM J. Control Optim.*, **28** (1990), 966–979.

[23] M. A. Rami, X. Chen and X. Y. Zhou, Discrete-time indefinite LQ control with state and control dependent noises, *J. Global Optim.*, **23** (2002), 245–265.

[24] H. Y. Sun, L. Y. Jiang and W. H. Zhang, Infinite horizon linear quadratic differential games for discrete-time stochastic systems, *J. Optim. Theory Appl.*, **10** (2012), 391–396.

[25] G. C. Wang and Z. Y. Yu, A Pontryagin’s maximum principle for non-zero sum differential games of BSDEs with applications, *IEEE Trans. Automat. Control*, **55** (2010), 1742–1747.

[26] H. X. Wang, H. S. Zhang and X. Wang, Optimal control for stochastic discrete-time systems with multiple input-delays, in *Proc. 10th World Congress on Intelligent Control and Automation*, Beijing, (2012), 1529–1534.

[27] Z. Wu, A general maximum principle for optimal control of forward-backward stochastic systems, *Automatica*, **49** (2013), 1473–1480.

[28] H. S. Zhang and X. Zhang, Second-order necessary conditions for stochastic optimal control problems, *SIAM Rev.*, **60** (2018), 139–178.

[29] W. H. Zhang, Y. L. Huang and H. S. Zhang, Stochastic $H_2/H_\infty$ control for discrete-time systems with state and disturbance dependent noise, *Automatica*, **43** (2007), 513–521.

[30] X. Zhang, R. J. Elliott and T. K. Siu, A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance, *SIAM J. Control Optim.*, **50** (2012), 964–990.

Received October 2020; revised February 2021.

E-mail address: wuzhen@sdu.edu.cn
E-mail address: zhangfeng1104@sdufe.edu.cn