On Completely Mixed Stochastic Games

Purba Das · T. Parthasarathy · G. Ravindran

Received: 8 June 2021 / Accepted: 6 June 2022 / Published online: 22 September 2022
© The Author(s) 2022

Abstract
In this paper, we consider a two-person finite state stochastic games with finite number of pure actions for both players in all the states. In particular, for a large number of results we also consider one-player controlled transition probability and show that if all the optimal strategies of the undiscounted stochastic game are completely mixed then for \(\beta\) sufficiently close to 1; all the optimal strategies of \(\beta\)-discounted stochastic games are also completely mixed. A counterexample is provided to show that the converse is not true. Further, for single-player controlled completely mixed stochastic games if the individual payoff matrices are symmetric in each state, then we show that the individual matrix games are also completely mixed. For the non-zerosum single-player controlled stochastic game under some non-singularity conditions, we show that if the undiscounted game is completely mixed, then the Nash equilibrium is unique. For non-zerosum \(\beta\)-discounted stochastic games when Nash equilibrium exists, we provide equalizer rules for corresponding value of the game.

Keywords Undiscounted stochastic game · \(\beta\)-discounted stochastic game · Limiting average payoff · \(\beta\)-discounted payoff · Completely mixed game · Value of a game · Single-player controlled transition · Zero-sum and non-zerosum stochastic game · Matrix game · Nash equilibrium

Purba Das
purba.das@maths.ox.ac.uk

T. Parthasarathy
pacha14@yahoo.com

G. Ravindran
ravi@isichennai.res.in

1 Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK
2 Chennai Mathematical Institute, Chennai, India
3 Indian Statistical Institute, Chennai Centre, Chennai, India
1 Introduction

Stochastic games (zero-sum) were first introduced by Shapley [1]. For zero-sum, $\beta$-discounted stochastic games he proved the existence of a “value” and showed the existence of stationary optimal strategies. The undiscounted (zero-sum) stochastic games was first introduced by Gillette [2]. He showed that, unlike $\beta$-discounted stochastic games, the undiscounted game may not process stationary optimal strategies. For further references look Blackwell and Ferguson [3]. Kaplansky [4] introduced the completely mixed matrix games and gave characterisations of such games. Later, Filar [5, 6] introduced the completely mixed stochastic games and extend some of the results of completely mixed matrix game to the completely mixed undiscounted stochastic game under the assumption of single-player controlled transition probabilities.

Nonzero sum discounted stochastic games were studied by Fink [7] and Takahashi [8] where they showed the existence of a stationary Nash equilibrium in the finite state case. Undiscounted zerosum stochastic games were introduced by Gillette [2] who showed the nonexistence of stationary optimal strategy using a nice example. For more details regarding the undiscounted problem see Blackwell-Ferguson [3] and Mertens-Neyman [9]. Parthasarathy and Raghavan introduced the notion of one-player controlled stochastic games [10] and have shown the existence of stationary optimal strategies for both discounted and undiscounted stochastic games. A similar result is also shown for the nonzero-sum bimatrix games. Kaplansky [4] introduced the concept of completely mixed matrix games. He showed such matrix games possess unique completely mixed optimal strategies for both players. An analogous result was shown by Raghavan for completely mixed bimatrix games [11].

Outline We consider a two-person finite state stochastic game with a finite number of pure actions for both players in all the states. For a large number of results in this paper, we consider a single-player controlled stochastic game.

In Sect. 2 we provide the related definitions of stochastic games and Nash equilibrium. For further details, one can refer to [10, 12].

In Sect. 3, we consider zero-sum stochastic games and show under the assumption of one-player controlled transition if the undiscounted stochastic game is completely mixed then for all $\beta$ sufficiently close to 1, the $\beta$-discounted stochastic games are completely mixed (Theorem 1). Example 1 shows converse is not true in general.

In Sect. 4 we consider non-zerosum two-person stochastic games. Under some regularity assumptions on the individual payoff matrices, the undiscounted completely mixed stochastic game possesses a unique Nash equilibrium (Theorem 7). Further, we show both for undiscounted and for $\beta$-discounted stochastic games if there exists a completely mixed Nash equilibrium (NE) then the value of the game follows an equalizer rule (Theorems 7, 8, and 10). We provide examples to illustrate our results. The paper ends with some open problems.
2 Definitions and Preliminaries

For a finite state and finite action two players stochastic game, the game is played with two players (known as player-1 and player-2) and every day. Each day, the game is in a specific state \( s \) and both players will play a matrix game \( R(s) \) (only depends on the state \( s \)) and will get some reward, which can be positive, negative or zero. The game will then move to a new state the next day and continue indefinitely.

Definition 1 (Finite two-person stochastic game) A two-person finite state stochastic game is a 6 tuple \( G := (S, A_1, A_2, r^1, r^2, q) \). The game is played as follows. \( S \) is the (finite) set of all states in the stochastic game. As state space is finite wlog \( S = \{s_1, s_2, \cdots, s_K\} \). \( A_1 \) (respectively \( A_2 \)) is the set of all pure actions available for player-1 (respectively player-2) in each state. That is, in state \( s \in S \) player \( i \) for \( i = \{1, 2\} \) has pure actions (finite) \( A_i(s) = \{1, 2, \cdots, m_i\} \). \( r^1(s, i, j) \) and \( r^2(s, i, j) \) are the reward functions for respective players and \( q(s' | s, i, j) \) is the corresponding transition probability function. If in state \( s \in S \), player-1 and player-2 choose pure actions \( i \) and \( j \) respectively, then the payoff for player-1 on that specific day is \( r^1(s, i, j) \) and the payoff for player-2 is \( r^2(s, i, j) \). With transition probability \( q(s' | s, i, j) \) the game moves to state \( s' \) on the next day and the game continues.

For the zero-sum stochastic game \( r^2(s, i, j) = -r^1(s, i, j) \) for all \( s \in S \), and \( i, j \). (in such a case we assume \( r^1 = -r^2 = r \).

The payoff matrices for player-1 and player-2 respectively in state \( s \in S \) are denoted as follows.

\[
R^1(s) = (r^1(s, i, j))_{m_1 \times m_2}, \quad \text{and} \quad R^2(s) = (r^2(s, i, j))_{m_1 \times m_2}.
\]

In general, the strategy for a player may depend on the whole history up to today, but we will only be considering those strategies which do not depend on the previous history.

Definition 2 (Stationary strategy) Denote \( P_{A_1} \) (respectively \( P_{A_2} \)), be the set of all probability distribution of player 1’s (respectively player-2) action space \( A_1 \) (respectively \( A_2 \)). Then a stationary strategy of player \( k \in \{1, 2\} \) is a function from state space \( S \) to the probability space \( P_{A_k} \), independent of where it might occur in the infinite game.

Definition 3 (Undiscounted payoffs for stochastic game) For \( i \in \{1, 2\} \) we denote \( r_i^{(n)}(f, g, s_0) \) as the expected immediate reward for player \( i \) on \( n^{th} \) day, if the game starts in state \( s_0 \) and player-1, player-2 choose stationary strategies \( f \) and \( g \) respectively.

So the undiscounted payoff player-1 and player-2 get in state \( s_0 \in S \) is \( \Phi^1(f, g)(s_0) \) and \( \Phi^2(f, g)(s_0) \) respectively. Where,
The “undiscounted payoff” is sometimes also known as limiting average payoff. For the zero-sum stochastic game $\Phi^2(f, g)(s_0) = -\Phi^1(f, g)(s_0)$ (For convenience denote $\Phi^1 = \Phi$).

**Definition 4 (β-discounted payoffs for general-sum stochastic game)** If player-1 and player-2 play stationary strategies $f$ and $g$ respectively then for $\beta \in [0, 1)$, the $\beta$-discounted payoff in state $s_0 \in S$ for player-1 and player-2 are $I^1_\beta(f, g)(s_0)$ and $I^2_\beta(f, g)(s_0)$ respectively. Where,

$$I^1_\beta(f, g)(s_0) = \sum_{n=0}^{\infty} \beta^n r^{(n)}_1(f, g, s_0),$$

$$I^2_\beta(f, g)(s_0) = \sum_{n=0}^{\infty} \beta^n r^{(n)}_2(f, g, s_0).$$

Where, $r_i^{(n)}(f, g, s_0)$ for $i \in \{1, 2\}$ is the expected immediate reward of player $i$ on $n^{th}$ day in state $s_0$, when player-1 and player-2 play the strategies $f$ and $g$, respectively.

For the zero-sum stochastic game we have $I^2_\beta(f, g)(s_0) = -I^1_\beta(f, g)(s_0)$ (For convenience assume $I^1_\beta = I_\beta$).

**Definition 5 (Optimal strategy and value of zero-sum stochastic game)** A pair of stationary strategies $(f^0, g^0)$ is said to be an optimal strategy in the zero-sum undiscounted stochastic game if, for all $f \in \mathbb{P}_1$ and $g \in \mathbb{P}_2$,

$$\Phi(f, g^0) \leq \Phi(f^0, g^0) \leq \Phi(f^0, g) \quad \text{coordinate-wise},$$

where $\Phi(f, g) = (\Phi(f, g)(s_1), \Phi(f, g)(s_2), \ldots, \Phi(f, g)(s_K))^T$. The value of the undiscounted stochastic game in state $s \in S$ is denoted as $v(s) = \Phi(f^0, g^0)(s)$.

Similarly a pair of stationary strategies $(f^0, g^0)$ is said to be an optimal stationary strategy for the zero-sum $\beta$-discounted stochastic game if, for all $f \in \mathbb{P}_{A_1}$ and for all $g \in \mathbb{P}_{A_2}$,

$$I_\beta(f, g^0) \leq I_\beta(f^0, g^0) \leq I_\beta(f^0, g) \quad \text{coordinate-wise}$$

where $I_\beta(f, g) = (I_\beta(f, g)(s_1), I_\beta(f, g)(s_2), \ldots, I_\beta(f, g)(s_K))^T$. The value of the $\beta$-discounted stochastic game in state $s \in S$ is denoted as $v_\beta(s) = I_\beta(f^0, g^0)(s)$.

We call $f^0$ (respectively $g^0$) optimal for player-1 (respectively player-2) in the $\beta$-discounted game if $I_\beta(f^0, g^0)(s) \geq \min_g \max_f I_\beta(f, g)$ (and $I_\beta(f, g^0)(s) \geq \max_g \min_f I_\beta$).
\((f, g)(s)\) for all \(g\) (for all \(f\)) and \(s \in S\). Analogous definition holds for undiscounted stochastic games.

The value of a zero-sum stochastic game (both discounted and undiscounted) always exists and is unique \([1, 2]\)), whereas the optimal strategy of a stochastic game is not necessarily unique.

**Definition 6 (Nash Equilibrium)** A pair of stationary strategy \((f^0, g^0)\) for player-1 and player-2 respectively is said to be Nash Equilibrium (NE) for a nonzero-sum undiscounted stochastic game if,

\[
\forall f \in P_{A_1}, \quad \Phi^1(f^0, g^0) \geq \Phi^1(f, g^0) \text{ coordinate-wise and,} \\
\forall g \in P_{A_2}, \quad \Phi^2(f^0, g^0) \geq \Phi^2(f^0, g) \text{ coordinate-wise},
\]

where \(\Phi^k(f, g) = (\Phi^k(f, g)(s_1), \Phi^k(f, g)(s_2), \ldots, \Phi^k(f, g)(s_K))^T\) for \(k \in \{1, 2\}\) (Assuming both player-1 and player-2 want to maximize their expected payoffs).

The value of the game associated with the NE \((f^0, g^0)\) in state \(s \in S\) is given by \(v^k_{f^0, g^0}(s) = \Phi^k(f^0, g^0)(s)\) for \(k \in \{1, 2\}\).

Similarly, a pair of strategies \((f^0, g^0)\) for player-1 and player-2 respectively is said to be a Nash equilibrium (NE) in the \(\beta\)-discounted stochastic game if

\[
I^1_\beta(f^0, g^0) \geq I^1_\beta(f, g^0) \text{ coordinate-wise } \forall f \in P_{A_1} \text{ and,} \\
I^2_\beta(f^0, g^0) \geq I^2_\beta(f^0, g) \text{ coordinate-wise } \forall g \in P_{A_2},
\]

where \(I^k_\beta(f, g) = (I^k_\beta(f, g)(s_1), I^k_\beta(f, g)(s_2), \ldots, I^k_\beta(f, g)(s_K))^T\) for \(k \in \{1, 2\}\) (Assuming both player-1 and player-2 want to maximize their expected payoff).

The corresponding value of the \(\beta\)-discounted game associated with the NE \((f^0, g^0)\) in state \(s \in S\) is given by \(v^k_{\beta, f^0, g^0}(s) = I^k_\beta(f^0, g^0)(s)\) for \(k \in \{1, 2\}\).

The value of a non-zero-sum stochastic game is not necessarily unique, it changes with the associated Nash Equilibrium.

**Definition 7 (Single-player controlled stochastic game [10])** A stochastic game is said to be a single-player controlled stochastic game if the transition probability is controlled by only one player. For a player-2 controlled stochastic games, we have \(q(s'|s, i) = q(s'|s, j)\) for all \(s, s' \in S\), pure action \(i\) of player-1 and \(j\) of player-2.

Realistic examples of one-player controlled transition can be found in \([12, Section 3.6, page 119]\). Throughout the rest of the paper, unless mentioned otherwise we will be considering a player-2 controlled stochastic game.

**Definition 8 (Completely mixed stochastic game [6])** A stochastic game is said to be completely mixed if every optimal strategy (Nash equilibrium) for both the players are completely mixed. That is for both player-1 and player-2 in each state \(s \in S\) all the pure actions \(i\) and \(j\) for player-1 and player-2 respectively are played with strictly positive probabilities.
The transition probability matrix \( Q(f, g) \) for some stationary strategy \( f \) of player-1 and \( g \) of player-2, is defined as
\[
Q(f, g) = (q(s'|s,f,g))_{K \times K}.
\]
Where, \( q(s'|s,f,g) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} f_i(s) q(s'|s,i,j) g_j(s) \). Under the assumption of player-2 controlled stochastic game, \( Q(f, g) \) becomes \( Q(g) = (q(s'|s,g))_{K \times K} \).

For a stationary strategy pair (or Nash equilibrium) \((f, g)\) the reward vector is defined as
\[
r^1(f, g) = (r^1(f, g, s_1), \ldots, r^1(f, g, s_K))^t
\]
\[
r^2(f, g) = (r^2(f, g, s_1), \ldots, r^2(f, g, s_K))^t
\]
with, \( r^k(f, g, s) = f(s)^t R^k(s) g(s) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} f_i(s) r^k(i,j,s) g_j(s) \) for \( k \in \{1, 2\} \).

Now the discounted and undiscounted payoffs for the stationary strategy (or Nash equilibrium) \((f, g)\) for player-1 and player-2 respectively in state \( s \in S \) can be written as-
\[
I^k_{\Gamma}(f, g)(s) = \{I - \beta Q(f, g)\}^{-1} r^k(f, g) \}
\]
\[
\Phi^k(f, g)(s) = \{Q^*(f, g) r^k(f, g)\}_{s}
\]
Where, \( k \in \{1, 2\} \), \( Q^0(f, g) = I \) and the markov matrix \( Q^*(f, g) \) is as follows.
\[
Q^*(f, g) = \lim_{n \to \infty} \frac{1}{N+1} \sum_{n=0}^{N} Q^n(f, g).
\]
And the notation \( \{.\}_{s} \) represents the \( s^{th} \) coordinate of the respective vector.

**Definition 9 (Uniform discount optimal)** A strategy \( g^0 \) for player-2 is said to be an uniform discount optimal if \( g^0 \) is optimal for player-2 in the \( \beta \)-discounted game \( \Gamma^\beta_\beta \) for all \( \beta \) close to 1. Similarly, we can extend the definition for player-1 also.

**Definition 10 (Auxiliary game)** For a \( \beta \)-discounted zero-sum stochastic game, the auxiliary game (starting) at state \( s \in S \) is defined by the matrix \( R^\beta_\beta(s) \). The \((i,j)^{th}\) entry of the matrix \( R^\beta_\beta(s) \) is given by
\[
r(i,j, s) + \beta \sum_{s' \in S} v^\beta(s') q(s'|s,i,j)
\]
where, \( v^\beta(s') \) is the value of the \( \beta \)-discounted stochastic game at state \( s' \in S \). The matrix \( R^\beta_\beta(s) \) is known as, **Shapley Matrix** [1].

The following results will be required to prove our results.

**Result 1 (Kaplansky [4], Theorem 1 page 475)** Consider a two-person zero sum matrix game with payoff matrix \( M \in \mathbb{R}^{mxn} \). Suppose player-2 has a completely mixed optimal strategy \( y' \) then for any optimal strategy \( x' \) for player-1, we have
\[ \sum_{i=1}^{m} m_{ij}x_i' \equiv v \text{ for all } j \in \{1, 2, \ldots, n\}, \text{ where } v \text{ is the value of the matrix game and } m_{ij} \text{ is the } (i,j)^{th} \text{ entry of the matrix } M. \]

Shapley showed that for \( \beta \)-discounted stochastic games, the value of a game exists and is unique.

**Result 2** (Shapley Matrix [1]) A two-player zero-sum \( \beta \)-discounted stochastic game \( \Gamma_\beta = (S, A_1, A_2, r, q, \beta) \) has an optimal value vector \( v_\beta \) which is obtained by

\[ v_\beta(s) = \text{val}(R_\beta(s)), \]

where \( R_\beta(s) \) is the shapley matrix in state \( s \).

For each state \( s \in S, (f_\beta(s), g_\beta(s)) \) is a pair of optimal strategies of the matrix game \( R_\beta(s) \), if and only if \( (f_\beta, g_\beta) \) is a pair of optimal strategies for the \( \beta \)-discounted stochastic game \( \Gamma_\beta \), where \( f_\beta = (f_\beta(s_1), f_\beta(s_2), \ldots, f_\beta(s_K)) \) and \( g_\beta = (g_\beta(s_1), g_\beta(s_2), \ldots, g_\beta(s_K)) \) and \( K \) is the number of states.

**Result 3** (Parthasarathy and Raghavan [10]) For player-2 controlled stochastic game (both zero-sum and non-zero-sum) if \( f_\beta \rightarrow f^0 \) component wise then, for \( k = \{1, 2\} \),

\[ \lim_{\beta \uparrow 1} (1 - \beta) I^k_\beta(f_\beta, g) \equiv \Phi^k(f^0, g) \]

where, \( g \) is any stationary strategy for player-2. Parthasarathy and Raghavan [10] also provides,

\[ v^k(s) = \lim_{\beta \uparrow 1} (1 - \beta) v^k_\beta(s) \text{ for all } s \in S \text{ and } k \in \{1, 2\}. \]

### 3 Zero-Sum Stochastic Games

Unless mentioned otherwise for Section 3, player-1 is the maximizing player and player-2 is the minimizing player. The following lemma is required to prove our main theorem.

**Lemma 1** Under player-2 controlled transition, suppose there exists \( \beta_0 \in [0, 1) \) and a completely mixed stationary strategy \( g^0 \) such that \( g^0 \) is optimal for player-2 (Minimizer) for every \( \beta \)-discounted stochastic game with \( \beta > \beta_0 \). Let \( \beta_n \in [\beta_0, 1) \) be such that \( \beta_n \uparrow 1 \). Let \( \{f_n\} \) be stationary optimal for player-1 (Maximizer) in the \( \beta_n \)-discounted stochastic games. Suppose \( f_n \rightarrow f_0 \) coordinate-wise, that is \( f_n(s) \rightarrow f_0(s) \) for each state \( s \in S \), then \( f_0 \) is optimal for player-1 in the undiscounted stochastic game.

**Proof** For player-2 controlled undiscounted stochastic game the value exists and is restricted to stationary strategy [10].
Since \((f_n, g^0)\) is an optimal strategy pair in the \(\beta_n\)-discounted stochastic game, using Results 1 and 2 we have:

\[
\forall \text{ stationary strategy } g \text{ of player-2; } I_{\beta_n}(f_n, g) \equiv v_{\beta_n}
\]

where, \(I_{\beta_n}(f_n, g) = (I_{\beta_n}(f_n, g)(s_1), \ldots, I_{\beta_n}(f_n, g)(s_K))', \) and \(v_{\beta_n} = (v_{\beta_n}(s_1), \ldots, v_{\beta_n}(s_K))'.\) Therefore, we have

\[
[I - \beta_n Q(g)]^{-1}r(f_n, g) \equiv v_{\beta_n}.
\]

Since, \([I - \beta_n Q(g)]^{-1} = \sum_{k=0}^{\infty} \beta_n^k Q^k(g)\) is a non-negative matrix, \(r(f_n, g)\) can be expresses as follows.

\[
r(f_n, g) \equiv [I - \beta_n Q(g)]v_{\beta_n}.
\]

\(f_n \to f_0\) point-wise and the reward function \(r(., .)\) is a continuous function on player-1’s strategy. Hence, for any given \(\epsilon > 0\) there exists \(N_0 \in \mathbb{N}\) such that for all \(n \geq N_0:\)

\[
[I - \beta_n Q(g)]v_{\beta_n} - e \epsilon \leq r(f_0, g) \leq [I - \beta_n Q(g)]v_{\beta_n} + e \epsilon
\]

coordinate-wise, where \(e\) is a suitable length column vector with all entry as 1. Therefore we have,

\[
v_{\beta_n} - \epsilon[I - \beta_n Q(g)]^{-1}e \leq [I - \beta_n Q(g)]^{-1}r(f_0, g) \leq v_{\beta_n} + \epsilon[I - \beta_n Q(g)]^{-1}e.
\]

As \((1 - \beta_n)\) is always non-negative for all \(\beta_n \in (0, 1),\) we have:

\[
(1 - \beta_n)v_{\beta_n} - (1 - \beta_n)e[I - \beta_n Q(g)]^{-1}e \leq (1 - \beta_n)[I - \beta_n Q(g)]^{-1}r(f_0, g) \leq (1 - \beta_n)v_{\beta_n} + (1 - \beta_n)e[I - \beta_n Q(g)]^{-1}e.
\]

We have, \([I - \beta_n Q(g)]^{-1}e = [\sum_{k=0}^{\infty} \beta_n^k Q^k(g)]e = \sum_{k=0}^{\infty} \beta_n^k Q^k(g)e = \sum_{k=0}^{\infty} \beta_n^k e\) as \(Q^k\) is a stochastic matrix for each \(k.\) Therefore the above inequality reduces to the following inequality.

\[
(1 - \beta_n)v_{\beta_n} - \epsilon e \leq (1 - \beta_n)[I - \beta_n Q(g)]^{-1}r(f_0, g) \leq (1 - \beta_n)v_{\beta_n} + \epsilon e.
\]

Now if we let \(\beta_n \uparrow 1,\) Using Parthasarathy and Raghavan (Result 3) we can conclude the following.

\[
v - \epsilon e \leq \Phi(f_0, g) \leq v + \epsilon e,
\]

where \(v = (v(s_1), \ldots, v(s_K))'\) is the value of the undiscounted stochastic game. Since \(\epsilon > 0\) is arbitrary, \(f_0 = \lim f_n\) is an optimal strategy for player-1 in the undiscounted stochastic game, concluding the proof of the lemma. \(\blacksquare\)

If an undiscounted single-player stochastic game is completely mixed then we conclude that for all \(\beta\) sufficiently close to 1, the \(\beta\)-discounted stochastic games are completely mixed.
**Theorem 1 (Main theorem)** Consider a finite, undiscounted, zero-sum, single-player controlled, completely mixed stochastic game $\Gamma$. Then there exists $\beta_0 \in [0, 1)$ such that for all $\beta > \beta_0$, the $\beta$-discounted stochastic games $\Gamma_\beta$ (obtained from the same payoff matrices) are completely mixed.

**Proof** We will prove the statement by contradiction. If possible let us assume, the undiscounted stochastic game is completely mixed but there does not exist any $0 \leq \beta_0 < 1$ such that all the $\beta$-discounted games are completely mixed for all $\beta > \beta_0$.

Then given any $\beta \in [0, 1)$ we can find a $\beta_1 \in (\beta_0, 1)$ such that there exists an optimal strategy $f_1$, that is not completely mixed for the $\beta_1$-discounted stochastic game. Similarly, we can find a $\beta_2 \in (\beta_1, 1)$ such that $f_2$ is not a completely mixed optimal strategy for the $\beta_2$-discounted stochastic game. Hence we will obtain a sequence $\beta_n \uparrow 1$ such that for all $\beta_n$-discounted games we have a sequence of optimal strategy $f_n$ which are not completely mixed. But the stochastic game has finitely many states and finitely many pure actions, so as a consequence there exists at-least one state $\bar{s} \in S$ and a pure action $i$ for player-1 such that the $i^{th}$ coordinate of $f_n(\bar{s})$ is zero for infinitely many $f_n$. That is, we can find a sub-sequence $f_{n_k}$ of $f_n$ for which a fixed state $\bar{s} \in S$ and a fixed pure action $i$ can be found, such that the $i^{th}$ pure strategy in the state $\bar{s}$ is always played with probability zero in the optimal strategy. Hence applying the Lemma 1, limit of this sub-sequence $f_{n_k}$ of $f_n$ (denote as $f_0$) is optimal strategy for player-1 in the undiscounted game. However $f^0$ is not completely mixed, yet it is optimal in the undiscounted stochastic game. This is a contradiction. So there exists $\beta_0 \in [0, 1)$ such that for all $\beta > \beta_0$ the $\beta$-discounted stochastic game $\Gamma_\beta$ obtained from the same payoff matrices are completely mixed.

The converse of the above theorem is not true. The following is an example of finite player-2 controlled zero-sum stochastic game where the $\beta$-discounted stochastic game is completely mixed for all $\beta \in [0, 1)$ but the undiscounted stochastic game is not completely mixed.

**Example 1**

\[
R(s_1) = \begin{bmatrix} 0/(0, 1) & 2/(0, 1) \\ 3/(0, 1) & 1/(0, 1) \end{bmatrix}, \quad R(s_2) = \begin{bmatrix} 2/(0, 1) & 0/(0, 1) \\ 0/(0, 1) & 2/(0, 1) \end{bmatrix}
\]

**Note:** $s_2$ is an absorbing state.

Consider the $\beta$-discounted game $\Gamma_\beta$ with the mentioned states and actions.

The Shapley matrix [1] $R_\beta(\cdot)$ is given by,

\[
R_\beta(s_1) = \begin{bmatrix} 0 + \frac{\beta}{1-\beta} & 2 + \frac{\beta}{1-\beta} \\ 3 + \frac{\beta}{1-\beta} & 1 + \frac{\beta}{1-\beta} \end{bmatrix}, \quad R_\beta(s_2) = \begin{bmatrix} 2 + \frac{\beta}{1-\beta} & 0 + \frac{\beta}{1-\beta} \\ 0 + \frac{\beta}{1-\beta} & 2 + \frac{\beta}{1-\beta} \end{bmatrix}
\]

Clearly, the matrix $R_\beta(s_1)$ is completely mixed for all $\beta \in [0, 1)$.

Furthermore, $R_\beta(s_2)$ is also completely mixed. Hence the game $\Gamma_\beta$ is completely mixed for all $\beta \in [0, 1)$.
But if we consider the undiscounted stochastic game $\Gamma$. Consider the strategy $f$ with $f(s_1) = (1, 0), f(s_2) = (\frac{1}{2}, \frac{1}{2})$. $f$ is a stationary optimal strategy for player-1 in the undiscounted stochastic game $\Gamma$ but not completely mixed. Hence the undiscounted game $\Gamma$ is not completely mixed.

**Lemma 2** Let, $A = (a_{ij})$ be a symmetric matrix of order $n$ with $a_{ij} > 0$ for every $i$ and $j$. Let $b' = (b_1, b_2, \ldots, b_n)$ be a non-negative vector. Let, $C = (c_{ij})$ where, $c_{ij} = a_{ij} + b_j$. Then $C$ completely mixed matrix game implies $A$ is also a completely mixed matrix game.

**Proof** Suppose, $C$ is completely mixed. Since the value $v$ of $C$ is positive, from [4] det($C$) $\neq 0$. Also from ([10], Lemma 4.1, page 381), we have det($A$) $\neq 0$.

Let $y$ be a completely mixed optimal strategy for $C$. Then $Cy = ve$, where $e$ is a vector with all coordinates equals to 1 and $v$ denotes the value of the matrix game $C$. It follows that $Ay = (v - b'y)e = \delta e$, where $\delta = v - b'y$. Since $A$ is symmetric, it follows that $\delta$ is the value of the matrix game $A$ and $(y, y)$ is an optimal strategy of $A$.

To complete the proof we will show that $y$ is the only optimal strategy for both players in the matrix game $A$. Let, $z$ be any other optimal strategy for player-2 in $A$. Then $Az = \delta e$. Since $y$ is a completely mixed optimal for both players (using Result 1) we also have $Ay = \delta e$. Thus $A(y - z)$ equals to zero vector. Since $A$ is non-singular, $y = z$. In other words, every optimal strategy for player-2 in $A$ coincides with $y$. A similar argument holds for player-1 in $A$. This completes the proof.

A similar type of proof is also provided in [13]. We now give a simple counter-example to show that if $A$ is a completely mixed matrix game, then $C$ need not be a completely mixed matrix game.

**Example 2** Let, $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and let $b = (b_1, b_2)' = (1, 2)'$, then $C = \begin{bmatrix} 2 & 4 \\ 3 & 3 \end{bmatrix}$. $A$ is a completely mixed matrix game but $C$ is not a completely mixed matrix game. So the converse of Lemma 2 is not true.

**Remark** Lemma 2 also holds under both the following circumstance.

1. a pair of completely mixed optimal strategies exist for the two players in $C$ instead of $C$ to be a completely mixed matrix game.
2. $A$ to be a symmetric, non-negative and irreducible instead of $A$ to be a strictly positive matrix.

Undiscounted completely mixed stochastic games not only proceed completely mixed $\beta$-discounted stochastic games for large enough $\beta$ (see Theorem 1) but also process completely mixed matrix games under some symmetry assumption.

**Theorem 2** Consider a finite, undiscounted, zero-sum, single-player (player-2) controlled completely mixed stochastic game $\Gamma$. Assume all the individual payoff matrices $R(s)$ are symmetric then the individual matrix game $R(s)$ is completely mixed for all $s \in S$. 

Springer
**Proof** Assume without loss of generality $R(s)$ is positive, that is every entry in $R(s)$ is positive. Take any $\beta > \beta_0$ where, $\beta_0$ is from Theorem 1. Then the matrix game $R_\beta(s)$ (also known as Shapley matrix) is completely mixed for all $s \in S$, where $R_\beta(s) = R(s) + b(s)$, with $b_j(s) = \beta \sum_{s' \in S} v_\beta(s') g(s'|s,j)$. Now the result is immediate from Theorem 1 and Lemma 2. □

Examples of completely mixed undiscounted stochastic games are far from obvious. The following is a finite single-player controlled completely mixed stochastic game.

**Example 3**

$$R(s_1) = \begin{bmatrix} 2/(1,0) & 0/(0,1) \\ 0/(1,0) & 2/(0,1) \end{bmatrix}, \quad R(s_2) = \begin{bmatrix} 3/(1,0) & -1/(0,1) \\ -1/(1,0) & 3/(0,1) \end{bmatrix}$$

This is a player-2 controlled stochastic game with 2 states $s_1$ and $s_2$. In both state $s_1$ and $s_2$ if player-2 chooses action 1 (column 1), then the game moves to state $s_1$ in the next day and if player-2 chooses action 2 (column 2), then the game moves to state $s_2$ in the next day.

The unique optimal strategy for player-1 is $f = \{(1/2,1/2)\} \in \text{above-mentioned undiscounted game}$. That is, choosing row 1 and row 2 is state $s_1$ with probability $1/2$ and $1/2$ respectively and choosing row 1 and row 2 is state $s_2$ with probability $1/2$ and $1/2$ respectively. The player-2’s unique optimal strategy is $g = \{(1/2,1/2)\}$. So $(f,g)$ is the unique optimal strategy for the undiscounted game mentioned before.

In Example 3, we can see that for all $\beta \in [0,1)$ the $\beta$-discounted stochastic game $\Gamma_\beta$ are completely mixed. As both $s_1$ and $s_2$ are symmetric matrix Theorem 2 says individual matrix games are completely mixed, which is easy to see in the above example.

Similar to Theorem 1, if we know the value of an undiscounted stochastic game is non-zero then we can conclude the value for the $\beta$-discounted games are also non-zero.

**Theorem 3** Assume a player-2 controlled transition. Let, $v(s)$ and $v_\beta(s)$ be the value of an undiscounted and $\beta$-discounted stochastic game in state $s \in S$ respectively. If $v(s) \neq 0$ for some $s \in S$ then there exists $\beta_0 \in [0,1)$ such that for all $\beta > \beta_0$, the discounted value $v_\beta(s) \neq 0$.

**Proof** Since $v(s) \neq 0$ from [10] there exist $(f^0, g^0)$ optimal strategy in $\Gamma$ such that $\lim_{\beta \downarrow 1} f^\beta(s) = f^0(s)$. And $(f^\beta, g^0)$ is optimal stationary strategy in $\Gamma_\beta$ for all $\beta > \beta_0$ for some $\beta_0 \in [0,1)$. Now without loss of generality we assume the reward $r(i,j,s)$ > 0 for all $i \in P_1, j \in P_2$ and $s \in S$. Hence, $R(s) > 0$ which implies $v(s) > 0$ for all $s \in S$. Now,

$$0 < v(s) = \Phi(f^0, g^0)(s) \leq \Phi(f^0, g)(s)$$
for all $g \in \mathbb{P}_A$. (Since we assume player-1 is maximizer and player-2 is minimizer). Therefore,

$$0 < \Phi(f^0, g^0) = \lim_{\beta \to 1} (1 - \beta) I_\beta(f^0, g^0)(s).$$

Implies, $\lim_{\beta \to 1} (1 - \beta) I_\beta((\lim_{\beta \to 1} f^\beta), g^0)(s) > 0$. This follows, $\lim_{\beta \to 1} (1 - \beta) I_\beta(f^\beta, g^0)(s) > 0$ (as for player-2 controlled stochastic game $I_\beta$ is continuous and linear in player-1’s strategy). Then, there exists $\beta_0 \in [0, 1)$ such that for all $\beta > \beta_0$ we have $I_\beta(f^\beta, g^0)(s) > 0$. Hence the Lemma follows. \hfill \Box

Note: The above theorem is true for non-zero-sum stochastic games also. For the non-zero-sum stochastic game, the proof will be similar to the above proof using results of [10].

Note: The converse of the Theorem 3 is not true. The following example shows that.

Example 4 Consider the following player-2 controlled stochastic game:

$$R(s_1) = \begin{bmatrix} 4/(0, 1, 0) & 2/(0, 0, 1) \\ 3/(0, 1, 0) & 1/(0, 0, 1) \end{bmatrix}, \quad R(s_2) = \begin{bmatrix} 2/(0, 1, 0) & 0/(0, 1, 0) \\ 0/(0, 1, 0) & 2/(0, 1, 0) \end{bmatrix},$$

$$R(s_3) = \begin{bmatrix} 1/(0, 0, 1) & -1/(0, 0, 1) \\ -1/(0, 0, 1) & 1/(0, 0, 1) \end{bmatrix}$$

States $s_2$ and $s_3$ are absorbing states. $R_\beta(s_1)$ is the auxiliary game (also known as Shapley matrix) in state 1 (denoted as $s_1$). For all $\beta \in [0, 1)$ $\text{val}(R_\beta(s_1)) = 2$, where $\text{val}(.)$ denotes the value of the corresponding matrix game. But the undiscounted value in state 1 is given by- $v(s_1) = \lim_{\beta \to 1} (1 - \beta) v_\beta(s_1) = 0$.

Lemma 3 Let $C$ be a completely mixed matrix game. $c_{ij} = a_{ij} + b_j$ where $a_{ij}, b_j > 0 \forall i, j$. If $\text{val}(A) = \text{val}(A')$ then $\text{val}(A) = \text{val}(C) - b^t \gamma_0$ where, $\gamma_0$ is the optimal strategy for player-2 in the matrix game $C$. Also, player-1 has an unique optimal strategy in the matrix game $A$.

Proof We skip the proof as it is straightforward. \hfill \Box

The following theorem shows under the symmetric assumption, the undiscounted value satisfies a linear equation.

Theorem 4 Let $\Gamma$ be a player-2 controlled completely mixed undiscounted stochastic game. Further assume, for all $s \in S$, the individual payoff matrices $R(s)$ satisfies $\text{val}(R(s)) = \text{val}(R(s)^t)$ where, $\text{val}(.)$ denotes the value of the corresponding matrix game. Then, the following equality holds.

$$v(s) = \sum_{s' \in S} v(s') q(s'|s, g^0),$$

for all $s \in S$. The following equality holds.
where \( g^0 \) is the unique optimal for player-2 in undiscounted game \( \Gamma \).

**Proof** For the \( \beta \)-discounted stochastic game \( \Gamma_\beta \) with payoff matrix \( R(s) \), denote the Shapley matrix in state \( s \), as \( R_\beta(s) \). So:

\[
R_\beta(s) = \left[ r(i, j, s) + \beta \sum_{s' \in S} v_\beta(s') q(s'|s, j) \right]_{ij}
\]

Now from Theorem 1 we can conclude that \( \exists \beta_0 \in (0, 1) \) such that \( \forall \beta > \beta_0 \), the \( \beta \)-discounted stochastic game \( \Gamma_\beta \) with same payoff matrix is completely mixed. This implies \( \forall s \in S \) and \( \forall \beta \) sufficiently close to 1 (say \( \beta > \beta_0 \)) all the Shapley matrix \( R_\beta(s) \) are completely mixed as well. Assume the unique optimal [6] in the undiscounted game to be \( (f^0, g^0) \). We know that, \( g^0 \) is also a uniform discount optimal for player-2 in the undiscounted game \( \Gamma \).

Without loss of generality we will assume the matrix \( R(s) > 0 \ \forall s \in S \). Taking \( A = R(s) \) and \( C = R_\beta(s) \) in Lemma 3 gives,

\[
val(R(s)) = val(R_\beta(s)) - \sum_j \beta \sum_{s' \in S} v_\beta(s') q(s'|s, j) g^0(s)_{ij},
\]

where \( g^0(s)_{ij} \) is the \( j \)th coordinate of \( g^0(s) \). So we have the following equality.

\[
val(R(s)) = v_\beta(s) - \beta \sum_{s' \in S} v_\beta(s') \sum_j q(s'|s, j) g^0(s)_{ij}.
\]

\[
\implies val(R(s)) = v_\beta(s) - \beta \sum_{s' \in S} v_\beta(s') q(s'|s, g^0).
\]

Multiplying by \( (1 - \beta) \) in both the sides we obtain:

\[
(1 - \beta) val(R(s)) = (1 - \beta) v_\beta(s) - \beta (1 - \beta) \sum_{s' \in S} v_\beta(s') q(s'|s, g^0).
\]

The above equality is true for all \( \beta \in [0, 1) \). Taking limit as \( n \to \infty \) we have the following equality.

\[
\lim_{\beta \uparrow 1} (1 - \beta) val(R(s)) = \lim_{\beta \uparrow 1} (1 - \beta) v_\beta(s) - \lim_{\beta \uparrow 1} \beta (1 - \beta) \sum_{s' \in S} v_\beta(s') q(s'|s, g^0)
\]

\[
\implies 0 = v(s) - \sum_{s' \in S} v(s') q(s'|s, g^0).
\]

**Remark 1** If \( \Gamma \) is an undiscounted non-zerosum single-player (player-2) controlled completely mixed stochastic game then all the strictly positive strategies for player-2 partitions the state space \( S \) of undiscounted game \( \Gamma \) into \( k \) sets of ergodic chains \( C_1, C_2, \ldots, C_k \) and a set \( H \) of transient states. Furthermore, the stochastic game \( \Gamma \) can
also be divided into \( k \) sub-games \( \Gamma_1, \Gamma_2, \ldots, \Gamma_k \), with value that is independent of states in each of the game \( \Gamma_s \) corresponding to states \( C_s \).

For the zero-sum games, if \( s \in H \) then both player-1 and player-2 both has exactly one action in the state \( s \in S \) [14].

Using Remark 1 and the fact that for player-2 controlled the stochastic game the value is a linear function of player-1’s strategy we can show that for an undiscounted stochastic game every completely mixed optimal strategy proceeds an equalizer rule involving the undiscounted value.

**Theorem 5** Let \( \Gamma \) be a zero-sum, undiscounted, player-2 controlled stochastic game. If \((f^0, g^0)\) is a completely mixed optimal strategy in the game \( \Gamma \) then:
\[
\forall f \in \mathbb{P}_1 \text{ and } \forall g \in \mathbb{P}_2 \quad \Phi(f^0, g^0) = \Phi(f^0, g) = \Phi(f, g^0).
\]

**Proof** For a completely mixed game, Filar [6] has shown a similar result of the theorem, when the strategies \( f \) and \( g \) are restricted to only the pure strategies of the respective players.

Since \( \Gamma \) is completely mixed from Remark 1 we can conclude that it is sufficient to look at the stochastic games \( \Gamma_c \) separately. Now all the sub-game \( \Gamma_c \) with states restricted to \( C_c \) are completely mixed. Also, if \((f^0, g^0)\) is optimal for the stochastic game \( \Gamma \) then \((f^0, g^0)\) restricted to state \( C_c \) (denoted as \((f^0, g^0)_c\)) is an optimal strategy in the stochastic game \( \Gamma_c \).

If we fixed player-1’s (unique) optimal strategy \( f^0_c \) then there exists vector \( \gamma^0 \) along with value \( v_c \mathbf{1} \) of the game \( \Gamma_c \), which satisfy the following equality [15]:
\[
v_c + \gamma^0(s) = r_c(f^0, j, s) + \sum_{s' \in S} q(s'|s, j)\gamma^0(s'),
\]
for all pure strategy \( j \) for player-2 and for all \( s \in C_c \). This implies,
\[
v_c g_j(s) + \gamma^0(s)g_j(s) = r_c(f^0, j, s)g_j(s) + \sum_{s' \in S} q(s'|s, j)g_j(s)\gamma^0(s')
\]
for all player-2’s pure strategy \( j \), for all \( s \in C_c \) and \( g \in \mathbb{P}_{A_j} \). So,
\[
\sum_j v_c g_j(s) + \sum_j \gamma^0(s)g_j(s) = \sum_j r_c(f^0, j, s)g_j(s) + \sum_j \sum_{s' \in S} q(s'|s, j)g_j(s)\gamma^0(s')
\]
for all \( s \in C_c \). So for all \( s \in C_c \):
\[
v_c + \gamma^0(s) = r_c(f^0, g, s) + \sum_{s' \in S} q(s'|s, g)\gamma^0(s').
\]
Writing the above equality in a vector notation we obtain:
\[
\forall g \in \mathbb{P}_{A_j}; \quad v_c \mathbf{1} + \gamma^0 = r_c(f^0, g) + Q(g)\gamma^0.
\]
Multiplying the Markov matrix \( Q^*(g) \) in both the sides of the above equality we get:
\[ v_c 1 + Q^*(g) \gamma^0 = \Phi_c(f^0, g) + Q^*(g) \gamma^0. \]

Which indeed follows,

\[ \text{for all } g \in \mathbb{P}_{A_2} \quad \Phi_c(f^0, g) = \Phi_c(f^0, g^0). \]

The above argument is true for all the individual completely mixed games \( \Gamma_1, \Gamma_2, \ldots, \Gamma_k \). Now if \( s \in H \) (\( H \) is the set of all transient states) then define \( p_g(s, s') \) be the probability that \( s' \in S \) is the first state reached outside \( H \). So for all \( s \in H \),

\[ \Phi(f^0, g)(s) = \sum_{s \in H} p_g(s, s') \Phi(f^0, g)(s'). \]

We already have \( \Phi_c(f^0, g)(s) = \Phi_c(f^0, g^0)(s) \) for \( s \in H^c \). Hence we conclude the proof.

The other side of the equality directly follows from ([6], Proposition 3.3) and the fact that \( \Phi \) is continuous on player-1’s strategy. \( \blacksquare \)

Let \( \Gamma' \) be a player-2 controlled zero sum undiscounted stochastic game with individual payoff matrix \( R(s) \) for all \( s \in S \). Let \( \Gamma_\beta \) be the corresponding \( \beta \)-discounted stochastic game with same payoff matrix \( R(s) \).

We say a stationary strategy \( g^0 \) is uniform discount optimal for player-2 in the game \( \Gamma \) if there exists \( \beta_0 \in (0, 1) \) such that \( g^0 \) is optimal for player-2 in the corresponding \( \beta \)-discounted game \( \Gamma_\beta \) for all \( \beta > \beta_0 \).

**Remark 2** (Parthasarathy and Raghavan [10]) For \( \beta \)-discounted stochastic game for each \( s \in S \) there exists a non-singular sub-matrix \( \bar{R}_\beta(s) \) of Shapley matrix \( R_\beta(s) \) such that if we define for all \( \beta > \beta_0 \)

\[ f^\beta(s) = v_\beta(s) e[\bar{R}_\beta(s)]^{-1}, \quad \text{and} \]
\[ g^\beta(s) = v_\beta(s)[\bar{R}_\beta(s)]^{-1} e, \]

where, \( e \) is the column vector of all 1’s; then the set of pair \( (f^\beta, g^\beta) \) obtained from \( (f^\beta(s), g^\beta(s)) \) by adding zero in the place corresponding to the rows/ columns of \( \bar{R}(s) \) which are not in \( R(s) \) from an optimal stationary strategy pair \( (f^\beta, g^\beta) \) in \( \Gamma_\beta \) for all \( \beta > \beta_0 \).

Furthermore, for all \( s \in S \) it is also shown that, \( \bar{R}(s) \) is non-singular and \( \bar{g}^\beta_1(s) = \bar{g}^\beta_2(s) \) for all \( \beta_1, \beta_2 > \beta_0 \). Denote, \( \bar{g}^\beta_1(s) = g^0(s) \quad \forall \beta > \beta_0 \). Then \( g^0 = (g^0(1), g^0(2), \ldots, g^0(S))' \) where \( g^0(s) \) is obtained by completing \( g(s) \) with 0’s is an uniformly discount optimal for player-2 in \( \Gamma \). It also turns out that \( (f^0, g^0) \), where \( f^0(s) = \lim_{\beta \downarrow 0} f^\beta(s) \) is an optimal strategy pair in the undiscounted game \( \Gamma \).

**Definition 11** (Filar and Raghavan [14]) A game (matrix, bimatrix, stochastic) is called “CM-I” if player-1 only has completely mixed optimal strategies. Similarly, we can define “CM-II” likewise.
For an undiscounted stochastic game, this definition was first introduced by Filar [5]. For matrix game, bimatrix game and $\beta$-discounted stochastic game if both players have the same number of pure strategies then CM-I (respectively CM-II) implies completely mixed game and hence CM-II (respectively CM-I). But for the undiscounted stochastic game, this is far from reality.

**Lemma 4** If an undiscounted player-2 controlled stochastic game is CM-I then player-2 has a completely mixed optimal in $\Gamma$ and there exists an $\beta_0 \in [0, 1)$ such that for all $\beta > \beta_0$ the corresponding $\beta$-discounted stochastic game $\Gamma_{\beta}$ is completely mixed.

**Proof** The proof is similar to the proof in Theorem 1. Since for $\beta$ closed to 1 $\Gamma_{\beta}$ is completely mixed, the discount optimal for player-2 is also completely mixed. Hence the player-2’s optimal is a completely mixed optimal in the undiscounted stochastic game as well. □

**Open problem:** Lemma 4 only says the $\beta$-discounted stochastic games are completely mixed. It is an open question whether the undiscounted stochastic game is completely mixed or not.

### 4 Non-zerosum Stochastic Games

Unless mentioned otherwise throughout Section 4 both player-1 and player-2 are maximizers. For the non-zerosum discounted stochastic game we have the following result which is needed to prove the main result.

**Theorem 6** (Filar and Vrieze [12]) The following assertions are equivalent:

(i) $(f^0, g^0)$ is an equilibrium point (EP) in the discounted stochastic game with equilibrium payoffs $(v_1^1(f^0, g^0), v_2^2(f^0, g^0))$.

(ii) For each $s \in S$, the pair $(f^0(s), g^0(s))$ constitutes an equilibrium point in the static bimatrix game $(B_1^1(s), B_2^2(s))$ with equilibrium payoffs $(v_1^1(f^0, g^0)(s), v_2^2(f^0, g^0)(s))$, where

$$B_1^1(s) = \left[(1 - \beta)r^1(i, j, s) + \beta \sum_{s' \in S} v_1^1(s') q(s'|s, i, j)\right]_{ij}$$

$$B_2^2(s) = \left[(1 - \beta)r^2(i, j, s) + \beta \sum_{s' \in S} v_2^2(s') q(s'|s, i, j)\right]_{ij}$$

For the bimatrix game, we have an equalizer rule for the payoff under certain completely mixed assumptions.
Lemma 5 Let \((A, B)\) be a bimatrix game. If \((x^0, y^0)\) be a EP in the bimatrix game with \(y^0\) completely mixed then \(r^2(x^0, y) = r^2(x^0, y^0)\) for all player-2’s stationary strategy \(y\).

Proof Since \((x^0, y^0)\) is an equilibrium point for all player-2’s stationary strategy \(y\) we have; \((x^0)^tBy^0 \geq (x^0)^tBy\). Taking \(y\) as \((1, 0, \cdots, 0), (0, 1, \cdots, 0), \cdots, (0, 0, \cdots, 1)\) respectively we get \((x^0)^tB \leq v_2\) where, \(v_2 = (x^0)^tBy^0\), the value of the game for player-2.

Since \(y^0\) is completely mixed, we have \(y^0 > 0\) coordinate-wise. Hence, \(x^0B = v_2e\) as otherwise, \(v_2 = (x^0)^tBy^0 < v_2e^ty_0 = v_2\)

which is a contradiction. Hence the lemma follows. □ 

A corresponding non-zerosum version of Lemma 1 is provided below.

Lemma 6 Assume player-2 controlled transition. Suppose, \(\exists \beta_0 \in [0, 1)\) and NE \((f^\beta, g^0)\) in \(\Gamma_p\) such that \(g^0\) is completely mixed, in every \(\beta\)-discounted non-zerosum stochastic game for all \(\beta > \beta_0\). Let, \(\beta_n \in [\beta_0, 1)\) be such that \(\beta_n \uparrow 1\). Let \((f_n, g^0)\) be NE for \(\beta_n\)-discounted non-zerosum stochastic game. Suppose \(f_n \rightarrow f_0\) coordinate-wise, that is \(f_n(s) \rightarrow f_0(s)\) for each state \(s \in S\), then, \((f_0, g^0)\) is an equilibrium pair in the undiscounted non-zerosum stochastic game \(\Gamma\).

Proof Under one player (for us player-2) controlled transition probability, an undiscounted non-zerosum stochastic game has value restricted to stationary strategies [10]. Suppose, \((f_n, g^0)\) is a NE for \(\beta_n\)-discounted stochastic game. As we have a completely mixed strategy \(g_0\) for player-2 in the \(\beta_n\)-discounted stochastic game from Theorem 6 and Lemma 5 we have

\[I^2_{\beta_n}(f_n, g) \equiv v^2_{\beta_n}\]

coordinate-wise for any stationary strategy \(g\) of player-2, where:

\[I^2_{\beta_n}(f_n, g) = (I^2_{\beta_n}(f_n, g)(s_1), \cdots, I^2_{\beta_n}(f_n, g)(s_K))^t, \text{ and } v^2_{\beta_n} = (v^2_{\beta_n}(s_1), \cdots, v^2_{\beta_n}(s_K)).\]

Therefore we have \([I - \beta_nQ(g)]^{-1}r^2(f_n, g) \equiv v^2_{\beta_n}\). Since, \([I - \beta_nQ(g)]^{-1} = \sum_{\varepsilon=0}^{\infty} \beta_n^\varepsilonQ^\varepsilon(g)\) is a non-negative matrix, we have the expression of \(r^2(f_n, g)\) as follows.

\[r^2(f_n, g) \equiv [I - \beta_nQ(g)]v^2_{\beta_n}\]

Now we have \(f_n \rightarrow f_0\) point-wise and the reward function \(r(\cdot, \cdot)\) is a continuous function on the strategy of player-1. Hence, for any given \(\epsilon > 0, \exists N_0 < \infty\) such that,

\[|I - \beta_nQ(g)|v^2_{\beta_n} - \epsilon \leq r^2(f_0, g) \leq |I - \beta_nQ(g)|v^2_{\beta_n} + \epsilon\]
coordinate-wise for all \( n \geq N_0 \). Where \( e \) is column vector with all entry as 1. Therefore we have,
\[
v_{\beta_n}^2 - e [I - \beta_n Q(g)]^{-1} e \leq [I - \beta_n Q(g)]^{-1} r^2(f_0, g) \leq v_{\beta_n}^2 + e [I - \beta_n Q(g)]^{-1} e.
\]
As, \((1 - \beta_n)\) is always non-negative for all \( \beta_n \in [0, 1) \), we have,
\[
(1 - \beta_n)v_{\beta_n}^2 - (1 - \beta_n)e [I - \beta_n Q(g)]^{-1} e \leq (1 - \beta_n)[I - \beta_n Q(g)]^{-1} r^2(f_0, g)
\leq (1 - \beta_n)v_{\beta_n}^2 + (1 - \beta_n)e [I - \beta_n Q(g)]^{-1} e.
\]
Since \( Q^k \) is a stochastic matrix for each \( k \), we have \([I - \beta_n Q(g)]^{-1} e = \sum_{k=0}^{\infty} \beta_n Q^k(g)e\) \(= \sum_{k=0}^{\infty} \beta_n e\). Therefore the above inequality reduces to the following inequality.
\[
(1 - \beta_n)v_{\beta_n}^2 - ee \leq (1 - \beta_n)[I - \beta_n Q(g)]^{-1} r^2(f_0, g) \leq (1 - \beta_n)v_{\beta_n}^2 + ee.
\]
Now if we let \( \beta_n \uparrow 1 \), using the result form [10], the above inequality can be further reduced as follows.
\[
v^2 - ee \leq \Phi^2(f_0, g) \leq v^2 + ee.
\]
Where, \( v = (v^2(s_1), \cdots, v^2(s_K))^T \) is the value of the undiscounted stochastic game. This is true for any \( \epsilon > 0 \). Hence \((f_0, g^0)\) constriicted above is a NE in the non-zerosum undiscounted stochastic game. ■

The following is an example of a non-zerosum completely mixed undiscounted stochastic game. This is the corresponding non-zerosum version of Example 3.

**Example 5**

\[
R(s_1) = \begin{bmatrix}
0 & 1/(0, 1) & 2, -1/(1, 0) \\
2, -1/(0, 1) & 0, 1/(1, 0)
\end{bmatrix}, \quad R(s_2) = \begin{bmatrix}
4, -3/(1, 0) & -2, 3/(0, 1) \\
-2, 3/(1, 0) & 4, -3/(0, 1)
\end{bmatrix}
\]

This example is a player-2 controlled stochastic game with states \( s_1 \) and \( s_2 \). In both state \( s_1 \) and \( s_2 \) if player-2 chooses action 1 (column 1) then the game moves to state \( s_1 \) in the next day and if player-2 chooses action 2 (column 2) then the game moves to state \( s_2 \) in the next day irrespective of player-1’s action.

The unique equilibrium strategy for player-1 is \( \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\} \) in the above-mentioned game, i.e., choosing row 1 and row 2 with probability \( \frac{1}{2} \) and \( \frac{1}{2} \) respectively both in state \( s_1 \) and in state \( s_2 \). The unique equilibrium strategy for player-2 is also \( \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\} \) for the above-mentioned game.

In Example 5, we can see that for all \( \beta \in [0, 1) \) the \( \beta \)-discounted non-zerosum stochastic game \( \Gamma_\beta \) are completely mixed.
\[
B^1_\beta(s_1) = \begin{bmatrix}
\beta & 2 - \beta \\
2 - \beta & \beta 
\end{bmatrix}, \quad B^2_\beta(s_1) = \begin{bmatrix}
1 - \beta & \beta - 1 \\
\beta - 1 & 1 - \beta 
\end{bmatrix}
\]

\[
B^1_\beta(s_2) = \begin{bmatrix}
4 - 3\beta & 3\beta - 2 \\
3\beta - 2 & 4 - 3\beta 
\end{bmatrix}, \quad B^2_\beta(s_2) = \begin{bmatrix}
3(\beta - 1) & 3(1 - \beta) \\
3(1 - \beta) & 3(\beta - 1) 
\end{bmatrix}
\]

The following Lemma is required to prove the later Theorem stating equalizer rule for a general undiscounted stochastic game. For the following lemma and theorem we do not have the assumption that only one player controls the transition probability.

**Lemma 7** Let \( \Gamma \) be a non-zero-sum undiscounted stochastic game (not necessarily single player controlled). Let \((f^0, g^0)\) be a completely mixed NE for the undiscounted stochastic game \( \Gamma \). Then there exists two vectors \( \gamma^1 = (\gamma^1(s_1), \ldots, \gamma^1(s_K))^t \) and \( \gamma^2 = (\gamma^2(s_1), \ldots, \gamma^2(s_K))^t \) such that

\[
v^2_{f^0, g^0}(s) + \gamma^2(s) = r^2(f^0, j, s) + \sum_{s' \in S} q(s'|s, f^0, j)\gamma^2(s')
\]

for all pure action \( j \) for player-2 and \( \forall s \in S \). Similarly,

\[
v^1_{f^0, g^0}(s) + \gamma^1(s) = r^1(i, g^0, s) + \sum_{s' \in S} q(s'|s, i, g^0)\gamma^1(s')
\]

for all pure action \( i \) of player-1, and \( \forall s \in S \), where, \( v^k_{f^0, g^0}(s) \) for \( k = \{1, 2\} \) is the value in the undiscounted stochastic game for player \( k \) corresponding to the NE \((f^0, g^0)\) in state \( s \in S \).

**Proof** From Markov decision process ([12], Theorem 3.8.4), we have existence of vectors \( \gamma^1 \) and \( \gamma^2 \) such that

\[
v^2_{f^0, g^0}(s) + \gamma^2(s) = \max_{\sigma} \left[ r^2(f^0, \sigma, s) + \sum_{s' \in S} q(s'|s, f^0, \sigma)\gamma^2(s') \right],
\]

for all stationary strategy \( \sigma \in \mathcal{P}_2 \) and \( s \in S \). And

\[
v^1_{f^0, g^0}(s) + \gamma^1(s) = \max_{\mu} \left[ r^1(\mu, g^0, s) + \sum_{s' \in S} q(s'|s, \mu, g^0)\gamma^1(s') \right],
\]

for all stationary strategy \( \mu \in \mathcal{P}_1 \) and \( s \in S \). Therefore for all pure action \( i \) for player-1, for all pure action \( j \) for player-2 and \( s \in S \):

\[
v^2_{f^0, g^0}(s) + \gamma^2(s) \geq r^2(f^0, j, s) + \sum_{s' \in S} q(s'|s, f^0, j)\gamma^2(s') \quad \text{and,}
\]

\[
v^1_{f^0, g^0}(s) + \gamma^1(s) \geq r^1(i, g^0, s) + \sum_{s' \in S} q(s'|s, i, g^0)\gamma^1(s').
\]

We will finish the proof by contradiction. So if possible let us assume, we have strict inequality for some pure action \( j_0 \) of player-2. That is-
\[
v^2_{f_0, g_0}(s) + \gamma^2(s) > r^2(f_0', j_0, s) + \sum_{s' \in S} q(s' | s, f_0', j_0) \gamma^2(s').
\]

Therefore,
\[
v^2_{f_0, g_0}(s) + \gamma^2(s) = v^2_{f_0, g_0}(s) \sum_j g_j^0(s) + \gamma^2(s) \sum_j g_j^0(s)
\]
\[
= \sum_j g_j^0(s)[v^2_{f_0, g_0}(s) + \gamma^2(s)] > \sum_j g_j^0(s)[r^2(f_0', j, s) + \sum_{s' \in S} q(s' | s, f_0', j) \gamma^2(s')]
\]
\[
= \sum_j r^2(f_0', j, s) g_j^0(s) + \sum_j \sum_{s' \in S} q(s' | s, f_0', j) g_j^0(s) \gamma^2(s')
\]
\[
= r^2(f_0', g_0', s) + \sum_{s' \in S} q(s' | s, f_0', g_0') \gamma^2(s') = v^2_{f_0, g_0}(s) + \gamma^2(s),
\]

which is a contradiction. Hence equality holds for all pure stationary strategy \(j\) for player-2. Exactly in the similar way we can show equality for pure action of player-1 as well. □

Using Lemma 7 we can show that for an undiscounted stochastic game with one completely mixed NE the game process equalizer rule for undiscounted payoff. Unlike the zerosum case (see Theorem 5) we have an equalizer for the payoff for both players.

**Theorem 7** Let \(\Gamma\) be a non-zerosum undiscounted stochastic game (not necessarily single player controlled). Also assume that, the undiscounted stochastic game is completely mixed. Let \((f_0, g_0)\) be a NE in the game, then we have the following equalizer rule,

\[
\Phi^{(2)}(f_0', g_0) = \Phi^{(2)}(f_0', g)
\]
\[
\Phi^{(1)}(f_0', g_0) = \Phi^{(1)}(f', g_0)
\]

for all stationary strategy \(f' \in \mathbb{P}_1\) and stationary strategy \(g \in \mathbb{P}_2\).

**Proof** From Remark 1 we have \((f_0', g_0)\), which is \((f_0', g_0)\) restricted to states \(s \in C_c\), is also a NE in the restricted game \(\Gamma_c\). Since the states \(C_c\) are irreducible, the value for both player-1 and player-2 corresponding to NE \((f_0', g_0)\) is independent of states (denote the value as \(v^1_c\) and \(v^2_c\) respectively). Therefore from Lemma 7 we conclude, there exists vector \(\gamma^2\) such that:

\[
v^2_c + \gamma^2(s) = r^2_c(f_0', j, s) + \sum_{s' \in S} q(s' | s, f_0', j) \gamma^2(s')
\]

for all pure strategy \(j\) of player-2’s and \(\forall s \in C_c\). Now for any stationary strategy \(g\) of player-2 we have the following equality.

\[
v^2_c g_j(s) + \gamma^2(s) g_j(s) = r^2_c(f_0', j, s) g_j(s) + \sum_{s' \in S} q(s' | s, f_0', j) g_j(s) \gamma^2(s')
\]
for all pure strategy \( j \) of player-2’s and \( \forall s \in C_c \). This implies,
\[
\sum_j v^2_c g_j(s) + \sum_j \gamma^2(s) g_j(s) = \sum_j r^2_c(f^0, j, s) g_j(s) + \sum_j \sum_{s' \in S} q(s'|s, f^0, j) g_j(s) \gamma^2(s')
\]
for all \( s \in C_c \). Then it follows:
\[
\text{for all } s \in C_c; \quad v^2_c + \gamma^2(s) = r^2_c(f^0, g, s) + \sum_{s' \in S} q(s'|s, f^0, g) \gamma^2(s').
\]

Multiplying be the Markov matrix \( Q^*(f^0, g) \) in both the sides of the above equality we obtain:
\[
v^2_c 1 + Q^*(f^0, g) \gamma^2 = \Phi^2_c(f^0, g) + Q^*(f^0, g) \gamma^2.
\]
Therefore we indeed have:
\[
\Phi^2_c(f^0, g^0) = \Phi^2_c(f^0, g).
\]

The above argument is true for all the individual completely mixed games \( \Gamma_1, \Gamma_2, \ldots, \Gamma_k \). Define \( p_g(s, s') \) be the probability that starting in state \( s \in H \), the first state reached outside \( H \) is \( s' \). So we have,
\[
\Phi^2(f^0, g)(s) = \sum_{s \in H^c} p_g(s, s') \Phi^2(f^0, g)(s'),
\]
where, \( s \in H \). We already have \( \Phi^2_c(f^0, g^0) = \Phi^2_c(f^0, g^0) \) for \( s \in H^c \).

The other side of the equality is similar to the proof for player-1. Using Lemma 7 we can derive the following for each \( \Gamma_c \).
\[
v^1_c(s) + \gamma^1(s) = r^1(f, g^0, s) + \sum_{s' \in S} q(s'|s, f, g^0) \gamma^1(s')
\]
for all \( f \in P_1 \). Now multiplying the Markov matrix \( Q^*(f, g^0) \) on both the sides we have the desired equality. 

\textbf{Theorem 8} Let \( \Gamma \) be a non-zerosum player-2 controlled undiscounted stochastic game. Let \( (f^0, g^0) \) be a NE with \( g^0 \) as an uniformly discount equilibrium. Then:
\[
v^1_{f^0, g^0}(s) = \sum_{s'} q(s'|s, g^0) v^1_{f^0, g^0}(s')
\]
\[
v^2_{f^0, g^0}(s) = \sum_{s'} q(s'|s, g^0) v^2_{f^0, g^0}(s').
\]

\textbf{Proof} The proof is easy. So we are skipping the proof. 

\textbf{Remark 3} For the undiscounted player-2 controlled non-zerosum stochastic game \( \Gamma \) the existence of uniform discount equilibrium \( (f^0, g^0) \) has been shown in [10]. Also
it is shown that \((f^\beta, g^0)\) with \(f^0(s) = \lim_{\beta \to 0} f^\beta(s)\) is a NE in the \(\beta\)-discounted stochastic game \(\Gamma_\beta\) for all \(\beta > \beta_0\). If the Undiscounted non-zero-sum stochastic game is completely mixed then the uniformly discount equilibrium \((f^0, g^0)\) is completely mixed hence for \(\beta\) closed to 1 \((f^0, g^0)\) is completely NE in \(\Gamma_\beta\). Now from Theorem 6 we have \((f^0(s), g^0(s))\) is a NE in the bimatrix game \((B^1_\beta(s), B^2_\beta(s))\). So we have \(B^1_\beta(s)g^0(s) = ce\). Therefore from ([10], Lemma 4.1), we have \(R^1(s)g^0(s) = c_1e\). It is also shown that for completely mixed games \(R^1(s)\) and \(R^2(s)\) are square matrix.

**Proposition 1** Let \(\Gamma\) be a non-zero-sum player-2 controlled undiscounted completely mixed stochastic game. Then for all \(s \in S\) the payoff matrix \(R^2(s)\) is non-singular.

**Proof** In [10] they showed existence of an uniform discount NE \((f^0, g^0)\) for the undiscounted game \(\Gamma\), such that \(f^0(s) = \lim_{\beta \to 1} f^\beta(s)\) for all \(s \in S\) and \((f^\beta, g^0)\) is a NE in the discounted game \(\Gamma_\beta\) for all \(\beta > \beta_0\). Without loss of generality assume \(r^k(i, j, s) > 0\) for all \(k = 1, 2\) and for all \(s \in S\). Now since \(\Gamma\) is completely mixed \((f^0, g^0)\) is also completely mixed. Hence \(\exists \beta_1 \in (0, 1)\) such that \((f^{\beta_1}, g^0)\) is also completely mixed for all \(\beta > \beta_1\). If possible let us assume for some \(s_0 \in S\); the corresponding \(R^2(s_0)\) is singular. From Theorem 6 we already have:

\[
B^1_\beta(s_0)g^0(s_0) = v^1_{f^\beta, g^0} 1,
\]

\[
f^\beta(s_0)B^2_\beta(s_0) = v^2_{f^\beta, g^0} 1^T.
\]

Using ([10], Lemma 4.1), we can conclude \(B^2_\beta(s_0)\) is also singular. Hence have some \(f^*_s \in \mathbb{P}_2\) with \(f^*_s(s_0)B^2_\beta(s_0) = v^2_{f^*, g^0} 1^T\). So, \((f^*_s, g^0)\) is another NE in \(\Gamma_\beta\). From ([10], Lemma 5.2), we further have \(S(g^0) = \{f(s, g^0)\\} \text{ NE in } \Gamma_\beta\}\) is convex. Hence we can choose \(\lambda\) such that \((\lambda f^*_s + (1 - \lambda)f^\beta, g^0)\) is a non-completely mixed equilibrium. Now a sub-sequence of \((1 + \lambda)f^*_s + \lambda f^\beta\) will converge to some \(\tilde{f}\) such that \((\tilde{f}, g^0)\) will be a non-completely mixed NE (from Lemma 6) in \(\Gamma\). Which is a contradiction and hence the proposition follows. 

**Proposition 2** Let \(\Gamma\) be non-zero-sum player-2 controlled undiscounted completely mixed stochastic game. If \(R^1(s)\) is non-singular for all \(s \in S\). Then \(T = \{g\mid (f, g)\text{ is a NE in } \Gamma\text{ for some } f \in \mathbb{P}_1\}\) is singleton.

**Proof** The proof is on the same way of Filar’s argument ([6], Proposition 3.4) for zero-sum games. (Parthasarathy and Raghavan [10], page 390) provides the existence of uniform discount equilibrium in \(\Gamma\). Let \((f^0, g^0)\) be NE in \(\Gamma\) with \(g^0\) uniform discount equilibrium. If possible assume that we have some other NE \((f^*, g^*)\) (Note: \(f^*\) may be equal with \(f^0\)). From Remark 1 the sets \(C_1, \cdots, C_k\) and \(H\) are same for \(g^0\) and \(g^*\). Also for all \(s \in H\) the number of action available in \(s\) is exactly 1. So \(g^0(s) = g^*(s)\) for all \(s \in H\). Now it is enough to consider the sub-games \(\Gamma_1, \cdots, \Gamma_k\) separately. Without loss of generality we consider only \(\Gamma_1\) and assume that \(C_1\) has \(S_1\) states. Also, associate with a NE \((f, g)\), the value \(v^0 = v^1(s)\) is a constant for all \(s \in C_1\) and \(k \in \{1, 2\}\). The stationary matrices \(Q^*(g^0)\) and \(Q^*(g^*)\) each have identical rows \(u^0 = (u^0(I), \cdots, u^0(S_1))\) and \(u^* = (u^*(1), \cdots, u^*(S_1))\) with \(u^0(s)\) and \(u^*(s) > 0\) for
all \( s \in C_1 \). By Theorem 8, \( \forall s \in C_1 \), for all player-1’s pure stationary strategy in \( \Gamma_1 \), we have,

\[
v^1_{(\sigma, g^0)} = \Phi^1(\sigma, g^0)(s) = [Q^r(g^0)R^1(\sigma, g)]_s
= \sum_{s' = 1}^{S_1} u^0(s')[\sigma(s')R^1(s')g^0(s')]
= \sum_{s' = 1}^{S_1} \sum_{j = 1}^{n_j} r^1(\sigma, j, s')u^0_j(s')
\tag{1}
\]

where \( u^0_j(s') = u^0(s')g^0_j(s') \) for all \( j = 1, \ldots, n_j \), and \( s' \in C_1 \). The above equality holds with \( g^* \) in place of \( g^0 \) and with \( u^*(s') = u^*(s')g^*(s') \) in place of \( u^0(s') \). Now let \( z_j(s) = u^0_j(s) - u^j_j(s) \) for all \( j = 1, \ldots, n_j \) and \( s \in C_1 \). Then from above equation we obtain

\[
\sum_{s' = 1}^{S_1} \sum_{j = 1}^{n_j} r^1(\sigma, j, s')z_j(s') = v^1_{(\sigma, g^0)} - v^1_{(\sigma, g^*)} = \bar{c}
\]

for every pure stationary strategy \( \sigma \) for player-1 in \( \Gamma_1 \). Let

\[
Z = (Z_1 : Z_2 : \cdots : Z_{S_1})^f
\]

be a column vector such that \( Z_s = (Z_1(s), Z_2(s), \ldots, Z_{n_j}(s)) \) for each \( s \in C_1 \), and let \( t_1 \), be the number of pure stationary strategies for player-1 in \( \Gamma_1 \). Fix \( \sigma(s) \) for each \( s > 1 \) and consider the \( n_1 \) equations extracted from Eq. 1 by letting \( \sigma(1) \) range over the \( n_1 \)-dimensional unit basis vectors (Since \( R^1(\sigma) \) is square from Remark 3). These are equivalent to

\[
R^1(\sigma)Z_1 = a1.
\]

Now, if \( \alpha = 0 \) the argument follows exactly in the same way of filar’s argument for zero sum games with help of Assumption. If \( \alpha \neq 0 \) the argument again goes in the same line as Filar’s argument with the help of Remark 3.

**Proposition 3** Let \( \Gamma \) be non-zerosum player-2 controlled undiscounted completely mixed stochastic game. Then given a NE \( (f^*, g^*) \), \( S(g^*) = \{f|(f, g^*) \text{ is a NE}\} \) is singleton.

**Proof** From ([10], Lemma 5.2), we can conclude \( S(g^*) \) is convex. Since \( \Gamma \) is also completely mixed, all elements of \( S(g^*) \) has to be completely mixed as well. Assume if possible we have \( f^0 \neq f^* \in S(g^*) \). Therefore we can choose \( \lambda \) such the \( \lambda f^0(s) + (1 - \lambda)f^*(s) \) is not completely mixed for some \( s \in S \). Which is a contradiction. Hence the proposition is true.
For non-zerosum stochastic games, a completely mixed game does not imply unique Nash Equilibrium. But under some non-singularity assumption of the payoff matrix, the undiscounted game does possess a unique Nash equilibrium.

**Theorem 9** The two-person non-zerosum player-2 controlled undiscounted stochastic game $\Gamma$ is completely mixed. If $R^i(s)$ is non-singular for all $s \in S$. Then the game process a unique Nash Equilibrium.

**Proof** Let $(f^0, g^0)$ and $(f^*, g^*)$ be two NE is $\Gamma$. Then from Theorem 8 we have $(f^0, g^*)$ and $(f^*, g^0)$ are also NE of the stochastic game $\Gamma$. Which contradicts the Proposition 2 and 3. Hence, $\Gamma$ process an unique NE. $\blacksquare$

**Theorem 10** Let $\Gamma_p$ be a non-zerosum $\beta$-discounted stochastic game. The game is not necessarily single player controlled. Also assume that the $\beta$-discounted stochastic game has a completely mixed NE $(f^0, g^0)$. Then we have the following equalizer rules,

$$
I^{(2)}_\beta(f^0, g^0) = I^{(2)}(f^0, g) \quad \text{and} \quad I^{(1)}_\beta(f^0, g^0) = I^{(1)}(f, g^0)
$$

for all stationary strategy $f \in P_{A_1}$ and stationary strategy $g \in P_{A_2}$.

**Proof** From markov decision process we have the following results-

$$
v^1_\beta(s) = \max_{\mu \in P_1} \left[ r^1(\mu, g^0, s) + \beta \sum_{s' \in S} q(s'|s, \mu, g^0)v^1_\beta(s') \right]
$$

$$
v^2_\beta(s) = \max_{\sigma \in P_2} \left[ r^2(f^0, \sigma, s) + \beta \sum_{s' \in S} q(s'|s, f^0, \sigma)v^2_\beta(s') \right]
$$

Now using the similar technique as used in Lemma 7 we can show that equality holds for all pure strategies (hence for all strategies) of player-1 and player-2. Therefore we have

$$
v^1_\beta(s) = r^1(f, g^0, s) + \beta \sum_{s' \in S} q(s'|s, f, g^0)v^1_\beta(s')
$$

$$
v^2_\beta(s) = r^2(f^0, g, s) + \beta \sum_{s' \in S} q(s'|s, f^0, g)v^2_\beta(s')
$$

for all $f \in P_1$ and $g \in P_2$. Writing the last equation in a vector from we have

$$
v^2_\beta = r^2(f^0, g) + \beta Q(f^0, g)v^2_\beta
$$

Now pre-multiplying the equation by $[I - \beta Q(f^0, g)]^{-1}$ we get-

$$
[I - \beta Q(f^0, g)]^{-1}v^2_\beta = [I - \beta Q(f^0, g)]^{-1}r^2(f^0, g) + \beta[I - \beta Q(f^0, g)]^{-1}Q(f^0, g)v^2_\beta.
$$

This implies,
\[ [I - \beta Q(f^0, g)]^{-1}v^2 = [I - \beta Q(f^0, g)]^{-1}r^2(f^0, g) + \beta \sum_{k=0}^{\infty} \beta^k Q^k(f^0, g)Q(f^0, g)v^2. \]

\[ \implies [I - \beta Q(f^0, g)]^{-1}v^2 = [I - \beta Q(f^0, g)]^{-1}r^2(f^0, g) + \sum_{k=1}^{\infty} \beta^k Q^k(f^0, g)v^2. \]

\[ \implies [I - \beta Q(f^0, g)]^{-1}v^2 = I^2_\beta(f^0, g) + [-I + \sum_{k=1}^{\infty} \beta^k Q^k(f^0, g)]v^2. \]

So finally we get:

\[ v^2 = I^2_\beta(f^0, g) \quad \forall g \in \mathbb{P}_2. \]

Similarly we can show \( I^{(1)}_\beta(f^0, g^0) = I^{(1)}_\beta(f^0, g^0) \) for all \( f \in \mathbb{P}_1. \)

## 5 Open Problem

In a switching control Stochastic game [5] we get a partition \( S_1 \) and \( S_2 \) of state space \( S \), such that in any states of \( S_1 \) player-1 alone controls the transition probability and in any states of \( S_2 \) player-2 alone controls the transition probability. The transition probabilities are as follows.

\[ q(s' | s, i, j) = q(s | s, i) \text{ for all } s, s' \in S_1, i \in A_1 \text{ and } j \in A_2. \]
\[ q(s' | s, i, j) = q(s | s, j) \text{ for all } s, s' \in S_2, i \in A_1 \text{ and } j \in A_2. \]

It is not known whether similar results of Theorems 1 and 2 holds for Switching control undiscounted stochastic games.

### Availability of Data, Material and Code

On behalf of all authors, the corresponding author states that there is no data/codes used for this manuscript.

### Declarations

#### Conflict of Interest

The authors declare no competing interests.

#### Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
References

1. Shapley LS (1953) Stochastic games. Proceedings of the National Academy of Sciences of the United States of America 39(10):1095–1100. https://doi.org/10.1073/pnas.39.10.1095
2. Gillette D (1957) Stochastic games with zero stop probabilities. Contributions to the Theory of Games (AM-39), Volume III, Princeton University Press
3. Blackwell D, Ferguson TS (1968) The big match. Ann Math Stat 39(1):159–163
4. Kaplansky I (1945) A contribution to Von Neumann’s theory of games. Annals of Mathematics Second Series 46(3):474–479
5. Filar JA (1981) Ordered field property for stochastic games when the player who controls transitions changes from state to state. J Optim Theory Appl 34(4):501–515
6. Filar JA (1985) The completely mixed single-controller stochastic game. Proc Am Math Soc 95(4):585–594
7. Fink AM (1964) equilibrium in a stochastic n-person game. J Science of Hiroshima University 28:89–93
8. Takahashi M (1964) Equilibrium points of stochastic non cooperative n-person games. J Science of Hiroshima University 28:95–99
9. Mertens J, Neyman A (1981) Stochastic games. Int J Game Theory 10(2):53–66
10. Parthasarathy T, Raghavan TES (1981) An orderfield property for stochastic games when one player controls transition probabilities. J Optim Theory Appl 33(issue 3):375–392. https://doi.org/10.1007/BF00935250
11. Raghavan TES (1970) Completely mixed strategies in bimatrix games. Manag Sci 2(4):709–712. https://doi.org/10.1112/jlms/2.Part_4.709
12. Filar J, Vrieze K (1996) Competitive Markov decision processes. Springer
13. Babu S, Krishnamurthy N, Parthasarathy T (2016) Stationary, completely mixed and symmetric optimal and equilibrium strategies in stochastic games. Int J Game Theory 1–22. https://doi.org/10.1007/s00182-016-0555-5
14. Filar JA, Raghavan TES (1984) A matrix game solution to single controller stochastic game. Math Oper Res 9(3):356–362
15. Hordijk A, Kallenberg LCM (1979) Linear programming and Markov decision chains. Manag Sci 25(4):352–362. https://doi.org/10.1287/mnsc.25.4.352

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.