Mathematical Programming Computation (2019) 11:341–380
https://doi.org/10.1007/s12532-018-0151-8

FULL LENGTH PAPER

Optimizing over pure stationary equilibria in consensus stopping games

Amin Dehghanian[1] · Murat Kurt[3] · Andrew J. Schaefer[2]

Received: 28 June 2016 / Accepted: 18 October 2018 / Published online: 2 November 2018

Abstract

Consensus decision-making, a widely utilized group decision-making process, requires the consent of all participants. We consider *consensus stopping games*, a class of stochastic games arising in the context of consensus decision-making that require the consent of all players to terminate the game. We show that a consensus stopping game may have many pure stationary equilibria, which in turn raises the question of equilibrium selection. Given an objective criterion, we study the NP-hard problem of finding a best pure stationary equilibrium. We characterize the pure stationary equilibria, show that they form an independence system, and develop several families of valid inequalities. We then solve the equilibrium selection problem as a mixed-integer linear program by a branch-and-cut approach. Our computational results demonstrate the effectiveness of our approach over a commercial solver.

Keywords Stochastic stopping games · Equilibrium selection · Consensus decision-making · Veto players · Independence system · Branch-and-cut

Mathematics Subject Classification 90C11 · 90C06 · 90C40 · 91A15

The authors would like to thank Gabriel L. Zenarosa of the University of North Carolina at Charlotte, Sajad Modaresi of the University of North Carolina at Chapel Hill, Sina Modaresi of Discover Financial Services, Saed Alizamir of Yale University, and Osman Ozaltin of North Carolina State University for their assistance and feedback. We are grateful to ILOG-CPLEX developers, especially Tobias Achterberg and Daniel Junglas, for their help during implementation of our algorithm. We are also grateful to anonymous reviewers for constructive comments. This paper was supported by National Science Foundation Grants CMMI-1100082, CMMI-1333758, and CMMI-1400009. The second author contributed to this paper before joining Bristol-Myers Squibb.

✉ Murat Kurt
murat.kurt7@gmail.com

Extended author information available on the last page of the article

1 Introduction

Discrete time stochastic games have modeled dynamic competitive interactions among multiple players since they were introduced by Shapley [36]. A stochastic game consists of periods, states, actions, rewards, players, and transition probabilities. In each period, the game occupies a state, and all players choose their actions simultaneously and independently. Subsequently, each player receives a reward that depends on the current state of the game and the actions of all players. The game transitions to the next state according to a discrete probability distribution conditioned on the current state and the chosen actions. Each player seeks to maximize his own reward criterion, e.g., his total expected discounted reward.

A *strategy* for each player specifies a probability distribution over the feasible actions in each period conditioned on the current state and the history of the game up to that period. If this distribution depends only on the current state, then the strategy is called *stationary*. A strategy is called *pure* when all the probabilities come from the binary set $\{0, 1\}$. A (stationary) *strategy profile* is a collection of (stationary) strategies of all players that fully specifies all actions in the game, and it must include one and only one (stationary) strategy for each player [41]. To analyze stochastic games, there are several solution concepts such as Nash equilibrium, subgame-perfect equilibrium, and stationary equilibrium. A Nash equilibrium is a strategy profile in which no player is better off by unilaterally changing her strategy. A *subgame-perfect equilibrium* is a strategy profile that is a Nash equilibrium of every subgame of the original game. A *stationary equilibrium* is a stationary strategy profile that maximizes every player's reward criterion in each state (among all stationary and non-stationary strategy profiles) given the strategies of the other players. Equivalently, a stationary equilibrium is a subgame-perfect equilibrium that is stationary. The existence of a stationary equilibrium for a discounted stochastic game with finite state and action spaces has long been established (e.g., [20,40]). Stationary equilibrium has attracted considerable attention in applications and theory of stochastic games (see, e.g., [24, 30]), and hence, we restrict our attention to this equilibrium concept in this paper.

Stochastic games are very challenging since it is often difficult to characterize even stationary equilibria [11,25]. Solan [41] notes that "Unfortunately, to date there are no efficient algorithms to calculate either the value in zero-sum stochastic games, or equilibria in non-zero-sum games". In many contexts, players may prefer to use pure stationary strategies; however, the characterization of pure stationary equilibria is even more challenging. Another issue that has limited the applicability of stochastic games is the existence of multiple stationary equilibria, and it is difficult to find all such equilibria [1]. Two major complications arise as consequences of the existence of multiple equilibria [27]: First, it becomes more difficult to predict players' behavior, and players may not reach an equilibrium at all. Second, many existing methods find one equilibrium and provide no systematic methodology to find all equilibria. Such complications raise an even more challenging question of finding a best stationary equilibrium with respect to a given criterion.

We consider *consensus stopping games*, a broad class of stochastic stopping games defined over a finite set of players, states, actions, and rewards. Such a game dynamically evolves over an infinite-horizon setting. In each period of the game, each player decides whether to offer to stop or continue the game. If all players offer to stop, the game terminates, and each player receives a lump-sum stopping reward. Conversely, if at least one player decides to continue, each player receives an immediate continuation reward, the game moves into the next state according to a Markovian transition, and the rest of the players must continue regardless of their decisions. In this paper, we study the problem of finding a best pure stationary equilibrium for this class of games. Our primary motivations for studying this problem include, but not limited to: (1) Optimizing over (pure stationary) equilibria rather than identifying such equilibria is inherently interesting, and we are the first who investigate it comprehensively. For a consensus stopping game, of interest equilibria are: (i) Maximizing a linear combination of payoffs; (ii) Minimizing the time until the termination of this game; (iii) Minimizing probability that the game never terminates. (2) Consensus stopping games are an important class of stochastic games with various applications, as discussed in the rest of this paper.

1.1 Summary of contributions

Motivated by consensus decision-makings and their applications, we consider consensus stopping games. Our study reveals a rich structure of these games that allows us to investigate the problem of finding a best pure stationary equilibrium. Such a problem has not been comprehensively investigated before for a class of stochastic games. The specific contributions of this paper are as follows: First, we characterize the pure stationary equilibria, and show that they form an independence system. This is used as a basis to derive two families of combinatorial valid inequalities for an mixed-integer linear program (MILP) developed for the problem of finding a best pure stationary equilibrium. Second, we develop an efficient branch-and-cut algorithm to solve the MILP by applying these valid inequalities. We also develop a family of Pareto-optimal optimality cuts, a family of upper bounds, and several algorithmic enhancements. Our extensive computational experiments show that the algorithm greatly outperforms a state-of-the-art commercial MILP solver.

This paper is the first to provide combinatorial characterizations of stationary equilibria for a class of stochastic games. It is also the first attempt to develop a novel cutting plane approach for the problem of finding a best pure stationary equilibrium for a class of stochastic games.

1.2 Outline of the paper

The remainder of this paper is organized as follows. In Sect. 2 we review the literature. In Sect. 3 we define consensus stopping games and the problem of finding a best pure stationary equilibrium. In Sect. 4, we elaborate on two applications of consensus stopping games. We characterize the pure stationary equilibria and develop two families of valid inequalities in Sect. 5. In Sect. 6 we develop an MILP formulation that optimizes over the set of pure stationary equilibria. We develop a branch-and-cut approach in Sect. 7, and describe our computational experiments in Sect. 8. All proofs are in the "Appendix".

2 Literature review

The stochastic game literature is vast (see the recent survey by Solan [41]). In the economics literature, stochastic games typically model dynamic interactions among firms [12,15]. To find a stationary equilibrium of a stochastic game, a common approach is to apply the homotopy method [7,24]. Weintraub et al. [46] introduce the oblivious equilibrium notion for Ericson and Pakes [15]-style models, and show that it can approximate stationary equilibria under some conditions. Weintraub et al. [47] develop an algorithm for computing an oblivious equilibrium. In the operations research literature, mathematical programming has been used to compute a stationary equilibrium of stochastic games. For instance, Filar and Vrieze [17] and Raghavan and Syed [35] compute a stationary equilibrium for certain classes of stochastic games. Filar et al. [18] present a nonlinear programming formulation whose global optima are the stationary equilibria of a finite discounted stochastic game; however, there was no computational study. Note that these approaches attempt to identify a single stationary equilibrium, and because of the multiplicity of stationary equilibria in stochastic games [1,12,24], they may be viewed as heuristics for finding a best stationary equilibrium.

This paper is also related to the literature of stopping games (for a survey, see, e.g., [33, Part III]). The literature includes different variants for the definition of stopping games [23]. This stream of research usually addresses the existence of an equilibrium (see, e.g., [23,37,42]).

What distinguishes our work from the literature of stochastic games and stopping games is that we focus on the more challenging question of finding a *best* pure stationary equilibrium, compared to the question of finding a stationary equilibrium or the question of establishing the existence of an equilibrium because for consensus stopping games, (1) finding a stationary equilibrium is trivial; (2) the game may possess many pure stationary equilibria, many of which may be Pareto-inefficient with respect to the players' payoff profile; and (3) we are able to optimize over pure stationary equilibria by providing effective algorithmic approaches to choose among those equilibria.

3 Consensus stopping games

We provide a detailed description of consensus stopping games, and present necessary and sufficient conditions for a strategy profile to be a pure stationary equilibrium. We define a consensus stopping game, $\mathcal{G}$, as follows: Let $\mathcal{N} = \{1, 2, \ldots, N\}$ be a set of players, and $\mathscr{S}$ represent the finite state space of $\mathcal{G}$. In each period, each player decides whether to offer to stop or continue based on the current state $s \in \mathscr{S}$. All players make their decisions independently and simultaneously, and $a_i(s) \in \mathbb{B}$ denotes player i's action at state s, where $a_i(s)$ is 1 if he offers to stop, and 0 otherwise. Because we only focus on pure stationary strategies, $a_i(s)$ is sufficient to characterize action of player i at state s in each period. If all players offer to stop at state s (i.e., if $\prod_{i \in \mathcal{N}} a_i(s) = 1$), $\mathcal{G}$ terminates and each player $i \in \mathcal{N}$ receives a lump-sum stopping reward $u_i(s, 1)$. If at least one player decides to continue (i.e., if $\prod_{i \in \mathcal{N}} a_i(s) = 0$), $\mathcal{G}$ moves into the next state s' under a Markovian transition probability $\mathscr{P}(s'|s)$ while each player $i \in \mathcal{N}$

receives an immediate continuation reward $u_i(s, 0)$. Each player $i \in \mathcal{N}$ has a periodic discount factor $\lambda_i \in [0, 1)$, and he seeks to maximize his total expected discounted reward. $\mathcal{G}$ is an almost perfect information game, i.e., at the beginning of each period, all players are perfectly informed of: (1) the current state along with all the actions and states that have already been realized, (2) the Markovian transition probabilities, and (3) their own reward, and the other players' rewards in each state.

We follow the convention that a term in bold refers to a real-valued vector; i.e., $\mathbf{v}$ refers to the vector $\langle v(s)\rangle_{s\in\mathscr{S}}$. Given a vector $\mathbf{v} \in \mathbb{R}^{|\mathscr{S}|}$, define $F_i(s, \mathbf{v}) := u_i(s, 0) + \lambda_i \sum_{s'\in\mathscr{S}} \mathscr{P}(s'|s)v(s')$ for any $s \in \mathscr{S}, i \in \mathcal{N}$, which we interpret as player i's expected continuation payoff starting from state s where $\mathbf{v}$ represents the payoffs in all possible states in the next period. Note that for all $s \in \mathscr{S}$, $F_i(s, \mathbf{v})$ is a monotonically increasing operator in $\mathbf{v}$ in the sense that if $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{|\mathscr{S}|}$ and $\mathbf{v}_1 \geq \mathbf{v}_2$ (i.e., $v_1(\bar{s}) \geq v_2(\bar{s})$ for all $\bar{s} \in \mathscr{S}$), then $F_i(s, \mathbf{v}_1) \geq F_i(s, \mathbf{v}_2)$ since $\mathscr{P}(s'|s) \geq 0$ for all $s' \in \mathscr{S}$. In the sequel of this paper, unless otherwise stated, we use the terms *strategy* and *equilibrium* to refer to pure stationary strategy and pure stationary equilibrium, respectively. Let $\mathbf{a}_i := \langle a_i(s)\rangle_{s\in\mathscr{S}}$ denote a strategy of player i for each $i \in \mathcal{N}$, $\mathbf{a} := \langle \mathbf{a}_i\rangle_{i\in\mathcal{N}}$ denote the resulting strategy profile, and $\mathbf{a}_{-i} := \langle \mathbf{a}_j\rangle_{j\in\mathcal{N}\setminus i}$ denote the collection of strategies of all players except for i. Moreover, let $w_i^{\mathbf{a}}(s)$ denote the total expected discounted reward of player i at state s under strategy profile $\mathbf{a}$. To formalize this notion, let s_t denote the state of the game in period t, and $r_i(s_t, \mathbf{a}(s_t))$ denote the reward of player $i \in \mathcal{N}$ under strategy profile $\mathbf{a}$ when state s_t is realized. Then,

$$w_i^{\mathbf{a}}(s) = \mathbb{E}\left\{ \lim_{n\to+\infty} \sum_{t=0}^{n} \lambda_i^t r_i(s_t, \mathbf{a}(s_t)) \middle| s \right\},$$

where $\mathbb{E}(\cdot)$ represents the expectation operator under the probability distribution induced by strategy profile $\mathbf{a}$ over the evolution of the states when the game is initialized at state s. Accordingly, $\mathbf{w}_i^{\mathbf{a}} := \left\langle w_i^{\mathbf{a}}(s)\right\rangle_{s\in\mathscr{S}}$ represents the payoff vector for player i under strategy profile $\mathbf{a}$. In a slight abuse of notation, let $\mathbf{w}^{\mathbf{a}} := \left\langle \mathbf{w}_i^{\mathbf{a}}\right\rangle_{i\in\mathcal{N}}$ represent the payoff profile under strategy profile $\mathbf{a}$. The outcome of the players' decisions (continuation or termination of the game) at each state s is uniquely characterized by a binary variable $x(s) := \prod_{i\in\mathcal{N}} a_i(s)$, where $\mathbf{x} := \langle x(s)\rangle_{s\in\mathscr{S}} \in \mathbb{B}^{|\mathscr{S}|}$ is the corresponding vector. Note that if all players offer to stop, then $x(s) = 1$; otherwise, $x(s) = 0$. For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$, let $A_{\mathbf{x}} := \left\{\mathbf{a} \in \mathbb{B}^{|\mathscr{S}|\times|N|} | \prod_{i\in\mathcal{N}} a_i(s) = x(s)\ \forall s \in \mathscr{S}\right\}$ which represents the set of strategy profiles inducing the same outcome of the players' decisions for all $s \in \mathscr{S}$.

Definition 1 A strategy profile $\mathbf{a}$ with associated payoff profile $\mathbf{w}^{\mathbf{a}}$ is an equilibrium if for all $i \in \mathcal{N}$, $\mathbf{a}_i$ is a best response to $\mathbf{a}_{-i}$; or mathematically, for all $i \in \mathcal{N}$ and $s \in \mathscr{S}$

$$w_i^{\langle \mathbf{a}_i, \mathbf{a}_{-i}\rangle}(s) \geq w_i^{\langle \mathbf{a}', \mathbf{a}_{-i}\rangle}(s) \qquad \forall \mathbf{a}'. \tag{1}$$

Equation (1) presents a true but practically hard to verify definition of an equilibrium. Hence, we present a practical statement of equilibrium in the next proposition.

Proposition 1 *Given a strategy profile* $\mathbf{a}$ *with associated payoff profile* $\mathbf{w}^{\mathbf{a}}$*:*

(i) *For all* $s \in \mathscr{S}, i \in \mathcal{N}$,

$$w_i^{\mathbf{a}}(s) = \left(\prod_{i \in \mathcal{N}} a_i(s)\right) u_i(s, 1) + \left(1 - \prod_{i \in \mathcal{N}} a_i(s)\right) F_i(s, \mathbf{w}_i^{\mathbf{a}}). \tag{2}$$

Furthermore, $\mathbf{w}^{\mathbf{a}}$ *is the unique solution for this set of equations.*

(ii) $\mathbf{a}$ *is an equilibrium if and only if for all* $s \in \mathscr{S}, i \in \mathcal{N}$:

$$w_i^{\mathbf{a}}(s) = \max \left\{ \left(\prod_{\substack{j \in \mathcal{N}, \\ j \neq i}} a_j(s)\right) u_i(s, 1) + \left(1 - \prod_{\substack{j \in \mathcal{N}, \\ j \neq i}} a_j(s)\right) F_i(s, \mathbf{w}_i^{\mathbf{a}}), F_i(s, \mathbf{w}_i^{\mathbf{a}}) \right\}. \tag{3}$$

(iii) *For each* $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$*, there exists an equilibrium in* $A_{\mathbf{x}}$ *if and only if strategy profile* $\bar{\mathbf{a}}$*, defined by* $\bar{\mathbf{a}}_i := \mathbf{x}$ *for all* $i \in \mathcal{N}$*, is an equilibrium.*

Parts (i) and (ii) are standard results in the literature of discounted stochastic games (see, e.g., [20]). We provide the proof for part (iii) in the "Appendix". For each $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$, Proposition 1 (i) implies that all strategy profiles in $A_{\mathbf{x}}$ have the same payoff profile. In other words, $\mathbf{x}$ is necessary and sufficient information for characterizing the payoff profile of a strategy profile $\mathbf{a}$. As we are interested in studying the set of equilibrium payoff profiles, by Proposition 1 (iii), it is sufficient to only focus on the set of strategy profiles in which $\mathbf{a}_i = \mathbf{x}$ for all $i \in \mathcal{N}$. We define such a set of strategy profiles as the set of *unanimous* strategy profiles, since all players take the same action at each state $s \in \mathscr{S}$. Hereafter, we restrict our attention to the set of unanimous strategy profiles, and with a slight abuse of notation, $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$ represents a unanimous strategy profile. Accordingly, we define $w_i^{\mathbf{x}}(s)$, $\mathbf{w}_i^{\mathbf{x}} := \langle w_i^{\mathbf{x}}(s) \rangle_{s \in \mathscr{S}}$, and $\mathbf{w}^{\mathbf{x}} := \langle \mathbf{w}_i^{\mathbf{x}} \rangle_{i \in \mathcal{N}}$ for unanimous strategy profile $\mathbf{x}$. We may restate Proposition 1 (i) and (ii) for unanimous strategy profiles as follows (for convenience we drop the word unanimous hereafter).

Proposition 2 *Given a strategy profile* $\mathbf{x}$ *with associated payoff profile* $\mathbf{w}^{\mathbf{x}}$*:*

(i) *For all* $s \in \mathscr{S}, i \in \mathcal{N}$,

$$w_i^{\mathbf{x}}(s) = x(s) u_i(s, 1) + (1 - x(s)) F_i(s, \mathbf{w}_i^{\mathbf{x}}). \tag{4}$$

Furthermore, $\mathbf{w}^{\mathbf{x}}$ *is the unique solution for this set of equations.*

(ii) $\mathbf{x}$ *is an equilibrium if and only if for all* $s \in \mathscr{S}, i \in \mathcal{N}$:

$$w_i^{\mathbf{x}}(s) = \max \left\{ x(s) u_i(s, 1) + (1 - x(s)) F_i(s, \mathbf{w}_i^{\mathbf{x}}), F_i(s, \mathbf{w}_i^{\mathbf{x}}) \right\}. \tag{5}$$

Proposition 2 follows immediately from Proposition 1, and the fact that strategy profile $\mathbf{x}$ is unanimous. Part (i) describes how to calculate the payoff profile of a strategy profile, and Part (ii) describes the Bellman–Shapley equations for $\mathcal{G}$.

In this paper, we study the problem of finding a best equilibrium, with respect to a given linear objective function of payoffs, for $\mathcal{G}$. Such an objective function is well accepted in the literature of group decision analysis and may easily capture various practical criterion (see, e.g., [13] and the references therein). Let $\Psi := \{\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|} | \, \mathbf{x}$ is an equilibrium of $\mathcal{G}\}$, $c_i(s) \in \mathbb{R}$ be an objective function coefficient for all $s \in \mathscr{S}$, $i \in \mathcal{N}$, and $\mathbf{c} := \langle c_i(s) \rangle_{s \in \mathscr{S}, i \in \mathcal{N}}$. Therefore, the problem of finding a best equilibrium is:

$$(\mathbf{P}) : \max\{\mathbf{c}^\top \mathbf{w}^\mathbf{x} | \, \mathbf{x} \in \Psi\}. \tag{6}$$

We present an MILP formulation for (**P**) in Sect. 6. Dehghanian [9] shows that (**P**) is NP-hard.

4 Application areas

Consensus stopping games arise in the context of consensus decision-making, which is indeed used by many international organizations in policy making. For instance, the Council of the European Union requires unanimity in some policy areas such as membership, taxation, social security, foreign and common defense policy and operational police cooperation among the Member States [16]. The World Trade Organization, the Association of Southeast Asian Nations, the North Atlantic Treaty Organization, the Conference on Security and Cooperation in Europe, the Executive Committee of the International Monetary Fund, and the Organization for Economic Cooperation and Development all make decisions by consensus [43]. Indeed, in consensus decision-making, every player has a veto in the sense that he can prevent a change from the status quo [44].

Consensus stopping games may model many dynamic noncooperative consensus decision-making environments to reach a permanent agreement. In the following, we elaborate on the elements of a consensus stopping game, e.g., the players, periods, actions, states, rewards, and transition probabilities, for two specific applications.

War termination Consider two countries fighting a war against each other until reaching peace or one side's definite victory, whichever occurs first. Periodically, each country chooses between continuing the war and offering peace. The war ends if and only if both countries offer peace simultaneously. This situation may be modeled by a consensus stopping game in which the players are countries, the decision epochs are, for example, daily, and the actions are whether to continue the war or offer peace. Political scientists have utilized the Composite Index of National Capabilities (CINC) score to measure power of a country and explain war outcomes. In calculating CINC scores, geopolitical factors such as a country's relative military, economic, and demographic capabilities are considered (e.g., [38,39]). Therefore, we may consider the CINC scores of both countries as the state of the game. The expected increase in the area of the occupied territories and the total area of the occupied territories since the

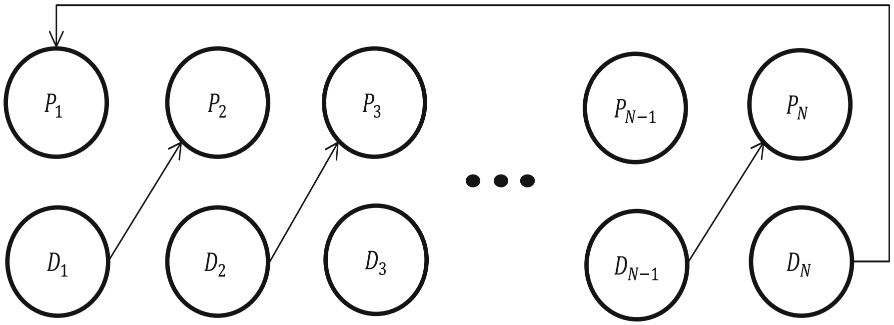

Fig. 1 A cycle of N patient–donor pairs for a PKE, where D_i refers to donor of intended recipient (or patient) P_i. Directed arcs in the cycle represent compatibility between patients and donors; that is, no donor is compatible with his intended recipient but only compatible with the intended recipient of the next donor in the cycle

beginning of the war may be regarded as the immediate continuation and stopping rewards, respectively. A transition probability matrix for the dynamic evolution of the CINC scores in the case of continuing the war reflects exogenous factors, which are out of the control of both countries, such as natural events, third parties' actions, etc. Other models of war termination may be found in the political science literature. For instance, Filson and Werner [19] present a two-stage asymmetric information bargaining game by which they provide several insights on the onset and termination of war. As another instance, Cunningham [8] studies the correlation between the duration of a civil war and the number of veto players. This statistical analysis shows that the more the number of veto players, the longer the conflict.

Organ exchange End-stage renal disease (ESRD) is the final stage of chronic kidney disease in which both kidneys are failing. The preferred choice of treatment for ESRD is living-donor transplantation, in which a living-donor donates one of his kidneys to the patient. A significant barrier to greater use of living-donors is that at least one-third of the patients with a willing living-donor are unable to receive the donor kidney due to blood type and/or tissue incompatibilities [31]. To mitigate this barrier, an emerging clinical practice is a paired kidney exchange (PKE) in which $N \geq 2$ self-interested incompatible patient–donor pairs for whom the only compatible exchange of kidneys is cyclical, swap their donors (see Fig. 1).

Consider a central PKE program in which many incompatible patient–donor pairs are enrolled. An important problem in such an environment is which incompatible patient–donor pairs should be assigned to the same PKE cycle subject to biological compatibility of those pairs. This assignment problem has been studied by many researchers (see, e.g., [3,22,45]), where its objective function is equal to the sum over values of all planned PKE cycles within the central PKE program. For this purpose, we need to value each potential PKE cycle, and the total number of transplants within a PKE cycle is considered as the value of the cycle in the literature, e.g., [3,22,45]. Kurt [28] notes that this valuation approach misses significant factors such as patient–donor autonomy and the effects of disease severity of patient in his/her decision. To overcome this shortcoming, Kurt [28] proposes a new stochastic game valuation approach in which a set of patient–donor pairs are assigned to the same PKE cycle. Moreover, he assumes that the PKE cycle is constant over time, in the sense that patient–donor pairs cannot change their PKE cycle from period to period. Once the PKE's cycles are planned, the patient–donor pairs in each planned cycle play the game of deciding the best time for the exchange to occur. More specifically, each patient–donor pair periodically decides whether to offer to exchange or not. If at least one patient–donor pair decides not to offer to exchange, the transplant exchange cycle breaks as each donor is willing to donate a kidney only in return for receiving a kidney for his intended recipient. Therefore, the transplants in the cycle are accomplished if and only if all patient–donor pairs offer to exchange consensually. Hence, this is a consensus decision-making environment, and it may be modeled as a consensus stopping game as patient–donor pairs cannot change their PKE cycle from period to period.

In Kurt [28], the players are the patient–donor pairs, the decision epochs are biweekly, and the players' actions are whether to offer to exchange or not. Glomerular filtration rates (GFRs), a measure of kidney functionality, of all patients are considered as the state of the game. Note that as the GFR of a patient may improve over time, the patient–donor pair does not necessarily offer to exchange as soon as the game starts [48]. Roughly speaking, the immediate continuation and stopping rewards are estimated by the expected number of days until the next decision epoch (14 days) and the expected number of post-transplant survival days for each patient, respectively. A transition probability matrix for the case of not offering to exchange describes the Markovian progression of the GFRs. Finally, Kurt [28] represents the problem of finding a best equilibrium of the game, with respect to a given linear objective function of payoffs, as an MILP. Our work contributes to this line of research by providing a solution method that is able to solve much larger instances of the MILP. More specifically, Kurt [28] solves the MILP directly by a commercial MILP solver. He is able to solve only two player instances with 11 states per player, and three player instances with 4 states per player. In Sect. 8, we demonstrate that our solution method can handle instances including two players with 60 states per player and instances including three players with 15 states per player. Solving larger instances improves modeling accuracy of kidney functionality, which is clinically important. Consensus stopping games may be applied to model the timing of transplants for other organ exchange problems, e.g., liver and lung exchanges [14].

5 Characterizing equilibria and combinatorial valid inequalities

In this section, we characterize the equilibria of $\mathcal{G}$ and develop two families of combinatorial valid inequalities for (**P**) to improve its representation. We first need to define several functions from the space of strategy profiles to the collection of all subsets of $\mathscr{S}$ as follows. Given a strategy profile $\mathbf{x}$, let:

$$\mathscr{S}^k(\mathbf{x}) = \{s \in \mathscr{S} \mid x(s) = k\} \qquad \forall k \in \{0,1\}, \tag{7a}$$

$$\mathscr{S}^k_{i,nv}(\mathbf{x}) = \{s \in \mathscr{S}^k(\mathbf{x}) \mid u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})\} \qquad \forall k \in \{0,1\}, i \in \mathscr{N}, \tag{7b}$$

$$\mathscr{S}^k_{i,v}(\mathbf{x}) = \{s \in \mathscr{S}^k(\mathbf{x}) \mid u_i(s,1) < F_i(s, \mathbf{w}_i^{\mathbf{x}})\} \qquad \forall k \in \{0,1\}, i \in \mathscr{N}. \tag{7c}$$

It can easily be seen that $\{\mathscr{S}^k_{i,v}(\mathbf{x}), \mathscr{S}^k_{i,nv}(\mathbf{x})\}$ is a partition of $\mathscr{S}^k(\mathbf{x})$ for all $i \in \mathscr{N}$, $k \in \{0,1\}$, and $\{\mathscr{S}^0(\mathbf{x}), \mathscr{S}^1(\mathbf{x})\}$ is a partition of $\mathscr{S}$ for each strategy profile $\mathbf{x}$. $\mathscr{S}^0(\mathbf{x})$ and $\mathscr{S}^1(\mathbf{x})$ are the sets of *continuing* and *stopping* states under $\mathbf{x}$, respectively. $\mathscr{S}^1_{i,v}(\mathbf{x})$ represents the set of stopping states in which the Bellman–Shapley equation (5) is violated for player i under $\mathbf{x}$; conversely, $\mathscr{S}^1_{i,nv}(\mathbf{x})$ represents the set of stopping states in which the Bellman–Shapley equation (5) is satisfied for player i under $\mathbf{x}$. The sets $\mathscr{S}^0_{i,v}(\mathbf{x})$ and $\mathscr{S}^0_{i,nv}(\mathbf{x})$ can similarly be interpreted for player i under $\mathbf{x}$.

Proposition 3 *Given strategy profiles* $\mathbf{x}$ *and* $\bar{\mathbf{x}}$ *with respective payoff profiles* $\mathbf{w}^{\mathbf{x}}$ *and* $\mathbf{w}^{\bar{\mathbf{x}}}$*:*

(i) *If there exists* $i \in \mathscr{N}$ *such that* $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$, *then* $w_i^{\mathbf{x}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ *for all* $s \in \mathscr{S}$.
(ii) *If there exists* $i \in \mathscr{N}$ *such that* $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$, *then* $w_i^{\mathbf{x}}(s) \leq w_i^{\bar{\mathbf{x}}}(s)$ *for all* $s \in \mathscr{S}$.

Lemma 1 *A strategy profile* $\mathbf{x}$ *is an equilibrium if and only if* $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ *for all* $i \in \mathscr{N}$.

Proposition 4 *Suppose a strategy profile* $\bar{\mathbf{x}}$ *is an equilibrium. Then, the following holds for any strategy profile* $\mathbf{x}$ *satisfying* $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$*:*

(i) *For all* $s \in \mathscr{S}, i \in \mathscr{N}$, $w_i^{\mathbf{x}}(s) \leq w_i^{\bar{\mathbf{x}}}(s)$.
(ii) $\mathbf{x}$ *is an equilibrium.*

The next corollary is an immediate consequence of Proposition 4 (ii).

Corollary 1 *If a strategy profile* $\bar{\mathbf{x}}$ *is not an equilibrium, then any strategy profile* $\mathbf{x}$ *where* $\mathscr{S}^1(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$ *is not an equilibrium.*

An *independence system* is a well-known combinatorial structure which is composed of a collection of subsets of a ground set, called independent sets, that satisfy two conditions: First, the empty set is an independent set. Second, every subset of an independent set is also an independent set [32].

Corollary 2 *Let* $\mathcal{J}$ *be the collection of all* $\mathscr{S}^1(\mathbf{x})$ *such that* $\mathbf{x}$ *is an equilibrium. The pair* $(\mathscr{S}, \mathcal{J})$ *is an independence system.*

Definition 2 An equilibrium $\bar{\mathbf{x}}$ is *maximal* if there does not exist any equilibrium $\mathbf{x}$ such that $\mathscr{S}^1(\bar{\mathbf{x}}) \subset \mathscr{S}^1(\mathbf{x})$.

Proposition 4 demonstrates that an equilibrium $\bar{\mathbf{x}}$ yields $2^{\sum_{s \in \mathscr{S}} \bar{x}(s)} - 1$ additional equilibria (i.e., any strategy profile $\mathbf{x}$ satisfying $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$), and this property describes why $\mathcal{G}$ may possess many equilibria in general. We illustrate this point by the following example: Consider a consensus stopping game with three states $\{s_1, s_2, s_3\}$, and suppose that strategy profile $\big(x(s_1), x(s_2), x(s_3)\big) = (1, 1, 0)$ is an equilibrium. Proposition 4 implies that the three strategy profiles, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 0)$ are equilibria as well.

Given an equilibrium $\bar{\mathbf{x}}$, all equilibria $\mathbf{x}$ satisfying $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$ are payoff-wise dominated by $\bar{\mathbf{x}}$ according to Proposition 4, and if $c_i(s) \geq 0$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$, they can be eliminated from consideration when we search for a best equilibrium. Based on this dominance, we may restrict our search to maximal equilibria and develop the following valid inequality for the optimality of a solution to (**P**) when the objective function coefficients are non-negative.

Proposition 5 (i) *If $c_i(s) \geq 0$ for all $s \in \mathscr{S}, i \in \mathcal{N}$, then there exists an optimal equilibrium that is maximal.*

(ii) *If a strategy profile $\bar{\mathbf{x}}$ is an equilibrium, then the inequality*

$$\sum_{s \in \mathscr{S}^1(\bar{\mathbf{x}})} [1 - x(s)] \leq |\mathscr{S}^1(\bar{\mathbf{x}})| \sum_{s \in \mathscr{S}^0(\bar{\mathbf{x}})} x(s) \tag{8}$$

is satisfied by every maximal equilibrium $\mathbf{x}$.

Proposition 6 *Given a strategy profile $\bar{\mathbf{x}}$, a strategy profile $\mathbf{x}$ is not an equilibrium if there exists $i \in \mathcal{N}$ such that $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$ and $\mathscr{S}^1(\mathbf{x}) \cap (\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})) \neq \emptyset$.*

An interesting feature of Proposition 6 is that it can provide insights about equilibria irrespective of whether strategy profile $\bar{\mathbf{x}}$ is an equilibrium or not.

Proposition 7 *Given a strategy profile $\bar{\mathbf{x}}$, the following inequalities are valid for Ψ:*

$$\sum_{s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})} x(s) \leq (|\mathscr{S}^0_{i,v}(\bar{\mathbf{x}})| + |\mathscr{S}^1_{i,v}(\bar{\mathbf{x}})|) \sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})} [1 - x(s)] \quad \forall i \in \mathcal{N}. \tag{9}$$

An advantage of Proposition 7 is that for each player $i \in \mathcal{N}$ where $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, we can derive a nontrivial valid inequality.

6 Equilibrium selection formulation

In this section, we present an MILP formulation for (**P**). Let coefficient $V_i(s)$ be an upper bound for equilibria payoffs of player i at state s. Kurt [28] suggests that for each player $i \in \mathcal{N}$, $\mathbf{V}_i := \langle V_i(s)\rangle_{s \in \mathscr{S}}$ may be calculated as a solution of the Markov decision process (MDP) equations $V_i(s) = \max\{u_i(s, 1), F_i(s, \mathbf{V}_i)\}$ for all $s \in \mathscr{S}$, and shows it is a valid upper bound as it represents optimal value function of player i in an MDP where he is maximizing his own payoffs when the autonomy of the other players is suppressed. By using this set of parameters, he proposes an MILP to represent the set of equilibria of $\mathcal{G}$, relying on the assumption that $u_i(s, 0), u_i(s, 1) \geq 0$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$. Let $\mathbf{0}$ be the strategy profile in which $x(s) = 0$ for all $s \in \mathscr{S}$, and

$\mathbf{d} := \langle d_i(s)\rangle_{s\in\mathscr{S},i\in\mathcal{N}}$ be its associated payoff profile. We propose a similar formulation to represent the set of equilibria of $\mathcal{G}$ in the following:

$$
\begin{aligned}
& w_i(s) \geq F_i(s, \mathbf{w}_i) && \forall s \in \mathscr{S}, i \in \mathcal{N}, && \text{(10a)}\\
& w_i(s) \leq F_i(s, \mathbf{w}_i) + [u_i(s,1) - d_i(s)]x(s) && \forall s \in \mathscr{S}, i \in \mathcal{N}, && \text{(10b)}\\
& w_i(s) \geq [u_i(s,1) - d_i(s)]x(s) + d_i(s) && \forall s \in \mathscr{S}, i \in \mathcal{N}, && \text{(10c)}\\
& w_i(s) \leq u_i(s,1)x(s) + F_i(s, \mathbf{V}_i)[1 - x(s)] && \forall s \in \mathscr{S}, i \in \mathcal{N}, && \text{(10d)}\\
& w_i(s) \text{ unrestricted} && \forall s \in \mathscr{S}, i \in \mathcal{N}, && \text{(10e)}\\
& x(s) \in \{0, 1\} && \forall s \in \mathscr{S}. && \text{(10f)}
\end{aligned}
$$

Let $\Delta := \{\langle \mathbf{x}, \mathbf{w}\rangle \,|(10a)\text{–}(10f)\}$.

Proposition 8 (i) $\langle \mathbf{x}, \mathbf{w}\rangle \in \Delta$ *if and only if strategy profile* $\mathbf{x}$ *is an equilibrium with associated payoff profile* $\mathbf{w}$.

(ii) *If* $u_i(s,0), u_i(s,1) \geq 0$ *for all* $s \in \mathscr{S}, i \in \mathcal{N}$, *then* Δ *is at least as strong as the MILP formulation proposed by Kurt* [28] *for pure stationary equilibria.*

Proposition 8 (i) implies that Ψ is the projection of Δ onto the $\mathbf{x}$-space and (**P**) may be reformulated as the following MILP:

$$
\begin{aligned}
\max \sum_{s\in\mathscr{S}, i\in\mathcal{N}} & c_i(s)w_i(s) && \text{(11a)}\\
s.t. & \\
& \langle \mathbf{x}, \mathbf{w}\rangle \in \Delta. && \text{(11b)}
\end{aligned}
$$

7 Algorithmic approach

In this section we develop an algorithm to solve (**P**) efficiently. (**P**) has a so-called "block-ladder" structure where $\mathbf{x}$ are the linking variables, and is amenable to Benders' decomposition [5]. However, our computational experiments showed that even a state-of-the-art implementation of Benders' decomposition [21] is ineffective. In this section, our goal is to develop a cutting plane approach that is carefully tailored to solve (**P**). Let θ_i be an artificial variable in the relaxed master problem that approximates $\sum_{s\in\mathscr{S}} c_i(s)w_i(s)$ for any $i \in \mathcal{N}$ during the algorithm, and let $\boldsymbol{\theta} := \langle\theta_i\rangle_{i\in\mathcal{N}}$. Before we present our decomposition algorithm, let *RMP*, LB, UB, $\mathbf{x}_{incum}$, and ϵ denote the restricted master problem, lower bound, upper bound, the incumbent solution, and a termination tolerance, respectively.

Decomposition algorithm

0. Initialization: Let $LB := -\infty$, $UB := +\infty$, $\mathbf{x}_{incum} := \emptyset$ and *RMP* be as follows:

$$\max \sum_{i \in \mathcal{N}} \theta_i \tag{12a}$$

$$s.t.$$

$$\theta_i \leq \sum_{s \in \mathscr{S}} c_i(s) V_i(s) \qquad \forall i \in \mathcal{N}, \tag{12b}$$

$$\theta_i \text{ unrestricted} \qquad \forall i \in \mathcal{N}, \tag{12c}$$

$$x(s) \in \{0, 1\} \qquad \forall s \in \mathscr{S}. \tag{12d}$$

1. Restricted Master: Solve *RMP* and obtain an optimal solution $(\bar{\mathbf{x}}, \bar{\boldsymbol{\theta}})$, and let $UB := \sum_{i \in \mathcal{N}} \bar{\theta}_i$.

2. Separation:

(a) Calculate payoff profile $\mathbf{w}^{\bar{\mathbf{x}}}$ associated with $\bar{\mathbf{x}}$ using Proposition 2 (i), and characterize the sets $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}})$, $\mathscr{S}^0_{i,nv}(\bar{\mathbf{x}})$, $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$ for all $i \in \mathcal{N}$.

(b) For any i for which $\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, add the valid inequality (9) to *RMP*.

(c) If $\bar{\mathbf{x}}$ is an equilibrium (i.e., $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathcal{N}$) and $\bar{\theta}_i > \sum_{s \in \mathscr{S}} c_i(s) w_i^{\bar{\mathbf{x}}}(s)$ for some $i \in \mathcal{N}$, then:

 (i) Add a Benders' optimality cut [5] to *RMP* for $i \in \mathcal{N}$ for which $\bar{\theta}_i > \sum_{s \in \mathscr{S}} c_i(s) w_i^{\bar{\mathbf{x}}}(s)$;

 (ii) Add the valid inequality (8) to *RMP* if $c_i(s) \geq 0$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$;

 (iii) If $\mathbf{c}^\top \mathbf{w}^{\bar{\mathbf{x}}} \geq LB$, let $\mathbf{x}_{incum} := \bar{\mathbf{x}}$, $LB := \mathbf{c}^\top \mathbf{w}^{\bar{\mathbf{x}}}$.

3. Termination: If $UB - LB \leq \epsilon$, then terminate the algorithm. Otherwise, go to Step 1.

If $\bar{\mathbf{x}}$ is not an equilibrium, there exists some $i \in \mathcal{N}$ such that $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \neq \emptyset$, and hence (9) cuts off $\bar{\mathbf{x}}$ in Step (2b). Hence, the Decomposition Algorithm finitely converges for any $\epsilon \in \mathbb{R}_+$. There are several differences between our proposed algorithm and Benders' decomposition: First, we solve the set of equations (4) as the subproblem. Second, we use the combinatorial feasibility cut (9). Third, we use the combinatorial optimality cut (8) in addition to a Benders' optimality cut.

In order to calculate $\mathbf{w}_i^{\bar{\mathbf{x}}}$ in Step (2a) for each $i \in \mathcal{N}$, by Proposition 2 (i) we solve the following linear program (LP):

$$\mathcal{R}_i(\bar{\mathbf{x}}) : \max \sum_{s \in \mathscr{S}} c_i(s) w_i(s)$$

$$s.t.$$

$$w_i(s) = F_i(s, \mathbf{w}_i) \qquad \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}),$$

$$w_i(s) = u_i(s, 1) \qquad \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}),$$

$$w_i(s) \text{ unrestricted} \qquad \forall s \in \mathscr{S}.$$

For each $i \in \mathcal{N}$, let $\gamma_i(s)$ be optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$ for all $s \in \mathscr{S}$, $y^+ := \max\{0, y\}$ and $y^- := -\min\{0, y\}$ for any $y \in \mathbb{R}$. To enhance the Decomposition Algorithm, we seek a Pareto-optimal Benders' optimality cut [29].

Proposition 9 *The following Benders' optimality cut is Pareto-optimal.*

$$\begin{aligned}\theta_i \leq & \sum_{s \in \mathscr{S}^0(\bar{\mathbf{x}})} \gamma_i(s) u_i(s, 0) + \big(\gamma_i^+(s)[u_i(s, 1) - d_i(s)]^+ x(s)\big) \\ & + \sum_{s \in \mathscr{S}^1(\bar{\mathbf{x}})} \gamma_i(s)\, u_i(s, 1) x(s) + \big(-\gamma_i(s)^- d_i(s) \\ & + \gamma_i(s)^+ F_i(s, \mathbf{V}_i)\big)\, [1 - x(s)] \quad \forall i \in \mathcal{N}. \end{aligned} \tag{13}$$

It is worth noting that cut (13) requires only optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$.

7.1 Branch-and-cut

In the Decomposition Algorithm, we need to solve the restricted master problem repeatedly. However, solving a mixed-integer restricted master repeatedly may be time-consuming. In order to alleviate this difficulty, we use a branch-and-cut framework. At the root node, we start with the MILP in Step 0 of the Decomposition Algorithm. Then, we solve the LP relaxation at each node and generate a violated cut at each integral node by using the separation procedure described in Step 2 of the Decomposition Algorithm. We use the default branching and node selection strategies of the MILP solver, [26].

Dynamic variable fixing and pruning At node t of the branch-and-cut tree, let J_t^k be the set of all states in which $x(s)$ is fixed to k, for any $k = \{0, 1\}$. A strategy profile $\mathbf{x}^t := \langle x^t(s) \rangle_{s \in \mathscr{S}}$ can be assigned to each node t as follows:

$$x^t(s) = \begin{cases} 1 & \text{if} \quad s \in J_t^1, \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathbf{x}^t$ is not an equilibrium, node t is pruned because all possible strategy profiles at this node (admissible with respect to the binary variables being fixed so far) are not equilibria by Corollary 1. On the other hand, if $\mathbf{x}^t$ is an equilibrium, we can apply Proposition 7 to fix some of unfixed binary variables at node t and its offspring. For this reason, $x(s)$ is set to 0 at node t for all $s \in \cup_{i \in \mathcal{N}} \mathscr{S}^0_{i,v}(\mathbf{x}^t)$.

Dynamic coefficient strengthening Recall that $\langle V_i(s) \rangle_{s \in \mathscr{S}}$ is the optimal value function of player i when the autonomy of the other players is suppressed, but indeed coefficient $V_i(s)$ may be replaced with any other upper bound of $w_i(s)$. In general, $V_i(s)$ may be relatively large, which weakens the LP relaxation of Δ. To address this, we dynamically tighten upper bounds for $w_i(s)$ in progress of the branch-and-cut tree rather than using a fixed value as an upper bound for $w_i(s)$.

At leaf node t, we can obtain a tightened upper bound of $w_i(s)$ by using the optimal value function of player i for a certain MDP. In particular, consider an MDP in which the autonomy of all players except for i is suppressed, while player i is restricted to strategies admissible with respect to the binary variables which are fixed at node t. Since the set of fixed binary variables increases as we go down further in the search tree, the tightened upper bound of $w_i(s)$ does not increase. Let $V_i^t(s)$ be the tightened upper bound for $w_i(s)$ at leaf node t. For each player i, $\mathbf{V}_i^t := \langle V_i^t(s)\rangle_{s\in\mathscr{S}}$ can be calculated as the unique solution of the following MDP equations:

$$\begin{aligned}
&V_i^t(s) = F_i(s, V_i^t) && \forall s \in J_t^0,\\
&V_i^t(s) = u_i(s, 1) && \forall s \in J_t^1,\\
&V_i^t(s) = \max\{u_i(s, 1), F_i(s, \mathbf{V}_i^t)\} && \forall s \in \mathscr{S}/J_t^0 \cup J_t^1.
\end{aligned}$$

Similarly, the coefficient $d_i(s)$ is a lower bound for $w_i(s)$, and we may tighten it as we go down in the search tree. At leaf node t, consider payoff profile $\mathbf{w}^{\mathbf{x}^t}$ associated with strategy profile $\mathbf{x}^t$, as defined in the previous part. As noted earlier, if $\mathbf{x}^t$ is not an equilibrium, node t is pruned. Otherwise, $\mathbf{x}^t$ is an equilibrium, and by Proposition 4 (i), $w_i^{\mathbf{x}^t}(s)$ is a lower bound for $w_i(s)$ at node t and its offspring. Compared to $d_i(s)$, $w_i^{\mathbf{x}^t}(s)$ is a tighter lower bound by Proposition 4. The new set of upper and lower bounds is applied to generate a Benders' optimality cut (13), which is only locally valid. Since the upper and lower bounds get tighter as we go down further in the search tree, deeper Benders' optimality cuts will be generated. Needless to say, this idea should only be implemented at nodes at which we generate a Benders' optimality cut. The idea of coefficient strengthening has received some recent attention in the optimization community (e.g., [34]).

Dynamic player-aggregated upper bounds In stochastic games, it is natural to assume that the players compete in the same environment, and therefore share the same discount factor (see, e.g., [24,25]). Suppose that the discount factors are equal for all players, i.e., $\lambda_i = \lambda$ for all $i \in \mathcal{N}$. We present a family of upper bounds for the objective function at each leaf node t. Let $\alpha_i \in \mathbb{R}$ for all $i \in \mathcal{N}$, and $\alpha := \langle \alpha_i \rangle_{i\in\mathcal{N}}$. We define an aggregated MDP, $\mathcal{G}^{\alpha,t}$, over the state space $\mathscr{S}$ as follows. In each state $s \in \mathscr{S}\backslash(J_t^0 \cup J_t^1)$, we may decide whether to stop or continue. In each state $s \in J_t^1$ (J_t^0), we have to stop (continue). If we decide to stop, then the MDP terminates and we receive a stopping reward $u^\alpha(s, 1) := \sum_{i\in\mathcal{N}} \alpha_i u_i(s, 1)$. Conversely, if we decide to continue, then the MDP moves into a new state s' under the Markovian transition probability $\mathscr{P}(s'|s)$ and we receive an immediate continuation reward $u^\alpha(s, 0) := \sum_{i\in\mathcal{N}} \alpha_i u_i(s, 0)$. All future rewards are discounted at rate λ. Therefore, $\mathcal{G}^{\alpha,t}$ has the same dynamic evolution as $\mathcal{G}$, but the players' rewards are aggregated in $\mathcal{G}^{\alpha,t}$ and its action space is admissible with respect to the binary variables which are fixed at node t. Similar to those of $\mathcal{G}$, we can define strategy $\mathbf{x}$ and its associated payoffs $\mathbf{w}^{\mathbf{x},\alpha,t}$ for $\mathcal{G}^{\alpha,t}$. For each strategy $\mathbf{x}$ with associated payoffs $\mathbf{w}^{\mathbf{x},\alpha,t}$ for $\mathcal{G}^{\alpha,t}$ and associated payoff profile $\mathbf{w}^{\mathbf{x}}$ for $\mathcal{G}$, it can easily be seen that

$$w^{\mathbf{x},\alpha,t}(s) = \sum_{i\in\mathcal{N}} \alpha_i w_i^{\mathbf{x}}(s) \qquad \forall s \in \mathscr{S}. \tag{14}$$

Let $\mathbf{V}^{\alpha,t} := \langle V^{\alpha,t}(s)\rangle_{s\in\mathscr{S}}$ be the optimal value function of $\mathcal{G}^{\alpha,t}$, which is calculated as the unique solution of the following MDP equations:

$$V^{\alpha,t}(s) = u^\alpha(s,0) + \lambda \sum_{s' \in \mathscr{S}} \mathscr{P}(s'|s) V^{\alpha,t}(s') \qquad \forall s \in J_t^0,$$
$$V^{\alpha,t}(s) = u^\alpha(s,1) \qquad \forall s \in J_t^1,$$
$$V^{\alpha,t}(s) = \max\{u^\alpha(s,1), u^\alpha(s,0) + \lambda \sum_{s' \in \mathscr{S}} \mathscr{P}(s'|s) V^{\alpha,t}(s')\} \quad \forall s \in \mathscr{S}/J_t^0 \cup J_t^1.$$

Consider the following LP:

$$\bar{\theta}^t := \max \sum_{s \in \mathscr{S}, i \in \mathcal{N}} c_i(s) w_i(s) \tag{15a}$$
$$s.t.$$
$$w_i(s) = F_i(s, \mathbf{w}_i) \qquad \forall s \in J_t^0, i \in \mathcal{N}, \tag{15b}$$
$$w_i(s) \geq F_i(s, \mathbf{w}_i) \qquad \forall s \in \mathscr{S}/J_t^0, i \in \mathcal{N}, \tag{15c}$$
$$w_i(s) = u_i(s,1) \qquad \forall s \in J_t^1, i \in \mathcal{N}, \tag{15d}$$
$$\sum_{i \in \mathcal{N}} \alpha_i w_i(s) \leq V^{\alpha,t}(s) \qquad \forall s \in \mathscr{S}, \tag{15e}$$
$$w_i(s) \leq V_i^t(s) \qquad \forall s \in \mathscr{S}, i \in \mathcal{N}, \tag{15f}$$
$$w_i(s) \text{ unrestricted} \qquad \forall s \in \mathscr{S}, i \in \mathcal{N}. \tag{15g}$$

If a strategy profile $\mathbf{x}$ is an equilibrium and admissible at node t, then $\mathbf{w}^{\mathbf{x}}$ satisfies (15b)–(15d) by Proposition 2, and satisfies (15e)–(15f) by (14) and the definitions of $\mathbf{V}^{\alpha,t}$ and $\mathbf{V}_i^t$. Hence, the following inequality is locally valid at node t and its offspring:

$$\sum_{i \in \mathcal{N}} \theta_i \leq \bar{\theta}^t. \tag{16}$$

Of special interest is the case when α is equal to e_i, the ith unit vector in $\mathbb{R}^{|\mathcal{N}|}$. In this case, we do not need the above-mentioned assumption about equality of the discount factors, and we use λ_i as the discount factor of $\mathcal{G}^{e_i,t}$. Consider the following LP for each $i \in \mathcal{N}$:

$$\bar{\theta}_i^t := \max \sum_{s \in \mathscr{S}} c_i(s) w_i(s) \tag{17a}$$
$$s.t.$$
$$w_i(s) = F_i(s, \mathbf{w}_i) \qquad \forall s \in J_t^0, \tag{17b}$$
$$w_i(s) \geq F_i(s, \mathbf{w}_i) \qquad \forall s \in \mathscr{S}/J_t^0, \tag{17c}$$
$$w_i(s) = u_i(s,1) \qquad \forall s \in J_t^1, \tag{17d}$$
$$w_i(s) \leq V_i^t(s) \qquad \forall s \in \mathscr{S}, \tag{17e}$$
$$w_i(s) \text{ unrestricted} \qquad \forall s \in \mathscr{S}. \tag{17f}$$

A reasoning similar to that for validity of the inequality (16), shows that following inequality is locally valid at node t and its offspring:

$$\theta_i \leq \bar{\theta}_i^t \qquad \forall i \in \mathcal{N}. \tag{18}$$

8 Computational experiments

8.1 Implementation and test instances

We implemented the branch-and-cut algorithm described in Sect. 7, using the ILOG-CPLEX 12.6 Callable Library embedded in C++ under Microsoft Visual Studio 2010. We conducted our computational experiments on an Intel Xeon PC with 3.7 GHz CPU, 32 GB RAM, and Windows 7 (64-bit) operating system. Each instance of our test bed was processed three times within a 4-h time limit: First by our implementation of the branch-and-cut algorithm described in Sect. 7 within CPLEX; second by solving the extensive formulation [i.e., the MILP formulation (11a)–(11b)] through CPLEX (with default settings); and third by our implementation of a state-of-the-art Benders' decomposition [21] within CPLEX. Benders' decomposition was ineffective, and for brevity we do not report its computational results here.

In the implementation of our branch-and-cut algorithm, all cuts are generated at integral nodes and added locally. The valid inequalities (16) and (18) are added locally in non-integral nodes as well. Dynamic variable fixing and pruning routines are implemented via a branch callback routine.

We restrict our attention to consensus stopping game instances in which each player $i \in \mathcal{N}$ has an individual state $s_i \in S_i$ representing his competitive advantage, where S_i denotes his state space. Also, each player $i \in \mathcal{N}$ has an individual Markovian transition probability matrix P_i, where $P_i(s_i'|s_i)$ shows the probability that player i will be in state $s_i' \in S_i$ at the next period given that he is now in state $s_i \in S_i$. As a result, the game state $s \in \mathscr{S}$ is $(s_1, s_2, \ldots, s_N)$, and the game state space $\mathscr{S}$ is the Cartesian product of $S_1, S_2, \ldots, S_N$ so that $|\mathscr{S}| = \prod_{i \in \mathcal{N}} |S_i|$. Moreover, the game transition probability matrix $\mathscr{P}$ is the Kronecker product of transition probability matrices of all players, i.e, $\mathscr{P}(s_1', \ldots, s_N'|s_1, \ldots, s_N) = \prod_{i \in \mathcal{N}} P_i(s_i'|s_i)$. Clearly, the equilibrium selection MILP grows rapidly as either the number of players or the size of S_i increases. For each instance, there is an initial state, denoted by $\hat{s}$, and the objective function (11a) is set to $\sum_{i \in \mathcal{N}} w_i(\hat{s})$. This is a reasonable objective since it is the sum of all players' expected reward-to-go from the initial state. Moreover, the discount factors are equal for all players, and hence we use both valid inequalities (16) and (18) in the branch-and-cut algorithm. In computation of the valid inequality (16), we let $\alpha \in \mathbb{R}^{|\mathcal{N}|}$ be a vector whose components are all equal to 1.

In order to enhance the performance, we provide all approaches with strategy profiles $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ as warm start solutions such that for each $s \in \mathscr{S}$,

$$\bar{x}_1(s) = \begin{cases} 1 & \text{if} \quad s = \hat{s}, \\ 0 & \text{otherwise,} \end{cases} \qquad \bar{x}_2(s) = \begin{cases} 1 & \text{if} \quad V_i(s) = u_i(s, 1)\, \forall i \in \mathcal{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 10 *Strategy profile* $\bar{\mathbf{x}}_2$ *is an equilibrium.*

Note that in the warm start solutions presented above, unlike $\mathbf{x}_2$, $\mathbf{x}_1$ does not need to be necessarily an equilibrium, and in such a case it is automatically eliminated from consideration by the MILP solver.

8.2 Computational results 1: real clinical instances

We used the method described in [28], to generate three categories of consensus stopping game instances for his kidney exchange problem. In the first (second) category, there are two players such that the size of S_i is equal to 40 (60) for both players. In the third category, there are three players such that the size of S_i is equal to 15 for all players.

Tables 1, 2 and 3 include computational results for the first, second, and third category of instances, respectively, and for ease of exposition, we let Branch-and-Cut refer to the branch-and-cut algorithm described in Sect. 7, CPLEX refer to solving the extensive formulation (11a)–(11b) through ILOG-CPLEX 12.6. In Tables 1, 2 and 3, we report the number of cuts and number of explored nodes in the respective branch-and-cut tree, the best solution, optimality gap (in %), setup time, and running time (both in the hour: minute: second time format) for both approaches. Note that $t.lim.$ denotes that the approach has reached its 4-h time limit. Moreover, setup time is the time required to compute necessary parameters and construct associated MILPs and LPs. More specifically, setup time for Branch-and-Cut is the time needed to compute parameters $\langle \mathbf{V}_i \rangle_{\forall i \in \mathcal{N}}$ and $\mathbf{d}$, plus the time needed to construct initial LPs which will be used during the algorithm to generate valid inequalities (16) and (18). The setup time for CPLEX is the time needed to compute parameters $\langle \mathbf{V}_i \rangle_{\forall i \in \mathcal{N}}$ and $\mathbf{d}$, plus the time needed to construct the MILP formulation (11a)–(11b). For convenience of exposition, we arranged the instances in the tables with respect to the optimality gap of Branch-and-Cut and CPLEX.

In Table 1, both Branch-and-Cut and CPLEX need almost the same setup time for each instance. Both approaches can solve instances $a1$–$a12$. On the majority of these instances, Branch-and-Cut is at least an order of magnitude faster than CPLEX. Branch-and-Cut needs about 16 minutes total to solve all instances of this subset while CPLEX needs almost 9.5 h total, i.e., Branch-and-Cut is 34 times faster in processing the whole subset. Branch-and-Cut can solve each instance of $a13$–$a18$ in less than 17 minutes while CPLEX cannot solve any of them. Neither Branch-and-Cut nor CPLEX is able to solve instances $a19$–$a30$. However, Branch-and-Cut provides smaller optimality gaps for all instances. Overall, Branch-and-Cut outperforms CPLEX on all instances. In some instances, we observe that the number of explored nodes and the optimality gap are 0, meaning that those instances are solved at the root node. In such instances, $\bar{\mathbf{x}}_1$, i.e., the provided warm start solution, is indeed the optimal solution; however, establishing optimality can be very challenging, e.g., instances $a13$–$a18$. Moreover, we observe that Branch-and-Cut solves some instances without adding any cut at the root node, meaning that the inequalities (12b) are enough to close the optimality gap. CPLEX also solves some instances without adding any cut at the root node by preprocessing and probing techniques. For instances $a20$–$a30$, no warm-start

Table 1 Two player with 40 state-per-player instances

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$a1$	18	0	11,954.3	0.00	0:00:45	0:01:43	0	0	11,954.3	0.00	0:00:38	0:11:35
$a2$	0	0	11,684.2	0.00	0:00:46	0:00:00	1526	0	11,684.2	0.00	0:00:39	0:22:58
$a3$	25	0	11,472.6	0.00	0:00:46	0:01:45	0	0	11,472.6	0.00	0:00:39	0:08:48
$a4$	0	0	13,057.1	0.00	0:00:45	0:00:00	0	0	13,057.1	0.00	0:00:38	0:02:45
$a5$	0	0	11,230.3	0.00	0:00:47	0:00:00	0	0	11,230.3	0.00	0:00:39	0:20:14
$a6$	0	0	11,765	0.00	0:00:48	0:00:00	0	0	11,765	0.00	0:00:39	0:02:59
$a7$	0	0	11,840.4	0.00	0:00:44	0:00:00	0	0	11,840.4	0.00	0:00:37	0:02:02
$a8$	18	0	12,125.8	0.00	0:00:45	0:01:43	0	0	12,125.8	0.00	0:00:39	1:29:17
$a9$	22	0	13,309.1	0.00	0:00:44	0:02:02	0	0	13,309.1	0.00	0:00:38	0:44:21
$a10$	124	0	11,648.1	0.00	0:00:45	0:04:32	0	0	11,648.1	0.00	0:00:38	0:21:40
$a11$	20	0	11,045.8	0.00	0:00:46	0:01:40	2123	31	11,045.8	0.00	0:00:42	1:58:44
$a12$	63	0	10,429.5	0.00	0:00:52	0:02:51	1276	67	10,429.5	0.00	0:00:43	3:40:14
$a13$	20	0	11,682.5	0.00	0:00:46	0:01:41	2138	182	11,682.5	0.50	0:00:44	$t.lim.$
$a14$	18	0	11,833.3	0.00	0:00:44	0:01:44	1891	68	11,833.3	1.35	0:00:38	$t.lim.$
$a15$	18	0	11,911.5	0.00	0:00:45	0:01:41	1858	122	11,911.5	2.75	0:00:39	$t.lim.$
$a16$	125	0	11,540.9	0.00	0:00:46	0:04:26	1546	53	11,540.9	2.72	0:00:39	$t.lim.$
$a17$	18	0	11,395.2	0.00	0:00:47	0:01:38	926	109	11,395.2	3.35	0:00:40	$t.lim.$
$a18$	172	0	11,447	0.00	0:00:46	0:05:42	1833	55	11,447	3.45	0:00:39	$t.lim.$
$a19$	1354	291	12,096.9	0.38	0:00:45	$t.lim.$	1895	22	12,096.9	5.22	0:00:39	$t.lim.$
$a20$	862	194	13,190	2.31	0:00:45	$t.lim.$	1447	4	13,124.2	6.12	0:00:38	$t.lim.$

Table 1 continued

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$a21$	651	154	12,312.1	2.43	0:00:46	$t.lim.$	1475	195	12,276.2	3.61	0:00:39	$t.lim.$
$a22$	699	162	12,272.6	2.81	0:00:47	$t.lim.$	1589	300	12,226	4.24	0:00:40	$t.lim.$
$a23$	471	99	11,650.6	3.78	0:00:46	$t.lim.$	1574	291	11,619	4.80	0:00:39	$t.lim.$
$a24$	849	189	11,938.4	3.35	0:00:46	$t.lim.$	1686	650	12,012	3.43	0:00:39	$t.lim.$
$a25$	696	156	10,350.9	3.14	0:00:55	$t.lim.$	1104	398	10,264.7	6.83	0:00:43	$t.lim.$
$a26$	647	149	11,957.8	3.50	0:00:46	$t.lim.$	1901	252	11,937.8	5.85	0:00:39	$t.lim.$
$a27$	824	192	10,954.1	3.56	0:00:48	$t.lim.$	1217	298	10,940.4	5.23	0:00:40	$t.lim.$
$a28$	611	130	11,544.9	4.25	0:00:47	$t.lim.$	1508	94	11,521.1	7.75	0:00:39	$t.lim.$
$a29$	763	165	9063.89	4.51	0:00:58	$t.lim.$	807	503	9046.38	5.52	0:00:46	$t.lim.$
$a30$	619	137	11,577.9	5.93	0:00:46	$t.lim.$	1202	44	11,558.7	8.62	0:00:38	$t.lim.$

Table 2 Two player with 60 state-per-player instances

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$b1$	0	0	11,736.6	0.00	0:12:38	0:00:00	0	0	11,736.6	0.00	0:11:28	0:21:08
$b2$	0	0	11,821	0.00	0:12:38	0:00:00	0	0	11,821	0.00	0:11:23	0:20:37
$b3$	0	0	11,699.1	0.00	0:12:50	0:00:01	0	0	11,699.1	0.00	0:11:27	0:23:47
$b4$	0	0	11,189.1	0.00	0:13:38	0:00:01	0	0	11,189.1	0.00	0:12:55	0:15:11
$b5$	0	0	13,358.9	0.00	0:13:17	0:00:01	0	0	13,358.9	0.00	0:11:58	0:15:22
$b6$	0	0	11,540.9	0.00	0:12:50	0:00:00	0	0	11,540.9	0.00	0:11:37	0:14:46
$b7$	0	0	11,699.1	0.00	0:12:39	0:00:00	0	0	11,699.1	0.00	0:11:22	0:17:35
$b8$	37	0	10,823	0.00	0:13:48	0:29:19	0	0	10,823	0.00	0:12:23	0:59:01
$b9$	27	0	11,163.8	0.00	0:14:04	0:28:48	69	0	11,163.8	0.00	0:12:22	0:54:26
$b10$	282	0	10,593	0.00	0:13:03	1:38:10	0	0	10,593	0.31	0:11:42	$t.lim.$
$b11$	0	0	12,049.7	0.00	0:12:49	0:00:01	0	0	12,049.7	0.55	0:11:20	$t.lim.$
$b12$	18	0	11,627.7	0.00	0:12:44	0:24:03	0	0	11,627.7	0.55	0:11:27	$t.lim.$
$b13$	100	0	12,071.5	0.00	0:12:39	0:54:19	0	0	12,071.5	3.69	0:11:30	$t.lim.$
$b14$	106	0	11,304.7	0.00	0:13:19	0:57:25	0	0	11,304.7	2.62	0:12:06	$t.lim.$
$b15$	25	0	11,117.6	0.00	0:14:11	0:26:07	0	0	11,117.6	3.75	0:12:05	$t.lim.$
$b16$	151	0	11,964.8	0.00	0:13:27	1:12:15	0	0	11,964.8	4.36	0:12:11	$t.lim.$
$b17$	236	0	11,466.9	0.00	0:13:37	1:43:29	0	0	11,466.9	5.37	0:12:19	$t.lim.$
$b18$	489	0	11,014.6	0.00	0:12:58	3:17:02	0	0	11,014.6	5.89	0:11:46	$t.lim.$
$b19$	425	0	10,352.7	0.00	0:14:00	2:52:16	0	0	10,352.7	6.45	0:12:42	$t.lim.$
$b20$	190	28	12,287.9	0.19	0:13:15	$t.lim.$	0	0	12,287.9	2.72	0:12:27	$t.lim.$

Table 2 continued

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
*b*21	81	15	11,831.1	1.26	0:12:36	*t.lim.*	0	0	11,831.1	7.31	0:11:23	*t.lim.*
*b*22	95	17	11,623.6	2.44	0:14:00	*t.lim.*	0	0	11,623.6	9.94	0:12:14	*t.lim.*
*b*23	191	38	11,609.8	4.04	0:13:32	*t.lim.*	3271	0	11,598.9	5.44	0:12:31	*t.lim.*
*b*24	76	14	10,959.8	3.49	0:12:59	*t.lim.*	0	0	10,959.8	11.28	0:11:55	*t.lim.*
*b*25	66	9	9621.96	3.56	0:13:22	*t.lim.*	3942	0	9606.76	4.59	0:11:33	*t.lim.*
*b*26	42	3	9712.26	3.57	0:12:54	*t.lim.*	4117	0	9709.84	4.65	0:11:34	*t.lim.*
*b*27	64	10	9996.52	4.41	0:13:02	*t.lim.*	0	0	9991.23	5.79	0:11:39	*t.lim.*
*b*28	94	15	10,103.9	6.51	0:13:41	*t.lim.*	0	0	10,089.9	8.90	0:12:20	*t.lim.*
*b*29	51	5	7982.41	6.10	0:14:28	*t.lim.*	0	0	7961.32	8.88	0:12:50	*t.lim.*
*b*30	42	4	9873.91	10.11	0:12:58	*t.lim.*	3692	0	9852.2	12.35	0:11:45	*t.lim.*

Table 3 Three player with 15 state-per-player instances

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$c1$	0	0	16,647	0.00	0:14:02	0:00:01	0	0	16,647	0.00	0:12:08	0:54:39
$c2$	0	0	18,230.9	0.00	0:11:38	0:00:01	0	0	18,230.9	0.00	0:11:12	0:08:47
$c3$	38	0	17,862.1	0.00	0:11:06	0:13:14	0	0	17,862.1	0.00	0:10:27	0:28:17
$c4$	444	0	15,437	0.00	0:14:16	2:05:55	0	0	15,437	0.00	0:11:39	0:33:11
$c5$	26	0	17,088.9	0.00	0:13:25	0:10:45	0	0	17,088.9	1.84	0:11:41	$t.lim.$
$c6$	38	0	16,345.6	0.00	0:13:58	0:12:22	0	0	16,345.6	6.42	0:11:58	$t.lim.$
$c7$	26	0	17,582	0.00	0:11:34	0:11:10	0	0	17,582	1.08	0:11:10	$t.lim.$
$c8$	597	0	16,176	0.00	0:13:59	2:14:16	0	0	16,176	2.02	0:12:13	$t.lim.$
$c9$	29	0	16,919.5	0.00	0:11:56	0:11:27	0	0	16,919.5	1.23	0:11:18	$t.lim.$
$c10$	26	0	17,886.8	0.00	0:11:25	0:11:03	0	0	17,886.8	1.72	0:10:45	$t.lim.$
$c11$	26	0	17,638.8	0.00	0:11:24	0:11:28	0	0	17,638.8	1.77	0:11:10	$t.lim.$
$c12$	29	0	17,802.4	0.00	0:11:17	0:12:03	0	0	17,802.4	1.73	0:11:02	$t.lim.$
$c13$	738	0	16,755.1	0.00	0:12:29	3:44:09	0	0	16,755.1	1.99	0:11:51	$t.lim.$
$c14$	26	0	17,251.9	0.00	0:12:10	0:11:00	0	0	17,251.9	2.46	0:11:44	$t.lim.$
$c15$	566	0	17,528.5	0.00	0:12:03	2:52:21	0	0	17,528.5	3.25	0:11:44	$t.lim.$
$c16$	330	0	17,362.6	0.00	0:11:37	1:42:47	0	0	17,362.6	1.67	0:11:20	$t.lim.$
$c17$	320	43	17,817.3	0.47	0:12:27	$t.lim.$	0	0	17,817.3	6.42	0:11:07	$t.lim.$
$c18$	319	50	18,727.3	0.64	0:12:45	$t.lim.$	0	0	18,692.3	0.94	0:11:36	$t.lim.$
$c19$	336	50	15,716.6	0.87	0:13:28	$t.lim.$	2064	0	15,677	1.04	0:12:14	$t.lim.$
$c20$	345	50	16,954.5	1.18	0:13:22	$t.lim.$	0	0	16,900.7	2.16	0:12:02	$t.lim.$

Table 3 continued

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$c21$	477	50	11,975.8	1.71	0:13:16	$t.lim.$	240	0	11,508.1	5.95	0:11:17	$t.lim.$
$c22$	283	32	14,279.9	2.25	0:12:46	$t.lim.$	0	0	14,094.7	3.79	0:11:46	$t.lim.$
$c23$	270	50	15,908.4	3.46	0:14:12	$t.lim.$	0	0	15,871.2	5.57	0:12:13	$t.lim.$
$c24$	322	50	13,821.9	3.01	0:13:29	$t.lim.$	0	0	13,776.8	3.58	0:11:56	$t.lim.$
$c25$	385	60	12,365.2	4.55	0:11:54	$t.lim.$	278	21	12,295.4	3.98	0:11:03	$t.lim.$
$c26$	434	60	15,251.7	4.62	0:12:06	$t.lim.$	0	0	15,100.6	6.46	0:11:06	$t.lim.$
$c27$	329	44	15,250.1	5.26	0:11:34	$t.lim.$	4219	0	14,793.1	9.20	0:11:03	$t.lim.$
$c28$	415	60	17,590	4.38	0:11:38	$t.lim.$	0	0	17,174	10.33	0:10:59	$t.lim.$
$c29$	293	41	12,955.9	10.80	0:12:07	$t.lim.$	369	0	12,416.1	13.84	0:10:53	$t.lim.$
$c30$	349	54	12,828.7	16.51	0:12:34	$t.lim.$	0	0	12,740.1	18.91	0:11:42	$t.lim.$

solution is optimal. For these instances, finding a better solution than the warm-start solution $\bar{\mathbf{x}}_2$ is quite challenging, and CPLEX is unable to find a better solution while Branch-and-Cut is able to find better solutions and provides smaller optimality gaps. Moreover, we let CPLEX explore these instances in a 24-h time limit, and observed that CPLEX was only able to find a better solution than $\bar{\mathbf{x}}_2$ just for one instance, and its performance in terms of optimality gap was dominated by that of Branch-and-Cut for a 4-h time limit.

Table 2 shows a similar pattern. Both Branch-and-Cut and CPLEX can solve instances $b1$–$b9$, and Branch-and-Cut is several orders of magnitude faster than CPLEX. Branch-and-Cut can solve instances $b10$–$b19$ while CPLEX cannot solve any of them. Instances $b20$–$b30$ cannot be solved by Branch-and-Cut nor by CPLEX. However, Branch-and-Cut provides us with better solutions and smaller optimality gaps. For instances of $b1$–$b19$, $\bar{\mathbf{x}}_1$ is optimal, and establishing optimality is the primary challenge in which Branch-and-Cut does remarkably well. For instances $b20$–$b30$, both finding a better solution than $\bar{\mathbf{x}}_2$ and establishing optimality are extremely challenging. For all of these instances, CPLEX is unable to find a better solution than $\bar{\mathbf{x}}_2$, even within 1 day. Moreover, its optimality gap within 1 day time limit is still larger than that of Branch-and-Cut within a 4-h time limit. Generally speaking, the advantage of Branch-and-Cut over CPLEX is even more apparent with the larger instances. Table 3 shows that Branch-and-Cut outperforms CPLEX on the third category of instances as well.

It is worth to underline that our Branch-and-Cut algorithm operates on a series of callback routines some of which appear to be computationally expensive as they require the solution of large-scale LPs. Particularly, excessive solution times of such LPs may keep the algorithm running over the time limit which may appear to show the improper functioning of the time-limit parameter (CPX_PARAM_TILIM). To circumvent this issue, the algorithm was terminated via the node-limit parameter (CPX_PARAM_NODELIM) instead of the time-limit parameter for instances $b20$–$b30$ and $c17$–$c30$ for which the issue appeared. For these instances, the appropriate value of the associated node-limit parameter to terminate the algorithm before hitting the 4-h time limit was identified on a trial and error basis.

In order to evaluate efficiency of the five families of valid inequalities applied in Branch-and-Cut, we deactivated each type of valid inequality in Branch-and-Cut one at a time and recollected the numerical results. In Table 4, we report the number of solved problems and average of the optimality gaps in rows 1–2, 3–4, and 5–6 for instances of Tables 1, 2 and 3, respectively. In column 1, we report the statistics for Branch-and-Cut. In columns 2–6, we report the statistics for Branch-and-Cut after deactivation of the associated valid inequalities. This table shows that the valid inequalities (16) and (18) have a significant effect in closing the optimality gap.

8.3 Computational results 2: more general instances

All instances studied in the preceding subsection had non-negative rewards. In this subsection, we investigate a set of random instances including both negative and non-negative rewards to test the computational performance of our approach under this

Table 4 Performance of Branch-and-cut when each type of valid inequality is deactivated

		Original Branch-and-cut	Ineq. (8)	Ineq. (9)	Ineq. (13)	Ineq. (16)	Ineq. (18)
Instances of Table 1	Number of solved problems	18	18	18	18	5	18
	Average of gap (%)	1.33	1.28	1.37	1.38	3.00	1.36
Instances of Table 2	Number of solved problems	19	19	19	19	10	17
	Average of gap (%)	1.52	1.41	1.54	1.54	2.89	1.78
Instances of Table 3	Number of solved problems	16	16	16	13	2	14
	Average of gap (%)	1.99	1.76	2.36	2.73	2.91	2.02

Table 5 Negative reward instances with two players and 40 states per player

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$d1$	0	0	−43,791.2	0.00	0:00:36	0:00:00	0	0	−43,791.2	UNB	0:00:23	0:00:05
$d2$	12,841	394	−42,298.1	0.19	0:00:37	$t.lim.$	0	0	−42,315.4	0.00	0:00:24	0:15:06
$d3$	229	0	−50,115	0.00	0:00:39	0:03:14	0	0	INF	N/A	0:00:29	0:00:03
$d4$	4654	309	−43,864.4	0.18	0:00:38	$t.lim.$	0	0	−44,002.2	UNB	0:00:27	0:00:03
$d5$	2030	175	−61,003.2	0.01	0:00:41	2:14:05	0	0	INF	N/A	0:00:30	0:00:07
$d6$	38	0	−41,204.1	0.00	0:00:49	0:00:42	0	0	INF	N/A	0:00:37	0:00:03
$d7$	6205	239	−40,912.4	0.14	0:00:40	$t.lim.$	0	0	INF	N/A	0:00:28	0:00:03
$d8$	10,377	428	−38,364.6	0.69	0:00:38	$t.lim.$	288	0	−38,414.4	0.61	0:00:26	$t.lim.$
$d9$	44	0	−46,703.5	0.00	0:00:39	0:02:22	0	0	−46,703.7	UNB	0:00:31	0:00:03
$d10$	140	0	−47,275.9	0.01	0:00:40	0:01:57	0	0	−47,277.6	0.00	0:00:28	0:00:28
$d11$	374	12	−69,717	0.01	0:00:45	0:06:44	0	0	−69,850.1	UNB	0:00:30	0:00:05
$d12$	86	0	−36,144.1	0.00	0:00:40	0:02:26	0	0	INF	N/A	0:00:31	0:00:03
$d13$	1582	64	−56,391.4	0.00	0:00:38	0:32:28	0	0	−56,431.1	UNB	0:00:27	0:00:03
$d14$	5851	466	−38,646.8	0.12	0:00:42	$t.lim.$	0	0	−38,672.8	UNB	0:00:29	0:00:03
$d15$	7853	396	−55,747.4	0.07	0:00:41	$t.lim.$	0	0	INF	N/A	0:00:27	0:00:04
$d16$	10,132	437	−33,180.6	0.88	0:00:39	$t.lim.$	180	0	−33,621.2	1.96	0:00:28	$t.lim.$
$d17$	1341	74	−60,506.6	0.01	0:00:46	1:05:46	0	0	−62,539.4	UNB	0:00:29	0:00:03
$d18$	51	0	−55,398.4	0.00	0:00:44	0:00:55	0	0	−55,398.6	0.00	0:00:29	0:00:06
$d19$	454	45	−49,544.6	0.00	0:00:39	0:22:04	0	0	INF	N/A	0:00:25	0:00:03

Table 5 continued

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$d20$	3569	263	− 50,500	0.01	0:00:35	1:57:46	0	0	− 50,662.4	UNB	0:00:24	0:00:03
$d21$	0	0	− 54,442.8	0.00	0:00:42	0:00:00	0	0	− 54,442.8	UNB	0:00:28	0:00:07
$d22$	94	0	− 26,155	0.00	0:00:43	0:02:40	0	0	INF	N/A	0:00:32	0:00:04
$d23$	103	0	− 42,945.8	0.00	0:00:51	0:01:44	0	0	− 43,032.7	UNB	0:00:35	0:00:03
$d24$	31	0	− 52,390.7	0.00	0:00:40	0:00:46	0	0	INF	N/A	0:00:30	0:00:03
$d25$	98	0	− 40,318	0.01	0:00:43	0:02:16	0	0	− 40,327.2	UNB	0:00:31	0:00:03
$d26$	10,398	408	− 74,470.6	0.50	0:00:42	$t.lim.$	0	0	− 74,506.1	0.00	0:00:28	0:12:39
$d27$	130	0	− 50,124.8	0.00	0:00:39	0:03:09	0	0	− 54,372.4	UNB	0:00:28	0:00:03
$d28$	0	0	− 47,153	0.00	0:00:38	0:00:00	0	0	INF	N/A	0:00:23	0:00:04
$d29$	3652	146	− 71,747.4	0.01	0:00:39	2:06:29	0	0	INF	N/A	0:00:26	0:00:06
$d30$	108	0	− 32,777.3	0.00	0:00:50	0:03:35	0	0	INF	N/A	0:00:34	0:00:03

INF infeasible, *UNB* unbounded, N/A not applicable

Table 6 Negative reward instances with two players and 60 states per player

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
$e1$	0	0	− 9859.9	0.00	0:08:10	0:00:01	0	0	− 9859.9	UNB	0:05:48	0:00:09
$e2$	363	0	− 40,168.4	2.13	0:06:37	$t.lim.$	0	0	INF	N/A	0:04:18	0:00:04
$e3$	959	43	− 15,920.3	4.92	0:04:26	$t.lim.$	0	0	INF	N/A	0:03:01	0:00:10
$e4$	0	0	5491.23	0.00	0:05:26	0:00:00	0	0	5491.23	0.00	0:04:10	0:00:29
$e5$	1902	113	− 14,832.3	2.25	0:02:36	$t.lim.$	0	0	INF	N/A	0:01:43	0:00:03
$e6$	494	17	− 36,524.4	7.55	0:06:57	$t.lim.$	0	0	INF	N/A	0:05:25	$t.lim.$
$e7$	0	0	− 16,584.5	0.00	0:06:05	0:00:00	3839	0	INF	N/A	0:04:40	$t.lim.$
$e8$	665	30	− 21,916.7	1.68	0:04:38	$t.lim.$	0	0	INF	N/A	0:03:23	0:00:04
$e9$	520	12	− 23,243.7	0.20	0:05:48	$t.lim.$	0	0	INF	N/A	0:03:57	0:00:49
$e10$	0	0	− 16,321.9	0.00	0:06:01	0:00:00	0	0	INF	N/A	0:04:07	0:00:10
$e11$	962	36	− 28,669.2	4.74	0:04:31	$t.lim.$	0	0	INF	N/A	0:02:55	0:00:10
$e12$	44	0	− 8071.9	0.00	0:06:08	0:12:57	0	0	INF	N/A	0:04:58	$t.lim.$
$e13$	1335	97	− 23,466.3	2.57	0:02:49	$t.lim.$	0	0	INF	N/A	0:01:56	0:00:03
$e14$	95	0	− 8939.3	0.00	0:05:50	0:30:57	0	0	INF	N/A	0:04:50	0:00:04
$e15$	536	39	− 9643.13	19.25	0:05:59	$t.lim.$	0	0	INF	N/A	0:04:42	$t.lim.$
$e16$	110	0	− 18,861.3	0.00	0:06:21	0:45:50	0	0	− 27,315.1	39.21	0:05:03	$t.lim.$
$e17$	0	0	− 911.841	0.00	0:09:26	0:00:00	0	0	INF	N/A	0:03:05	0:00:11
$e18$	0	0	5865.58	0.00	0:05:05	0:00:00	879	0	INF	N/A	0:03:47	$t.lim.$
$e19$	0	0	− 4283.5	0.00	0:04:10	0:00:00	0	0	− 4283.5	0.00	0:03:14	0:00:13
$e20$	462	15	− 25,861.9	5.56	0:07:39	$t.lim.$	0	0	− 27,120.3	10.80	0:05:58	$t.lim.$

Table 6 continued

Name	Branch-and-cut						CPLEX					
	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time	# of cuts	# of nodes	Best solution	Gap (%)	Setup time	Solution time
*e*21	0	0	− 10,039.5	0.00	0:04:19	0:00:00	0	0	INF	N/A	0:03:17	0:00:04
*e*22	105	0	1717.15	0.00	0:04:46	0:37:55	0	0	INF	N/A	0:03:47	*t.lim.*
*e*23	0	0	208.367	0.00	0:06:31	0:00:01	0	0	INF	N/A	0:04:14	0:00:10
*e*24	100	0	− 9698.95	0.00	0:04:41	0:36:39	0	0	− 17,128.8	69.72	0:03:28	*t.lim.*
*e*25	887	36	− 15,209.1	11.06	0:04:10	*t.lim.*	0	0	− 19,753.8	0.00	0:02:56	1:18:27
*e*26	0	0	6784.35	0.00	0:02:32	0:00:00	0	0	INF	N/A	0:01:47	0:00:03
*e*27	0	0	8851.23	0.00	0:06:21	0:00:00	0	0	INF	N/A	0:05:07	*t.lim.*
*e*28	399	12	− 14,376.5	17.56	0:07:29	*t.lim.*	0	0	− 14,376.5	0.00	0:06:04	2:07:20
*e*29	496	11	− 34,220.7	0.71	0:06:24	*t.lim.*	0	0	INF	N/A	0:04:41	0:00:11
*e*30	562	32	− 39,354.9	7.64	0:06:29	*t.lim.*	0	0	INF	N/A	0:04:48	*t.lim.*

INF infeasible, *UNB* unbounded, *N/A* not applicable setting. In our synthetic test bed, each instance consists of four components: Players' individual transition matrices, rewards, discount factors for each player, and the initial state of the game. All these components, but discount factors, are randomly generated for two categories of instances as described in the "Appendix".

We report our computational results for the first and second categories of instances in Tables 5 and 6, respectively. These tables illustrate that Branch-and-Cut greatly outperforms CPLEX. Note that CPLEX incorrectly finds many instances infeasible or unbounded. It also misses the optimal solution for instances $d2$, $d26$, and $e25$ in Table 5. We attempted to circumvent these numerical failures by setting the feasibility tolerance parameter (CPX_PARAM_EPRHS) to its highest allowed value (0.1), but this did not solve these numerical failures. There are a couple of reasons behind these failures: (1) Large variability in the coefficients of the formulation Δ due to the existence of the transition probability matrix as well as the big-M type coefficients $\langle V_i(s)\rangle_{s\in\mathscr{S},i\in\mathcal{N}}$ and $\langle d_i(s)\rangle_{s\in\mathscr{S},i\in\mathcal{N}}$ raises the possibility of numerical instability. (2) For each $\mathbf{x}\in\mathbb{B}^{|\mathscr{S}|}$, if $x(s)=0$ ($x(s)=1$), then inequalities (10a)–(10b) [(10c)–(10d)] must hold as equalities, which are more numerically unstable. In summary, numerical results of CPLEX are unreliable for instances of Tables 5 and 6.

9 Conclusions

We consider consensus stopping games, a broad class of stochastic stopping games. For the NP-hard problem of finding a best equilibrium for this important class of games, we propose an MILP formulation and develop several valid inequalities to tackle it efficiently. We demonstrate the effect and strength of these inequalities by invoking them in a tailored branch-and-cut framework which also utilizes other features such as dynamic variable fixing and coefficient tightening along the solution process.

The majority of results in this paper can be applied to nonlinear objective functions of payoffs. In particular, the valid inequalities (8) and (9) may be applied to any nondecreasing and general nonlinear objective function of payoffs, respectively. The approach of this paper might also be amenable to analyze equilibria of other types of stochastic games. We leave this extension as a topic for future research.

Appendix

Proof of Proposition 1 (iii) ($\Rightarrow$) Suppose that there exists an equilibrium in $A_{\mathbf{x}}$, denoted by $\tilde{\mathbf{a}}$. We show that strategy profile $\bar{\mathbf{a}}$ is an equilibrium as well. Since $\bar{\mathbf{a}}, \tilde{\mathbf{a}}\in A_{\mathbf{x}}$, we must have $\mathbf{w}^{\tilde{\mathbf{a}}}=\mathbf{w}^{\bar{\mathbf{a}}}$ by part (i). There are two cases:

If $x(s) = 1$, then for all $i \in \mathcal{N}$, $\tilde{a}_i(s) = \bar{a}_i(s) = 1$, and

$$
\begin{aligned}
w_i^{\bar{\mathbf{a}}}(s) = w_i^{\tilde{\mathbf{a}}}(s) &= \max\left\{\left(\prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \tilde{a}_j(s)\right) u_i(s,1)\right.\\
&\qquad\left. + \left(1 - \prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \tilde{a}_j(s)\right) F_i(s, \mathbf{w}_i^{\tilde{\mathbf{a}}}),\ F_i(s, \mathbf{w}_i^{\tilde{\mathbf{a}}})\right\}\\
&= \max\left\{\left(\prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \bar{a}_j(s)\right) u_i(s,1)\right.\\
&\qquad\left. + \left(1 - \prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \bar{a}_j(s)\right) F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}),\ F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}})\right\},
\end{aligned}
$$

where the second equality follows from the fact that $\tilde{\mathbf{a}}$ is an equilibrium and part (ii). If $x(s) = 0$, then for all $i \in \mathcal{N}$, $\bar{a}_i(s) = 0$, and

$$
\begin{aligned}
w_i^{\bar{\mathbf{a}}}(s) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}) &= \max\left\{\left(\prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \bar{a}_j(s)\right) u_i(s,1)\right.\\
&\qquad\left. + \left(1 - \prod_{\substack{j\in\mathcal{N},\\ j\neq i}} \bar{a}_j(s)\right) F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}}),\ F_i(s, \mathbf{w}_i^{\bar{\mathbf{a}}})\right\},
\end{aligned}
$$

where the first equality follows from part (i).

In summary, the Eq. (3) is satisfied for all $s \in \mathcal{S}, i \in \mathcal{N}$. Therefore, $\bar{\mathbf{a}}$ is an equilibrium by part (ii).

($\Leftarrow$) Follows directly from the definitions of $\bar{\mathbf{a}}$ and $A_{\mathbf{x}}$. □

Proof of Proposition 3 Recall that under a fixed strategy profile $\mathbf{x}$, the payoffs of each player represent a stationary Markov reward process, and hence can be calculated by value iteration [10]. Let $[w_i^{\mathbf{x}}(s)]^n$ denote the value associated with state $s \in \mathcal{S}$ at the nth iteration of the value iteration algorithm under strategy profile $\mathbf{x}$ for player i. Furthermore, we initialize our value iteration with payoffs of player i under $\bar{\mathbf{x}}$, i.e., $[w_i^{\mathbf{x}}(s)]^0 = w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathcal{S}$.

(i) From Proposition 2 (i) and the hypothesis about the relation between $\mathbf{x}$ and $\bar{\mathbf{x}}$, there are four mutually exclusive cases for s:

- If $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$, then $x(s) = 1$ and $[w_i^{\mathbf{x}}(s)]^1 = u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.
- If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$, then $x(s) = 0$ and $[w_i^{\mathbf{x}}(s)]^1 = F_i(s, [\mathbf{w}_i^{\mathbf{x}}]^0) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.
- If $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ and $x(s) = 1$, then $[w_i^{\mathbf{x}}(s)]^1 = u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$.
- If $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ and $x(s) = 0$, then $[w_i^{\mathbf{x}}(s)]^1 = F_i(s, [\mathbf{w}_i^{\mathbf{x}}]^0) = F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq u_i(s,1) = w_i^{\bar{\mathbf{x}}}(s) = [w_i^{\mathbf{x}}(s)]^0$ where the inequality follows from $s \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$.

From all four cases, it follows that $[w_i^{\mathbf{x}}(s)]^1 \geq [w_i^{\mathbf{x}}(s)]^0$ for all $s \in \mathscr{S}$. From the monotonicity of the dynamic programming operator induced by strategy profile $\mathbf{x}$ for player i [6], it follows that for any n, $[w_i^{\mathbf{x}}(s)]^{n+1} \geq [w_i^{\mathbf{x}}(s)]^n$ for all $s \in \mathscr{S}$. As a result, $w_i^{\mathbf{x}}(s) = \lim_{n\to\infty} [w_i^{\mathbf{x}}(s)]^n \geq [w_i^{\mathbf{x}}(s)]^0 = w_i^{\bar{\mathbf{x}}}(s)$.

(ii) The proof is similar to that of part (i). □

Remark 1 A nice property of Proposition 2 (ii) is that when we assess equilibrium conditions for a strategy profile $\mathbf{x}$, we only need to check if the Bellman–Shapley equation (5) is satisfied for each $s \in \mathscr{S}$, $i \in \mathscr{N}$ in which $x(s) = 1$ since (5) is trivially satisfied for each $s \in \mathscr{S}$, $i \in \mathscr{N}$ in which $x(s) = 0$. In particular, this implies that the strategy profile $\mathbf{0}$ is an equilibrium.

Proof of Lemma 1 ($\Rightarrow$) If $s \in \mathscr{S}^1(\mathbf{x})$, then $w_i^{\mathbf{x}}(s) = u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ by Proposition 2. Therefore, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}^1(\mathbf{x})$, $i \in \mathscr{N}$. It follows that $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$.

($\Leftarrow$) $\mathscr{S}^1(\mathbf{x}) = \mathscr{S}^1_{i,nv}(\mathbf{x}) \cup \mathscr{S}^1_{i,v}(\mathbf{x})$ and $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$. It follows that $\mathscr{S}^1(\mathbf{x}) = \mathscr{S}^1_{i,nv}(\mathbf{x})$ for all $i \in \mathscr{N}$. Therefore, if $s \in \mathscr{S}^1(\mathbf{x})$, then for all $i \in \mathscr{N}$, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ and Proposition 2 (i) implies that $w_i^{\mathbf{x}}(s) = u_i(s,1)$. Hence, the Bellman–Shapley equation (5) is satisfied for all $s \in \mathscr{S}^1(\mathbf{x})$, $i \in \mathscr{N}$. As noted in Remark 1, the Bellman–Shapley equation (5) is trivially satisfied for all $s \in \mathscr{S}^0(\mathbf{x})$, $i \in \mathscr{N}$. Consequently, $\mathbf{x}$ is an equilibrium by Proposition 2 (ii). □

Proof of Proposition 4 (i) As $\bar{\mathbf{x}}$ is an equilibrium, $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathscr{N}$ by Lemma 1, and the hypothesis states that $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$. Therefore, $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$ for all $i \in \mathscr{N}$. The result follows from Proposition 3 (ii).

(ii) If $s \in \mathscr{S}^1(\mathbf{x})$, then $s \in \mathscr{S}^1(\bar{\mathbf{x}}) = \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$ for all $i \in \mathscr{N}$ since $\mathscr{S}^1_{i,v}(\bar{\mathbf{x}}) = \emptyset$ for all $i \in \mathscr{N}$ by Lemma 1. This implies that $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}})$ for all $s \in \mathscr{S}^1(\mathbf{x})$, $i \in \mathscr{N}$. Moreover, $w_i^{\bar{\mathbf{x}}}(s) \geq w_i^{\mathbf{x}}(s)$ for all $s \in \mathscr{S}$, $i \in \mathscr{N}$ by part (i), and hence by monotonicity of $F_i(s, \mathbf{v})$ with respect to $\mathbf{v}$, it follows that $F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}$, $i \in \mathscr{N}$. Therefore, $u_i(s,1) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\mathbf{x}})$ for all $s \in \mathscr{S}^1(\mathbf{x})$, $i \in \mathscr{N}$. So, $\mathscr{S}^1_{i,v}(\mathbf{x}) = \emptyset$ for all $i \in \mathscr{N}$. It follows from Lemma 1 that $\mathbf{x}$ is an equilibrium. □

Proof of Corollary 2 Immediate from Remark 1 and Proposition 4 (ii). □

Proof of Proposition 5 (i) Immediate from Remark 1, Proposition 4 (i), and the assumption that $c_i(s) \geq 0$ for all $s \in \mathscr{S}$, $i \in \mathscr{N}$.

(ii) If $\sum_{s \in \mathscr{S}^0(\bar{\mathbf{x}})} x(s) > 0$, (8) is redundant. Otherwise, $\sum_{s \in \mathscr{S}^0(\bar{\mathbf{x}})} x(s) = 0$, implying $\mathscr{S}^1(\mathbf{x}) \subseteq \mathscr{S}^1(\bar{\mathbf{x}})$, so (8) cuts off an equilibrium $\mathbf{x}$ if and only if $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$.

Moreover, if $\mathscr{S}^1(\mathbf{x}) \subset \mathscr{S}^1(\bar{\mathbf{x}})$, then $\mathbf{x}$ is a non-maximal equilibrium since $\bar{\mathbf{x}}$ is an equilibrium. □

Proof of Proposition 6 Let $\Xi_{\bar{\mathbf{x}}}$ be the set of strategy profiles $\tilde{\mathbf{x}}$ for which there exists an state $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$ such that:

- $\tilde{x}(s) = 1$ for all $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$,
- $\tilde{x}(s) = 0$ for all $s \in \mathscr{S}^0_{i,nv}(\bar{\mathbf{x}})$,
- $\tilde{x}(\hat{s}) = 1$ for some $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$,
- $\tilde{x}(s) = 0$ for all $s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})/\{\hat{s}\}$.

We show that each $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ is not an equilibrium. Since $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, there are two mutually exclusive cases for $\hat{s}$:

Case 1 If $\hat{s} \in \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$, then $w_i^{\tilde{\mathbf{x}}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$ by Proposition 3 (i). As a result, $F_i(s, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}})$ for all $s \in \mathscr{S}$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(\hat{s}, \mathbf{w}_i^{\bar{\mathbf{x}}}) > u_i(\hat{s}, 1)$. Since $F_i(\hat{\mathbf{s}}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) > u_i(\hat{s}, 1)$ and $\tilde{x}(\hat{s}) = 1$, the Bellman–Shapley equation (5) is violated at state $\hat{s}$ under $\tilde{\mathbf{x}}$, so it is not an equilibrium.

Case 2 If $\hat{s} \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}})$, then consider strategy profile $\check{\mathbf{x}}$ defined as follows:

- $\check{x}(s) = 1$ for all $s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$,
- $\check{x}(s) = 0$ for all $s \in \mathscr{S}/\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})$.

Therefore, $\tilde{\mathbf{x}}$ and $\check{\mathbf{x}}$ take the same value for all $s \in \mathscr{S}/\{\hat{s}\}$. By Corollary 1, if the strategy profile $\check{\mathbf{x}}$ is not an equilibrium, $\tilde{\mathbf{x}}$ cannot be an equilibrium since $\mathscr{S}^1(\check{\mathbf{x}}) \subseteq \mathscr{S}^1(\tilde{\mathbf{x}})$.

Suppose $\check{\mathbf{x}}$ is an equilibrium. By Proposition 3 (i), $w_i^{\check{\mathbf{x}}}(s) \geq w_i^{\bar{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$. Therefore, $F_i(s, \mathbf{w}_i^{\check{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\bar{\mathbf{x}}}) > u_i(s, 1)$ for all $s \in \mathscr{S}^0_{i,v}(\bar{x})$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\check{\mathbf{x}}}) > u_i(\hat{s}, 1)$. Moreover, suppose that $\tilde{\mathbf{x}}$ is an equilibrium; thus $w_i^{\tilde{\mathbf{x}}}(s) \geq w_i^{\check{\mathbf{x}}}(s)$ for all $s \in \mathscr{S}$ by Proposition 4 (i). As a result, $F_i(s, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(s, \mathbf{w}_i^{\check{\mathbf{x}}})$ for all $s \in \mathscr{S}$. In particular, $F_i(\hat{s}, \mathbf{w}_i^{\tilde{\mathbf{x}}}) \geq F_i(\hat{s}, \mathbf{w}_i^{\check{\mathbf{x}}}) > u_i(\hat{s}, 1)$, and this means that the Bellman–Shapley equation (5) is violated at state $\hat{s}$ under $\tilde{\mathbf{x}}$. Therefore, $\tilde{\mathbf{x}}$ is not an equilibrium, which is a contradiction.

So far, we have shown that each $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ is not an equilibrium. For any strategy profile $\mathbf{x}$, satisfying the conditions of Proposition 6, there exists a strategy profile $\tilde{\mathbf{x}} \in \Xi_{\bar{\mathbf{x}}}$ such that $\mathscr{S}^1(\tilde{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$, so $\mathbf{x}$ cannot be an equilibrium by Corollary 1. □

Proof of Proposition 7 If $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})}[1 - x(s)] > 0$, then (9) is redundant. Now, consider the other case in which $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})}[1 - x(s)] = 0$. This implies $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$. If $\mathbf{x}$ is an equilibrium and $\mathscr{S}^1_{i,nv}(\bar{\mathbf{x}}) \subseteq \mathscr{S}^1(\mathbf{x})$, then $\mathscr{S}^1(\mathbf{x}) \cap (\mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})) = \emptyset$ by Proposition 6. This is equivalent to saying that if $\mathbf{x}$ is an equilibrium and $\sum_{s \in \mathscr{S}^1_{i,nv}(\bar{\mathbf{x}})}[1 - x(s)] = 0$, then $x(s) = 0$ for all $s \in \mathscr{S}^0_{i,v}(\bar{\mathbf{x}}) \cup \mathscr{S}^1_{i,v}(\bar{\mathbf{x}})$. □

Recall that $\mathbf{d}$ is the payoff profile of the strategy profile $\mathbf{0}$.

Lemma 2 (i) *$d_i(s) = F_i(s, \mathbf{d}_i)$ for all $s \in \mathscr{S}$, $i \in \mathscr{N}$.*
(ii) *If a strategy profile $\mathbf{x}$ is an equilibrium, then $w_i^{\mathbf{x}}(s) \geq d_i(s)$ for all $s \in \mathscr{S}$, $i \in \mathscr{N}$.*
(iii) *If $u_i(\bar{s}, 1) < d_i(\bar{s})$ for some $\bar{s} \in \mathscr{S}$, $i \in \mathscr{N}$, then $x(\bar{s}) = 0$ for each equilibrium $\mathbf{x}$.*

Proof (i) Immediate from Proposition 2 (i).

(ii) Note that $\mathscr{S}^1(\mathbf{0}) = \emptyset \subseteq \mathscr{S}^1(\mathbf{x})$ for each equilibrium $\mathbf{x}$. The result follows from Proposition 4 (i).

(iii) Substituting $\mathbf{0}$ for $\bar{\mathbf{x}}$ in Proposition 7 implies that the following set of inequalities are valid for Ψ.

$$\sum_{s\in \mathscr{S}^0_{i,v}(\mathbf{0})\cup \mathscr{S}^1_{i,v}(\mathbf{0})} x(s) \le \left(\left|\mathscr{S}^0_{i,v}(\mathbf{0})\right| + \left|\mathscr{S}^1_{i,v}(\mathbf{0})\right|\right) \sum_{s\in \mathscr{S}^1_{i,nv}(\mathbf{0})} [1 - x(s)] = 0 \quad \forall i \in \mathcal{N},$$

where we have made use of the fact that $\mathscr{S}^1_{i,nv}(\mathbf{0}) \subseteq \mathscr{S}^1(\mathbf{0}) = \emptyset$, to write the equality. Note that $\bar{s} \in \mathscr{S}^0_{i,v}(\mathbf{0})$ since $u_i(\bar{s}, 1) < d_i(\bar{s})$. By the above set of inequalities, $x(\bar{s}) = 0$ is valid for Ψ, and hence $x(\bar{s}) = 0$ for each equilibrium $\mathbf{x}$. □

Proof of Proposition 8 (i) ($\Rightarrow$) $\mathbf{x} \in \mathbb{B}^{|\mathscr{S}|}$ by (10f). For each $s \in \mathscr{S}$, there are two cases: If $x(s) = 0$, $w_i(s) = F_i(s, \mathbf{w}_i)$ for all $i \in \mathcal{N}$ by (10a)–(10b).

If $x(s) = 1$, $w_i(s) = u_i(s, 1)$ for all $i \in \mathcal{N}$ by (10c)–(10d).

As a result, $\mathbf{w}$ is associated payoff profile of $\mathbf{x}$ by Proposition 2 (i). It also follows from (10a) that the Bellman–Shapley equation (5) is satisfied for all $s \in \mathscr{S}^1(\mathbf{x})$. Therefore, $\mathbf{x}$ is an equilibrium by Proposition 2 (ii).

($\Leftarrow$) Suppose $\mathbf{x}$ is an equilibrium with associated payoff profile $\mathbf{w}$. Constraint (10f) is satisfied since $\mathbf{x}$ is a (pure) strategy profile. Constraint (10a) is satisfied by Proposition 2 (ii). For each $s \in \mathscr{S}$, there are two cases:
If $x(s) = 0$, then $w_i(s) = F_i(s, \mathbf{w}_i)$ by Proposition 2 (i). Constraint (10b) is obviously satisfied, and (10c) is satisfied by Lemma 2 (ii). Constraint (10d) is satisfied since:

$$w_i(s) = F_i(s, \mathbf{w}_i) \le F_i(s, \mathbf{V}_i),$$

where the inequality follows from the definition of $\mathbf{V}_i$.

If $x(s) = 1$, then $w_i(s) = u_i(s, 1)$ by Proposition 2 (i). Constraint (10b) is satisfied since:

$$\begin{aligned} w_i(s) = u_i(s, 1) &\le u_i(s, 1) + F_i(s, \mathbf{w}_i) - F_i(s, \mathbf{d}_i) \\ &= u_i(s, 1) + F_i(s, \mathbf{w}_i) - d_i(s), \end{aligned}$$

where the first inequality follows from Lemma 2 (ii), and the second equality follows from Lemma 2 (i). Constraints (10c)–(10d) are obviously satisfied.

(ii) If $u_i(s, 0), u_i(s, 1) \ge 0$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$, it can easily be shown by value iteration that $d_i(s) \ge 0$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$. The rest of the proof is straightforward. □

Proof of Proposition 9 Recall that $\gamma_i(s)$ for all $s \in \mathscr{S}$, $i \in \mathcal{N}$, are optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$. Let $\mathcal{P}_i(\bar{\mathbf{x}})$ denote the subproblem associated with each player i in (11a)–(11b) when $\mathbf{x}$ is set to $\bar{\mathbf{x}}$:

$$
\begin{aligned}
\mathcal{P}_i(\bar{\mathbf{x}}) : \max \sum_{s \in \mathscr{S}} & c_i(s) w_i(s) \\
s.t. & \\
& w_i(s) \geq F_i(s, \mathbf{w}_i) && \forall s \in \mathscr{S}, \\
& w_i(s) \leq F_i(s, \mathbf{w}_i) + [u_i(s,1) - d_i(s)]\bar{x}(s) && \forall s \in \mathscr{S}, \\
& w_i(s) \geq [u_i(s,1) - d_i(s)]\bar{x}(s) + d_i(s) && \forall s \in \mathscr{S}, \\
& w_i(s) \leq u_i(s,1)\bar{x}(s) + F_i(s, \mathbf{V}_i)[1 - \bar{x}(s)] && \forall s \in \mathscr{S}, \\
& w_i(s) \text{ unrestricted} && \forall s \in \mathscr{S}.
\end{aligned}
$$

In order to generate a Benders' optimality cut in Step (2c), optimal dual multipliers of $\mathcal{P}_i(\bar{\mathbf{x}})$ are needed while we only know optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$. In fact, optimal dual multipliers of $\mathcal{R}_i(\bar{\mathbf{x}})$ and $\mathcal{P}_i(\bar{\mathbf{x}})$ are closely related. Let $\pi_{i,1}(s), \pi_{i,2}(s), \pi_{i,3}(s), \pi_{i,4}(s)$ be optimal dual multipliers of $\mathcal{P}_i(\bar{\mathbf{x}})$ associated with equations (10a)–(10d), for all $s \in \mathscr{S}, i \in \mathcal{N}$, respectively. It can easily be seen that for all $i \in \mathcal{N}$:

$$
\begin{aligned}
& \pi_{i,1}(s) = \pi_{i,2}(s) = 0 && \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}), && (19a) \\
& \pi_{i,1}(s) + \pi_{i,2}(s) = \gamma_i(s) && \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}), && (19b) \\
& \pi_{i,3}(s) = \pi_{i,4}(s) = 0 && \forall s \in \mathscr{S}^0(\bar{\mathbf{x}}), && (19c) \\
& \pi_{i,3}(s) + \pi_{i,4}(s) = \gamma_i(s) && \forall s \in \mathscr{S}^1(\bar{\mathbf{x}}), && (19d) \\
& \pi_{i,1}(s) \leq 0, \pi_{i,2}(s) \geq 0, \pi_{i,3}(s) \leq 0, \pi_{i,4}(s) \geq 0 && \forall s \in \mathscr{S}. && (19e)
\end{aligned}
$$

At each iteration of the Decomposition Algorithm, $\mathcal{R}_i(\bar{\mathbf{x}})$ is solved, and $\gamma_i(s)$ is obtained for all $s \in \mathscr{S}, i \in \mathcal{N}$. There are multiple dual optimal solutions for $\mathcal{P}_i(\bar{\mathbf{x}})$, and we may use any $\pi_{i,1}(s), \pi_{i,2}(s), \pi_{i,3}(s), \pi_{i,4}(s)$, satisfying (19a)–(19e), to generate a Benders' optimality cut. A Benders' optimality cut is as follows:

$$
\begin{aligned}
\theta_i \leq \sum_{s \in \mathscr{S}} & \pi_{i,1}(s)\, u_i(s,0) + \pi_{i,2}(s) \Big(u_i(s,0) + [u_i(s,1) - d_i(s)]x(s) \Big) \\
& + \pi_{i,3}(s) \Big([u_i(s,1) - d_i(s)]x(s) + d_i(s) \Big) \\
& + \pi_{i,4}(s) \Big(u_i(s,1)x(s) + F_i(s, \mathbf{V}_i)[1 - x(s)] \Big) \qquad \forall i \in \mathcal{N}.
\end{aligned} \tag{20}
$$

Since $\mathcal{R}_i(\bar{\mathbf{x}})$ is not degenerate, $\langle \gamma_i(s) \rangle_{s \in \mathscr{S}}$ is the unique dual solution of $\mathcal{R}_i(\bar{\mathbf{x}})$. In order to generate a Pareto-optimal Benders' optimality cut, we need to find $\pi^*_{i,1}(s), \pi^*_{i,2}(s), \pi^*_{i,3}(s), \pi^*_{i,4}(s)$ for all $s \in \mathscr{S}$ such that they minimize the term on the right-hand side of (20) subject to (19a)–(19e) for each equilibrium $\mathbf{x}$. The right-hand side of (20) subject to (19a)–(19e) may be rewritten as follows:

$$
\sum_{s \in \mathscr{S}} \pi_{i,1}(s)\, u_i(s,0) + \pi_{i,2}(s) \Big(u_i(s,0) + [u_i(s,1) - d_i(s)]x(s) \Big)
$$

$$
\begin{aligned}
&+ i_{i,3}(s)\left([u_i(s,1) - d_i(s)]x(s) + d_i(s)\right)\\
&+ \pi_{i,4}(s)\left(u_i(s,1)x(s) + F_i(s,\mathbf{V}_i)[1-x(s)]\right)\\
= &\sum_{s\in\mathscr{S}^0(\bar{\mathbf{x}})} [\pi_{i,1}(s) + \pi_{i,2}(s)]\, u_i(s,0)\\
&+ \pi_{i,2}(s)\,[u_i(s,1) - d_i(s)]x(s) + \sum_{s\in\mathscr{S}^1(\bar{\mathbf{x}})} [\pi_{i,3}(s) + \pi_{i,4}(s)]\, u_i(s,1)x(s)\\
&+ [\pi_{i,3}(s)\, d_i(s) + \pi_{i,4}(s)\, F_i(s,\mathbf{V}_i)][1-x(s)]\\
= &\sum_{s\mathscr{S}^0(\bar{\mathbf{x}})} \gamma_i(s)\, u_i(s,0) + \pi_{i,2}(s)\,[u_i(s,1) - d_i(s)]x(s)\\
&+ \sum_{s\in\mathscr{S}^1(\bar{\mathbf{x}})} \gamma_i(s)\, u_i(s,1)x(s) + [\pi_{i,3}(s)\, d_i(s) + \pi_{i,4}(s)\, F_i(s,\mathbf{V}_i)][1-x(s)],
\end{aligned}
\tag{21}
$$

where the first and second equality follow from (19a), (19c) and (19b), (19d), respectively. In order to minimize (21) subject to (19a)–(19e), we may seek to minimize (21) for each $s \in \mathscr{S}$ separately. There are three mutually exclusive cases for s:

1. If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$ and $u_i(s,1) \geq d_i(s)$, then $\pi^*_{i,1}(s) := \min\{0, \gamma_i(s)\} = -\gamma_i^-(s), \pi^*_{i,2}(s) := \max\{0, \gamma_i(s)\} = \gamma_i^+(s), \pi^*_{i,3}(s) := 0, \pi^*_{i,4}(s) := 0$ minimizes (21) subject to (19a)–(19e).
2. If $s \in \mathscr{S}^0(\bar{\mathbf{x}})$ and $u_i(s,1) < d_i(s)$, then $x(s)$ is equal to 0 for any equilibrium $\mathbf{x}$ by Lemma 2. As a result, $\pi^*_{i,1}(s) := -\gamma_i^-(s), \pi^*_{i,2}(s) := \gamma_i^+(s), \pi^*_{i,3}(s) := 0, \pi^*_{i,4}(s) := 0$ minimizes (21) subject to (19a)–(19e).
3. If $s \in \mathscr{S}^1(\bar{\mathbf{x}})$, then $\pi^*_{i,1}(s) := 0, \pi^*_{i,2}(s) := 0, \pi^*_{i,3}(s) := -\gamma_i^-(s), \pi^*_{i,4}(s) := \gamma_i^+(s)$ minimizes (21) subject to (19a)–(19e) since $F_i(s,\mathbf{V}_i) \geq F_i(s,\mathbf{d}_i) = d_i(s)$.

□

Proof of Proposition 10 If $s \in \mathscr{S}^0(\bar{\mathbf{x}}_2)$, the Bellman–Shapley equation (5) is obviously satisfied for all $i \in \mathscr{N}$. Conversely, if $s \in \mathscr{S}^1(\bar{\mathbf{x}}_2)$, then

$$
w_i^{\bar{\mathbf{x}}_2}(s) = u_i(s,1) = V_i(s) \geq F_i(s,\mathbf{V}_i) \geq F_i(s,\mathbf{w}_i^{\bar{\mathbf{x}}_2}) \qquad \forall i \in \mathscr{N},
$$

where the first equality follows from Proposition 2 (i), the second equality follows from the definition of $\bar{\mathbf{x}}^2$, and the first and second inequalities follow from the definition of $\mathbf{V}_i$. Hence, the Bellman–Shapley equation (5) is satisfied for all $s \in \mathscr{S}^1(\bar{\mathbf{x}}_2)$, $i \in \mathscr{N}$.

□

Random instance generation We generated two categories of two-player consensus stopping game instances as follows. For the first (second) category, the size of S_i is equal to 40 (60) for both players. Recall from Sect. 8.1 that (1) the game state is $s = (s_1, s_2)$, i.e., $\mathscr{S} = S_1 \times S_2$, and (2) the game transition probability matrix $\mathscr{P}$ is the Kronecker product of the individual transition probability matrices, i.e.,

$\mathscr{P}(s_1', s_2'|s_1, s_2) = P_i(s_1'|s_1)P_i(s_2'|s_2)$. In many practical applications of consensus stopping games such as war termination and organ exchange, which were discussed in Sect. 1, transition in state of the game is slow, i.e., the game most likely remains in the same state at the next period as that of the current period. For this reason, we randomly generated a set of individual transition probability matrices that are highly diagonal, i.e., the diagonal entries are close to 1. Specifically, in generating our transition matrices we used the notion of increasing failure rate (IFR) property. The IFR property has its origins in maintenance optimization and reliability literature [4], but it has been recently shown that data in varying real contexts, primarily in healthcare and service operations, empirically exhibit IFR property [2]. The transition matrices we generated are designed to be moderately sparse but do not have any diagonal entry that is less than 0.99. Such transition matrices can be encountered in real-life dynamic settings where decision epochs are spaced very close to each other so that leaving the state of the system in one period is not very likely. It is also common to see sparse transition matrices with large diagonal entries when solutions of a large-scale dynamic decision-making problem are approximated through state aggregation. To ensure our transition matrices have the IFR property and the specified threshold probabilities in their diagonals, we simulated their entries iteratively starting from the top row. In each particular row, we simulated the entries from left to right in column order after randomly fixing the diagonal entry in the specified range. All entries of the same row are generated from a uniform distribution whose boundaries are imposed by a corresponding partial row sum from the previous row due to the IFR property. While such order restrictions can disallow some entries to be positive, we also allowed each nondiagonal entry to be 0 with probability 0.01. Across all transition matrices we generated, on average, the transition matrices for 40-state-per-player instances were 53% sparse whereas the transition matrices for 60-state-per-player instances were 63% sparse.

For each $i \in \{1, 2\}$ and $s \in \mathscr{S}$, the rewards $u_i(s, 0)$ and $u_i(s, 1)$ only depend on s_i. For each $i \in \{1, 2\}$ and $s \in \{(s_1, s_2) \in \mathscr{S} : s_i \neq |S_i|\}$, $u_i(s, 0)$ is generated according to a uniform distribution on the interval $[-150, 50]$, and $u_i(s, 1)$ is equal to $\min_{s\in\mathscr{S}} d_i(s) + rand(s_i)\frac{|S_i|-s_i}{|S_i|} \max_{s\in\mathscr{S}} d_i(s)$, where $rand(s_i)$ is a random number from a uniform distribution on [0, 1]. In addition, the last state (i.e., $s_i = |S_i|$) is absorbing, and its continuation and stopping rewards are 0 and $-\infty$, respectively. Finally, the initial state of the game $\hat{s}$ is generated by a discrete uniform distribution between $\frac{|S_i|}{4}$ and $\frac{|S_i|}{2}$.

References

1. Aguirregabiria, V., Mira, P.: Sequential estimation of dynamic discrete games. Econometrica **75**(1), 1–53 (2007)
2. Alagoz, O.: Optimal Policies for the Acceptance of Living- and Cadaveric-donor Livers. Ph.D. thesis, University of Pittsburgh (2004)
3. Awasthi, P., Sandholm, T.: Online stochastic optimization in the large: application to kidney exchange. In: Proceedings of the International Joint Conference on Artificial Intelligence, vol. 9, pp. 405–411 (2009)
4. Barlow, R.E., Proschan, F.: Mathematical Theory of Reliability. Wiley, Hoboken (1965)
5. Benders, J.F.: Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik **4**(1), 238–252 (1962)

6. Blackwell, D.: Discounted dynamic programming. Ann. Math. Stat. **36**(1), 226–235 (1965)
7. Borkovsky, R.N., Doraszelski, U., Kryukov, Y.: A user's guide to solving dynamic stochastic games using the homotopy method. Oper. Res. **58**(4–Part–2), 1116–1132 (2010)
8. Cunningham, D.E.: Veto players and civil war duration. Am. J. Polit. Sci. **50**(4), 875–892 (2006)
9. Dehghanian, A.: Integer Programming Approaches to Stochastic Games Arising in Paired Kidney Exchange and Industrial Organization. Ph.D. thesis, University of Pittsburgh (2015)
10. Denardo, E.V.: Contraction mappings in the theory underlying dynamic programming. SIAM Rev. **9**(2), 165–177 (1967)
11. Doraszelski, U., Escobar, J.F.: A theory of regular Markov perfect equilibria in dynamic stochastic games: genericity, stability, and purification. Theor. Econ. **5**(3), 369–402 (2010)
12. Doraszelski, U., Satterthwaite, M.: Computable Markov-perfect industry dynamics. RAND J. Econ. **41**(2), 215–243 (2010)
13. Eliashberg, J., Winkler, R.L.: Risk sharing and group decision making. Manag. Sci. **27**(11), 1221–1235 (1981)
14. Ergin, H., Sonmez, T., Unver, M.U.: Lung exchange. Technical report, Boston College, Department of Economics (2015)
15. Ericson, R., Pakes, A.: Markov-perfect industry dynamics: a framework for empirical work. Rev. Econ. Stud. **62**(1), 53–82 (1995)
16. European Communities: The Treaty of Lisbon. European Communities (2007)
17. Filar, J.A., Vrieze, K.: Competitive Markov Decision Processes. Springer, New York (1997)
18. Filar, J.A., Schultz, T.A., Thuijsman, F., Vrieze, O.J.: Nonlinear programming and stationary equilibria in stochastic games. Math. Program. **50**(1–3), 227–237 (1991)
19. Filson, D., Werner, S.: A bargaining model of war and peace: anticipating the onset, duration, and outcome of war. Am. J. Polit. Sci. **46**(4), 819–837 (2002)
20. Fink, A.M.: Equilibrium in a stochastic n-person game. J. Sci. Hiroshima Univ. Ser. AI **28**(1), 89–93 (1964)
21. Fischetti, M., Salvagnin, D., Zanette, A.: A note on the selection of Benders' cuts. Math. Program. **124**(1–2), 175–182 (2010)
22. Glorie, K.M., van de Klundert, J.J., Wagelmans, A.P.M.: Kidney exchange with long chains: an efficient pricing algorithm for clearing barter exchanges with branch-and-price. Manuf. Serv. Oper. Manag. **16**(4), 498–512 (2014)
23. Heller, Y.: Sequential correlated equilibria in stopping games. Oper. Res. **60**(1), 209–224 (2012)
24. Herings, P., Peeters, R.J.A.P.: Stationary equilibria in stochastic games: structure, selection, and computation. J. Econ. Theory **118**(1), 32–60 (2004)
25. Hörner, J., Sugaya, T., Takahashi, S., Vieille, N.: Recursive methods in discounted stochastic games: an algorithm for $\delta \to 1$ and a folk theorem. Econometrica **79**(4), 1277–1318 (2011)
26. ILOG-CPLEX 12.4.: IBM ILOG CPLEX Optimization Studio 12.4 Information Center (2012)
27. Köppe, M., Ryan, C.T., Queyranne, M.: Rational generating functions and integer programming games. Oper. Res. **59**(6), 1445–1460 (2011)
28. Kurt, M.: Dynamic Decision Models for Managing the Major Complications of Diabetes. Ph.D. thesis, University of Pittsburgh (2012)
29. Magnanti, T.L., Wong, R.T.: Accelerating Benders decomposition: algorithmic enhancement and model selection criteria. Oper. Res. **29**(3), 464–484 (1981)
30. Maskin, E., Tirole, J.: Markov perfect equilibrium: I. Observable actions. J. Econ. Theory **100**(2), 191–219 (2001)
31. Montgomery, R.A., Zachary, A.A., Ratner, L.E., Segev, D.L., Hiller, J.M., Houp, J., Cooper, M., Kavoussi, L., Jarrett, T., Burdick, J., Maley, W.R., Melancon, J.K., Kozlowski, T., Simpkins, C.E., Phillips, M., Desai, A., Reeb, B., Kraus, E., Rabb, H., Leffell, M.S., Warren, D.S.: Clinical results from transplanting incompatible live kidney donor/recipient pairs using kidney paired donation. J. Am. Med. Assoc. **294**(13), 1655–1663 (2005)
32. Nemhauser, G.L., Trotter Jr., L.E.: Properties of vertex packing and independence system polyhedra. Math. Program. **6**(1), 48–61 (1974)
33. Nowak, A.S., Szajowski, K.: Advances in Dynamic Games. Springer, Berlin (2005)
34. Qiu, F., Ahmed, S., Dey, S.S., Wolsey, L.A.: Covering linear programming with violations. INFORMS J. Comput. **26**(3), 531–546 (2014)
35. Raghavan, T.E.S., Syed, Z.: A policy-improvement type algorithm for solving zero-sum two-person stochastic games of perfect information. Math. Program. **95**(3), 513–532 (2003)

36. Shapley, L.S.: Stochastic games. Proc. Natl. Acad. Sci. U. S. A. **39**(10), 1095–1100 (1953)
37. Shmaya, E., Solan, E.: Two-player nonzero-sum stopping games in discrete time. Ann. Probab. **32**(3), 2733–2764 (2004)
38. Singer, J.D.: Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985. Int. Interact. **14**(2), 115–132 (1988)
39. Singer, J.D., Bremer, S., Stuckey, J.: Capability distribution, uncertainty, and major power war, 1820–1965. In: Russett, B. (ed.) Peace, War, and Numbers, pp. 19–48. Sage Beverly Hills, Beverly Hills (1972)
40. Sobel, M.J.: Noncooperative stochastic games. Ann. Math. Stat. **42**(6), 1930–1935 (1971)
41. Solan, E.: Stochastic games. In: Meyers, R.A. (ed.) Computational Complexity, pp. 3064–3074. Springer, New York (2012)
42. Solan, E., Vieille, N.: Quitting games. Math. Oper. Res. **26**(2), 265–285 (2001)
43. Steinberg, R.H.: In the shadow of law or power? Consensus-based bargaining and outcomes in the GATT/WTO. Int. Organ. **56**(2), 339–374 (2002)
44. Tsebelis, G.: Veto Players: How Political Institutions Work. Princeton University Press, Princeton (2002)
45. Ünver, M.U.: Dynamic kidney exchange. Rev. Econ. Stud. **77**(1), 372–414 (2010)
46. Weintraub, G.Y., Benkard, C.L., Van Roy, B.: Markov perfect industry dynamics with many firms. Econometrica **76**(6), 1375–1411 (2008)
47. Weintraub, G.Y., Benkard, C.L., Van Roy, B.: Computational methods for oblivious equilibrium. Oper. Res. **58**(4), 1247–1265 (2010)
48. Weis, L., Metzger, M., Haymann, J.P., Thervet, E., Flamant, M., Vrtovsnik, F., Gauci, C., Houillier, P., Froissart, M., Letavernier, E., et al.: Renal function can improve at any stage of chronic kidney disease. PLoS One **8**(12), e81835 (2013)

Affiliations

Amin Dehghanian[1] · Murat Kurt[3] · Andrew J. Schaefer[2]

Amin Dehghanian
amin.dehghanian@aut.ac.ir

Andrew J. Schaefer
andrew.schaefer@rice.edu

[1] Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

[2] Department of Computational and Applied Mathematics, Rice University, Houston, USA

[3] Princeton, USA