Dynamic Games and Applications (2021) 11:363–388
https://doi.org/10.1007/s13235-020-00364-x

Upper and Lower Values in Zero-Sum Stochastic Games with Asymmetric Information

Dhruva Kartik[1] · **Ashutosh Nayyar**[1]

Published online: 27 July 2020

Abstract
A general model for zero-sum stochastic games with asymmetric information is considered. In this model, each player's information at each time can be divided into a common information part and a private information part. Under certain conditions on the evolution of the common and private information, a dynamic programming characterization of the value of the game (if it exists) is presented. If the value of the zero-sum game does not exist, then the dynamic program provides bounds on the upper and lower values of the game.

Keywords Dynamic games · Asymmetric information · Upper and lower values

1 Introduction

Zero-sum games have been widely used as a model of strategic decision making in the presence of adversaries. Such decision-making scenarios arise in a range of domains including (i) security of cyber-physical and infrastructure systems such as the power grid and water networks in the presence of cyber or physical attacks [2,3,38,41,42,44], (ii) cyber-security of networked computing and communication systems [1,41], (iii) designing anti-poaching measures [7–9], (iv) military operations in the presence of hostile agents [15] and (v) competitive markets and geopolitical interactions [4,24]. In many cases, the adversarial interactions occur over time in a dynamic and uncertain environment. Zero-sum stochastic games provide a useful model for these situations. In these games, two players may jointly control the evolution of the state of a stochastic dynamic system with one player trying to minimize the total cost while the other trying to maximize it. In stochastic games with symmetric information, all players have the same information about the state and action histories. Such games have been

Preliminary version of this paper appears in the proceedings of the 58th Conference on Decision and Control (CDC), 2019 [17].

✉ Dhruva Kartik
mokhasun@usc.edu

Ashutosh Nayyar
ashutosh.nayyar@usc.edu

[1] Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90007, USA

extensively studied in the literature in both zero-sum and nonzero-sum settings [6,10,11]. In many situations of interest, however, the players may have different information about the state and action histories. A potential attacker of a cyber-physical system, for example, may not have the same information as the defender; adversaries in a battlefield may have different information about the surroundings and about each other. The focus of this paper is on such *asymmetric information* settings.

We adopt a model of asymmetric information that was originally developed for decentralized stochastic control [28]. This model partitions each player's information at each time into a common information part and a private information part. The common information at time t is known to all players at that time and at all times in the future. In addition to the common information, each player may have some private information. It has been noted in the existing literature that this model subsumes a wide range of information structures [26,28].

In our model, it may be the case that no player knows the current state of the underlying stochastic system perfectly. Further, since each player may have some private information, one player's information is not necessarily included in the other player's information. The partial observability of the state, the asymmetry of information and the fact that each player may have some private information complicate the characterization and computation of the equilibrium cost (value) and equilibrium strategies. We provide two results for this general model of zero-sum stochastic game with asymmetric information: (i) If the game has a Nash equilibrium in behavioral strategies, then our result provides dynamic programming-based characterizations of the value of the game. Each step of these programs involves a min–max (or a max-min) problem over the space of *prescriptions* which are functions from players' private information to actions. (ii) If the game does not have a Nash equilibrium, then our dynamic programs provide a lower bound on the upper value of the game and an upper bound on the lower value of the game.

1.1 Related Work and Our Contributions

1. *Stochastic games of symmetric information:* In this stochastic game model, the players have access to the same information. Thus, at any time t, each player has no uncertainty regarding other players' information and makes a decision anticipating the other players' strategies. Such games of symmetric information have been extensively studied in the literature [6,10,11]. Because of this symmetry, players' shared information (or a function of it) can be treated as a state and utilized to decompose a dynamic game into simpler single-stage Bayesian games. These single-stage games can then be solved in a backward-inductive manner to obtain the value and Nash equilibria (if any exist). In this paper, we focus on models in which players have different pieces of information and thus the methodology described above for symmetric information games is not directly applicable to our model. We provide a backward-inductive characterization of value for a general model of zero-sum stochastic games of asymmetric information. Our model subsumes, as a simple special case, zero-sum stochastic games of symmetric information.
2. *Zero-sum games with a more informed player:* Stochastic zero-sum games in which one player knows the other player's information have been investigated before with varying degrees of generality. Some of these works [12,20,23,32,33,35] considered both finite-horizon and infinite-horizon games and focused on studying the existence of a *uniform value* [23]. Other works [19,43] focused on the computational aspects associated with finding the value and Nash equilibrium strategies in these games. The most general treatment of these games can be found in [12] and [20]. All of these works assume

that the players have perfect recall. We make a weaker assumption of perfect recall of common information for our results. Further, in our model (described in Sect. 2), each player may have some information that the other does not.

3. *Incomplete information on both sides:* Stochastic zero-sum game models in which *both* players have private information have been considered in [4,13,34]. In these works, the private information either does not change with time [4,34] or evolves in an uncontrolled manner [13]. Further, players in these works have perfect recall and their actions are publicly observed. Our model described in Sect. 2 makes substantially weaker assumptions on players' information structure.
4. *General zero-sum games:* A general model of zero-sum games is described in Section IV.3 of [23]. A characterization of the value is provided in [23] using a recursive formula. The recursive formula in [23] is in terms of certain probability distributions on an abstract space referred to as the universal belief space. Our results differ from the results in [23] in two key respects. Firstly, the model in [23, Section IV.3] assumes that the players have perfect recall, whereas we do not make this assumption. Secondly, the minimizations and maximizations in our dynamic program are over smaller spaces than those in the recursive formula of [23, Section IV.3]. We elaborate on these differences in Sect. 4.3.
5. *Stochastic games of asymmetric information with strategy-independent common information beliefs:* In [26], a common information-based dynamic program was developed for finding Nash equilibria in general (i.e., not necessarily zero-sum) stochastic games of asymmetric information. The key idea in this approach is to first convert the game of asymmetric information into a virtual game of symmetric information. This virtual game of symmetric information is then solved using a common information-based dynamic program. However, this approach relies on an assumption on the players' information (see Assumption 2 in [26]). This assumption holds only for certain classes of information structures and may not necessarily be true for the asymmetric information games described in Sect. 2. A key contribution of our work is to obtain a characterization of value for models not covered by [26].
6. *Common information-based perfect Bayesian equilibria in stochastic games of asymmetric information:* Authors in [30] consider a stochastic game model in which the system state can be decomposed into a public state that is commonly observed by all players and a private state that is privately observed by each player. In this model, all the players' past actions are commonly observed and, additionally, an imperfect version of players' private state may be disclosed to all the players at each time. A special case of this model has been considered in [40]. For the models in [30] and [40], the authors provide characterizations of perfect Bayesian equilibria under some assumptions on the evolution of players' private state. In this paper, we focus only on two-player zero-sum games. However, the system dynamics and the information structure in our model are more general than those in the model of [30,40]. For instance, unlike in [30,40], players' actions may not be fully observed in our model. Further, the solutions in [30,40] rely on strong existence assumptions that may not be true in general. Our result provides a characterization of upper and lower values for a wide class of games under a mild assumption on the information structure and no assumption on the existence of any particular kind of equilibrium.

Our work is most closely related to [25] and [26]. We follow the approach in [25] and build on its results. The system model in [25] conformed to a specific structure; that is, the system state could be decomposed into three components: a public state that is commonly observed (perhaps partially) and a privately observed component for each player. The model

in our paper is substantially more general than in [25]. Another major restriction in [25] was that the players were allowed to play only pure strategies. In this paper, we allow the players to play behavioral strategies. Our model is similar to [26] but we do not make the critical assumption made in [26] that the common information-based beliefs are strategy-independent (see Assumption 2 of [26]). Removing this assumption makes our model much more widely applicable than the model in [26].

We note that a common assumption in much of the work outlined above is that players have prefect recall, i.e., each player remembers its past observations and actions at each time. The model we study in the paper (and the model in [26]) does not make this assumption. Instead, we make the weaker assumption that only the common information among the players (defined precisely in Sect. 2) should be perfectly recalled. Our results show that even under this weaker assumption, a dynamic programming characterization of value can be obtained. This suggests that nondecreasing common information among players as opposed to nondecreasing total information of each player is the more fundamental condition for characterizing value. Games without perfect recall but with nondecreasing common information could provide useful models for strategic interactions when the players are in fact devices with limited memories. These players may be able to access public information that is located on a public database (e.g., a cloud server) and is perfectly recalled but can only use limited amount of private information due to their memory constraints. We note that several bounded memory models have been investigated for single-agent [14,21,37] and cooperative multiagent [39] decision problems.

1.2 Notation

Random variables/vectors are denoted by uppercase letters and their realizations by the corresponding lowercase letters. In general, subscripts are used as time index, while superscripts are used to index decision-making agents. For time indices $t_1 \leq t_2$, $X_{t_1:t_2}$ (resp. $g_{t_1:t_2}$) is the shorthand notation for the variables $(X_{t_1}, X_{t_1+1}, ..., X_{t_2})$ (resp. functions $(g_{t_1}, \ldots, g_{t_2})$). Similarly, $X^{1:2}$ is the shorthand notation for the collection of variables (X^1, X^2). Operators $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ denote the probability of an event and the expectation of a random variable, respectively. For random variables/vectors X and Y, $\mathbb{P}(\cdot|Y = y)$, $\mathbb{E}[X|Y = y]$ and $\mathbb{P}(X = x \mid Y = y)$ are denoted by $\mathbb{P}(\cdot|y)$, $\mathbb{E}[X|y]$ and $\mathbb{P}(x \mid y)$, respectively. For a strategy g, we use $\mathbb{P}^g(\cdot)$ (resp. $\mathbb{E}^g[\cdot]$) to indicate that the probability (resp. expectation) depends on the choice of g. For any finite set $\mathcal{A}$, $\Delta\mathcal{A}$ denotes the probability simplex over the set $\mathcal{A}$.

1.3 Organization

The rest of the paper is organized as follows. We formulate the game in Sect. 2 and construct a virtual game with symmetric information in Sect. 3. In Sect. 4, we construct an expanded virtual game and use it provide a dynamic programming characterization of the value. We conclude the paper in Sect. 5. Proofs of key results are provided in Appendices.

2 Problem Formulation

Consider a dynamic system with two players. The system operates in discrete time over a horizon T. Let $X_t \in \mathcal{X}_t$ be the state of the system at time t, and let $U_t^i \in \mathcal{U}_t^i$ be the action of player i at time t, where $i = 1, 2$. The state of the system evolves in a controlled Markovian

manner as

$$X_{t+1} = f_t(X_t, U_t^1, U_t^2, W_t^s), \tag{1}$$

where W_t^s is the system noise. There are two observation processes $Y_t^1 \in \mathcal{Y}_t^1$ and $Y_t^2 \in \mathcal{Y}_t^2$ given as

$$Y_t^i = h_t^i(X_t, U_{t-1}^1, U_{t-1}^2, W_t^i),\ i = 1, 2, \tag{2}$$

where W_t^1 and W_t^2 are observation noises. We assume that the sets $\mathcal{X}_t, \mathcal{U}_t^i$ and $\mathcal{Y}_t^i$ are finite for all i and t. Further, the random variables X_1, W_t^s, W_t^i (referred to as *the primitive random variables*) can take finitely many values and are mutually independent.

2.1 Information Structure

The collection of variables (i.e., observations, actions) available to player i at time t is denoted by I_t^i. I_t^i is a subset of all observations until time t and actions until $t-1$, i.e., $I_t^i \subseteq \{Y_{1:t}^{1:2}, U_{1:t-1}^{1:2}\}$. The set of all possible realizations of I_t^i is denoted by $\mathcal{I}_t^i$.

Information I_t^i can be decomposed into *private* and *common* information, i.e., $I_t^i = C_t \cup P_t^i$. Common information C_t is the set of variables known to both players at time t, while variables in the private information P_t^i are known only to player i. Let $\mathcal{C}_t$ be the set of all realizations of common information at time t, and let $\mathcal{P}_t^i$ be the set of all realizations of private information for player i at time t. We make the following assumption on the evolution of common and private information. This is similar to Assumption 1 of [26][1].

Assumption 1 *The evolution of common and private information available to the players is as follows:*

1. *The common information C_t is nondecreasing with time, i.e., $C_t \subset C_{t+1}$. Let $Z_{t+1} := C_{t+1} \setminus C_t$ be the increment in common information. Thus, $C_{t+1} = \{C_t, Z_{t+1}\}$. Furthermore,*

$$Z_{t+1} = \zeta_{t+1}(P_t^{1:2}, U_t^{1:2}, Y_{t+1}^{1:2}), \tag{3}$$

   where ζ_{t+1} is a fixed transformation.
2. *The private information evolves as*

$$P_{t+1}^i = \xi_{t+1}^i(P_t^i, U_t^i, Y_{t+1}^i), \tag{4}$$

   where ξ_{t+1}^i is a fixed transformation.

2.2 Examples of Information Structures that Satisfy Assumption 1

As noted in [28] and [26], a number of information structures satisfy the above assumption. We briefly mention a few below:

1. *No common information:* Consider the case where each player only has access to its own observations and actions, i.e., $I_t^i = \{Y_{1:t}^i, U_{1:t-1}^i\}$, $i = 1, 2$. In this case, there is no common information, i.e., $C_t = \varnothing$. It is easy to verify that Assumption 1 is valid in this case.

[1] Note that we do not impose Assumption 2 of [26].

2. *No private information:* Consider the case where all observations and actions are public, i.e., $I_t^i = \{Y_{1:t}^{1:2}, U_{1:t-1}^{1:2}\}$. In this case, players do not have any private information, i.e., $P_t^i = \varnothing$. Once again, Assumption 1 is true.
3. *Player 1 is more informed:* Consider the case where Player 1 knows both players' observations and actions, i.e., $I_t^1 = \{Y_{1:t}^{1:2}, U_{1:t-1}^{1:2}\}$. On the other hand, player 2 has access only to its own observations and actions, i.e., $I_t^2 = \{Y_{1:t}^2, U_{1:t-1}^2\}$. Thus, the common information $C_t = I_t^2$. Player 1's private information $P_t^1 = \{Y_{1:t}^1, U_{1:t-1}^1\}$, and Player 2 does not have any private information. Clearly, the common information is increasing with time and the private information satisfies Assumption 1. This model has been considered in [12]. With additional restrictions, it has also been considered in [19,32,33,35,43].
4. *Full state information on one side and quantized state information on the other:* Consider the model in which player 1 knows the state X_t and player 2 sees a quantized version of X_t. That is, $Y_t^1 = X_t$ and $Y_t^2 = q(X_t)$ where q is an arbitrary function. Both players' actions are commonly observed. Since player 1 knows the state X_t and Y_t^2 is a deterministic function of the state, player 1 also knows player 2's observation Y_t^2. Thus, in this case, $I_t^1 = \{Y_{1:t}^{1:2}, U_{1:t-1}^{1:2}\}$ and $I_t^2 = \{Y_{1:t}^2, U_{1:t-1}^{1:2}\}$. Therefore, $C_t = \{Y_{1:t}^2, U_{1:t-1}^{1:2}\}$, $P_t^1 = Y_{1:t}^1$ and $P_t^2 = \varnothing$. Clearly, this model satisfies Assumption 1.
5. *Delayed sharing:* In this model, players' actions and observations become public with a delay of $s > 1$ time steps. Thus, player i's information at time t is given by $I_t^i = \{Y_{1:t-s}^{1:2}, U_{1:t-s}^{1:2}, Y_{t-s+1:t}^i, U_{t-s+1:t-1}^i\}$. The common information $C_t = \{Y_{1:t-s}^{1:2}, U_{1:t-s}^{1:2}\}$, which is increasing with time. Player i's private information is $P_t^i = \{Y_{t-s+1:t}^i, U_{t-s+1:t-1}^i\}$ which satisfies Assumption 1. Notice that the new information received by Player 1 at time t is $Y_t^1, Y_{t-s}^2, U_{t-s}^2$. Thus, in our model, players may receive *signals* that depend on past states and actions.
6. *Bounded private memory:* Consider a setup in which players' actions are common information. Each player stores its observations on a private device with bounded memory. Let the size of this memory be s. Player i's information is given by $I_t^i = \{Y_{t-s:t}^i, U_{1:t-1}^{1:2}\}$. In this case, the common information $C_t = U_{1:t-1}^{1:2}$ and for player i, private information $P_t^i = Y_{t-s:t}^i$. Clearly, this information structure satisfies Assumption 1. Also, note that players do not have perfect recall in this game.
7. *Bounded private memory with strategic memory updates:* Consider a setup where there is no common information and players do not have perfect recall. Instead, at each time t, player i has a private memory state M_t^i that can take values in a finite set $\mathcal{M}$. Based on its current memory state, the player picks an action U_t^i to influence the state evolution and a memory update action L_t^i. Here, L_t^i is an action that affects the memory update through the following fixed transformation: $M_{t+1}^i = \xi_{t+1}^i(M_t^i, Y_{t+1}^i, L_t^i)$. This model can be viewed as an instance of our general model with M_t^i as player i's private information and (U_t^i, L_t^i) as the player's actions.
8. *Information structure in* [40]: In this model, player i has a private state X_t^i, $i = 1, 2$. Player i knows its private state and both players' actions are commonly observed. This information structure can be seen as a special case of our model in the following manner: Let the state $X_t := (X_t^1, X_t^2)$ and the observation processes $Y_t^i = X_t^i$ for $i = 1, 2$. Define the information sets at time t as $I_t^i = \{Y_{1:t}^i, U_{1:t-1}^{1:2}\}$. In this case, $C_t = \{U_{1:t-1}^1, U_{1:t-1}^2\}$ and $P_t^i = Y_{1:t}^i$. Clearly, this information structure satisfies Assumption 1. Similarly, information structures in [30] and [25] can also be seen as special cases of our model.

2.3 Strategies and Values

Players can use any information available to them to select their actions, and we allow behavioral strategies for both players. Thus, player i chooses a distribution δU_t^i over its action space using a *control law* $g_t^i : \mathcal{I}_t^i \to \Delta\mathcal{U}_t^i$, i.e.

$$\delta U_t^i = g_t^i(I_t^i) = g_t^i(P_t^i, C_t). \tag{5}$$

Player i's action at time t is randomly chosen from $\mathcal{U}_t^i$ according to the distribution δU_t^i. We will at times refer to δU_t^i as player i's *behavioral action* at time t. It will be helpful for our analysis to explicitly describe the randomization procedure used by the players. To do so, we assume that player i has access to i.i.d. random variables $K_{1:T}^i$ that are uniformly distributed over the interval (0, 1]. The variables $K_{1:T}^1$, $K_{1:T}^2$ are independent of each other and of the primitive random variables. Further, player i has access to a mechanism κ that takes as input K_t^i and a distribution over $\mathcal{U}_t^i$ and generates a random action with the input distribution. Thus, player i's action at time t can be written as $U_t^i = \kappa(g_t^i(I_t^i), K_t^i)$.

Remark 1 One choice of the mechanism κ can be described as follows: Suppose $\mathcal{U}_t^i = \{1, 2, ..n\}$ and the input distribution is $(p_1, ...p_n)$. We can *partition* the interval (0, 1] into n intervals $(a_i, b_i]$ such that the length of ith interval is $b_i - a_i = p_i$. Then, $U_t^i = k$ if $K_t^i \in (a_k, b_k]$ for $k = 1, \ldots, n$.

The collection of control laws $g^i = (g_1^i, \ldots, g_T^i)$ is referred to as the *control strategy* of player i, and the pair of control strategies (g^1, g^2) is referred to as a *strategy profile*. Let the set of all possible control strategies for player i be $\mathcal{G}^i$.

The total expected cost associated with a strategy profile (g^1, g^2) is

$$J(g^1, g^2) := \mathbb{E}^{(g^1, g^2)}\left[\sum_{t=1}^{T} c_t(X_t, U_t^1, U_t^2)\right], \tag{6}$$

where $c_t : \mathcal{X}_t \times \mathcal{U}_t^1 \times \mathcal{U}_t^2 \to \mathbb{R}$ is the cost function at time t. Player 1 wants to minimize the total expected cost, while Player 2 wants to maximize it. We refer to this zero-sum game as Game $\mathscr{G}$.

Definition 1 The upper value of the game $\mathscr{G}$ is defined as

$$S^u(\mathscr{G}) := \inf_{g^1 \in \mathcal{G}^1} \sup_{g^2 \in \mathcal{G}^2} J(g^1, g^2). \tag{7}$$

The lower value of the game $\mathscr{G}$ is defined as

$$S^l(\mathscr{G}) := \sup_{g^2 \in \mathcal{G}^2} \inf_{g^1 \in \mathcal{G}^1} J(g^1, g^2). \tag{8}$$

If the upper and lower values are the same, they are referred to as the value of the game and denoted by $S(\mathscr{G})$.

A Nash equilibrium of the zero-sum game $\mathscr{G}$ is a strategy profile (g^{1*}, g^{2*}) such that for every $g^1 \in \mathcal{G}^1$ and $g^2 \in \mathcal{G}^2$, we have

$$J(g^{1*}, g^2) \le J(g^{1*}, g^{2*}) \le J(g^1, g^{2*}). \tag{9}$$

Nash equilibria in zero-sum games satisfy the following property [29].

Proposition 1 *If a Nash equilibrium in Game $\mathscr{G}$ exists, then for every Nash equilibrium (g^{1*}, g^{2*}) in Game $\mathscr{G}$, we have*

$$J(g^{1*}, g^{2*}) = S^l(\mathscr{G}) = S^u(\mathscr{G}) = S(\mathscr{G}). \tag{10}$$

Remark 2 Note that the existence of a Nash equilibrium is not guaranteed in general. However, if players have perfect recall, i.e.

$$\{U^i_{1:t-1}\} \cup I^i_{t-1} \subseteq I^i_t \tag{11}$$

for every i and t, then the existence of a behavioral strategy equilibrium is guaranteed by Kuhn's theorem [22].

The objective of this work is to characterize the upper and lower values $S^u(\mathscr{G})$ and $S^l(\mathscr{G})$ of Game $\mathscr{G}$. To this end, we will define a virtual game $\mathscr{G}_v$ and an "expanded" virtual game $\mathscr{G}_e$. These virtual games will be used to obtain bounds on the upper and lower values of the original game $\mathscr{G}$.

3 Virtual Game $\mathscr{G}_v$

The virtual game $\mathscr{G}_v$ is constructed using the methodology in [26]. This game involves the same set of primitive random variables as in Game $\mathscr{G}$. The two players of game $\mathscr{G}$ are replaced by two virtual players in $\mathscr{G}_v$. The virtual players operate as follows. At each time t, virtual player i selects a function Γ^i_t that maps private information P^i_t to a distribution δU^i_t over the space $\mathcal{U}^i_t$. We refer to these functions as *prescriptions*. Let $\mathcal{B}^i_t$ be the set of all possible prescriptions for virtual player i at time t (i.e. $\mathcal{B}^i_t$ is the set of all mappings from $\mathcal{P}^i_t$ to $\Delta\mathcal{U}^i_t$).

Once the virtual players select their prescriptions, the action U^i_t is randomly generated according to distribution $\Gamma^i_t(P^i_t)$. More precisely, the system dynamics for this game are given by:

$$X_{t+1} = f_t(X_t, U^{1:2}_t, W^s_t) \tag{12}$$

$$P^i_{t+1} = \xi^i_{t+1}(P^i_t, U^i_t, Y^i_{t+1}) \qquad i = 1, 2, \tag{13}$$

$$Y^i_{t+1} = h^i_{t+1}(X_{t+1}, U^{1:2}_t, W^i_{t+1}) \qquad i = 1, 2, \tag{14}$$

$$U^i_t = \kappa(\Gamma^i_t(P^i_t), K^i_t) \qquad i = 1, 2, \tag{15}$$

$$Z_{t+1} = \zeta_{t+1}(P^{1:2}_t, U^{1:2}_t, Y^{1:2}_{t+1}), \tag{16}$$

where the functions $f_t, h^i_t, \xi^i_t, \kappa$ and ζ_t are the same as in $\mathscr{G}$.

In the virtual game, virtual players use the common information C_t to select their prescriptions at time t. The ith virtual player selects its prescription according to a control law χ^i_t, i.e., $\Gamma^i_t = \chi^i_t(C_t)$. For virtual player i, the collection of control laws over the entire time horizon $\chi^i = (\chi^i_1, \ldots, \chi^i_T)$ is referred to as its control strategy. Let $\mathcal{H}^i_t$ be the set of all possible control laws for virtual player i at time t, and let $\mathcal{H}^i$ be the set of all possible control strategies for virtual player i, i.e., $\mathcal{H}^i = \mathcal{H}^i_1 \times \cdots \times \mathcal{H}^i_T$. The total cost associated with the game for a strategy profile (χ^1, χ^2) is

$$\mathcal{J}(\chi^1, \chi^2) = \mathbb{E}^{(\chi^1, \chi^2)}\left[\sum_{t=1}^{T} c_t(X_t, U^1_t, U^2_t)\right], \tag{17}$$

where the function c_t is the same as in Game $\mathscr{G}$.

The following lemma establishes a connection between the original game $\mathscr{G}$ and the virtual game $\mathscr{G}_v$ constructed above.

Lemma 1 *Let $S^u(\mathscr{G}_v)$ and $S^l(\mathscr{G}_v)$ be, respectively, the upper and lower values of the virtual game $\mathscr{G}_v$. Then,*

$$S^l(\mathscr{G}) = S^l(\mathscr{G}_v) \text{ and } S^u(\mathscr{G}) = S^u(\mathscr{G}_v).$$

Consequently, if a Nash equilibrium exists in the original game $\mathscr{G}$, then $S(\mathscr{G}) = S^l(\mathscr{G}_v) = S^u(\mathscr{G}_v)$.

Proof See "Appendix 1." □

The authors in [26] use the virtual game to find equilibrium costs and strategies for a stochastic dynamic game of asymmetric information. However, the methodology in [26] is applicable *only under the assumption that the posterior beliefs on state X_t and private information $P_t^{1,2}$ given the common information C_t do not depend on the strategy profile being used* (see Assumption 2 in [26]). We will refer to this assumption as the *strategy-independent beliefs* (SIB) assumption. As pointed out in [26], the SIB assumption is satisfied by some special system models and information structures but is not true for general stochastic dynamic games. A simple example which does not satisfy the SIB assumption is the following delayed sharing information structure [27]: Consider game $\mathscr{G}$ with common information $C_t = \{Y^{1,2}_{1:t-2}, U^{1,2}_{1:t-2}\}$ and $P^i_t = \{Y^i_t, Y^i_{t-1}, U^i_{t-1}\}$.

Thus, we are faced with the following situation: If our zero-sum game satisfies the SIB assumption, we can adopt the results in [26] to find equilibrium costs (i.e., the value) of our game. However, if the zero-sum game does not satisfy the SIB assumption, then the methodology of [26] is inapplicable. In the next section, we will develop a methodology to bound the upper and lower values of the zero-sum game $\mathscr{G}$ even when the game does not satisfy the SIB assumption.

4 Expanded Virtual Game $\mathscr{G}_e$ with Prescription History

In order to circumvent the SIB assumption, we now construct an expanded virtual game $\mathscr{G}_e$ by increasing the amount of information available to virtual players in game $\mathscr{G}_v$. In this new game $\mathscr{G}_e$, the state dynamics, observation processes, primitive random variables and cost function are all the same as in the game $\mathscr{G}_v$. The only difference is in the information used by the virtual players to select their prescriptions. The virtual players now have access to the common information C_t *as well as* all the past prescriptions of both players, i.e., $\Gamma^{1:2}_{1:t-1}$. Virtual player i selects its prescription at time t using a control law $\tilde{\chi}^i_t$, i.e., $\Gamma^i_t = \tilde{\chi}^i_t(C_t, \Gamma^{1:2}_{1:t-1})$. Let $\tilde{\mathcal{H}}^i_t$ be the set of all such (measurable) control laws at time t for virtual player i. $\tilde{\mathcal{H}}^i := \tilde{\mathcal{H}}^i_1 \times \cdots \times \tilde{\mathcal{H}}^i_T$ is the set of all control strategies for player i. The total cost associated with the game for a strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ is

$$\mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) = \mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}\left[\sum_{t=1}^{T} c_t(X_t, U^1_t, U^2_t)\right]. \tag{18}$$

Remark 3 Note that any strategy $\chi^i \in \mathcal{H}^i$ is equivalent to the strategy $\tilde{\chi}^i \in \tilde{\mathcal{H}}^i$ that satisfies the following condition: For each time t and for each realization of common information

$c_t \in \mathcal{C}_t$,

$$\tilde{\chi}_t^i(c_t, \gamma_{1:t-1}^{1:2}) = \chi_t^i(c_t) \quad \forall\, \gamma_{1:t-1}^{1:2} \in \mathcal{B}_{1:t-1}^{1:2}. \tag{19}$$

Hence, with slight abuse of notation, we can say that the strategy space $\mathcal{H}^i$ in the virtual game $\mathcal{G}_v$ is a subset of the strategy space $\tilde{\mathcal{H}}^i$ in the expanded game $\mathcal{G}_e$. For this reason, the function $\mathcal{J}$ in (18) can be thought of as an extension of the function $\mathcal{J}$ in (17).

Remark 4 Expansion of information structures has been used in prior work to find equilibrium costs/strategies. See, for example, [5] which studies a linear stochastic differential game where both players have a common noisy observation of the state. Similar virtual games with expanded information structures referred to as *auxiliary games* have also been used in [12,23,32,33,35].

4.1 Upper and Lower Values of Games $\mathcal{G}_v$ and $\mathcal{G}_e$

We will now establish the relationship between the upper and lower values of the expanded game $\mathcal{G}_e$ and the virtual game $\mathcal{G}_v$. To do so, we define the following mappings between the strategies in games $\mathcal{G}_v$ and $\mathcal{G}_e$.

Definition 2 Let $\varrho^i : \tilde{\mathcal{H}}^1 \times \tilde{\mathcal{H}}^2 \to \mathcal{H}^i$ be an operator that maps a strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ in virtual game $\mathcal{G}_e$ to a strategy χ^i for virtual player i in game $\mathcal{G}_v$ as follows: For $t = 1, 2, \ldots, T$,

$$\chi_t^i(c_t) := \tilde{\chi}_t^i(c_t, \tilde{\gamma}_{1:t-1}^{1:2}), \tag{20}$$

where $\tilde{\gamma}_s^j = \tilde{\chi}_s^j(c_s, \tilde{\gamma}_{1:s-1}^{1:2})$ for every $1 \le s \le t-1$ and $j = 1, 2$. We denote the ordered pair (ϱ^1, ϱ^2) by ϱ.

The mapping ϱ is defined in such a way that the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ and the strategy profile $\varrho(\tilde{\chi}^1, \tilde{\chi}^2)$ induce identical dynamics in the respective games $\mathcal{G}_e$ and $\mathcal{G}_v$.

Lemma 2 *Let (χ^1, χ^2) and $(\tilde{\chi}^1, \tilde{\chi}^2)$ be strategy profiles for games $\mathcal{G}_v$ and $\mathcal{G}_e$, such that $\chi^i = \varrho^i(\tilde{\chi}^1, \tilde{\chi}^2)$, $i = 1, 2$. Then,*

$$\mathcal{J}(\chi^1, \chi^2) = \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2). \tag{21}$$

Proof See "Appendix 2." □

The following theorem connects the upper and lower values of the two virtual games and the original game.

Theorem 1 *The lower and upper values of the three games defined above satisfy the following:*

$$S^l(\mathcal{G}) = S^l(\mathcal{G}_v) \le S^l(\mathcal{G}_e) \le S^u(\mathcal{G}_e) \le S^u(\mathcal{G}_v) = S^u(\mathcal{G}).$$

Consequently, if a Nash equilibrium exists in the original game $\mathcal{G}$, then $S(\mathcal{G}) = S^l(\mathcal{G}_e) = S^u(\mathcal{G}_e)$.

Proof See "Appendix 3. □

Using Theorem 1, we can obtain bounds on the upper and lower values of the original game by computing the upper and lower values of the expanded game $\mathcal{G}_e$.

4.2 The Dynamic Programming Characterization

We now describe a methodology for finding the upper and lower values of the expanded game $\mathscr{G}_e$. Suppose the virtual players are using the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ in the expanded game $\mathscr{G}_e$. Let Π_t be the virtual players' belief on the state and private information based on their information in game $\mathscr{G}_e$. Thus, Π_t is defined as

$$\Pi_t(x_t, p_t^{1:2}) := \mathbb{P}^{(\tilde{\chi}^1, \tilde{\chi}^2)}(X_t = x_t, P_t^{1:2} = p_t^{1:2} \mid C_t, \Gamma_{1:t-1}^{1:2}), \ \forall x_t, p_t^1, p_t^2.$$

We refer to Π_t as the *common information belief* (CIB). Π_t takes values in the set $\mathcal{S}_t := \Delta(\mathcal{X}_t \times \mathcal{P}_t^1 \times \mathcal{P}_t^2)$.

Definition 3 Given a belief π on the state information and private information at time t and mappings γ^i, $i = 1, 2$, from $\mathcal{P}_t^i$ to $\Delta\mathcal{U}_t^i$, we define $\gamma^i(p_t^i; u)$ as the probability assigned to action u under the probability distribution $\gamma^i(p_t^i)$. Also, define

$$\tilde{c}_t(\pi, \gamma^1, \gamma^2) := \sum_{x_t, p_t^{1:2}, u_t^{1:2}} c_t(x_t, u_t^1, u_t^2)\pi(x_t, p_t^1, p_t^2)\gamma^1(p_t^1; u_t^1)\gamma^2(p_t^2; u_t^2). \tag{22}$$

$\tilde{c}_t(\pi, \gamma^1, \gamma^2)$ is the expected value of the cost at time t if the state information and private information have π as their probability distribution and γ^1, γ^2 are the prescriptions chosen by the virtual players.

Lemma 3 *For any strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$, the common information-based belief Π_t evolves almost surely as*

$$\Pi_{t+1} = F_t(\Pi_t, \Gamma_t^{1:2}, Z_{t+1}), \quad t \geq 1, \tag{23}$$

where F_t is a fixed transformation that does not depend on the virtual players' strategies. Further, the total expected cost can be expressed as

$$\mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) = \mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}\left[\sum_{t=1}^{T} \tilde{c}_t(\Pi_t, \Gamma_t^1, \Gamma_t^2)\right], \tag{24}$$

where $\tilde{c}_t$ is as defined in Eq. (22).

Proof See "Appendix 4." □

Remark 5 Because (23) is an almost sure equality, the transformation F_t in Lemma 3 is not necessarily unique. In "Appendix 4," we identify a class of transformations such that for any transformation F_t in this class, Lemma 3 holds. We denote this class by $\mathscr{B}$.

We now describe two dynamic programs, one for each virtual player in $\mathscr{G}_e$.

4.2.1 The Min–Max Dynamic Program

The minimizing virtual player (virtual player 1) in game $\mathscr{G}_e$ solves the following dynamic program. Define $V_{T+1}^u(\pi_{T+1}) = 0$ for every π_{T+1}. In a backward-inductive manner, at each time $t \leq T$ and for each possible common information belief π_t and prescriptions γ_t^1, γ_t^2, define

$$w_t^u(\pi_t, \gamma_t^1, \gamma_t^2) := \tilde{c}_t(\pi_t, \gamma_t^1, \gamma_t^2) + \mathbb{E}[V_{t+1}^u(F_t(\pi_t, \gamma_t^{1:2}, Z_{t+1})) \mid \pi_t, \gamma_t^{1:2}] \tag{25}$$

$$V_t^u(\pi_t) := \inf_{\gamma_t^1} \sup_{\gamma_t^2} w_t^u(\pi_t, \gamma_t^1, \gamma_t^2). \tag{26}$$

4.2.2 The Max-Min Dynamic Program

The maximizing virtual player (virtual player 2) in game $\mathscr{G}_e$ solves the following dynamic program. Define $V^l_{T+1}(\pi_{T+1}) = 0$ for every π_{T+1}. In a backward-inductive manner, at each time $t \leq T$ and for each possible common information belief π_t and prescriptions γ_t^1, γ_t^2, define

$$w^l_t(\pi_t, \gamma_t^1, \gamma_t^2) := \tilde{c}_t(\pi_t, \gamma_t^1, \gamma_t^2) + \mathbb{E}[V^l_{t+1}(F_t(\pi_t, \gamma_t^{1:2}, Z_{t+1})) \mid \pi_t, \gamma_t^{1:2}] \tag{27}$$

$$V^l_t(\pi_t) := \sup_{\gamma_t^2} \inf_{\gamma_t^1} w^l_t(\pi_t, \gamma_t^1, \gamma_t^2). \tag{28}$$

Lemma 4 *For any realization of common information-based belief π_t, the* inf *and* sup *in* (26) *are achieved; i.e. there exists a measurable mapping $\Xi^1_t : \mathcal{S}_t \to \mathcal{B}^1_t$ such that*

$$V^u_t(\pi_t) = \min_{\gamma_t^1} \max_{\gamma_t^2} w^u_t(\pi_t, \gamma_t^1, \gamma_t^2) = \max_{\gamma_t^2} w^u_t(\pi_t, \Xi^1_t(\pi_t), \gamma_t^2). \tag{29}$$

Similarly, for any realization of common information-based belief π_t, the sup *and* inf *in* (28) *are achieved, i.e., there exists a measurable mapping $\Xi^2_t : \mathcal{S}_t \to \mathcal{B}^2_t$ such that*

$$V^l_t(\pi_t) = \max_{\gamma_t^2} \min_{\gamma_t^1} w^l_t(\pi_t, \gamma_t^1, \gamma_t^2) = \min_{\gamma_t^1} w^l_t(\pi_t, \gamma_t^1, \Xi^2_t(\pi_t)). \tag{30}$$

Proof See "Appendix 6." □

Definition 4 Define strategies $\tilde{\chi}^{1*}$ and $\tilde{\chi}^{2*}$ for virtual players 1 and 2, respectively, as follows: For each instance of common information c_t and prescription history $\gamma^{1:2}_{1:t-1}$, let

$$\tilde{\chi}^{1*}_t(c_t, \gamma^{1:2}_{1:t-1}) := \Xi^1_t(\pi_t) \tag{31}$$

$$\tilde{\chi}^{2*}_t(c_t, \gamma^{1:2}_{1:t-1}) := \Xi^1_2(\pi_t), \tag{32}$$

where Ξ^1_t and Ξ^2_t are the mappings defined in Lemma 4 and π_t (which is a function of $c_t, \gamma^{1:2}_{1:t-1}$) is obtained in a forward-inductive manner using the relation

$$\pi_1(x_1, p_1^1, p_1^2) = \mathbb{P}[X_1 = x_1, P_1^1 = p_1^1, P_1^2 = p_1^2 \mid C_1 = c_1] \; \forall \, x_1, p_1^1, p_1^2, \tag{33}$$

$$\pi_{\tau+1} = F_\tau(\pi_\tau, \gamma_\tau^1, \gamma_\tau^2, z_{\tau+1}), \; 1 \leq \tau < t. \tag{34}$$

Note that F_τ is the common information belief update function defined in Lemma 3.

The following theorem establishes that the two dynamic programs described above characterize the upper and lower values of game $\mathscr{G}_e$.

Theorem 2 *The upper and lower values of the expanded virtual game $\mathscr{G}_e$ are given by*

$$S^u(\mathscr{G}_e) = \mathbb{E}[V^u_1(\Pi_1)], \tag{35}$$

$$S^l(\mathscr{G}_e) = \mathbb{E}[V^l_1(\Pi_1)]. \tag{36}$$

Further, the strategies $\tilde{\chi}^{1*}$ and $\tilde{\chi}^{2*}$ as defined in Definition 4 are, respectively, min–max and max-min strategies in the expanded virtual game $\mathscr{G}_e$.

Proof See "Appendix 7." □

Theorem 2 gives us a dynamic programming characterization of the upper and lower values of the expanded game. As mentioned in Theorem 1, the upper and lower values of the expanded game provide bounds on the corresponding values of the original game. Further, if the original game has a Nash equilibrium, the dynamic programs of Theorem 2 characterize the value of the game. Note that this applies to any dynamic game of the form in Section 2 where the common information is nondecreasing in time and the private information has a "state-like" update equation (see Assumption 1). As noted before, a variety of information structures satisfy this assumption [26,28].

The computational burden of solving the dynamic programs of Theorem 2 would depend on the specific information structure being considered, i.e., on the exact nature of common and private information. At one extreme, we can consider the following instance of the original game $\mathscr{G}$: $C_t = (X_{1:t})$, $P_t^1 = P_t^2 = \emptyset$. It is easy to see that in this case, the common information belief can be replaced by the current state in the dynamic programs and the prescriptions are simply distributions on the players' finite action sets. Also, in this case, w_t^u and w_t^l are bilinear functions of the prescriptions and the min–max/max–min problems at each stage of the dynamic program can be solved by a linear program [31]. On the other extreme, we can consider an instance of game $\mathscr{G}$ with $C_t = \emptyset$, $P_t^i = Y_{1:t}^i$, $i = 1, 2$. In this case, the common information belief will be on the current state and observation histories of the two players and the prescriptions will take values in a large-dimensional space. Also, the functions w_t^u and w_t^l (for $t < T$) in this case do not have any apparent structure that can be exploited for efficient computation of the min–max and max-min values in the dynamic program. One general approach that can be used for any instance of game $\mathscr{G}$ is to discretize the CIB belief space and compute approximate value functions V_t^u and V_t^l in a backward-inductive manner. However, we believe that significant structural and computational insights can be obtained by specializing the dynamic programs of Theorem 2 to the specific instance of the game being considered. We demonstrate this in [18] where we consider an information structure in which one player has complete information while the other player has only partial information. For this information structure, it is shown in [18] that the functions w_t^u and w_t^l turn out to be identical at all times t and they satisfy some structural properties that can be leveraged for computation.

Comparison with [30] *and* [40]: In [30], the authors considered an n-player stochastic game model which can potentially be nonzero sum. In this model, each player has a private state that is privately observed by the corresponding player and a public state that is commonly observed by all the players. The model in [30] additionally allows players' private information to be partially revealed in the form of common observations. The actions of all the players in this model are commonly observed. The authors also make the assumption that the evolution of the private states of the players is conditionally independent. The model in [40] can be viewed as a special case of the model in [30]. For these models, backward-inductive algorithms were presented to compute perfect Bayesian equilibria. Consider the case when the number of players in the games of [30] and [40] is two and the games are zero-sum. Then:

1. The models in [30] and [40] can be viewed as special cases of our model in Sect. 2.
2. The players in these games have perfect recall. Hence, we can use Kuhn's theorem to conclude that a Nash equilibrium and, thus, the value exist for these zero-sum games. Therefore, we can use the dynamic programs in Sect. 4.2 to the characterize the value of these zero-sum games. This characterization does not make any additional assumptions. The backward-inductive algorithms in [30] and [40], however, require the existence of a particular kind of *fixed point* solution at each stage. This fixed-point solution is not guaranteed to exist in general. Thus, there may be instances where the approaches in

[30] and [40] fail to characterize the value of the game, while our dynamic program in Sect. 4.2 can always characterize it.

4.3 Connections with the Recursive Formula in [23]

A general model of zero-sum games is described in Section IV.3 of [23]. The state dynamics and the observation model in [23, Section IV.3] are similar to the state dynamics (Eq. (1)) and observation model (Eq. (2)) we described in Sect. 2. We would like to highlight two key differences between our model and the model in [23, Section IV.3]. Firstly, the model in [23, Section IV.3] assumes that the players have perfect recall, whereas we do not make this assumption. In our setup, players may or may not have perfect recall. We only require the common information to be recalled perfectly by the players (see Assumption 1). Secondly, in the model of [23, Section IV.3], the new information (or the signal) a player gets at each stage can be seen as a function of the current and previous states, the actions taken at previous stage and some random noise. In our model, the new information that a player gets at each stage includes the new private information and an increment in common information that could depend on the history of past states and actions (through the term P_t^i in equations (3) and (4))[2]. In order to accommodate this feature in the state dynamics and observation model of [23, Section IV.3], one needs to redefine the state in [23, Section IV.3] to consist of both the original system state and the state of players' private information.

Since the players in the model of [23, Section IV.3] have perfect recall, the value of the game exists. A characterization of this value is provided in [23] using a recursive formula. The recursive formula in [23] is in terms of certain probability distributions $\mathbf{P}_t$ and has the following structure[3]:

$$V_{T+1}(\mathbf{P}_{T+1}) = 0 \tag{37}$$

$$V_t(\mathbf{P}_t) = \min_{g_t^1} \max_{g_t^2} \left(\mathbb{E}^{(g_t^1, g_t^2)} \left[c_t(X_t, U_t^1, U_t^2) \right] + V_{t+1}(\mathbf{P}_{t+1}) \right). \tag{38}$$

Here, g_t^i is player i's behavioral strategy at time t. While this recursive formula appears to be similar to our dynamic programming characterization, it has some key differences which are listed below:

1. The minimization and maximization in the recursive formula are over behavioral strategies at each time. Note that the behavioral strategy at each time is a mapping from a player's entire information at that time to probability distributions over its action sets. In our dynamic program, the minimization and maximization are over prescription spaces $\mathcal{B}_t^1$ and $\mathcal{B}_t^2$, respectively. Recall that a prescription is a mapping from a player's *private* information to probability distributions over its action sets. Since a player's private information may be much smaller than the entire information available to it, the prescription spaces in our dynamic program may be much smaller than behavioral strategy spaces. This conceptual difference in the two characterizations may also have computational implications since minimizing/maximizing over the smaller space of prescriptions would generally be easier than minimizing/maximizing over the larger space of behavioral strategies at each time.
2. The *information state* in the recursive formula (i.e., the arguments of the value functions above) is a probability distribution on an abstract space referred to as the universal belief

[2] For example, see the delayed sharing information structure in Sect. 2.2.

[3] Our notation is different from [23, Section IV.3].

space[4]. The information state in our dynamic program is the common information belief, which is more tangible and is explicitly defined as a posterior probability distribution on the current state and players' private information based on the common information. We believe that this explicit description of the information state as a common belief is both conceptually more illuminating and computationally more useful.

5 Conclusion

In this paper, we considered a general model of zero-sum stochastic games with asymmetric information. We model each player's information as consisting of a private information part and a common information part. In our model, players need not have perfect recall. We only need the common information to be perfectly recalled. For this general model, we provided a dynamic programming approach for characterizing the value (if it exists). This dynamic programming characterization of value relies on our construction of two virtual games that have the same value as our original game. If the value does not exist in the original game, then our dynamic program provides bounds on the upper and lower values of the original game.

Proof of Lemma 1

It was shown in [26] that there exist *bijective* mappings $\mathcal{M}^i : \mathcal{G}^i \to \mathcal{H}^i$, $i = 1, 2$, such that for every $g^1 \in \mathcal{G}^1$ and $g^2 \in \mathcal{G}^2$, we have

$$J(g^1, g^2) = \mathcal{J}(\mathcal{M}^1(g^1), \mathcal{M}^2(g^2)). \tag{39}$$

Therefore, for any strategy $g^1 \in \mathcal{G}^1$, we have

$$\sup_{g^2 \in \mathcal{G}^2} J(g^1, g^2) = \sup_{g^2 \in \mathcal{G}^2} \mathcal{J}(\mathcal{M}^1(g^1), \mathcal{M}^2(g^2)) \tag{40}$$

$$= \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\mathcal{M}^1(g^1), \chi^2). \tag{41}$$

Consequently,

$$\inf_{g^1 \in \mathcal{G}^1} \sup_{g^2 \in \mathcal{G}^2} J(g^1, g^2) = \inf_{g^1 \in \mathcal{G}^1} \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\mathcal{M}^1(g^1), \chi^2) \tag{42}$$

$$= \inf_{\chi^1 \in \mathcal{H}^1} \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\chi^1, \chi^2). \tag{43}$$

This implies that $S^u(\mathscr{G}) = S^u(\mathscr{G}_v)$. We can similarly prove that $S^l(\mathscr{G}) = S^l(\mathscr{G}_v)$.

Remark 6 We can also show that a strategy profile (g^1, g^2) is a Nash equilibrium in game $\mathscr{G}$ if and only if $(\mathcal{M}^1(g^1), \mathcal{M}^2(g^2))$ is a Nash equilibrium in game $\mathscr{G}_v$.

Proof of Lemma 2

Let us consider the evolution of the virtual game $\mathscr{G}_v$ under the strategy profile (χ^1, χ^2) and the expanded virtual game $\mathscr{G}_e$ under the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$. Let the primitive variables

[4] We refer the reader to [23, Chapter III] for a detailed discussion on the universal belief space.

and the randomization variables K_t^i in both games be identical. The variables such as the state, action and information variables in the expanded game $\mathscr{G}_e$ are distinguished from those in the virtual game $\mathscr{G}_v$ by means of a tilde. For instance, X_t is the state in game $\mathscr{G}_v$ and $\tilde{X}_t$ is the state in game $\mathscr{G}_e$.

We will prove by induction that the system evolution in both these games is identical over the entire horizon. This is clearly true at the end of time $t = 1$ because the state, observations and the common and private information variables are identical in both games. Moreover, since $\chi^i = \varrho^i(\tilde{\chi}^1, \tilde{\chi}^2)$, $i = 1, 2$, the strategies χ_1^i and $\tilde{\chi}_1^i$ are identical by definition (see Definition 2). Thus, the prescriptions and actions at $t = 1$ are also identical.

For induction, assume that the system evolution in both games is identical until the end of time t. Then,

$$X_{t+1} = f_t(X_t, U_t^{1:2}, W_t^s) = f_t(\tilde{X}_t, \tilde{U}_t^{1:2}, W_t^s) = \tilde{X}_{t+1}.$$

Using Eqs. (2), (4) and (3), we can similarly argue that $Y_{t+1}^i = \tilde{Y}_{t+1}^i$, $P_{t+1}^i = \tilde{P}_{t+1}^i$ and $C_{t+1} = \tilde{C}_{t+1}$. Since $\chi^i = \varrho^i(\tilde{\chi}^1, \tilde{\chi}^2)$, we also have

$$\tilde{\Gamma}_{t+1}^i = \tilde{\chi}_{t+1}^i(\tilde{C}_{t+1}, \tilde{\Gamma}_{1:t}^{1:2}) \overset{a}{=} \chi_{t+1}^i(\tilde{C}_{t+1}) \overset{b}{=} \Gamma_{t+1}^i. \tag{44}$$

Here, equality (a) follows from the construction of the mapping ϱ^i (see Definition 2) and equality (b) follows from the fact that $C_{t+1} = \tilde{C}_{t+1}$. Further,

$$U_{t+1}^i = \kappa(\Gamma_{t+1}^i(P_{t+1}^i), K_{t+1}^i) = \kappa(\tilde{\Gamma}_{t+1}^i(\tilde{P}_{t+1}^i), K_{t+1}^i) \tag{45}$$

$$= \tilde{U}_{t+1}^i. \tag{46}$$

Thus, by induction, the hypothesis is true for every $1 \le t \le T$. This proves that the virtual and expanded games have identical dynamics under strategy profiles (χ^1, χ^2) and $(\tilde{\chi}^1, \tilde{\chi}^2)$.

Since the virtual and expanded games have the same cost structure, having identical dynamics ensures that strategy profiles (χ^1, χ^2) and $(\tilde{\chi}^1, \tilde{\chi}^2)$ have the same expected cost in games $\mathscr{G}_v$ and $\mathscr{G}_e$, respectively. Therefore, $\mathcal{J}(\chi^1, \chi^2) = \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2)$.

Proof of Theorem 1

For any strategy $\chi^1 \in \mathcal{H}^1$, we have

$$\sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\chi^1, \tilde{\chi}^2) \ge \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\chi^1, \chi^2), \tag{47}$$

because $\mathcal{H}^2 \subseteq \tilde{\mathcal{H}}^2$. Further,

$$\sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\chi^1, \tilde{\chi}^2) = \sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\varrho^1(\chi^1, \tilde{\chi}^2), \varrho^2(\chi^1, \tilde{\chi}^2)). \tag{48}$$

$$= \sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\chi^1, \varrho^2(\chi^1, \tilde{\chi}^2)) \tag{49}$$

$$\le \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\chi^1, \chi^2), \tag{50}$$

where the first equality is due to Lemma 2, the second equality is because $\varrho^1(\chi^1, \tilde{\chi}^2) = \chi^1$ and the last inequality is due to the fact that $\varrho^2(\chi^1, \tilde{\chi}^2) \in \mathcal{H}^2$ for any $\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2$.

Combining (47) and (50), we obtain that

$$\sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\chi^1, \chi^2) = \sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\chi^1, \tilde{\chi}^2). \tag{51}$$

Now,

$$\begin{aligned} S^u(\mathcal{G}_e) &:= \inf_{\tilde{\chi}^1 \in \tilde{\mathcal{H}}^1} \sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) && (52)\\ &\leq \inf_{\chi^1 \in \mathcal{H}^1} \sup_{\tilde{\chi}^2 \in \tilde{\mathcal{H}}^2} \mathcal{J}(\chi^1, \tilde{\chi}^2) && (53)\\ &= \inf_{\chi^1 \in \mathcal{H}^1} \sup_{\chi^2 \in \mathcal{H}^2} \mathcal{J}(\chi^1, \chi^2), && (54)\\ &=: S^u(\mathcal{G}_v). && (55) \end{aligned}$$

where inequality (53) is true since $\mathcal{H}^1 \subseteq \tilde{\mathcal{H}}^1$ and the equality in (54) follows from (51). Therefore, $S^u(\mathcal{G}_e) \leq S^u(\mathcal{G}_v)$. We can use similar arguments to show that $S^l(\mathcal{G}_v) \leq S^l(\mathcal{G}_e)$.

Proof of Lemma 3

We begin with defining the following transformations for each time t. Recall that $\mathcal{S}_t$ is the set of all possible common information beliefs at time t and $\mathcal{B}_t^i$ is the prescription space for virtual player i at time t.

Definition 5 (i) Let $P_t^j : \mathcal{S}_t \times \mathcal{B}_t^1 \times \mathcal{B}_t^2 \to \Delta(\mathcal{Z}_{t+1} \times \mathcal{X}_{t+1} \times \mathcal{P}_{t+1}^1 \times \mathcal{P}_{t+1}^2)$ be defined as

$$\begin{aligned} &P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, x_{t+1}, p_{t+1}^{1:2}) && (56)\\ &:= \sum_{x_t, p_t^{1:2}, u_t^{1:2}} \pi_t(x_t, p_t^{1:2}) \gamma_t^1(p_t^1; u_t^1) \gamma_t^2(p_t^2; u_t^2) \mathbb{P}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid x_t, p_t^{1:2}, u_t^{1:2}]. && (57) \end{aligned}$$

We will use $P_t^j(\pi_t, \gamma_t^{1:2})$ as a shorthand for the probability distribution $P_t^j(\pi_t, \gamma_t^{1:2}; \cdot, \cdot, \cdot)$. The distribution $P_t^j(\pi_t, \gamma_t^{1:2})$ can be viewed as a joint distribution over the variables $Z_{t+1}, X_{t+1}, P_{t+1}^{1:2}$ if the distribution on $X_t, P_t^{1:2}$ is π_t and prescriptions $\gamma_t^{1:2}$ are chosen by the virtual players at time t.

(ii) Let $P_t^m : \mathcal{S}_t \times \mathcal{B}_t^1 \times \mathcal{B}_t^2 \to \Delta \mathcal{Z}_{t+1}$ be defined as

$$P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1}) = \sum_{x_{t+1}, p_{t+1}^{1:2}} P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, x_{t+1}, p_{t+1}^{1:2}). \tag{58}$$

The distribution $P_t^m(\pi_t, \gamma_t^{1:2})$ is the marginal distribution of the variable Z_{t+1} obtained from the joint distribution $P_t^j(\pi_t, \gamma_t^{1:2})$ defined above.

(iii) Let $F_t : \mathcal{S}_t \times \mathcal{B}_t^1 \times \mathcal{B}_t^2 \times \mathcal{Z}_{t+1} \to \mathcal{S}_{t+1}$ be defined as

$$F_t(\pi_t, \gamma_t^{1:2}, z_{t+1}) = \begin{cases} \frac{P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, \cdot, \cdot)}{P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1})} & \text{if } P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1}) > 0 \\ G_t(\pi_t, \gamma_t^{1:2}, z_{t+1}) & \text{otherwise,} \end{cases} \tag{59}$$

where $G_t : \mathcal{S}_t \times \mathcal{B}_t^1 \times \mathcal{B}_t^2 \times \mathcal{Z}_{t+1} \to \mathcal{S}_{t+1}$ can be any arbitrary measurable mapping. One such mapping is the one that maps every element $\pi_t, \gamma_t^{1:2}, z_{t+1}$ to the uniform distribution over the finite space $\mathcal{X}_{t+1} \times \mathcal{P}_{t+1}^1 \times \mathcal{P}_{t+1}^2$.

Let the collection of transformations F_t that can be constructed using the method described in Definition 5 be denoted by $\mathscr{B}$. Note that the transformations P_t^j, P_t^m and F_t do not depend on the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ because the term $\mathbb{P}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid x_t, p_t^{1:2}, u_t^{1:2}]$ in (57) depends only on the system dynamics (see Eqs. (12–16)) and not on the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$.

Consider a strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$. Note that the number of possible realizations of common information and prescription history under $(\tilde{\chi}^1, \tilde{\chi}^2)$ is finite. Let $c_{t+1}, \gamma_{1:t}^{1:2}$ be a realization of the common information and prescription history at time $t+1$ with nonzero probability of occurrence under $(\tilde{\chi}^1, \tilde{\chi}^2)$. For this realization of virtual players' information, the common information-based belief on the state and private information at time $t+1$ is given by

$$\begin{aligned}
&\pi_{t+1}(x_{t+1}, p_{t+1}^{1:2}) \\
&= \mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[X_{t+1} = x_{t+1}, P_{t+1}^{1:2} = p_{t+1}^{1:2} \mid c_{t+1}, \gamma_{1:t}^{1:2}] \\
&= \mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[X_{t+1} = x_{t+1}, P_{t+1}^{1:2} = p_{t+1}^{1:2} \mid c_t, \gamma_{1:t-1}^{1:2}, z_{t+1}, \gamma_t^{1:2}] \\
&= \frac{\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[X_{t+1} = x_{t+1}, P_{t+1}^{1:2} = p_{t+1}^{1:2}, Z_{t+1} = z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}]}{\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[Z_{t+1} = z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}]}.
\end{aligned} \tag{60}$$

Notice that expression (60) is well defined; that is, the denominator is nonzero because of our assumption that the realization $c_{t+1}, \gamma_{1:t}^{1:2}$ has nonzero probability of occurrence. Let us consider the numerator in expression (60). For convenience, we will denote it with $\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}]$. We have

$$\begin{aligned}
&\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}] \\
&= \sum_{x_t, p_t^{1:2}, u_t^{1:2}} \pi_t(x_t, p_t^{1:2}) \gamma_t^1(p_t^1; u_t^1) \gamma_t^2(p_t^2; u_t^2) \mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}, x_t, p_t^{1:2}, u_t^{1:2}]
\end{aligned} \tag{61}$$

$$= \sum_{x_t, p_t^{1:2}, u_t^{1:2}} \pi_t(x_t, p_t^{1:2}) \gamma_t^1(p_t^1; u_t^1) \gamma_t^2(p_t^2; u_t^2) \mathbb{P}[x_{t+1}, p_{t+1}^{1:2}, z_{t+1} \mid x_t, p_t^{1:2}, u_t^{1:2}] \tag{62}$$

$$= P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, x_{t+1}, p_{t+1}^{1:2}), \tag{63}$$

where π_t is the common information belief on X_t, P_t^1, P_t^2 at time t given the realization[5] $c_t, \gamma_{1:t-1}^{1:2}$ and P_t^j is as defined in Definition 5. The equality in (62) is due to the structure of the system dynamics in game $\mathscr{G}_e$ described by Eqs. (12–16). Similarly, the denominator in (60) satisfies

$$\begin{aligned}
0 < \mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[z_{t+1} \mid c_t, \gamma_{1:t}^{1:2}] &= \sum_{x_{t+1}, p_{t+1}^{1:2}} P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, x_{t+1}, p_{t+1}^{1:2}) \\
&= P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1}),
\end{aligned} \tag{64}$$

where P_t^m is as defined is Definition 5. Thus, from Eq. (60), we have

$$\pi_{t+1} = \frac{P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, \cdot, \cdot)}{P_t^m(\pi_t, \gamma_t^{1:2}, z_{t+1})} = F_t(\pi_t, \gamma_t^{1:2}; z_{t+1}), \tag{65}$$

[5] Note that the belief $\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[x_t, p_t^{1:2} \mid c_t, \gamma_{1:t-1}^{1:2}] = \mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[x_t, p_t^{1:2} \mid c_t, \gamma_{1:t}^{1:2}]$ because $\gamma_t^i = \tilde{\chi}_t^i(c_t, \gamma_{1:t-1}^{1:2}), i = 1, 2$.

where F_t is as defined in Definition 5. Since relation (65) holds for every realization $c_{t+1}, \gamma_{1:t}^{1:2}$ that has nonzero probability of occurrence under $(\tilde{\chi}^1, \tilde{\chi}^2)$, we can conclude that the common information belief Π_t evolves *almost surely* as

$$\Pi_{t+1} = F_t(\Pi_t, \Gamma_t^{1:2}, Z_{t+1}), \quad t \geq 1, \tag{66}$$

under the strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$.

The expected cost at time t can be expressed as follows

$$\mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}[c_t(X_t, U_t^1, U_t^2)] = \mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}[\mathbb{E}[c_t(X_t, U_t^1, U_t^2) \mid C_t, \Gamma_{1:t}^{1:2}]] \tag{67}$$

$$= \mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}[\tilde{c}_t(\Pi_t, \Gamma_t^1, \Gamma_t^2)], \tag{68}$$

where the function $\tilde{c}_t$ is as defined in Eq. (22). Therefore, the total cost can be expressed as

$$\mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}\left[\sum_{t=1}^{T} c_t(X_t, U_t^1, U_t^2)\right] = \mathbb{E}^{(\tilde{\chi}^1, \tilde{\chi}^2)}\left[\sum_{t=1}^{T} \tilde{c}_t(\Pi_t, \Gamma_t^1, \Gamma_t^2)\right]. \tag{69}$$

Some Continuity Results

In this section, we will state and prove some technical results that will be useful for proving Lemma 4.

Let $\mathcal{S}_t$ denote the set of all probability distributions over the finite set $\mathcal{X}_t \times \mathcal{P}_t^1 \times \mathcal{P}_t^2$. Thus, $\mathcal{S}_t$ is the set of all possible common information-based beliefs at time t. Define

$$\bar{\mathcal{S}}_t := \{\alpha \pi_t : 0 \leq \alpha \leq 1, \pi_t \in \mathcal{S}_t\}. \tag{70}$$

The functions $\tilde{c}_t$ in (22), P_t^j in (56), P_t^m in (58) and F_t in (65) were defined for any $\pi_t \in \mathcal{S}_t$. We will extend the domain of the argument π_t in these functions to $\bar{\mathcal{S}}_t$ as follows. For any $\gamma_t^i \in \mathcal{B}_t^i, i = 1, 2, z_{t+1} \in \mathcal{Z}_{t+1}, 0 \leq \alpha \leq 1$ and $\pi_t \in \mathcal{S}_t$, let

(i) $\tilde{c}_t(\alpha \pi_t, \gamma_t^1, \gamma_t^2) := \alpha \tilde{c}_t(\pi_t, \gamma_t^1, \gamma_t^2)$

(ii) $P_t^j(\alpha \pi_t, \gamma_t^{1:2}) := \alpha P_t^j(\pi_t, \gamma_t^{1:2})$

(iii) $P_t^m(\alpha \pi_t, \gamma_t^{1:2}) := \alpha P_t^m(\pi_t, \gamma_t^{1:2})$

(iv) $F_t(\alpha \pi_t, \gamma_t^{1:2}, z_{t+1}) := \begin{cases} F_t(\pi_t, \gamma_t^{1:2}, z_{t+1}) & \text{if } \alpha > 0 \\ \mathbf{0} & \text{if } \alpha = 0, \end{cases}$

where $\mathbf{0}$ is a zero vector of size $|\mathcal{X}_t \times \mathcal{P}_t^1 \times \mathcal{P}_t^2|$.

Having extended the domain of the above functions, we can also extend the domain of the argument π_t in the functions $w_t^u(\cdot), w_t^l(\cdot), V_t^u(\cdot), V_t^l(\cdot)$ defined in the dynamic programs of Section 4.2. First, for any $0 \leq \alpha \leq 1$ and $\pi_{T+1} \in \mathcal{S}_{T+1}$, define $V_{T+1}^u(\alpha \pi_{T+1}) := 0$. We can then define the following functions for every $t \leq T$ in a backward-inductive manner: For any $\gamma_t^i \in \mathcal{B}_t^i, i = 1, 2, \ 0 \leq \alpha \leq 1$ and $\pi_t \in \mathcal{S}_t$, let

$$w_t^u(\alpha \pi_t, \gamma_t^1, \gamma_t^2) := \tilde{c}_t(\alpha \pi_t, \gamma_t^1, \gamma_t^2) + \sum_{z_{t+1}} \left[P_t^m(\alpha \pi_t, \gamma_t^{1:2}; z_{t+1}) V_{t+1}^u(F_t(\alpha \pi_t, \gamma_t^{1:2}, z_{t+1}))\right] \tag{71}$$

$$V_t^u(\alpha \pi_t) := \inf_{\gamma_t^1} \sup_{\gamma_t^2} w_t^u(\alpha \pi_t, \gamma_t^1, \gamma_t^2). \tag{72}$$

Note that when $\alpha = 1$, the above definition of w_t^u is equal to the definition of w_t^u in Eq. (26) of the dynamic program. We can similarly extend w_t^l and V_t^l. These extended value functions satisfy the following homogeneity property. A similar result was shown in [19, Lemma III.1] for a special case of our model.

Lemma 5 *For any constant $0 \leq \alpha \leq 1$ and any $\pi_t \in \bar{\mathcal{S}}_t$, we have $\alpha V_t^u(\pi_t) = V_t^u(\alpha\pi_t)$ and $\alpha V_t^l(\pi_t) = V_t^l(\alpha\pi_t)$.*

Proof The proof can be easily obtained from the above definitions of the extended functions.

The following lemmas will be used in "Appendix 6" to establish some useful properties of the extended functions.

Lemma 6 *Let $V : \bar{\mathcal{S}}_{t+1} \to \mathbb{R}$ be a continuous function satisfying $V(\alpha\pi) = \alpha V(\pi)$ for every $0 \leq \alpha \leq 1$ and $\pi \in \bar{\mathcal{S}}_{t+1}$. Define*

$$V'(\pi_t, \gamma_t^1, \gamma_t^2) := \sum_{z_{t+1}} P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1})[V(F_t(\pi_t, \gamma_t^{1:2}, z_{t+1}))].$$

For a fixed γ_t^1, γ_t^2, $V'(\cdot, \gamma_t^1, \gamma_t^2)$ is a function from $\bar{\mathcal{S}}_{t+1}$ to $\mathbb{R}$. Then, the family of functions

$$\mathscr{F}_1 := \{V'(\cdot, \gamma_t^1, \gamma_t^2) : \gamma_t^i \in \mathcal{B}_t^i, i = 1, 2\} \tag{73}$$

is equicontinuous. Similarly, the following families of functions

$$\mathscr{F}_2 := \{V'(\pi_t, \cdot, \gamma_t^2) : \gamma_t^2 \in \mathcal{B}_t^2, \pi_t \in \bar{\mathcal{S}}_t\} \tag{74}$$

$$\mathscr{F}_3 := \{V'(\pi_t, \gamma_t^1, \cdot) : \gamma_t^1 \in \mathcal{B}_t^1, \pi_t \in \bar{\mathcal{S}}_t\} \tag{75}$$

are equicontinuous in their respective arguments.

Proof A continuous function is bounded and uniformly continuous over a compact domain (see Theorem 4.19 in [36]). Therefore, V is bounded and uniformly continuous over $\bar{\mathcal{S}}_{t+1}$.

Using the fact that $V(\alpha\pi) = \alpha V(\pi)$ and the definition of F_t in Definition 5, the function V' can be written as

$$V'(\pi_t, \gamma_t^1, \gamma_t^2) = \sum_{z_{t+1}} V\left(P_t^j(\pi_t, \gamma_t^{1:2}; z_{t+1}, \cdot, \cdot)\right). \tag{76}$$

Recall that P_t^j is trilinear in π_t, γ_t^1 and γ_t^2 with bounded coefficients for a fixed value of z_{t+1} (see (56)). Therefore, for each z_{t+1}, $\{P_t^j(\cdot, \gamma_t^1, \gamma_t^2, z_{t+1})\}$ is an equicontinuous family of functions in the argument π_t, where $P_t^j(\pi_t, \gamma_t^1, \gamma_t^2, z_{t+1})$ is a shorthand notation for the measure $P_t^j(\pi_t, \gamma_t^1, \gamma_t^2, z_{t+1}, \cdot, \cdot)$ over the space $\mathcal{X}_{t+1} \times \mathcal{P}_{t+1}^1 \times \mathcal{P}_{t+1}^2$.

Also, since V is uniformly continuous, the family $\left\{V\left(P_t^j(\cdot, \gamma_t^{1:2}, z_{t+1})\right)\right\}$ is equicontinuous in π_t for each z_{t+1}. This is because composition with a uniformly continuous function preserves equicontinuity. Therefore, the family of functions $\mathscr{F}_1$ is equicontinuous in π_t. We can use similar arguments to prove equicontinuity of the other two families. □

Lemma 7 *Let $w : \mathcal{B}_t^1 \times \mathcal{B}_t^2 \to \mathbb{R}$ be a function such that (i) the family of functions $\{w(\cdot, \gamma^2) : \gamma^2 \in \mathcal{B}_t^2\}$ is equicontinuous in the first argument and (ii) the family of functions $\{w(\gamma^1, \cdot) : \gamma^1 \in \mathcal{B}_t^1\}$ is equicontinuous in the second argument. Then, $\sup_{\gamma^2} w(\gamma^1, \gamma^2)$ is a continuous function of γ^1 and, similarly, $\inf_{\gamma^1} w(\gamma^1, \gamma^2)$ is a continuous function of γ^2.*

Proof Let $\epsilon > 0$. For a given γ^1, there exists a $\delta > 0$ such that

$$|w(\gamma^1, \gamma^2) - w(\gamma'^1, \gamma^2)| \leq \epsilon \quad \forall \gamma^2, \forall ||\gamma^1 - \gamma'^1|| \leq \delta. \tag{77}$$

Let $\bar{\gamma}^2$ be a prescription such that

$$w(\gamma^1, \bar{\gamma}^2) = \sup_{\gamma^2} w(\gamma^1, \gamma^2). \tag{78}$$

Note that the existence of $\bar{\gamma}^2$ is guaranteed because of continuity of $w(\gamma^1, \cdot)$ in the second argument and compactness of $\mathcal{B}_t^2$. Pick any γ'^1 satisfying $||\gamma^1 - \gamma'^1|| \leq \delta$. Let $\bar{\gamma}'^2$ be a prescription such that

$$w(\gamma'^1, \bar{\gamma}'^2) = \sup_{\gamma^2} w(\gamma'^1, \gamma^2). \tag{79}$$

Using (77), we have

$$\begin{aligned} (i)\ \ w(\gamma^1, \bar{\gamma}^2) - w(\gamma'^1, \bar{\gamma}'^2) &\geq w(\gamma^1, \bar{\gamma}'^2) - w(\gamma'^1, \bar{\gamma}'^2) \\ &\geq -\epsilon, \end{aligned} \tag{80}$$

$$\begin{aligned} (ii)\ \ w(\gamma^1, \bar{\gamma}^2) - w(\gamma'^1, \bar{\gamma}'^2) &\leq w(\gamma^1, \bar{\gamma}^2) - w(\gamma'^1, \bar{\gamma}^2) \\ &\leq \epsilon. \end{aligned} \tag{81}$$

Equations (78)–(81) imply that $\sup_{\gamma^2} w(\gamma^1, \gamma^2)$ is a continuous function of γ^1. We can use a similar argument for showing continuity of $\inf_{\gamma^1} w(\gamma^1, \gamma^2)$ in γ^2. □

Proof of Lemma 4

We first use the definitions of extensions of w_t^u, w_t^l, V_t^u, V_t^l in "Appendix 5" and Lemmas 5 and 6 to establish the following equicontinuity result.

Lemma 8 *The families of functions*

$$\mathscr{F}_t^a := \{w_t^u(\cdot, \gamma_t^1, \gamma_t^2) : \gamma_t^i \in \mathcal{B}_t^i, i = 1, 2\} \tag{82}$$

$$\mathscr{F}_t^b := \{w_t^u(\pi_t, \cdot, \gamma_t^2) : \gamma_t^2 \in \mathcal{B}_t^2, \pi_t \in \bar{\mathcal{S}}_t\} \tag{83}$$

$$\mathscr{F}_t^c := \{w_t^u(\pi_t, \gamma_t^1, \cdot) : \gamma_t^1 \in \mathcal{B}_t^1, \pi_t \in \bar{\mathcal{S}}_t\} \tag{84}$$

are all equicontinuous in their arguments for every $t \leq T$. A similar statement holds for w_t^l.

Proof We use a backward-induction argument for the proof. For induction, assume that V_{t+1}^u is a continuous function for some $t \leq T$. This is clearly true for $t = T$. Using the continuity of V_{t+1}^u, we will establish the statement of the lemma for time t and also prove the continuity of V_t^u. This establishes the lemma for all $t \leq T$.

Equicontinuity of w_t^u: Since $\tilde{c}_t(\pi_t, \gamma_t^1, \gamma_t^2)$ is linear in π_t with uniformly bounded coefficients for any given $\gamma_t^{1:2}$ (see (22)), it is equicontinuous in the argument π_t. In Lemma 5, we showed that the value functions V_t^u satisfy the condition $V_t^u(\alpha\pi) = \alpha V_t^u(\pi)$ for every $0 \leq \alpha \leq 1, \pi \in \mathcal{S}_t$. Further, due to our induction hypothesis, V_{t+1}^u is continuous. Thus, using Lemma 6, the second term of w_t^u,

$$\sum_{z_{t+1}} P_t^m(\pi_t, \gamma_t^{1:2}; z_{t+1}) V_{t+1}^u(F_t(\pi_t, \gamma_t^{1:2}, z_{t+1})),$$

is also equicontinuous in π_t. Hence, the family $\mathscr{F}_t^a$ is equicontinuous in π_t.

Continuity of V_t^u: Due to the equicontinuity of the family $\mathscr{F}_t^a$, we have the following. For any given $\epsilon > 0$ and $\pi_t \in \bar{\mathcal{S}}_t$, there exists a $\delta > 0$ such that

$$|w_t^u(\pi_t, \gamma_t^1, \gamma_t^2) - w_t^u(\pi_t', \gamma_t^1, \gamma_t^2)| < \epsilon \tag{85}$$

for every γ_t^1, γ_t^2 and π_t' satisfying $||\pi_t - \pi_t'|| < \delta$. Therefore,

$$w_t^u(\pi_t, \gamma_t^1, \gamma_t^2) < w_t^u(\pi_t', \gamma_t^1, \gamma_t^2) + \epsilon \ \forall \gamma_t^1, \gamma_t^2 \tag{86}$$

$$\implies \sup_{\gamma_t^2} w_t^u(\pi_t, \gamma_t^1, \gamma_t^2) \leq \sup_{\gamma_t^2} w_t^u(\pi_t', \gamma_t^1, \gamma_t^2) + \epsilon \ \forall \gamma_t^1 \tag{87}$$

$$\implies \inf_{\gamma_t^1} \sup_{\gamma_t^2} w_t^u(\pi_t, \gamma_t^1, \gamma_t^2) \leq \inf_{\gamma_t^1} \sup_{\gamma_t^2} w_t^u(\pi_t', \gamma_t^1, \gamma_t^2) + \epsilon$$

$$\implies V_t^u(\pi_t) \leq V_t^u(\pi_t') + \epsilon, \tag{88}$$

for every π_t' that satisfies $||\pi_t - \pi_t'|| < \delta$. Similarly, $V_t^u(\pi_t) \geq V_t^u(\pi_t') - \epsilon$ for every π_t' that satisfies $||\pi_t - \pi_t'|| < \delta$. Therefore, $V_t^u(\pi_t)$ is continuous at π_t.

Hence, by induction, we can say that the family $\mathscr{F}_t^a$ is equicontinuous in π_t for every $t \leq T$. We can use similar arguments to prove the equicontinuity of the other families. □

The continuity of w_t^u established above implies that $\sup_{\gamma_t^2} w_t^u(\pi_t, \gamma_t^1, \gamma_t^2)$ is achieved for every π_t, γ_t^1. Further, Lemma 8 implies that w_t^u and w_t^l satisfy the equicontinuity conditions in Lemma 7 for any given realization of belief π_t. Therefore, we can use Lemma 7 to argue that $\sup_{\gamma_t^2} w_t^u(\pi_t, \gamma_t^1, \gamma_t^2)$ is continuous in γ_t^1. And since γ_t^1 lies in the compact space $\mathcal{B}_t^1$, a minmaximizer exists for the function w_t^u. Further, we can use the measurable selection condition (see Condition 3.3.2 in [16]) to prove the existence of measurable mapping $\Xi_t^1(\pi_t)$ as defined in Lemma 4. A similar argument can be made to establish the existence of a maxminimizer and a measurable mapping $\Xi_t^2(\pi_t)$ as defined in Lemma 4. This concludes the proof of Lemma 4.

Proof of Theorem 2

Let us first define a distribution $\tilde{\Pi}_t$ over the space $\mathcal{X}_t \times \mathcal{P}_t^1 \times \mathcal{P}_t^2$ in the following manner. The distribution $\tilde{\Pi}_t$, given $C_t, \Gamma_{1:t-1}^{1:2}$, is recursively obtained using the following relation

$$\tilde{\Pi}_1(x_1, p_1^1, p_1^2) = \mathbb{P}[X_1 = x_1, P_1^1 = p_1^1, P_1^2 = p_1^2 \mid C_1] \ \forall \ x_1, p_1^1, p_1^2, \tag{89}$$

$$\tilde{\Pi}_{\tau+1} = F_\tau(\tilde{\Pi}_\tau, \Gamma_\tau^1, \Gamma_\tau^2, Z_{\tau+1}), \quad \tau \geq 1, \tag{90}$$

where F_τ is as defined in Definition 5 in "Appendix 4." We refer to this distribution as the strategy-independent common information belief (SI-CIB).

Let $\tilde{\chi}^1 \in \tilde{\mathcal{H}}^1$ be any strategy for virtual player 1 in game $\mathscr{G}_e$. Consider the problem of obtaining virtual player 2's best response to the strategy $\tilde{\chi}^1$ with respect to the cost $\mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2)$ defined in (18). This problem can be formulated as a Markov decision process (MDP) with common information and prescription history $C_t, \Gamma_{1:t-1}^{1:2}$ as the state. The control action at time t in this MDP is Γ_t^2, which is selected based on the information $C_t, \Gamma_{1:t-1}^{1:2}$ using strategy $\tilde{\chi}^2 \in \mathcal{H}^2$. The evolution of the state $C_t, \Gamma_{1:t-1}^{1:2}$ of this MDP is as follows

$$\{C_{t+1}, \Gamma_{1:t}^{1:2}\} = \{C_t, Z_{t+1}, \Gamma_{1:t-1}^{1:2}, \tilde{\chi}_t^1(C_t, \Gamma_{1:t-1}^{1:2}), \Gamma_t^2\}, \tag{91}$$

where

$$\mathbb{P}^{(\tilde{\chi}^1,\tilde{\chi}^2)}[Z_{t+1} = z_{t+1} \mid C_t, \Gamma_{1:t-1}^{1:2}, \Gamma_t^2] = P_t^m[\tilde{\Pi}_t, \Gamma_t^1, \Gamma_t^2; z_{t+1}], \tag{92}$$

almost surely. Here, $\Gamma_t^1 = \tilde{\chi}_t^1(C_t, \Gamma_{1:t-1}^{1:2})$ and the transformation P_t^m is as defined in Definition 5 in "Appendix 4." Notice that due to Lemma 3, the common information belief Π_t associated with any strategy profile $(\tilde{\chi}^1, \tilde{\chi}^2)$ is equal to $\tilde{\Pi}_t$ almost surely. This results in the state evolution equation in (92). The objective of this MDP is to maximize, for a given $\tilde{\chi}^1$, the following cost

$$\mathbb{E}^{(\tilde{\chi}^1,\tilde{\chi}^2)}\left[\sum_{t=1}^{T} \tilde{c}_t(\tilde{\Pi}_t, \Gamma_t^1, \Gamma_t^2)\right], \tag{93}$$

where $\tilde{c}_t$ is as defined in Eq. (22). Due to Lemma 3, the total expected cost defined above is equal to the cost $\mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2)$ defined in (18).

The MDP described above can be solved using the following dynamic program. For every realization of virtual players' information $c_{T+1}, \gamma_{1:T}^{1:2}$, define

$$V_{T+1}^{\tilde{\chi}^1}(c_{T+1}, \gamma_{1:T}^{1:2}) := 0.$$

In a backward-inductive manner, for each time $t \leq T$ and each realization $c_t, \gamma_{1:t-1}^{1:2}$, define

$$V_t^{\tilde{\chi}^1}(c_t, \gamma_{1:t-1}^{1:2}) := \sup_{\gamma_t^2}[\tilde{c}_t(\tilde{\pi}_t, \gamma_t^1, \gamma_t^2) + \mathbb{E}[V_{t+1}^{\tilde{\chi}^1}(c_t, Z_{t+1}, \gamma_{1:t}^{1:2}) \mid c_t, \gamma_{1:t}^{1:2}]], \tag{94}$$

where $\gamma_t^1 = \tilde{\chi}_t^1(c_t, \gamma_{1:t-1}^{1:2})$ and $\tilde{\pi}_t$ is the SI-CIB associated with the information $c_t, \gamma_{1:t-1}^{1:2}$. Note that the measurable selection condition (see condition 3.3.2 in [16]) holds for the dynamic program described above. Thus, the value functions $V_t^{\tilde{\chi}^1}(\cdot)$ are measurable and there exists a measurable best-response strategy for player 2 which is a solution to the dynamic program described above. Therefore, we have

$$\sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) = \mathbb{E}V_1^{\tilde{\chi}^1}(C_1). \tag{95}$$

Claim 1 For any strategy $\tilde{\chi}^1 \in \tilde{\mathcal{H}}^1$ and for any realization of virtual players' information $c_t, \gamma_{1:t-1}^{1:2}$, we have

$$V_t^{\tilde{\chi}^1}(c_t, \gamma_{1:t-1}^{1:2}) \geq V_t^u(\tilde{\pi}_t), \tag{96}$$

where V_t^u is as defined in (26) and $\tilde{\pi}_t$ is the SI-CIB belief associated with the instance $c_t, \gamma_{1:t-1}^{1:2}$. As a consequence, we have

$$\sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) \geq \mathbb{E}V_1^u(\Pi_1). \tag{97}$$

Proof The proof is by backward induction. Clearly, the claim is true at time $t = T + 1$. Assume that the claim is true for all times greater than t. Then, we have

$$\begin{aligned} V_t^{\tilde{\chi}^1}(c_t, \gamma_{1:t-1}^{1:2}) &= \sup_{\gamma_t^2}[\tilde{c}_t(\tilde{\pi}_t, \gamma_t^1, \gamma_t^2) + \mathbb{E}[V_{t+1}^{\tilde{\chi}^1}(c_t, Z_{t+1}, \gamma_{1:t}^{1:2}) \mid c_t, \gamma_{1:t}^{1:2}]] \\ &\geq \sup_{\gamma_t^2}[\tilde{c}_t(\tilde{\pi}_t, \gamma_t^1, \gamma_t^2) + \mathbb{E}[V_{t+1}^u(F_t(\tilde{\pi}_t, \gamma_t^{1:2}, Z_{t+1})) \mid c_t, \gamma_{1:t}^{1:2}]] \\ &\geq V_t^u(\tilde{\pi}_t). \end{aligned}$$

The first equality follows from the definition in (94), and the inequality after that follows from the induction hypothesis. The last inequality is a consequence of the definition of the value function V_t^u. This completes the induction argument. Further, using Claim 1 and the result in (95), we have

$$\sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) = \mathbb{E}V_1^{\tilde{\chi}^1}(C_1) \geq \mathbb{E}V_1^u(\tilde{\Pi}_1) = \mathbb{E}V_1^u(\Pi_1).$$

□

We can therefore say that

$$S^u(\mathcal{G}_e) = \inf_{\tilde{\chi}^1} \sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) \geq \inf_{\tilde{\chi}^1} \mathbb{E}V_1^u(\Pi_1) = \mathbb{E}V_1^u(\Pi_1). \tag{98}$$

Further, for the strategy $\tilde{\chi}^{1*}$ defined in Definition 4, inequalities (96) and (97) hold with equality. We can prove this using an inductive argument similar to the one used to prove Claim 1. Therefore, we have

$$S^u(\mathcal{G}_e) = \inf_{\tilde{\chi}^1} \sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^1, \tilde{\chi}^2) \leq \sup_{\tilde{\chi}^2} \mathcal{J}(\tilde{\chi}^{1*}, \tilde{\chi}^2) = \mathbb{E}V_1^{\tilde{\chi}^{1*}}(C_1) = \mathbb{E}V_1^u(\Pi_1). \tag{99}$$

Combining (98) and (99), we have

$$S^u(\mathcal{G}_e) = \mathbb{E}V_1^u(\Pi_1).$$

Thus, the inequality in (99) holds with equality which leads us to the result that the strategy $\tilde{\chi}^{1*}$ is a min–max strategy in game $\mathcal{G}_e$. A similar argument can be used to show that

$$S^l(\mathcal{G}_e) = \mathbb{E}V_1^l(\Pi_1),$$

and that the strategy $\tilde{\chi}^{2*}$ defined in Definition 4 is a max–min strategy in game $\mathcal{G}_e$.

References

1. Alpcan T, Başar T (2010) Network security: a decision and game-theoretic approach. Cambridge University Press, Cambridge
2. Amin S, Litrico X, Sastry S, Bayen AM (2012) Cyber security of water scada systemspart I: analysis and experimentation of stealthy deception attacks. IEEE Trans Control Syst Technol 21(5):1963–1970
3. Amin S, Schwartz GA, Cardenas AA, Sastry SS (2015) Game-theoretic models of electricity theft detection in smart utility networks: providing new capabilities with advanced metering infrastructure. IEEE Control Syst Mag 35(1):66–81
4. Aumann RJ, Maschler M, Stearns RE (1995) Repeated games with incomplete information. MIT Press, Cambridge
5. Başar T (1981) On the saddle-point solution of a class of stochastic differential games. J Optim Theory Appl 33(4):539–556
6. Basar T, Olsder GJ (1999) Dynamic noncooperative game theory, vol 23. SIAM, Philadelphia
7. Bondi E, Oh H, Xu H, Fang F, Dilkina B, Tambe M (2019) Using game theory in real time in the real world: A conservation case study. In: Proceedings of the 18th international conference on autonomous agents and multiagent systems, pp. 2336–2338. International Foundation for Autonomous Agents and Multiagent Systems
8. Fang F, Nguyen TH, Pickles R, Lam WY, Clements GR, An B, Singh A, Schwedock BC, Tambe M, Lemieux A (2017) Paws-a deployed game-theoretic application to combat poaching. AI Mag 38(1):23–36
9. Fang F, Stone P, Tambe M (2015) When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In: Twenty-fourth international joint conference on artificial intelligence
10. Filar J, Vrieze K (2012) Competitive Markov decision processes. Springer, Berlin

11. Fudenberg D, Tirole J (1991) Game theory. MIT Press, Cambridge
12. Gensbittel F, Oliu-Barton M, Venel X (2014) Existence of the uniform value in zero-sum repeated games with a more informed controller. J Dyn Games 1(3):411–445
13. Gensbittel F, Renault J (2015) The value of Markov chain games with incomplete information on both sides. Math Oper Res 40(4):820–841
14. Hansen EA (1998) Solving pomdps by searching in policy space. In: Proceedings of the fourteenth conference on Uncertainty in artificial intelligence, pp 211–219
15. Haywood O Jr (1954) Military decision and game theory. J Oper Res Soc Am 2(4):365–385
16. Hernández-Lerma O, Lasserre JB (2012) Discrete-time Markov control processes: basic optimality criteria, vol 30. Springer, Berlin
17. Kartik D, Nayyar A (2019) Stochastic zero-sum games with asymmetric information. In: 58th IEEE conference on decision and control. IEEE
18. Kartik D, Nayyar A (2019) Zero-sum stochastic games with asymmetric information. arXiv preprint arXiv:1909.01445
19. Li L, Shamma J (2014) LP formulation of asymmetric zero-sum stochastic games. In: 53rd IEEE conference on decision and control, pp. 1930–1935. IEEE
20. Li X, Venel X (2016) Recursive games: uniform value, tauberian theorem and the mertens conjecture. Int J Game Theory 45(1–2):155–189
21. Loch J, Singh SP (1998) Using eligibility traces to find the best memoryless policy in partially observable markov decision processes. In: Proceedings of the fifteenth international conference on machine learning, pp. 323–331
22. Maschler M, Solan E, Zamir S (2013) Game theory. Cambridge University Press. https://doi.org/10.1017/CBO9780511794216
23. Mertens JF, Sorin S, Zamir S (2015) Repeated games, vol 55. Cambridge University Press, Cambridge
24. Morrow JD (1994) Game theory for political scientists. Princeton University Press, Princeton
25. Nayyar A, Gupta A (2017) Information structures and values in zero-sum stochastic games. In: American control conference (ACC), 2017, pp. 3658–3663. IEEE
26. Nayyar A, Gupta A, Langbort C, Başar T (2014) Common information based Markov perfect equilibria for stochastic games with asymmetric information: finite games. IEEE Trans Autom Control 59(3):555–570
27. Nayyar A, Mahajan A, Teneketzis D (2010) Optimal control strategies in delayed sharing information structures. IEEE Trans Autom Control 56(7):1606–1620
28. Nayyar A, Mahajan A, Teneketzis D (2013) Decentralized stochastic control with partial history sharing: a common information approach. IEEE Trans Autom Control 58(7):1644–1658
29. Osborne MJ, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge
30. Ouyang Y, Tavafoghi H, Teneketzis D (2017) Dynamic games with asymmetric information: common information based perfect bayesian equilibria and sequential decomposition. IEEE Trans Autom Control 62(1):222–237
31. Ponssard JP, Sorin S (1980) The lp formulation of finite zero-sum games with incomplete information. Int J Game Theory 9(2):99–105
32. Renault J (2006) The value of Markov chain games with lack of information on one side. Math Oper Res 31(3):490–512
33. Renault J (2012) The value of repeated games with an informed controller. Math Oper Res 37(1):154–179
34. Rosenberg D (1998) Duality and markovian strategies. Int J Game Theory 27(4):577–597
35. Rosenberg D, Solan E, Vieille N (2004) Stochastic games with a single controller and incomplete information. SIAM J Control Optim 43(1):86–110
36. Rudin W et al (1964) Principles of mathematical analysis, vol 3. McGraw-hill, New York
37. Sandell NR Jr (1974) Control of Finite-State, Finite-Memory Stochastic Systems, Massachusetts Institute of Technology, PhD Thesis. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740018646.pdf
38. Shelar D, Amin S (2017) Security assessment of electricity distribution networks under der node compromises. IEEE Trans Control Netw Syst 4(1):23–36. https://doi.org/10.1109/TCNS.2016.2598427
39. Teneketzis D (2006) On the structure of optimal real-time encoders and decoders in noisy communication. IEEE Trans Inf Theory 52(9):4017–4035
40. Vasal D, Sinha A, Anastasopoulos A (2019) A systematic process for evaluating structured perfect bayesian equilibria in dynamic games with asymmetric information. IEEE Trans Autom Control 64(1):78–93
41. Washburn A, Wood K (1995) Two-person zero-sum games for network interdiction. Oper Res 43(2):243–251
42. Wu M, Amin S (2018) Securing infrastructure facilities: When does proactive defense help? Dyn Games Appl 9:1–42

43. Zheng J, Castañón DA (2013) Decomposition techniques for Markov zero-sum games with nested information. In: 52nd IEEE conference on decision and control, pp. 574–581. IEEE
44. Zhu Q, Basar T (2015) Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: games-in-games principle for optimal cross-layer resilient control systems. IEEE Control Syst Mag 35(1):46–65

Dynamic Games and Applications (2021) 11:389
https://doi.org/10.1007/s13235-020-00366-9

CORRECTION

Correction to: Upper and Lower Values in Zero-Sum Stochastic Games with Asymmetric Information

Dhruva Kartik[1] · **Ashutosh Nayyar**[1]

Published online: 16 September 2020

Correction to: Dynamic Games and Applications https://doi.org/10.1007/s13235-020-00364-x

The original article has been published without funding information. This has been corrected with this erratum.

Acknowledgements This work was supported by NSF Grants ECCS 1509812 and ECCS 1750041 and ARO Award No. W911NF-17-1-0232.

The original article can be found online at https://doi.org/10.1007/s13235-020-00364-x.

✉ Dhruva Kartik
mokhasun@usc.edu

Ashutosh Nayyar
ashutosh.nayyar@usc.edu

[1] Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90007, USA