# Information concealing games in communication networks

Saswati Sarkar, Eitan Altman, Rachid El-Azouzi and Yezekael Hayel[‡]

**Abstract**

Consider the situation in which a decision maker (Actor) has to decide which of several available resources to use in the presence of an adversary (called Controller) that can prevent the Actor of receiving information on the state of some of the resources. The Controller has a limitation on the amount of information it can conceal. What information should it deny from the the Actor? How should the Actor choose a resource as a function of the statistics of the states of the resources and of the non-concealed information on the state of the others. We formulate this problem as a non-zero sum game and transform it into an equivalent zero-sum game. We then propose ways to compute the most harmful behavior of the Controller as well as the best choice of a resource for the Actor, and analyse their complexity. We identify cases in which the exact solution is computationally intractable, and provide approximate solutions with polynomial complexity. We present many motivating examples and explore numerically the performance of the approximations.

## I. Introduction

### A. Overview

Exchange of information among different entities forms the basis of most technological advances in the information era and also of social interactions. Several seminal advances in communication systems have lead to schemes that maximize the rate of exchange of information. An aspect that has received somewhat less attention, and is as important, is that of designing a framework for deciding what information should be revealed and what should be concealed during exchange of information among different entities so as to maximize their utilities. The main challenge towards developing such a framework is that oftentimes such decisions depend on the objective for exchange of information, and hence can only be determined on a case by case basis. The contribution of this paper is to develop a rigorous mathematical framework for deciding what information an entity should reveal when the objectives satisfy certain broad characterizations that capture the essence of several communication and social systems.

We consider a system with two entities. The state of the system is a random vector of dimension $n$. At any given time the first entity (*controller*) has complete information about the state of the system, and must reveal a certain "minimum" amount of information about the system state to the second entity. It can however choose the nature of the information it reveals subject to satisfying the above constraint. The second entity (*actor*) takes certain actions based on the information the controller reveals, and the actions are associated with certain utilities for both

Dept. of Electrical and Systems Eng., University of Pennsylvania 200 S. 33rd Street, Philadelphia 19104, USA

INRIA, 2004 Route des Lucioles, B.P.93, 06902 Sophia-Antipolis Cedex, France. The work of this author has been partially supported by the Eureopean Bionets project

LIA, Université d'Avignon 339, chemin des Meinajaries, Agroparc BP 1228, 84911 Avignon Cedex 9, France. E-mail: rachid.elazouzi@univ-avignon.fr, yezekael.hayel@univ-avignon.fr

the controller and the actor which also depend on the state of the system. The same actions and the system states fetch different utilities for the controller and the actor, and usually when one entity has a high utility the other has a low utility. We devise a framework that enables the controller to decide the information it would reveal, or equivalently conceal, so as to maximize its own utility, and the actor to determine its actions based on the information it has about the system so as to again maximize its utility.

### *B. Motivation*

We first establish that this information concealing problem forms the basis of several communication systems.

*1) Information concealing problems in wireless networks:*

*a) Cognitive radio networks:* Consider a transmitter with access to $n$ channels, whose qualities constitute the state of the system. The transmitter needs to select one channel for transmission, and the transmission quality of the selected channel determines the rate of successful transmission. Hence, the transmitter probes the channels in order to assess their qualities before it transmits any packet. A malicious entity, say a jammer, seeks to reduce the rate of successful transmission. The jammer is usually assumed to accomplish its goal by generating signals that interfere with the sender's communication; however the jammer may be able to deteriorate the transmission rate much more by preventing the transmitter from learning the states of the channels. This may cause the transmitter to make a wrong choice, that is, select a channel with a poor transmission quality, and thereby obtain a poor data rate for a while. Note that the jammer can prevent the sender from learning the states of some channels, possibly by generating signals that interfere with the corresponding probe packets or responses to these probes, and generating such signals may consume less energy as compared to those that jam the actual transmission since the probe packets are transmitted over shorter durations. We therefore consider the case where the jammer blocks the probe packets and not the actual transmission. Furthermore, we assume that the jammer knows the quality of the channels and can block the probes in at most $k$ channels since the blocking process consumes energy. Hence, the states of at most $k$ channels can be concealed from the transmitter. The transmitter selects the channel after it learns about the states of the channels the jammer does not conceal. Note that the transmitter may either select a channel whose state has been revealed or one whose state has been concealed; the latter may happen since the fact that the jammer has concealed the state of a channel may indicate that the transmission quality of the corresponding channel is good. The rate of successful transmission attained by the transmitter determines the utility of the transmitter and the jammer. The information concealing problem we described will enable the jammer (controller) to optimally determine which channels it would conceal, and the transmitter (actor) to select the channel.

*2) Information concealing problems in other information systems:*

*a) Query resolution networks:* We next describe another communication system in which the information concealing problem arises. Consider a client that needs to locate a desired information. It queries some data bases to determine which of them has the information. The responses constitute the state of the system and specify the probability with which the requested information is present in the data base (as the search in response to such preliminary queries may not be comprehensive and also the information may be dated). The responses reach the node through a gateway that has a malicious entity which blocks some of the responses in order to undermine the information location service. The client needs to determine which database it would request the information from based on the responses to its query, and again it may choose one it received a response from or one it did not receive a response from (the latter may happen if the responses it receives reveal low probabilities). The utility of the client and the malicious entity depends on the probability that the client obtains the information it is interested in. The information concealing problem we described will enable the malicious entity (controller) to optimally determine which responses it would suppress and the client (actor) to determine which database it would query.

*b) Buyer-Seller authentication in e-commerce:* Next, consider an e-commerce system where a buyer and a seller are bargaining. The authentication process between them proceeds in two stages. The buyer has $n$ pieces of information using which he can authenticate himself to the seller. He reveals limited information about $k$ of these pieces using which the seller can complete the first stage of the authentication successfully if the buyer is who he claims to be (e.g., using some proof verification methods). Next, the seller identifies himself to the buyer, and subsequently asks about complete information for one of the $n$ pieces which may or may not be among those that the buyer initially selects. The buyer provides the requested information and the authentication is successful if again he is who he claims to be. This two-stage authentication process allows each entity to identify himself once he has some (albeit incomplete) information about the other participant. Now, the complete information the buyer reveals about any one piece in the authenticating process may allow the seller to acquire more information about the buyer than that required for mere authentication, e.g., information about his previous transactions with other merchants, etc. This will for example allow him to bargain more effectively with the buyer once the authentication is successful. Now, the different pieces of information the buyer possesses about himself reveals different amount of information about him, and the buyer must select the $k$ pieces in the first stage so as to minimize the additional information he finally reveals to the seller. The seller must subsequently select the piece in the second stage to acquire maximum possible information about the buyer. The information concealing problem we described will enable the buyer (controller) and the seller (actor) to attain their respective objectives by optimally selecting the pieces in question.

*3) Information concealing problems in social context:*

*a) Gambling:* Consider a gambling game in which two gamblers have a common collection of $N$ cards each of which can have one of $m$ colors. They randomly select a number for each card and write the chosen number on one side of the corresponding card. Subsequently, they put all cards in a bin, and the second gambler draws $n$ cards randomly from the bin without observing the numbers on them. The first gambler then observes the colors and the numbers of the cards drawn and tells the second the numbers and the colors of $k$ of these cards, and only the colors of the rest of the cards. The second gambler needs to select one of these $n$ cards (either a card whose number it knows or one whose number it does not know), and the first pays him an amount that equals the number on the selected card (if this number is negative then the second pays the first). The first gambler (the controller) needs to select the $k$ cards so as to minimize the amount it pays, and the second needs to select a card so as to maximize the amount it receives.

*b) Security systems:* Consider a corrupt employee who sells secrets about the company's security system to some burglars. The building in which the company is located has $n$ gates, and the employee knows the efficacy of the security system at each gate (e.g., he may know the number of guards at each gate which may be a random variable owing to the company's security plan), and based on the price the burglar has offered or in order to conceal his collusion in the event of an enquiry, the employee informs the burglar information about only $k$ of these gates. He also decides to select the gates whose information he reveals so as to minimize the probability that the break in is successful since if there is a successful break-in a comprehensive enquiry is likely to be launched. The burglar can break in through one among the $n$ gates, and selects this gate based on the information he obtains from the employee so as to maximize his chance of success.

In both these examples, the information concealing problem we described will enable the controller (first gambler or employee) and the actor (second gambler or burgler) to attain their objectives by making appropriate selections.

### C. Contribution and Challenges

Our first contribution is to provide a framework for investigating information concealing problems. We formulate this problem as a stochastic leader-follower game (Section III). We demonstrate that the well-known Nash equilibrium solution concept can not be effectively used in this game since the utilities of the players turn out to be functions rather than numbers. Subsequently, we develop suitable solution concepts, that of point-wise Nash equilibria, that capture the subtleties of this game. For example, the actor can learn about the system not only from the information the controller reveals, but also from the choices of the controller regarding which sources of information it conceals, since the fact that an information has been concealed may provide important insight about its nature. Thus, the actor must determine its optimal action so as to exploit the information contained in both of the above, and the controller must determine what it should conceal considering that the actor will learn from both the above. For example, in cognitive radio networks, a naive policy for the jammer would be to conceal the states of the channels that have the $k$ best transmission qualities. But, if the transmitter knows the jammer's policy, then it knows that the transmission quality of any channel whose state has been concealed is at least as good as that of a channel whose state has been revealed, and thus, its best action is to select a channel whose state has been concealed. But, if the jammer reveals the states of some channels whose transmission qualities are better than those whose states it conceals, the transmitter will be confused regarding the choice of the channel even when it knows the jammer's policy, and is therefore more likely to make a poor selection. Our framework formally establishes that the naive policy described above is suboptimal for a controller (Lemma 3.1).

We next prove that there is one-to-one correspondence between the set of point-wise Nash equilibria in the above game and the set of saddle points in a two-person zero-sum game (Section IV-A), which we refer to as an equivalent game. This equivalence turns out to be very useful as it implies that a point-wise Nash equilibrium exists for the original game and can be computed using a linear program, since a saddle-point of any two-person zero-sum game can be computed using a linear program. The equivalence is however somewhat surprising as the controller and actor has different amount of information about the system, that is, the controller has complete information whereas the actor only has partial information about the system state. Since the policies and the utilities of each player depends on the information it has about the system, the utilities of the two players turn out to be functions with different domains, and hence their sum can not be defined, whereas the sum of the utilities in a two-person zero-sum game is always zero. Furthermore, both players act simultaneously in two-person zero-sum games, whereas in the information concealing problem, the sequence in which the two players can act is unique: the controller first needs to reveal information about the system state, and then the actor can determine its actions.

We next investigate the computational aspects of the information concealing games. Our results in this area constitute our second contribution since general results that can address the computational aspects in this case are not known in the game theory or approximation algorithm literature. We first observe that the number of variables and constraints in the standard linear program formulations for computing the saddle points of the equivalent games are super-exponential in $n$, where $n$ is the dimension of the state-space of the system. Thus, the linear program becomes computationally intractable even for moderate values of $n$. Exploiting specific characteristics of the game under consideration, we next obtain linear programs which compute the saddle-points of the equivalent game and the optimal policies for the two players while using exponential number of variables and constraints (Section IV-B). This significant reduction in computation time enables the computation of the optimal policies for moderate $n$. We next obtain linear time ($O(n)$) computable policies with provable performance guarantees for the two players (Section V). Specifically, these policies attain utilities that differ from the utilities of the saddle points by (a) constant factors in several important special cases, and (b) by factors that depend only on the amount of information that the controller reveals to the actor, and do not depend on $n$ in the most general case.

## II. Related literature

To the best of our knowledge, the information concealing game has not been investigated before. The closest game that has been investigated before is that introduced by noble-laureate P. Aumann *et. al.* [1]. They consider a family of $K$ two-player matrix games $\{G^k(i,j)\}$, each of size $I \times J$ and a probability $p$ over the discrete set $\{1,..,\}$. The set of games and $p$ are common knowledge. Nature chooses one of these games with probability $p$ and informs player 1 which of the games is played. The same game will now be played again and again. Player 2 is not informed on the game being played. At each time unit $t = 1, 2, ...$, player 1 chooses an action $i_t$ among $I$ and simultaneously, player 2 chooses action $j_t$ in $J$. Player 1 then pays $G^k(i_t, j_t)$ to player 2. Player 2 does not observe the payment but both players observe the actions of each other. Player 1's policies are sequences of probability measures over $I$ conditioned on the past actions of both players Player 2's policies are sequences of probability measures over $J$ conditioned on her past actions only. In this game, player 1 is confronted with the dilemma of whether to play optimally in the game chosen by nature; if he does that (and if player 2 knows which strategy is used by player 1), then player 2 will eventually be able to guess which is the game being played, so that player 1 looses her advantage of being informed. If on the contrary, he uses a strategy that does not utilize his knowledge of the game, then again he does not gain from being informed. The main difference between this setting and our setting is that in this setting the informed player does not directly control what information to reveal or to conceal to the other player. Also, an important aspect of this setting is that the information chosen by nature does not change with time and thus at any given time a player can exploit the knowledge he has acquired from past interactions. We assume that the nature's choice changes with time and the evolution is temporally independent. The temporal independence and also our specific context imply that the players can not exploit information acquired in the past, and hence the game effectively starts fresh at each instant (our solutions therefore do not consider any temporal relation at all). Thus, the formal questions that are answered and also the techniques used to obtain the answers substantially differ in the two cases.

Finally, information concealing has been extensively investigated in context of multi-media [2]. An example is the research on watermarking, where one tries to hide a signature in some picture or audio recording in order to be able to identify it later. Informally speaking, these scenarios consist of only one player who seeks to conceal as much information as possible. We consider a scenario with two players such that both players act sequentially and the first conceals information with the goal of degrading the performance of the second by decreasing the second's capabilities to make good action choices. Again, the formal questions that are answered and also the techniques used to obtain the answers substantially differ in the two cases.

## III. A Mathematical Framework

We formulate the information concealing problem as a stochastic leader-follower game and develop appropriate solution concepts for such games (Section III-A). We next elucidate the terminologies and the solution concepts using the motivating examples presented in the previous section (Section III-B). We finally demonstrate that the Nash-equilibrium for this game exhibits several counter-intuitive properties which indicate that the computation of such equilibrium may not be straight-forward (Section III-C).

### *A. Terminologies and Solution Concepts*

We start by modeling the information concealing game as a stochastic leader-follower game between two players: the controller and the actor. We describe the game in both the normal form as well as in the strategic form. Let $\mathcal{N} = \{1, \ldots, n\}$.

- **System state:** The state of the system is an $n$-dimensional vector $\vec{X}$ whose entries take values in $\mathcal{K} = \{0, ..., K-1\}$. The state space is $\mathcal{K}^n$. The random variables corresponding to the components of the state vector may be dependent and can be described by a joint probability distribution $\beta$.
- **Information of the Controller:** The controller knows the system state vector $\vec{X}$, and thereby has full information.
- **Actions of the Controller:** The controller conceals the values of at most $k$ components of the system state vector from the actor; it decides which components it would conceal based on its information. Thus, the controller's action is a subset of $\mathcal{N}$ with cardinality $k$ or lower. Note that each such action determines a sub-vector of the system state, with size $n-k$ or more, that the actor observes. Let $\mathcal{A}_c(\vec{x})$ denote the set of all such sub-vectors when the controller's information (i.e., the system state vector) is $\vec{x}$, and $\mathcal{A}_c = \cup_{\vec{x} \in \mathcal{K}^n} \mathcal{A}_c(\vec{x})$.
- **Information of the actor:** The actor knows the states of those components of the system state vector which the controller does not conceal. Specifically, if $c$ be the action taken by the controller and the system state is $\vec{x}$, then the actor's information $\vec{y}$ consists of the sub-vector of $\vec{x}$ with components in $N \setminus c$. Therefore, from its information $\vec{y}$, the actor knows the controller's action (i.e., the subset of components $a(\vec{y})$ the controller conceals). Let $\mathcal{I}_a$ be the set of all possible informations of the actor about the entries of the system state vector. It consists of at least $|\mathcal{A}_c| \times K^{n-k}$ elements.
- **Actions of the actor:** The actor selects one of the components of the system state vector. Thus, its action is an integer $l$ where $1 \leq l \leq n$. Thus, $\mathcal{N}$ is the set of all actions of the actor.

- **Payoff function:** If a component of the system state vector has value $i$, then the expected utility associated with that component is $r(i)$ such that $r(0) < r(1) < \ldots < r(K-1)$. If the system state is $\vec{x}$, and the actor selects component $l$, then the payoff for the actor is $r(x_l)$.
- **Common Knowledge:** Both the controller and the actor know $n, k, K$ and the joint probability distribution for the system state vector $\beta$. These parameters are determined based on goals and constraints of specific systems (e.g., $k$ may be determined based on resource constraints of the jammer in the cognitive radio network and the price the burglar has offered in the security system) - investigation of how these parameters are determined is beyond the scope of the current paper.
- **Strategies:**
  - **Behavioral strategies:** A behavioral strategy of a player is a function from its information set to the set of probability measures over its action space. More precisely, the controller can decide randomly which components to conceal based on the system state vector, and the actor can randomly select a component based on the revealed sub-vector. Let $u$ ($v$, respectively) be a behavioral policy of the controller (actor, respectively). Then, $u(\vec{x})$ ($v(\vec{y})$, respectively) is the probability distribution used by the controller (actor, respectively) for selecting its actions when its information is $\vec{x}$ ($\vec{y}$, respectively). Specifically, $u(\vec{x})_{\vec{y}}$ ($v(\vec{y})_i$, respectively) is the probability with which the controller (actor, respectively) conceals the sub-vector $\vec{y} \in \mathcal{A}_c(\vec{x})$ (selects the component $i \in \mathcal{N}$, respectively) when its information is $\vec{x}$ ($\vec{y}$, respectively). Let $\mathcal{U}$ ($\mathcal{V}$, respectively) be the set of behavioral strategies for the controller (actor, respectively).
  - **Pure policies:** Let $\mathcal{U}^p \subset \mathcal{U}$ ($\mathcal{V}^p \subset \mathcal{V}$, respectively) be the set of pure (deterministic) behavioral policies for the controller (actor, respectively). A controller's pure policy is a function from $\mathcal{K}^n$ to $\mathcal{A}_c$. An actor's pure policy is a function from $\mathcal{I}_a$ to $\mathcal{N}$.
  - **Mixed strategies:** A mixed strategy of a player is a probability measure over its pure policies. Let $\mathcal{U}^M$ ($\mathcal{V}^M$, respectively) be the set of mixed strategies for controller (actor, respectively).

  Note that behavioral and mixed are alternate representations of the randomized policies of the two players.
- **Probability space:** Any given joint probability distribution for the system state vector $\beta$ and strategies $u$ and $v$ for the controller and actor, respectively, define a probability $P_\beta^{u,v}$ measure over the state, actions and informations of the two players. Let $E_\beta^{u,v}$ be the corresponding expectation operator.
- **Utility:**
  - **Utility of the actor:** The actor's utility is its expected payoff conditioned on its information, and is therefore a function of its information. Specifically, when the actor's information is $\vec{y}$, the controller and the actor use (behavioral or mixed) strategies $u$ and $v$ respectively, and the joint probability distribution of the system states is $\beta$, the actor's utility $J_a^{u,v,\beta}(\vec{y})$ is given by

    $$J_a^{\beta,u,v}(\vec{y}) = E_\beta^{u,v}[r(X_B)|\vec{Y}_a = \vec{y}],$$

    where $\vec{Y}_a$ is the random information of the actor, $X_i$ is the random state of the $i$th component of the system state vector, $B$ is the action of the actor. Hence, $X_B$ is the random state of the component which is chosen possibly in a random way by the actor.

– **Utility of the controller:** The controller's utility is the negative of the expected payoff of the actor conditioned on the controller's information, and is therefore again a function of the controller's information. Specifically, when the system state vector is $\vec{x}$, and the controller and the actor use (behavioral or mixed) strategies $u$ and $v$ respectively, the controller's utility $J_c^{u,v}(\vec{x})$ is given by

$$J_c^{u,v}(\vec{x}) = -E^{u,v}[r(x_B)|\vec{X} = \vec{x}],$$

where $\vec{X}$ is the random system state vector, $x_B$ is the $B$th component of $\vec{x}$, $B$ is (potentially random) action of the actor when the system state is $\vec{x}$ and the controller and the actor use (behavioral or mixed) strategies $u$ and $v$. Note that $\beta$ is not used explicitly in computing the above expectation. If however $u, v$ depend on $\beta$, the value of this expectation may depend on $\beta$.

- **Controller's and Actor's goals:** The controller and the actor seek to maximize their respective utilities $J_c^{u,v}(\vec{x})$, $J_a^{\beta,u,v}(\vec{y})$ for all values of their respective informations $\vec{x}, \vec{y}$.

Since the controller's and actor's utilities are functions and not numbers, we can not use Nash equilibrium as a solution concept (unless we define an ordering between vectors). We however use related solution concepts, that of, *point-wise Nash equilibrium*, which we define next.

*Definition 3.1:* Let $u^*$ and $v^*$ be behavioral or mixed strategies of the controller and actor respectively. Then $(u^*, v^*)$ is a point-wise Nash equilibrium if the following two conditions hold:

- for each system state vector $\vec{x}$ such that $\beta(\vec{x}) > 0$, $u^*(\vec{x})$ is a best response of the controller against $v^*$ of the actor, i.e., $u^*(\vec{x})$ maximizes $J_c^{u,v^*}(\vec{x})$ among all strategies $u$ of the controller, and
- for each information $\vec{y}$ of the actor which occurs with positive probability under $\beta, u^*$, $v^*(\vec{y})$ is a best response of the actor against $u^*$ of the controller, i.e., $v^*(\vec{x})$ maximizes $J_a^{\beta,u^*,v}(\vec{y})$ among all strategies $v$ of the actor.

## B. Elucidating examples

We now elucidate the above terminologies using the examples in Section I-B.

In cognitive radio networks the system state vector constitutes the states of the channels, each channel can be in $K$ states, and $r(i)$ is the expected rate of successful transmission of the transmitter (actor) when it transmits in a channel that is in state $i$. The jammer's (controller's) action is to conceal the states some ($\leq k$) channels and the transmitter's action is to select a channel for transmission. An example pure behavioral strategy of the jammer (denoted as *Greedy for Controller* or GC), is to conceal the channels with $k$-best states, that is, those with $k$-best expected rates of successful transmission (ties are broken in any pre-determined order). An example pure policy of the transmitter, (denoted as *Best Among Revealed for Actor* or BRA), is to select the channel that has the highest state among the revealed channels (ties are broken in any pre-determined order). An example behavioral policy of the jammer that is not pure is to conceal the states of as many channels that are in state $K-1$ as possible (subject to concealing the states of at most $k$ channels), and if fewer than $k$ channels are in state $K-1$ then select the remaining channels whose states are to be concealed uniformly among the channels that are not in state $K-1$. An example behavioral policy of the transmitter (denoted as *Uniform among Concealed* *for Actor* or UCA) that is not pure is to select a channel for transmission uniformly among those whose states are concealed. Next, $J_c^{u,v}(\vec{x})$ is the negative of the expected rate of successful transmission of the transmitter when the channel state vector is $\vec{x}$ and the jammer and transmitter use policies $u, v$ respectively. Also, $J_a^{\beta,u,v}(\vec{y})$ is the expected rate of successful transmission of the transmitter when the jammer reveals $\vec{y}$ to the transmitter, jammer and transmitter use policies $u, v$ respectively and the joint distribution of the channel state vector is $\beta$. For example, let $u$ be GC and $v$ be UCA. Then $J_c^{u,v}(\vec{x}) = -\frac{\max S \subseteq N, |S|=k \sum_{i \in S} x_i}{k}$, and $J_a^{\beta,u,v}(\vec{y}) = \frac{\sum_{i \in a(\vec{y})} \mathbf{E}(X_i|\vec{y})}{k}$ (note that the conditional expectation in the latter depends on $\beta$). If the transmitter uses BRA, GC is the best response of the jammer, and if the state processes of the channels are independent and identically distributed, UCA is the best response of the transmitter against the GC policy of the jammer .

In the authentication example for e-commerce, the seller (actor) may have different bargaining powers associated with different informations it can learn about the buyer (controller), and the buyer may not know the seller's bargaining power associated with any piece even though he knows the details about the piece. This is because different sellers may have access to different data bases and therefore may be able to extract different amount of additional information about the buyer from the same content. The buyer may however know the expected bargaining power of the seller associated with each piece of information. This scenario can be modelled by assuming that each different piece of information of the buyer can be in one of $K$ states, and the bargaining power associated with a particular state, say $i$, of a piece of information is a random variable whose expectation $r(i)$ is known to both the buyer and the seller. The system state vector consists the states of the $n$ pieces of informations the buyer has about himself. The buyer knows the system state vector (note that the knowledge of the state of a piece of information implies that the buyer knows the expected and not the exact value of the bargaining power associated with that piece). The action of the buyer is to reveal limited information about some $(n-k)$ pieces of information in the first stage of the authentication: the seller can only determine $\vec{y}$ the states of these pieces of information from the limited information the buyer reveals (since although he knows what databases he can search he does not know the details about any of these pieces). The seller's action is to select one piece for which it requests details. Next, $J_c^{u,v}(\vec{x})$ is the negative of the expected bargaining power of the seller when the system state vector is $\vec{x}$ and the buyer and the seller use policies $u, v$ respectively. Also, $J_a^{\beta,u,v}(\vec{y})$ is the expected bargaining power of the seller when it observes $\vec{y}$ in the first stage, the buyer and seller use policies $u, v$ respectively and the joint distribution of the system state vector is $\beta$.

In the gambling game, $\beta$ can be obtained from the distribution that is simultaneously used to draw the random numbers, and $K$ is the cardinality of the support set of this original distribution. Note that the random numbers drawn may be negative; we enumerate them using $K$ positive integers, and each such enumeration constitutes the state of a card. Thus, each card has $K$ possible states, and $r(i)$ is the number associated with the $i$th state. The system state vector consists the random numbers on the cards drawn by the second gambler (actor), and is known only to the first. The action of the first gambler (controller) is to reveal the states of some $(\geq n-k)$ of these cards, which constitutes the information $\vec{y}$ for the second and the second gambler's action is to select one card for examination of the number among those that it selected initially. Next, $J_c^{u,v}(\vec{x})$ is the negative of the expectation of the random number on the card the second finally (potentially randomly) selects for examination when the system state vector is $\vec{x}$ and the gamblers use policies $u, v$ respectively. Also, $J_a^{\beta,u,v}(\vec{y})$ is the expectation of the number on the card the second finally selects for examination, when it observes $\vec{y}$, the gamblers use policies $u, v$ respectively and the joint distribution of the system state vector is $\beta$.

The query resolution network and the security systems are similar to the cognitive radio network. In the former, the system state vector constitutes the states of the databases, each database can be in $K$ states, and $r(i)$ is the probability that the information sought is in a database that is in state $i$. In the latter, the system state vector constitutes the states of the gates (e.g., the number of guards at each gate), each gate can be in $K$ states, each state represents a level of efficacy of the security system at the gate and $r(i)$ is the probability that the burglar will successfully break in through a gate that is in state $i$.

### C. Counter-intuitive properties of the point-wise Nash equilibrium

We now demonstrate that the point-wise Nash equilibrium exhibits several counter-intuitive properties which suggests that the point-wise Nash equilibrium may not always consist of simple policies that can be represented in closed form. This in turn motivates the design of efficient frameworks for computing the point-wise Nash equilibrium, which is the focus of the next two sections.

Consider the "Greedy for Controller" (GC) policy for the controller (Section III-B). This policy conceals the components with $k$ highest states. Intuitively, it seems that GC minimizes the efficacy of the actor and therefore there always exists some policy $v^*$ for the actor such that (GC, $v^*$) constitutes a point-wise nash equilibrium. The following lemma shows that this intuition is unfounded, even when the joint probability distribution $\beta$ is such that the state processes for different components are mutually independent and identically distributed (i.e., even when all channels are i.i.d. in cognitive radio networks).

*Lemma 3.1:* There may not exist any policy $v^*$ for the actor such that (GC, $v^*$) constitutes a point-wise Nash equilibrium, even in systems where the state processes for different components are mutually independent and identically distributed.

Next, consider a simple policy " Statistically Best for Actor" (SBA) for the actor under which it decides its action without exploiting any knowledge of the controller's policy. Specifically it selects the component $i$ for which the expectation of the utility ($r(X_i)$) conditioned on the states of channels whose states have been revealed is the maximum under $\beta$ (it uses only $\beta$ and not the controller's policy in determining the above conditional expectation). For example, when the state processes of all components are mutually independent, $K = 2$ (i.e., each component has $2$ states), if the state of a component that is in state $1$ has been revealed, SBA selects the component and otherwise SBA selects the component for which the expected reward is the maximum under the prior distribution $\beta$. It may seem that at least in simple special cases, i.e., when $K = 2$, there always exists some policy $u^*$ for the actor such that ($u^*$, SBA) constitutes a point-wise Nash equilibrium. The following lemma shows that such intuition is founded.

*Lemma 3.2:* There may not exist any policy $u^*$ for the controller such that ($u^*$, SBA) constitutes a point-wise Nash equilibrium, even in systems where the state processes for different components are mutually independent and $K = 2$.

We prove lemmas 3.1 and 3.2 after obtaining some additional properties of the point-wise Nash equilibrium (Section **??**).

## IV. A COMPUTATIONAL FRAMEWORK FOR POINT-WISE NASH EQUILIBRIUM

The stochastic leader-follower game formulated in the previous section is clearly not a two-person zero-sum game as the arguments of the controller's and actor's utility functions have different dimensions, and hence the sum of these functions is not well-defined. Nevertheless, we demonstrate that there exists an equivalent zero-sum game such that a policy pair $(u, v)$ of the controller and actor constitutes a point-wise Nash equilibrium in the original game if and only if it constitutes a saddle-point in the equivalent game (Section IV-A). This equivalence implies that there always exists a point-wise Nash equilibrium in the original game, and one such equilibrium can be determined by solving a pair of linear programs. We develop a framework for computing the point-wise Nash equilibrium using this equivalence (Section IV-B).

### A. An equivalent two-person zero-sum game

*Definition 4.1:* Consider a game with two players: the controller and the actor. The action of each player now is to select one of its pure behavioral policies in the stochastic leader-follower game described in the previous section. When the two players select policies $u, v$ respectively, the utility of the actor under the joint probability distribution $\beta$ for the system states is given by

$$R_\beta^{u,v} = E_\beta^{u,v}[r(X_B)] = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) E_\beta^{u,v}[r(x_B)|\vec{X} = \vec{x}]. \tag{1}$$

where $B$ is the action of the transmitter under policies $u, v$ and random system state vector $\vec{X}$. The actor seeks to maximize its utility and the controller seeks to minimize the actor's utility. The game is clearly a two-person zero-sum game.

Clearly,

$$R_\beta^{u,v} = -\sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) J_c^{u,v}(\vec{x}) \ \forall\ u, v, \beta, \tag{2}$$

$$\text{and } R_\beta^{u,v} = \sum_{\vec{y}\in\mathcal{K}^n} \Pr{}^{\beta,u}(\vec{y}) J_a^{\beta,u,v}(\vec{y}) \ \forall\ u, v, \beta. \tag{3}$$

*Definition 4.2:* Let $\mathcal{U}$ and $\mathcal{V}$ respectively be the set of behavioral strategies of the controller and actor in the two-person zero-sum game. The upper and lower values, $\overline{R}_\beta, \underline{R}_\beta$ of the above game are

$$\overline{R}_\beta = \sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v} \qquad \underline{R}_\beta = \inf_{u\in\mathcal{U}} \sup_{v\in\mathcal{V}} R_\beta^{u,v}.$$

For any $u^* \in \mathcal{U}$ and $v^* \in \mathcal{V}$ we have

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v^*} \le \underline{R}_\beta \le R_\beta^{u^*,v^*} \le \overline{R}_\beta \le \sup_{v\in\mathcal{V}} R_\beta^{u^*,v}. \tag{4}$$

*Definition 4.3:* If for some $u^* \in \mathcal{U}$ and $v^* \in \mathcal{V}$, $\inf_{u\in\mathcal{U}} R_\beta^{u,v^*} = \sup_{v\in\mathcal{V}} R_\beta^{u^*,v}$ then all inequalities in (4) hold with equality and $(u^*, v^*)$ are called saddle point policies, $u^*$ is the saddle-point policy of the controller, $v^*$ is the saddle-point policy of the actor, and $R_\beta^{u^*,v^*}$ is denoted as the value of the game.

Two-person zero-sum games, with finitely many pure strategies for each player, are known to ohave a saddle point within the mixed strategies which can be computed using standard linear programs (see footnote at Section IV-B) For each player, there is a one-to-one correspondence between the class of its behavioral and its mixed strategies [3] such that for any policy of the other player, the expected utility under the mixed strategy and the equivalent behavioral one is the same. Thus, a saddle point exists within the behavioral policies as well.

Note that a pure policy for any player in this game is to select a particular action which corresponds to a specific pure policy in the original game. Thus, there is a one-to-one correspondence between the sets of pure policies for each player in the two games such that for each pure policy for a player in a game the corresponding pure policy for the same player in the other game takes the same actions if presented with the same information. Since a mixed-policy for any player in any game is a probability distribution over the pure policies, there is a similar one-to-one correspondence between the sets of mixed policies for each player in the two games. Thus, it follows from the previous paragraph that there is a one-to-one correspondence between the sets of mixed policies in the two-person zero-sum game and behavioral policies in the original game, such that for each behavioral (mixed) policy for any given player in the original (two-person zero-sum) game, the corresponding mixed (behavioral) policy for the same player in the two-person zero-sum (original) game has the same utility [3]. It follows using same arguments that similar correspondence exists between the sets of the behavioral policies in the two games. Thus, for notational simplicity, we use the same notations (e.g., u, v, etc.) to denote the mixed or behavioral policies in both games. The following theorem proves that a pair of policies constitute a saddle point for the two-person zero-sum game if and only if it constitutes a point-wise Nash equilibrium of the original game.

*Theorem 4.1:* A mixed or behavioral policy pair $(u^*, v^*)$ is a point-wise Nash equilibrium in the original game if and only if the corresponding mixed or behavioral policy pair $(u^*, v^*)$ is a saddle point pair in the two-person zero-sum game.

*Proof:* Assume that $(u^*, v^*)$ is a point-wise Nash equilibrium. We show that it is a saddle-point pair. From definition 4.3 and since there always exists a saddle-point pair in the two-person zero-sum game [3], the above is indeed the case if (i) $u^*$ minimizes $R_\beta(u, v^*)$ and (ii) $v^*$ maximizes $R_\beta(u^*, v)$. We show that (i) holds. Assume it does not. Then for some $u$, $R_\beta(u, v^*) < R_\beta(u^*, v^*)$. Hence, from (2), there exists some $\vec{x} \in \mathcal{K}^n$ such that $J_c^{u,v^*}(\vec{x}) > J_c^{u^*,v^*}(\vec{x})$ and $\beta(\vec{x}) > 0$. This contradicts the assumption that $(u^*, v^*)$ is a point-wise Nash equilibrium. Thus, (i) holds. Using (3), it can be similarly shown that (ii) holds as well. Thus, $(u^*, v^*)$ is a saddle-point pair.

Conversely, assume that $(u^*, v^*)$ is a saddle-point pair. We show that (i) in Definition 3.1 holds. Assume it does not. Then for some $\vec{x}$ and $u$, $J_c^{u,v^*}(\vec{x}) > J_c^{u^*,v^*}(\vec{x})$ and $\beta(\vec{x}) > 0$. Define the policy $w$ for the jammer as the one that coincides with $u$ if the channel state is $\vec{x}$ and that coincides otherwise with $u^*$. Then $R_\beta(w, v^*) < R_\beta(u^*, v^*)$. This contradicts the assumption that $(u^*, v^*)$ is a saddle-point. Thus, (i) holds. It can be similarly shown that (ii) holds as well. Thus, $(u^*, v^*)$ is a point-wise Nash equilibrium. ■

*Corollary 4.1:* A point-wise Nash-equilibrium $(u^*, v^*)$ exists in the original game.

The above corollary follows from Theorem 4.1 and the discussion between Definition 4.3 and Theorem 4.1.

Henceforth, we focus on the properties and computations of the saddle-point.

### B. *Computation of the saddle point*

We now investigate the computation of a saddle-point. It follows from standard results that a saddle point of a two-person zero-sum game can be computed using a linear program whose number of variables equal the number of pure strategies of a player and the number of constraints equal the number of pure strategies of the other player.[1] This may sound quite encouraging at first since solving LPs has polynomial complexity as a function of the number of decision variables and constraints. Neverthless, the computation is intractable due to the huge number of pure strategies $N_c$ of the Controller and $N_a$ of the Actor, given by

$$\boxed{N_c = \left(\sum_{i=0}^{k}\binom{n}{i}\right)^{K^n} \quad \text{and} \quad N_a = n^{\sum_{i=0}^{k}\binom{n}{i}K^{n-i}}.} \tag{5}$$

(5) is obtained as follows.

- The Controller's information has $K^n$ possible values, and for each such information it can choose $\sum_{i=0}^{k}\binom{n}{i}$ actions (note that $\sum_{i=0}^{k}\binom{n}{i}$ is the number of subsets of the components of cardinality at most $k$).
- The Actor's information has $\sum_{i=0}^{k}\binom{n}{i}K^{n-i}$ possible values, and for each such information it can choose $n$ actions.

Simplifying (5), the number of pure strategies of the controller (actor, respectively) in the original game is at least $\binom{n}{k}^{K^n}$ ($n^{\min(\binom{n}{(n/2)},K^{n/2})}$, respectively). The computation is therefore intractable even for moderate values of $n, K$.

Next, exploiting system characteristics, we obtain linear programs for computing the saddle point strategies such that their computation times are polynomials in $(K^n + k)\binom{n}{k}$. Note that this substantially reduces the computation time.

Henceforth, $u$ ($v$, respectively) are the behavioral policies of the controller (actor respectively).

*1) Saddle point for the controller:* The following linear program obtains a saddle-point policy for the controller.

$$\boxed{\begin{array}{rcl}
\text{LP-CONTROLLER:} & & \text{Min} \sum_{\vec{y}\in\mathcal{A}_c} z(\vec{y}) \ s.t. \\
z(\vec{y}) & \geq & \sum_{\vec{x}:\vec{y}\in\mathcal{A}_c(\vec{x})} \beta(\vec{x}) r(x_i) u(\vec{x})_{\vec{y}} \\
 & & \forall\ i \in \mathcal{N}, \vec{y} \in \mathcal{A}_c \\
\sum_{\vec{y}\in\mathcal{A}_c(\vec{x})} u(\vec{x})_{\vec{y}} & = & 1 \ \text{ for all } \vec{x} \in \mathcal{K}^n \\
u(\vec{x})_{\vec{y}} & \geq & 0 \ \forall \ \ \vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_c(\vec{x})
\end{array}}$$

*Theorem 4.2:* The optimum solution $\{u(\vec{x})_{\vec{y}}\}_{\vec{y}\in\mathcal{A}_c(\vec{x}),\vec{x}\in\mathcal{K}^n}$ of LP-CONTROLLER constitutes the saddle-point policy $u^*$ for the controller.

[1]For example, consider a matrix game whose entries are $R^{u,v}$ where player I (minimizing) chooses a row $u$ and player 2 (maximizing) chooses a column $v$. A saddle point policy for player 2 is obtained by maximizing $\left(\inf_u R^{u,v}\right)$. Then for any $u$, the value $z$ is smaller than or equal to $\max_v \left(R^{u,v}\right)$. Moreover the value is the largest constant with this property. The LP is thus $\max_{p\in\Delta(\mathcal{V}^P),z} z \ s.t. \ \ z \leq \sum_{v\in\mathcal{V}^P} p(v)R^{u,v}, \ \forall u \in \mathcal{U}^P$. ($\Delta(\mathcal{V}^P)$ are all probability measures over $\mathcal{V}^P$).

*Proof:*

From (3), for any $u, v, \beta, \in \mathcal{K}^n$

$$R_\beta^{u,v} = \sum_{\vec{y} \in \mathcal{A}_c} \mathrm{Pr}^{\beta,u}(\vec{y}) E_\beta^{u,v}[r(X_B)|\vec{Y} = \vec{y}].$$

Given $u \in \mathcal{U}$, consider a policy $v_u \in \mathcal{V}$ such that for each $\vec{y} \in \mathcal{A}_c$, $v_u(\vec{y})_j = 1$ for some $j$ such that $E_\beta^u[r(X_j)|\vec{Y} = \vec{y}] = \max_{i \in \mathcal{N}} E_\beta^u[r(X_i)|\vec{Y} = \vec{y}]$, and $v_u(\vec{y})_j = 0$, for other values of $j$ (i.e., under $v_u$ w.p. $1$ $B$ is a component $i$ that attains the above maximum and hence $v_u$ is the actor's best response to controller's strategy $u$). Note that

$$\max_{v \in \mathcal{V}} E_\beta^{u,v}[r(X_B)|\vec{Y} = \vec{y}] = \max_{i \in \mathcal{N}} E_\beta^u[r(X_i)|\vec{Y} = \vec{y}] = E_\beta^{u,v_u}[r(X_B)|\vec{Y} = \vec{y}], \ \forall \ \vec{y} \in \mathcal{A}_c.$$

Thus,

$$\sup_{v \in \mathcal{V}} R_\beta^{u,v} = \sum_{\vec{y} \in \mathcal{A}_c} \mathrm{Pr}^{\beta,u}(\vec{Y} = \vec{y}) \max_{i \in \mathcal{N}} E_\beta^u[r(X_i)|\vec{Y} = \vec{y}] = R_\beta^{u,v_u}. \tag{6}$$

Thus,

$$\underline{R}_\beta = \inf_{u \in \mathcal{U}} R_\beta^{u,v_u}. \tag{7}$$

Next,

$$\begin{aligned} E_\beta^u[r(X_i)|\vec{Y} = \vec{y}] &= \sum_{\vec{x} \in \mathcal{K}^n} E_\beta^u[r(X_i)|\vec{Y} = \vec{y}, \vec{X} = \vec{x}] \mathrm{Pr}^{\beta,u}(\vec{X} = \vec{x}|\vec{Y} = \vec{y}) \\ &= \sum_{\vec{x} \in \mathcal{K}^n} r(x_i) \mathrm{Pr}^{\beta,u}(\vec{Y} = \vec{y}|\vec{X} = \vec{x}) \mathrm{Pr}^{\beta,u}(\vec{X} = \vec{x}) / \mathrm{Pr}^{\beta,u}(\vec{Y} = \vec{y}) \\ &= \sum_{\vec{x} \in \mathcal{K}^n} r(x_i) u(\vec{x})_{\vec{y}} \beta(\vec{x}) / \mathrm{Pr}^{\beta,u}(\vec{Y} = \vec{y}). \end{aligned}$$

From,

$$E_\beta^u[r(X_i)|\vec{Y} = \vec{y}] \mathrm{Pr}^{\beta,u}(\vec{Y} = \vec{y}) = \sum_{\vec{x} \in \mathcal{K}^n} r(x_i) u(\vec{x})_{\vec{y}} \beta(\vec{x}).$$

Thus, from (6) and (7),

$$R_\beta^{u,v_u} = \sum_{\vec{y} \in \mathcal{A}_c} \max_{i \in \mathcal{N}} \sum_{\vec{x} \in \mathcal{K}^n} r(x_i) u(\vec{x})_{\vec{y}} \beta(\vec{x})$$

$$\text{and } \underline{R}_\beta = \inf_{u \in \mathcal{U}} \sum_{\vec{y} \in \mathcal{A}_c} \max_{i \in \mathcal{N}} \sum_{\vec{x} \in \mathcal{K}^n} r(x_i) u(\vec{x})_{\vec{y}} \beta(\vec{x}).$$

Now, consider a feasible solution $(u, z)$ of the LP-CONTROLLER, such that $z$ is chosen so as to minimize the value of the objective function subject to choosing $u$. The value of the objective function is $R_\beta^{u,v_u}$ for any such pair.

Thus, if $u^O$ is the optimum solution of LP-CONTROLLER, $\underline{R}_\beta = R_\beta^{u^O,v_{u^O}}$. Thus, from (6), $\underline{R}_\beta = \sup_{v \in \mathcal{V}} R_\beta^{u^O,v}$. Now, since a saddle-point always exists, it follows from Definition 4.3 that any $u' \in \mathcal{U}$ for which $\underline{R}_\beta = \sup_{v \in \mathcal{V}} R_\beta^{u',v}$ constitutes a saddle-point policy of the controller. Thus, $u^O$ constitutes a saddle-point policy of the controller. ■

The following corollary proves an intuitive property of saddle point policies of the controller, and will help reduce the number of variables of LP-CONTROLLER.

*Corollary 4.2:* There exists a saddle-point policy $u^*$ of the controller in which it always conceals the states of $k$ components.

*Proof:* Consider an optimal solution $(u, z)$ for which there exists $\vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_c(\vec{x})$ such that $u(\vec{x})_{\vec{y}} > 0$ and $|a(\vec{y})| < k$. Clearly, $z(\vec{y}') = \max_{i \in \mathcal{N}} \sum_{\vec{x}' : \vec{y}' \in \mathcal{A}_c(\vec{x}')} \beta(\vec{x}') r(x'_i) u(\vec{x}')_{\vec{y}'} \ \forall \ \vec{y}' \in \mathcal{A}_c$. Consider a sub-vector of $\vec{y}$, $\vec{w}$, such that $|a(\vec{w})| = k$. Note that $\vec{w} \in \mathcal{A}_c(\vec{x}')$ for all $\vec{x}'$ such that $\vec{y} \in \mathcal{A}_c(\vec{x}')$.

Consider a new feasible solution $(u', z')$ such that $u'(\vec{x}')_{\vec{y}'} = u(\vec{x}')_{\vec{y}'}$ for all $\vec{x}', \vec{y}' \notin \{\vec{y}, \vec{w}\}$, $u'(\vec{x}')_{\vec{y}} = 0$, for all $\vec{x}'$, $u'(\vec{x}')_{\vec{w}} = u(\vec{x}')_{\vec{y}} + u(\vec{x}')_{\vec{w}}$ for all $\vec{x}'$, and
$z'(\vec{y}') = \max_{i \in \mathcal{N}} \sum_{\vec{x}' : \vec{y}' \in \mathcal{A}_c(\vec{x}')} \beta(\vec{x}') r(x'_i) u'(\vec{x}')_{\vec{y}'} \ \forall \ \vec{y}' \in \mathcal{A}_c$. (Here, $(u', z')$ is feasible since $\vec{w} \in \mathcal{A}_c(\vec{x}')$ for all $\vec{x}'$ such that $\vec{y} \in \mathcal{A}_c(\vec{x}')$). Also,

$$\left\{ \begin{array}{c} \vec{y}' : u'(\vec{x}')_{\vec{y}'} > 0 \text{ for some } \vec{x}' \in \mathcal{K}^n, \text{ and} \\ |a(\vec{y}')| < k \} \subset \{ \vec{y}' : u(\vec{x}')_{\vec{y}'} > 0 \text{ for some } \vec{x}' \in \mathcal{K}^n, \text{ and } |a(\vec{y}')| < k \end{array} \right\}. \tag{8}$$

Clearly, $z'(\vec{y}') = z(\vec{y}')$ for all $\vec{y}' \notin \{\vec{y}, \vec{w}\}$, $z'(\vec{y}) = 0$ and $z'(\vec{w}) \leq z(\vec{w}) + z(\vec{y})$. Thus, the value of the objective function under $(u', z')$ is not higher than that under $(u, z)$. Thus, $(u', z')$ is also an optimal solution of LP-CONTROLLER. Thus, due to (8), repeating this process we obtain an optimal solution $(u^*, z^*)$ of LP-CONTROLLER such that $\{\vec{y}' : u^*(\vec{x}')_{\vec{y}'} > 0 \text{ for some } \vec{x}' \in \mathcal{K}^n, \text{ and } a(\vec{y}') < k\} = \phi$. The result follows. ■

Now, consider the following definition.

*Definition 4.4:* Let $\mathcal{A}_{c,k} = \{\vec{y} : |a(\vec{y})| = k, \vec{y} \in \mathcal{A}_c\}$ and $\mathcal{A}_{c,k}(\vec{x}) = \mathcal{A}_{c,k} \cap \mathcal{A}_{c,k}(\vec{x})$.

Due to Corollary 4.2, we only need to consider the variables $z(\vec{y})$ such that $|a(\vec{y})| = k$. Also, since for any $\vec{y}$ and $\vec{x}$ such that $\vec{y} \in \mathcal{A}_c(\vec{x})$, $x_i = y_i$ for any $i \in \mathcal{N} \setminus a(\vec{y})$, for any $\vec{y}$, $i \in \mathcal{N} \setminus a(\vec{y})$, and $y_i < \max_{j \in \mathcal{N} \setminus a(\vec{y})} y_j$, the value of the right hand side of the lower bound constraint in LP-CONTROLLER is less than or equal to that for $\vec{y}$, $l \in \mathcal{N} \setminus a(\vec{y})$, and $y_l = \max_{j \in \mathcal{N} \setminus a(\vec{y})} y_j$ irrespective of the choice of $\beta, u$. Thus, these constraints can be ignored as well, and LP-CONTROLLER can be described as follows.

$$\begin{array}{llll}
\text{LP-CONTROLLER:} & & \text{Minimize } \sum_{\vec{y} \in \mathcal{A}_c} z(\vec{y}) \quad s.t. & \\
z(\vec{y}) & \geq & \max_{i \in \mathcal{N} \setminus a(\vec{y})} r(y_i) \sum_{\vec{x} : \vec{y} \in \mathcal{A}_c(\vec{x})} \beta(\vec{x}) u(\vec{x})_{\vec{y}}, & \forall \ \vec{y} \in \mathcal{A}_{c,k} \\
z(\vec{y}) & \geq & \sum_{\vec{x} : \vec{y} \in \mathcal{A}_c(\vec{x})} \beta(\vec{x}) r(x_i) u(\vec{x})_{\vec{y}}, & \forall \ i \in a(\vec{y}), \vec{y} \in \mathcal{A}_{c,k}, \\
\sum_{\vec{y} \in \mathcal{A}_{c,k}(\vec{x})} u(\vec{x})_{\vec{y}} & = & 1 & \forall \vec{x} \in \mathcal{K}^n \\
u(\vec{x})_{\vec{y}} & \geq & 0 & \forall \ \ \vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})
\end{array}$$

Henceforth, we will use this description of LP-CONTROLLER. Note that LP-CONTROLLER has $K^n \binom{n}{k}$ variables and $(k+1)\binom{n}{k} + K^n + K^n \binom{n}{k}$ constraints. Thus, the computation time of this linear program is polynomial in $(K^n + k)\binom{n}{k}$.

*2) Saddle point for the actor:* The following linear program obtains a saddle-point policy for the actor.

$$\begin{array}{rcl}
\text{LP-ACTOR:} & & \text{Maximize } \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x})z(\vec{x}) \\
z(\vec{x}) & \leq & \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i) \ \forall \vec{y}\in\mathcal{A}_c(\vec{x}), \vec{x}\in\mathcal{K}^n \\
v(\vec{y})_j & \geq & 0 \ \forall \ \vec{y}, j\in\mathcal{N} \\
\sum_{j\in\mathcal{N}} v(\vec{y})_j & = & 1 \ \forall \ \vec{y}\in\mathcal{A}_c
\end{array}$$

*Theorem 4.3:* The optimum solution $\{v(\vec{y})_i\}_{i\in\mathcal{N},\vec{y}\in\mathcal{A}_c}$ of LP-ACTOR constitutes the saddle-point policy $v^*$ for the actor.

*Proof:* From (2), for any $u, v, \beta$,

$$R_\beta^{u,v} = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) E_\beta^{u,v}[r(x_B)|\vec{X}=\vec{x}] = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) \sum_{\vec{y}\in\mathcal{A}_c(\vec{x})} u(\vec{x})_{\vec{y}} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i).$$

Consider a policy $u_v \in \mathcal{U}$ such that for each $\vec{x}\in\mathcal{K}^n$, $u_v(\vec{x})_{\vec{y}} = 1$ for some $\vec{y}\in\mathcal{A}_c(\vec{x})$ such that $\sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i) = \min_{\vec{t}\in\mathcal{A}_c(\vec{x})} \sum_{i\in\mathcal{N}} v(\vec{t})_i r(x_i)$ and $w(\vec{x})_{\vec{y}} = 0$, for all other $\vec{y}\in\mathcal{A}_c(\vec{x})$.
Since $u(\vec{x})$ is a probability distribution on $\mathcal{A}_c(\vec{x})$,

$$\inf_{u\in\mathcal{U}} \sum_{\vec{y}\in\mathcal{A}_c(\vec{x})} u(\vec{x})_{\vec{y}} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i) = \min_{\vec{y}\in\mathcal{A}_c(\vec{x})} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i) = \sum_{\vec{y}\in\mathcal{A}_c(\vec{x})} u_v(\vec{x})_{\vec{y}} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i).$$

Thus,

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v} = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) \min_{\vec{y}\in\mathcal{A}_c(\vec{x})} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i) = R_\beta^{u_v,v} \tag{9}$$

(i.e., $u_v$ is the controller's best response to actor's $v$). Now, $\overline{R}_\beta = \sup_{v\in\mathcal{V}} R_\beta^{u_v,v}$. Thus, from (9),

$$\overline{R}_\beta = \sup_{v\in\mathcal{V}} \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) \min_{\vec{y}\in\mathcal{A}_c(\vec{x})} \sum_{i\in\mathcal{N}} v(\vec{y})_i r(x_i).$$

Now, consider a feasible solution $(v, z)$ of the LP-ACTOR, such that $z$ is chosen so as to maximize the value of the objective function subject to choosing $v$. The value of the objective function is $R_\beta^{u_v,v}$ for any such pair.

Thus, if $v^O$ is the optimum solution of LP-ACTOR, $\overline{R}_\beta = R_\beta^{u_{v^O},v^O}$. Thus, from (9), $\overline{R}_\beta = \inf_{u\in\mathcal{U}} R_\beta^{u,v^O}$. Now, since a saddle-point always exists, it follows from Definition 4.3 that any $v' \in \mathcal{V}$ for which $\overline{R}_\beta = \inf_{u\in\mathcal{U}} R_\beta^{u,v'}$ constitutes a saddle-point policy of the actor. Thus, $v^O$ constitutes a saddle-point policy of the actor. ■

*Definition 4.5:* A policy $v\in\mathcal{V}$ of an actor is said to be *sensible* if it never selects a component whose state has been revealed and which is in a state that is lower than the highest state among the states of all components whose states have been revealed (i.e., $v(\vec{y})_i = 0$ if $i \notin a(\vec{y})$ and $y_i \neq \max_{j\in\mathcal{N}\setminus a(\vec{y})} y_j$).

*Observation 1:* Note that $R_\beta^{u,v^1} = R_\beta^{u,v^2}$ for any $u\in\mathcal{U}, v^1, v^2\in\mathcal{V}$ such that $v^1(\vec{y})_i = v^2(\vec{y})_i$ for any $i\in a(\vec{y})$ and $\sum_{i:i\notin a(\vec{y}),y_i=j} v^1(\vec{y})_i = \sum_{i:i\notin a(\vec{y}),y_i=j} v^2(\vec{y})_i$ for each $j\in\{0,\ldots,K-1\}$.

The following corollary proves an intuitive property of saddle point policies of the actor, and will help reduce the number of variables of LP-ACTOR.

*Corollary 4.3:* There exists a sensible saddle-point policy $v^*$ of the actor.

*Proof:* Note that for any $i \in \mathcal{N} \setminus a(\vec{y})$, $x_i = y_i$ if $\vec{x}$ is the system state vector (i.e., if $\vec{y} \in \mathcal{A}_c(\vec{x})$). Thus, the first constraint in LP-ACTOR can be written as $z(\vec{x}) \leq \gamma(\vec{y}) + \sum_{i\in a(\vec{y})} v(\vec{y})_i r(x_i)$ for all $\vec{y} \in \mathcal{A}_c(\vec{x})$, where $\gamma(\vec{y}) = \sum_{i\in\mathcal{N}\setminus a(\vec{y})} v(\vec{y})_i r(y_i)$. Given a feasible solution $v$, consider another feasible solution $v'$ such that $v(\vec{y})'_i = v(\vec{y})_i$ if $i \in a(\vec{y})$, $v(\vec{y})'_i = \sum_{j\in\mathcal{N}\setminus a(\vec{y})} v(\vec{y})_j$ for some $i$ such that $i \in \mathcal{N} \setminus a(\vec{y})$ and $y_i = \max_{j\in\mathcal{N}\setminus a(\vec{y})} y_j$, and $v(\vec{y})'_i = 0$ otherwise. Note that $v'$ is a feasible solution as well which satisfies the property required in the corollary, and the maximum value of the objective function for $v$ (the maximization is w.r.t. $z$) is not higher than that for $v'$. This is because $\gamma(\vec{y}') \geq \gamma(\vec{y})$ for each $\vec{y}$ and $\sum_{i\in\mathcal{N}\setminus a(\vec{y})} v(\vec{y}')_i r(x_i) = \sum_{i\in\mathcal{N}\setminus a(\vec{y})} v(\vec{y})_i r(x_i)$ for each $\vec{x}, \vec{y}$. The result follows. ■

.

Due to Corollaries 4.2 and 4.3 and the above observation, we only consider variables $v(\vec{y})$ such that $|a(\vec{y})| = k$ and need to determine the components $v(\vec{y})_j$ such that $j \in a(\vec{y}) \cup j \in (\mathcal{N} \setminus a(\vec{y})) \cap \{l : y_l = \max_{m\in\mathcal{N}\setminus a(\vec{y})} y_m\}$. Thus, LP-ACTOR can be re-written as follows.

LP-ACTOR:

$$
\begin{aligned}
&\text{Maximize} \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) z(\vec{x}) \\
z(\vec{x}) &\leq \left(1 - \sum_{i\in a(\vec{y})} v(\vec{y})_i\right) \max_{i\in\mathcal{N}\setminus a(\vec{y})} r(y_i) \\
&+ \sum_{i\in a(\vec{y})} v(\vec{y})_i r(x_i) \ \forall\ \vec{y} \in \mathcal{A}_{c,k}(\vec{x}), \vec{x} \in \mathcal{K}^n \\
v(\vec{y})_j &\geq 0, \quad \forall\ j \in a(\vec{y}),\ \vec{y}\ \in \mathcal{A}_{c,k} \\
\sum_{j\in a(\vec{y})} v(\vec{y})_j &\leq 1, \quad \forall\ \vec{y} \in \mathcal{A}_{c,k}
\end{aligned}
$$

## C. Discussion

Saddle point policies (and hence the related equilibrium strategies in the original game) enjoy some robustness properties related to the order of moves.

More precisely, we may interpret the upper value as corresponding to the problem where first the Controller chooses a strategy $u$ without knowing the Actor's strategy, and then the Actor chooses a strategy knowing the choice of the Controller.[2] We then say that the Actor has an information advantage over the Controller. A symmetric interpretation holds for the lower value. The fact that the upper and lower values are equal implies that the order of choices of strategies does not matter. A player thus does not gain by having an information advantage. $u^*$ (resp. $v^*$) is a saddle point policy for the Actor (resp. the Controller) *if and only if* it maximizes $\left(\inf_{u\in\mathcal{U}} R_\beta^{u,v}\right)$

[2]The latter is seen by the fact that the term $R_\beta^{u,v}$ maximized by the Actor in the definition of the upper value, is a function of the Controller's strategy $u$. On the other hand the term minimized by the Controller in the definition of the upper value is $\left(\sup_{v\in\mathcal{V}} R_\beta^{u,v}\right)$ and is not a function of $v$.

(it minimizes $\left(\sup_{v\in\mathcal{V}} R_\beta^{u,v}\right)$, resp.); it thus has to be optimal in the game in which the other player has the information advantage.

We illustrate the important implications of this property.. Consider the first motivating example concerning jamming. We show that the GC policy need not be a saddle point policy. If it were, then it should have the best performance even when the transmitter (i.e. the Actor) knows that policy before choosing its own. But knowing the policy of the jammer, the transmitter knows that the quality of any channel whose state has been concealed is at least as good as that of a channel whose state has been revealed, and thus, the best action for him is to select a channel whose state has been concealed. Now, if instead of using the GC policy, the jammer reveals the states of some channels whose transmission qualities are better than those whose states he conceals, the transmitter may be confused regarding the choice of the channel, and is therefore more likely to make a poor selection. Thus we may expect the GC policy not to be a saddle point. We shall illustrate this point in Section VI through numerical examples.

## V. Performance guarantees using polynomial time computation

We have proved that the saddle-point policies can be obtained by solving linear programs whose number of variables is exponential in $n$ and polynomial in $K$. Using fast algorithms for solving linear programs, the saddle points can now be computed for moderate values of $n$ but the computation will still be intractable for large $n$. We therefore focus on obtaining provable performance guarantees using polynomial time computable policies. We first consider the important special case where the system consists of few classes of components such that all components in each class are statistically identical and the number of states $K$ is small (note that each class may have a large number of components and therefore $n$ can be large). We prove that the saddle point policies can be computed in polynomial time in such systems (Section V-A). Specifically, when the system consists of $M$ classes of components, the saddle point policies can be obtained by solving linear programs with $O(n^{2KM})$ variables and $O(n^{2KM})$ constraints for arbitrary $n, K, k, M$. Thus, when all components are statistically identical ($M = 1$), the computation time is polynomial in $n$, but exponential in $K$ (note that $K$ is small in most systems). The result is interesting given that some intuitive policies do not constitute saddle point policies even when all components are statistically identical (Lemma 3.1). We next show that provable approximation guarantees can be obtained in arbitrary systems using some simple policies that can be computed in almost linear time (either $O(n)$ or $O(n\log n)$) time (Section V-B).

### *A. Polynomial time computation of saddle points in systems with constant number of classes of components and constant number of states*

We first formally define the notion of classes of components and motivate the investigation of the special case where the system consists of a few classes and few states for the components. We subsequently present a key technical property (Theorem 5.1) for systems with arbitrary number of classes of components and states (Section V-A.1). Using this property and some additional terminologies (Section V-A.2), we show how saddle point policies for the controller and actor can be computed in polynomial time when $K, M$ are constant (Sections V-A.3 and V-A.4).

*Definition 5.1:* Let $\vec{x}^{i,j} \in \mathcal{K}^n$ be obtained by interchanging the $i$th and the $j$th components of $\vec{x} \in \mathcal{K}^n$. Let $\vec{y}^{i,j} \in \mathcal{A}_c$ be obtained as follows: (a) if $i, j \notin a(\vec{y})$ $a(\vec{y}^{i,j}) = a(\vec{y})$, $y_i^{i,j} = y_j$,

$y^{i,j}_j = y_i$, $y^{i,j}_l = y_l$, $l \notin a(\vec{y}) \cup \{i,j\}$ (b) if $i \in a(\vec{y})$, $j \notin a(\vec{y})$, then $a(\vec{y}^{i,j}) = a(\vec{y}) \cup \{j\} \setminus \{i\}$, $y^{i,j}_i = y_j$, $y^{i,j}_l = y_l$, $l \notin a(\vec{y}^{i,j}) \cup \{i\}$, (c) if $i \notin a(\vec{y})$, $j \in a(\vec{y})$, then $a(\vec{y}^{i,j}) = a(\vec{y}) \cup \{i\} \setminus \{j\}$, $y^{i,j}_j = y_i$, $y^{i,j}_l = y_l$, $l \notin a(\vec{y}^{i,j}) \cup \{j\}$, (d) $\vec{y}^{i,j} = \vec{y}$, otherwise.

*Definition 5.2:* Components $i, j$ are said to be in the same class if $\beta(\vec{x}) = \beta(\vec{x}^{i,j})$ for all $\vec{x} \in \mathcal{K}^n$. Note that the membership in the same class constitutes an equivalence relation and hence the classes constitute a partition of $\mathcal{N}$. Let the system consist of $M$ classes, where $1 \le M \le n$. The classes are numbered as $1, \ldots, M$, and $n_i$ components are in class $i$ where $\sum_{i=1}^{M} n_i = n$. Let $a(\vec{y}, i)$ be the set of components in class $i$ that have been concealed when the actor's information is $\vec{y}$. Note that $a(\vec{y}) = \cup_{i=1}^{M} a(\vec{y}, i)$.

Note that $M$ can be determined from $\beta$ and hence is also known to both players.

Several systems have large number of components but small or moderate number of classes of components and states. For example, cognitive radio networks may have large number of channels, but often, many of these channels are statistically identical, and hence the number of classes of channels is often substantially less than the number of channels. Also, the total number of states of these channels is likely to be moderate as well. Next, consider the gambling example (Section I-B). The cards that have the same color constitute the same class as the distributions of the random numbers are statistically identical for all cards of the same color. Usually, the number of colors, or more generally number of types of cards (e.g., aces, jokers, etc.) is small although the number of cards can be large.

We first present a key property of systems with arbitrary number of classes of components.

*1) Symmetry among components in the same class:*

*Definition 5.3:* Let $u, v$ be behavioral policies of the controller and actor respectively and $i, j \in \mathcal{N}$. The mirror image w.r.t $(i, j)$ of the policy $u$ ($v$, respectively), $u^{i,j}$ ($v^{i,j}$, respectively) is a policy obtained as follows: $u^{i,j}(\vec{x})_{\vec{y}} = u(\vec{x}^{i,j})_{\vec{y}^{i,j}}$ ($v^{i,j}(\vec{y})_i = v(\vec{y}^{i,j})_j$ and $v^{i,j}(\vec{y})_j = v(\vec{y}^{i,j})_i$, respectively).

Intuitively, $u^{i,j}$ ($v^{i,j}$, respectively) treat $i$ as $j$ and $j$ as $i$.

*Definition 5.4:* A policy $u \in \mathcal{U}$ ($v \in \mathcal{V}$, respectively) is said to be symmetric w.r.t. $(i, j)$ if $u = u^{i,j}$ ($v = v^{i,j}$, respectively). A policy $u \in \mathcal{U}$ ($v \in \mathcal{V}$, respectively) is said to be symmetric if it is symmetric w.r.t. each pair of components that are in the same class. Let $\mathcal{U}^s \subset \mathcal{U}$ and $\mathcal{V}^s \subset \mathcal{V}$ be the classes of all symmetric policies of the controller and actor respectively.

The following theorem shows the existence of a symmetric saddle-point.

*Theorem 5.1:* There exists a symmetric policy $u \in \mathcal{U}^s$ ($v \in \mathcal{V}^s$, respectively) for the controller (actor, respectively) such that $u$ ($v$, respectively) constitutes a saddle-point of the controller (actor, respectively).

*Proof:* We prove the theorem for the controller, and the proof for the actor is similar. Let $\mathcal{S}^u \subseteq \mathcal{N} \times \mathcal{N}$ be the set of tuples $(a, b)$ such that $a, b$ are in the same class and $u$ is not symmetric w.r.t. $a, b$. Clearly, $\mathcal{S}^u = \phi$ iff $u \in \mathcal{U}^s$. If there exists an optimal solution $u$ of LP-CONTROLLER such that $\mathcal{S}^u = \phi$, then the result follows from Theorem 4.2. So, from Theorem 4.2, it is sufficient to prove that if there exists an optimal solution $u$ of LP-CONTROLLER such that $\mathcal{S}^u \neq \phi$, there exists an optimal solution $\hat{u}$ of LP-CONTROLLER such that $\mathcal{S}^{\hat{u}} \subset \mathcal{S}^u$. First note that $u^{a,b}$ is an optimal solution of LP-CONTROLLER for any pair of components $a, b$ that are in the same class. Now, consider an arbitrary pair of components $i, j \in \mathcal{S}^u$, and a policy $\hat{u} \in \mathcal{U}$ such that

$\hat{u}(\vec{x})_{\vec{y}} = u(\vec{x})_{\vec{y}} + u^{i,j}(\vec{x})_{\vec{y}}$ for each $\vec{x} \in \mathcal{K}^n$ and $\vec{y} \in \mathcal{A}_c(\vec{x})$. Thus, clearly, $\hat{u}$ is an optimal solution of LP-CONTROLLER. Also, note that $\mathcal{S}^{\hat{u}} = \mathcal{S}^u \setminus \{(i,j)\}$. The result follows. ■

Using Theorem 5.1, we show that the computation time for LP-CONTROLLER and LP-ACTOR can be substantially reduced when $M$ and $K$ are small.

*2) Additional Terminologies:*

*Definition 5.5:* Let $\mathbf{l}(\vec{x})$ be a matrix with $M$ rows and $K$ columns and entries in $0, \ldots, n$ such that $l(\vec{x})_{i,j}$ is the number of components of $\vec{x}$ that are in class $i$ and state $j$. Let $\mathcal{L} = \{\mathbf{l} : \mathbf{l}(\vec{x}) = \mathbf{l}, \vec{x} \in \mathcal{K}^n\}$. Let $\mathbf{m}(\vec{y})$ be a matrix with $M$ rows and $K$ columns with entries in $0, \ldots, a(\vec{y})$ such that $m(\vec{y})_{i,j}$ is the number of components of $\vec{y}$ that are in class $i$ and state $j$. Let $\mathcal{M}_{\vec{x}} = \{\mathbf{m} : \mathbf{m}(\vec{y}) = \mathbf{m}, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})\}$. Note that $\mathcal{M}_{\vec{x}^1} = \mathcal{M}_{\vec{x}^2}$ if $\mathbf{l}(\vec{x}^1) = \mathbf{l}(\vec{x}^2)$. Let $\mathcal{M}_{\mathbf{l}} = \cup_{\vec{x} \in \mathcal{K}^n, \mathbf{l}(\vec{x}) = \mathbf{l}} \mathcal{M}_{\vec{x}}$. Let $\mathcal{M} = \cup_{\mathbf{l} \in \mathcal{L}} \mathcal{M}_{\mathbf{l}}$.

With slight abuse of notation, we have used $\mathbf{l}, \mathbf{m}$ to denote both the functions and the values of the functions as well - the implication of specific usages are clear from the context.

Note that (a) $|\{\vec{y} : \mathbf{m}(\vec{y}) = \mathbf{m}, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})\}|$ depends on $\vec{x}$ only through $\mathbf{l}(\vec{x})$. and (b) $|\{\vec{x} : \mathbf{l}(\vec{x}) = \mathbf{l}, \vec{y} \in \mathcal{A}_c(\vec{x})\}|$ depends on $\vec{y}$ only through $\mathbf{m}(\vec{y})$. Thus, we can introduce the following definitions.

*Definition 5.6:* Let $\Theta_1(\mathbf{l}, \mathbf{m})$ denote for one (representative) $\vec{x}$ such that $\mathbf{l}(\vec{x}) = \mathbf{l}$ the number of $\vec{y}$ in $\mathcal{A}_{c,k}(\vec{x})$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$. Let $\Theta_2(\mathbf{l}, \mathbf{m})$ denote is the number of system state vectors $\vec{x}$ such that (a) $\mathbf{l}(\vec{x}) = \mathbf{l}$ and (b) $\vec{y} \in \mathcal{A}_c(\vec{x})$ for one (representative) $\vec{y}$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$. Let $\Theta_3(\mathbf{m}) = |\vec{y} \in \mathcal{A}_{c,k} : \mathbf{m}(\vec{y}) = \mathbf{m}|$, and $\Theta_4(\mathbf{l}) = |\vec{x} \in \mathcal{K}^n : \mathbf{l}(\vec{x}) = \mathbf{l}|$.

Note that both $\Theta_2(\mathbf{l}, \mathbf{m})\Theta_3(\mathbf{m})$ and $\Theta_1(\mathbf{l}, \mathbf{m})\Theta_4(\mathbf{l})$ constitute the number of tuples $(\vec{x}, \vec{y})$ such that $\vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})$ and $\mathbf{l}(\vec{x}) = \mathbf{l}, \mathbf{m}(\vec{y}) = \mathbf{m}$. Thus,

$$\Theta_2(\mathbf{l}, \mathbf{m})\Theta_3(\mathbf{m}) = \Theta_1(\mathbf{l}, \mathbf{m})\Theta_4(\mathbf{l})$$

*Definition 5.7:* Let

$$\begin{aligned} R_1(\mathbf{m}) &= \max_{j : m_{i,j} > 0} r(j), \\ \text{and } R_2(\mathbf{l}, \mathbf{m}, i) &= \sum_{j=0}^{K-1} r(j) \frac{l_{i,j} - m_{i,j}}{n_i - \sum_{j=0}^{K-1} m_{i,j}}. \end{aligned}$$

Note that $R_1(\mathbf{m})$ is the expected reward the actor obtains when its information is $\vec{y}$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$ and it selects a component whose state has been revealed and whose state is the highest among those of the components whose states have been revealed. Also, $R_2(\mathbf{l}, \mathbf{m}, i)$ is the expected reward the actor obtains when its information is $\vec{y}$ such that $m(\vec{y}) = \mathbf{m}$, the system state is $\vec{x}$ such that $\mathbf{l}(\vec{x}) = \mathbf{l}$ and it selects a component of class $i$ uniformly among $a(\vec{y}, i)$.

*Definition 5.8:* Let $\mathcal{C}(\mathbf{m})$, $1 \le |\mathcal{C}(\mathbf{m})| \le \min(k, M)$, be the set of classes for which at least one component's state has been concealed when the actor's information $\vec{y}$ is such that $\mathbf{m}(\vec{y}) = \mathbf{m}$. Let $\Phi(\mathbf{m}, i)$ be the number of components of class $i$ that have been concealed when the actor's information $\vec{y}$ is such that $\mathbf{m}(\vec{y}) = \mathbf{m}$. Note that $\Phi(\mathbf{m}, i) = \sum_{j=0}^{K-1} m_{i,j}$, and $|\mathcal{C}(\mathbf{m})| = \sum_{i=1}^{M} \min\left(\Phi(\mathbf{m}, i), 1\right)$.

Finally, note that since $\beta(\vec{x}) = \beta(\vec{x}^{i,j})$ for all $i, j$ that are in the same class, $\beta(\vec{x}^1) = \beta(\vec{x}^2)$ if $\mathbf{l}(\vec{x}^1) = \mathbf{l}(\vec{x}^2)$.

*Definition 5.9:* Let $\beta'(\mathbf{l})$ denote $\beta(\vec{x})$ for some (representative) $\vec{x} \in \mathcal{K}^n$ such that $\mathbf{l}(\vec{x}^1) = \mathbf{l}$, and $\beta''(\mathbf{l}) = \Theta_4(\vec{l})\beta'(\vec{l})$.

*3) Polynomial time computation of saddle point policy of controller for constant $K, M$:* We now consider the simplification of LP-CONTROLLER.

Note that $u$ is symmetric if and only if $u(\vec{x}^1)_{\vec{y}^1} = u(\vec{x}^2)_{\vec{y}^2}$ whenever the following conditions hold: (a) $\mathbf{l}(\vec{x}^1) = \mathbf{l}(\vec{x}^2)$, (b) $\mathbf{m}(\vec{y}^1) = \mathbf{m}(\vec{y}^2)$ (c) $\vec{y}^1 \in \mathcal{A}_c(\vec{x}^1)$, $\vec{y}^2 \in \mathcal{A}_c(\vec{x}^2)$. Let $u'(\mathbf{l})_{\mathbf{m}}$ denote $u(\vec{x})_{\vec{y}}$ for some (representative) $\vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})$ such that $\mathbf{l}(\vec{x}) = \mathbf{l}$, $\mathbf{m}(\vec{y}) = \mathbf{m}$. Thus, each $u \in \mathcal{U}^s$ is uniquely described by $u^s(\mathbf{l})_{\mathbf{m}}$ where $u^s(\mathbf{l})_{\mathbf{m}} = \Theta_1(\mathbf{l}, \mathbf{m})u'(\mathbf{l})_{\mathbf{m}}$.

We now state LP-CONTROLLER-CLASS that computes $\{u^s(\mathbf{l})_{\mathbf{m}}\}$ for the symmetric saddle point strategy of the controller.

$$
\boxed{
\begin{array}{rcl}
\text{LP-CONTROLLER-CLASS:} & & \text{Minimize} \sum_{\mathbf{m}\in\mathcal{M}} \eta(\vec{m}) \quad s.t. \\
\forall\ \mathbf{m} \in \mathcal{M},\ \eta(\mathbf{m}) & \geq & R_1(\mathbf{m} \in \mathcal{M}) \sum_{\mathbf{l}:\mathbf{m}\in\mathcal{M}_{\mathbf{l}}} \beta''(\mathbf{l})u^s(\mathbf{l})_{\mathbf{m}} \\
\eta(\mathbf{m}) & \geq & \sum_{\mathbf{l}:\mathbf{m}\in\mathcal{M}_{\mathbf{l}}} \beta''(\mathbf{l})u(\mathbf{l})_{\mathbf{m}} R_2(\mathbf{l}, \mathbf{m}, i) \\
 & & \forall\ \mathbf{m} \in \mathcal{M}, i \in \mathcal{C}(\mathbf{m}) \\
\sum_{\mathbf{m}\in\mathcal{M}_{\mathbf{l}}} u(\mathbf{l})_{\mathbf{m}} & = & 1 \quad \text{for all } \mathbf{l} \in \mathcal{L} \\
u(\mathbf{l})_{\mathbf{m}} & \geq & 0 \quad \forall \quad \mathbf{m} \in \mathcal{M}_{\mathbf{l}}, \mathbf{l} \in \mathcal{L}.
\end{array}
}
\tag{10}
$$

*Theorem 5.2:* The optimum solution $\{u^s(\mathbf{l})_{\mathbf{m}}\}_{\mathbf{m}\in\mathcal{M}_{\mathbf{l}},\mathbf{l}\in\mathcal{L}}$ of LP-CONTROLLER-CLASS constitutes a policy $u^*$ for the controller such that $u^* \in \mathcal{U}^s$ and $u^*$ is a saddle-point policy for the controller.

*Proof:* Consider the description of LP-CONTROLLER at the end of Section IV-B.1, and restrict the feasible solutions $u$ to $\mathcal{U}^s$. From Theorem 5.1, the optimal solution of LP-CONTROLLER constitutes a saddle point even with this restriction, and the optimal solution is clearly a symmetric strategy for the controller. It is therefore sufficient to show that there is a one-to-one correspondence between the set of optimal solutions of LP-CONTROLLER-CLASS to that of LP-CONTROLLER with the above restriction.

Consider LP-CONTROLLER and $u \in \mathcal{U}^s$. Let $\mathcal{L}(\vec{y}) = \{\mathbf{l} : \mathbf{l}(\vec{x}) = \mathbf{l}$ for some $\vec{x}$ s.t. $\vec{y} \in \mathcal{A}_c(\vec{x})\}$. Note that $\mathcal{L}(\vec{y})$ depends on $\vec{y}$ only through $\mathbf{m}(\vec{y})$, and can therefore be denoted as $\mathcal{L}(\mathbf{m}(\vec{y}))$ Note that for each $\vec{y} \in \mathcal{A}_{c,k}$, we can write the first constraint as

$$
\begin{aligned}
z(\vec{y}) &\geq R_1(\mathbf{m}(\vec{y})) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m}(\vec{y}))} \sum_{\vec{x}:\mathbf{l}(\vec{x})\in\mathcal{L}(\mathbf{m}(\vec{y})),\vec{y}\in\mathcal{A}_{c,k}(\vec{x})} \beta(\vec{x})u(\vec{x})_{\vec{y}} \\
&= R_1(\mathbf{m}(\vec{y})) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m}(\vec{y}))} \beta'(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}(\vec{y})} |\{\vec{x} : \mathbf{l}(\vec{x}) = \mathbf{l}, \vec{y} \in \mathcal{A}_c(\vec{x})\}| \quad (\text{since } u \in \mathcal{U}^s) \\
&= R_1(\mathbf{m}(\vec{y})) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m}(\vec{y}))} \beta'(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}(\vec{y})} \Theta_2(\mathbf{l}, \mathbf{m}(\vec{y})).
\end{aligned}
\tag{11}
$$

Now, note that for each $\vec{y} \in \mathcal{A}_{c,k}$, $i \in a(\vec{y})$, we can write the second constraint as

$$\begin{aligned} z(\vec{y}) &\geq \sum_{\mathbf{l} \in \mathcal{L}(\mathbf{m}(\vec{y}))} \sum_{\vec{x}: \mathbf{l}(\vec{x}) \in \mathcal{L}(\mathbf{m}(\vec{y})), \vec{y} \in \mathcal{A}_{c,k}(\vec{x})} \beta(\vec{x}) u(\vec{x})_{\vec{y}} r(x_i) \\ &\geq \sum_{\mathbf{l} \in \mathcal{L}(\mathbf{m}(\vec{y}))} \beta'(\mathbf{l}) u'(\mathbf{l})_{\mathbf{m}(\vec{y})} \Theta_2(\mathbf{l}, \mathbf{m}(\vec{y})) \frac{\sum_{\vec{x}: \mathbf{l}(\vec{x}) = \mathbf{l}, \vec{y} \in \mathcal{A}_c(\vec{x})} r(x_i)}{\Theta_2(\mathbf{l}, \mathbf{m}(\vec{y}))}. \end{aligned}$$

Let $\nu(i)$ denote the class of component $i$. Now, note that

$$\frac{\sum_{\vec{x}: \mathbf{l}(\vec{x}) = \mathbf{l}, \vec{y} \in \mathcal{A}_c(\vec{x})} r(x_i)}{\Theta_2(\mathbf{l}, \mathbf{m}(\vec{y}))} = R_2(\mathbf{l}, \mathbf{m}(\vec{y}), \nu(i)).$$

Thus, for each $\vec{y} \in \mathcal{A}_{c,k}$, $i \in a(\vec{y})$, we can write the second constraint as

$$z(\vec{y}) \geq \sum_{\mathbf{l} \in \mathcal{L}(\mathbf{m}(\vec{y}))} \beta'(\mathbf{l}) u'(\mathbf{l})_{\mathbf{m}(\vec{y})} \Theta_2(\mathbf{l}, \mathbf{m}(\vec{y})) R_2(\mathbf{l}, \mathbf{m}(\vec{y}), \nu(i)). \tag{12}$$

The third constraint can be written as

$$\sum_{\mathbf{m} \in \mathcal{M}_{\vec{x}}} |\{\vec{y} : \mathbf{m}(\vec{y}) = \mathbf{m}, \vec{y} \in \mathcal{A}_{c,k}(\vec{x})\}| u'(\mathbf{l}(\vec{x}))_{\mathbf{m}} = 1 \quad \text{for all } \vec{x} \in \mathcal{K}^n.$$

Since $\mathcal{M}_{\vec{x}}$ depends on $\vec{x}$ through $\mathbf{l}(\vec{x})$ and can be denoted by $\mathcal{M}_{\mathbf{l}(\vec{x})}$, the third constraint can be written as

$$\sum_{\mathbf{m} \in \mathcal{M}_{\mathbf{l}(\vec{x})}} \Theta_1(\mathbf{l}(\vec{x}), \mathbf{m}) u'(\mathbf{l}(\vec{x}))_{\mathbf{m}} = 1 \quad \text{for all} \quad \vec{x} \in \mathcal{K}^n. \tag{13}$$

The fourth constraint can be written as

$$u'(\mathbf{l}(\vec{x}))_{\mathbf{m}(\vec{y})} \geq 0 \quad \forall \quad \vec{x} \in \mathcal{K}^n, \vec{y} \in \mathcal{A}_{c,k}(\vec{x}). \tag{14}$$

We can write the objective function as $\sum_{\mathbf{m} \in \mathcal{M}} \sum_{\vec{y}: \mathbf{m}(\vec{y}) = \mathbf{m}} z(\vec{y})$.

Thus, from (11) to (14), there exists at least one optimal solution of LP-CONTROLLER-CLASS in which $z, u'$ depend on $\vec{x}, \vec{y}$ only through $\mathbf{m}(\vec{y})$ and $\mathbf{l}(\vec{x})$. Thus, we can rewrite the above optimization as follows.

LP-CONTROLLER:

$$\begin{array}{rcl} & & \text{Minimize} \quad \sum_{\mathbf{m} \in \mathcal{M}} \Theta_3(\mathbf{m}) z(\mathbf{m}) \\ z(\mathbf{m}) & \geq & R_1(\mathbf{m}) \sum_{\mathbf{l} \in \mathcal{L}(\mathbf{m})} \beta'(\mathbf{l}) u'(\mathbf{l})_{\mathbf{m}} \Theta_2(\mathbf{l}, \mathbf{m}) \ \forall \ \mathbf{m} \in \mathcal{M} \\ z(\mathbf{m}) & \geq & \sum_{\mathbf{l} \in \mathcal{L}(\mathbf{m})} \beta'(\mathbf{l}) u'(\mathbf{l})_{\mathbf{m}} \Theta_2(\mathbf{l}, \mathbf{m}) R_2(\mathbf{l}, \mathbf{m}, i) \ \forall \ \mathbf{m} \in \mathcal{M}, i \in \mathcal{C}(\vec{m}) \\ \sum_{\mathbf{m} \in \mathcal{M}_{\mathbf{l}}} \Theta_1(\mathbf{l}, \mathbf{m}) u'(\mathbf{l})_{\mathbf{m}} & = & 1 \quad \text{for all } \vec{l} \in \mathcal{L} \\ u'(\mathbf{l})_{\mathbf{m}} & \geq & 0 \quad \forall \quad \vec{l} \in \mathcal{L}, \mathbf{m} \in \mathcal{M}_{\mathbf{l}} \end{array}$$

We can write the first constraint as follows.

$$\begin{aligned}\Theta_3(\mathbf{m})z(\mathbf{m}) &\geq R_1(\mathbf{m}) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m})} \beta'(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}}\Theta_2(\mathbf{l},\mathbf{m})\Theta_3(\mathbf{m}) \ \forall\ \mathbf{m}\in\mathcal{M} \\ &= R_1(\mathbf{m}) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m})} \Theta_4(\mathbf{l})\beta'(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m}) \ \forall\ \mathbf{m}\in\mathcal{M} \\ &\qquad\qquad\qquad (\text{since } \Theta_2(\mathbf{l},\mathbf{m})\Theta_3(\mathbf{m}) = \Theta_1(\mathbf{l},\mathbf{m})\Theta_4(\mathbf{l})) \\ &= R_1(\mathbf{m}) \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m})} \beta''(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m}) \ \forall\ \mathbf{m}\in\mathcal{M} \\ &\qquad\qquad\qquad (\text{since } \Theta_4(\mathbf{l})\beta'(\mathbf{l}) = \beta''(\mathbf{l})).\end{aligned}$$

Similarly, the rest of the constraints can be written as

$$\begin{aligned}\Theta_3(\mathbf{m})z(\mathbf{m}) &\geq \sum_{\mathbf{l}\in\mathcal{L}(\mathbf{m})} \beta''(\mathbf{l})u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m})R_2(\mathbf{l},\mathbf{m},i) \ \forall\ \mathbf{m}\in\mathcal{M}, i\in\mathcal{C}(\mathbf{m}) \\ \sum_{\mathbf{m}\in\mathcal{M}_{\mathbf{l}}} u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m}) &= 1 \ \forall\ \mathbf{l}\in\mathcal{L} \\ u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m}) &\geq 0 \ \ \forall\ \ \mathbf{l}\in\mathcal{L}, \mathbf{m}\in\mathcal{M}_{\mathbf{l}}\end{aligned}$$

In the above linear program, we substitute (a) $\Theta_3(\mathbf{m})z(\mathbf{m})$ with $\eta(\mathbf{m})$ in the objective function and the first two constraints, and (b) $u'(\mathbf{l})_{\mathbf{m}}\Theta_1(\mathbf{l},\mathbf{m})$ with $u^s(\mathbf{l})_{\mathbf{m}}$ in all the constraints. Clearly, there is a one to one correspondence between the set of optimal solutions of LP-CONTROLLER and the resulting linear program which is LP-CONTROLLER-CLASS. ■

Thus, LP-CONTROLLER-CLASS has $O(n^{2KM})$ variables and $O(n^{2KM})$ constraints. Thus, the computation time of LP-CONTROLLER-CLASS is polynomial in $n$ and exponential in $K, M$, and hence polynomial in $n$ if $K, M$ are constants.

*4) Polynomial time computation of saddle point policy of actor for constant $K, M$:* We now consider the computation of a symmetric saddle point strategy for the actor. Note that the actor's policy $v$ is symmetric if and only if $v(\vec{y}^1)_i = v(\vec{y}^2)_j$ whenever the following conditions hold: (a) $\mathbf{m}(\vec{y}^1) = \mathbf{m}(\vec{y}^2)$ (b) $i, j$ are in the same class, and (b) either (i) $i \in a(\vec{y}^1), j \in a(\vec{y}^2)$, or (ii) $i \notin a(\vec{y}^1), j \notin a(\vec{y}^2), y_i^1 = y_j^2$.

Consider a $\mathbf{m} \in \mathcal{M}$ and a class $i \in \mathcal{C}(\mathbf{m})$. Then, let $v'(\mathbf{m})_i$ be the probability with which a symmetric policy $v$ selects one (representative) component, say $j$, that is in class $i$ and has been concealed, when the actor's information state is a (representative) $\vec{y}$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$ (i.e., $v'(\mathbf{m})_i = v(\vec{y})_j$). Let $v^s(\mathbf{m})_j = \Phi(\mathbf{m}, j)v'(\mathbf{m})_j$, $j \in \mathcal{C}(\mathbf{m})$, be the total probability with which a symmetric policy $v \in \mathcal{V}^s$ of the actor selects a component which is in class $j$ and whose state has been concealed, when the actor's information state is a (representative) $\vec{y}$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$. Thus, $v$ selects a component whose state has been revealed with probability $1 - \sum_{j\in\mathcal{C}(\vec{y})} v^s(\mathbf{m}(\vec{y}))_j$. From Corollary 4.3 it is sufficient to consider only sensible policies. Thus, from Observation 1, $v^s(\mathbf{m})_j$, $j \in \mathcal{C}(\mathbf{m})$ uniquely specify a symmetric saddle point strategy $v \in \mathcal{V}^s$. We prove that the following linear program, LP-ACTOR-CLASS, computes the above.

$$\begin{array}{rcl} \text{LP-ACTOR-CLASS:} & & \text{Maximize} \sum_{\mathbf{l}\in\mathcal{L}} \beta''(\mathbf{l})\eta(\mathbf{l}) \quad s.t. \\ \eta(\mathbf{l}) & \leq & (1 - \sum_{i\in\mathcal{C}(\mathbf{m})})v^s(\mathbf{m})_i R_1(\mathbf{m}) + \sum_{i\in\mathcal{C}(\mathbf{m})} v^s(\mathbf{m})_i R_2(\mathbf{l},\mathbf{m},i) \\ & & \forall \mathbf{m}\in\mathcal{M}_{\mathbf{l}}, \mathbf{l}\in\mathcal{L} \\ v^s(\mathbf{m})_i & \geq & 0 \ \ \forall\ i\in\mathcal{C}(\mathbf{m}), \mathbf{m}\in\mathcal{M} \\ \sum_{i\in\mathcal{C}(\mathbf{m})} v(\mathbf{m})_i & \leq & 1 \ \ \forall\ \mathbf{m}\in\mathcal{M} \end{array}$$

*Theorem 5.3:* The optimum solution $\{v^s(\mathbf{m})_j\}_{\mathbf{m}\in\mathcal{M},j\in\mathcal{C}(\mathbf{m})}$ of LP-ACTOR-CLASS constitutes a policy $v^*$ for the actor such that $v^*\in\mathcal{V}^s$ and $v^*$ is a saddle-point policy for the actor.

*Proof:* Consider the description of LP-ACTOR at the end of Section IV-B.2, and restrict the feasible solutions $v$ to $\mathcal{V}^s$. From Theorem 5.1, the optimal solution of LP-ACTOR constitutes a saddle point even with this restriction, and the optimal solution is clearly a symmetric strategy for the actor. It is therefore sufficient to show that there is a one-to-one correspondence between the set of optimal solutions of LP-ACTOR-CLASS to that of LP-ACTOR with the above restriction.

Consider LP-ACTOR and $v\in\mathcal{V}^s$.

For each $\vec{x}\in\mathcal{K}^n$ and $\vec{y}\in\mathcal{A}_{c,k}(\vec{x})$, we can write the first constraint as

$$\begin{aligned} z(\vec{x}) \ \ \geq\ \ & R_1\left(\mathbf{m}(\vec{y})\right)\left(1-\sum_{i\in C(\mathbf{m}(\vec{y}))}\sum_{j\in a(\vec{y},i)} v(\vec{y})_j\right) + \sum_{i\in C(\mathbf{m}(\vec{y}))}\sum_{j\in a(\vec{y},i)} v(\vec{y})_j r(x_j) \\ = \ \ & R_1\left(\mathbf{m}(\vec{y})\right)\left(1-\sum_{i\in C(\mathbf{m}(\vec{y}))} v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),i)\right) + \sum_{i\in C(\mathbf{m}(\vec{y}))} v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),i)\frac{\sum_{j\in a(\vec{y},i)} r(x_j)}{\Phi(\mathbf{m}(\vec{y}),i)} \\ & (\text{ since } v\in\mathcal{V}^s) \\ = \ \ & R_1(\mathbf{m}(\vec{y}))\left(1-\sum_{j\in C(\mathbf{m}(\vec{y}))} v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),i)\right) + \sum_{i\in C(\mathbf{m}(\vec{y}))} v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),j) R_2\left(\mathbf{l}(\vec{x}),\mathbf{m}(\vec{y}),i\right). \end{aligned}$$

We can write the second constraint as

$$v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),i) \geq 0,\ \forall\ i\in C\left(\mathbf{m}(\vec{y})\right), \vec{y}\in\mathcal{A}_{c,k}.$$

For each $\vec{y}\in\mathcal{A}_{c,k}$, we can write the third constraint as

$$\begin{aligned} \sum_{i\in C(\mathbf{m}(\vec{y}))}\sum_{j\in a(\vec{y},i)} v(\vec{y})_j \ \ & \leq \ \ 1 \\ \sum_{i\in C(\mathbf{m}(\vec{y}))} v'\left(\mathbf{m}(\vec{y})\right)_i \Phi(\mathbf{m}(\vec{y}),i) \ \ & \leq \ \ 1. \end{aligned}$$

The objective function can be written as $\sum_{\mathbf{l}\in\mathcal{L}}\sum_{\vec{x}:\mathbf{l}(\vec{x})=\mathbf{l}}\beta(\vec{x})z(\vec{x})$, which equals $\sum_{\mathbf{l}\in\mathcal{L}}\beta'(\mathbf{l})\sum_{\vec{x}:\mathbf{l}(\vec{x})=\mathbf{l}} z(\vec{x})$. Thus, clearly, there exists one optimum solution $v', z$ which depend on $\vec{y}$ and $\vec{x}$ only through $\mathbf{m}$ and $\mathbf{l}$ respectively. We can therefore rewrite LP-ACTOR as

$$
\begin{array}{|rcl|}
\hline
\text{LP-ACTOR:} & & \text{Maximize} \sum_{\mathbf{l}\in\mathcal{L}} \beta'(\mathbf{l})z(\mathbf{l})\Theta_4(\mathbf{l}) \\
z(\mathbf{l}) & \leq & (1-\sum_{i\in\mathcal{C}(\mathbf{m})})v'(\mathbf{m})_i\Phi(\mathbf{m},i)R_1(\mathbf{m}) + \sum_{i\in\mathcal{C}(\mathbf{m})} v'(\mathbf{m})_i\Phi(\mathbf{m},i)R_2(\mathbf{l},\mathbf{m},i) \\
 & & \quad \forall\ \mathbf{m}\in\mathcal{M}_1, \mathbf{l}\in\mathcal{L} \\
v'(\mathbf{m})_i\Phi(\mathbf{m},i) & \geq & 0 \ \ \forall\ i\in\mathcal{C}(\mathbf{m}), \mathbf{m}\in\mathcal{M} \\
\sum_{i\in\mathcal{C}(\mathbf{m})} v'(\mathbf{m})_i\Phi(\mathbf{m},i) & \leq & 1 \ \ \forall\ \mathbf{m}\in\mathcal{M} \\
\hline
\end{array}
\tag{15}
$$

Now, we replace $\beta'(\mathbf{l})\Theta_4(\mathbf{l})$ with $\beta''(\mathbf{l})$ in the objective function, and $v'(\mathbf{m})_i\Phi(\mathbf{m},i)$ with $v^s(\mathbf{m})_i$ in all the constraints. The set of optimal solutions of the resulting linear program, which is LP-ACTOR-CLASS, has one-to-one correspondence with that of the above linear program. ■

LP-ACTOR-CLASS has $O(n^{KM})$ variables and $O(n^{2KM})$ constraints. Thus, the computation time of LP-ACTOR-CLASS is polynomial in $n$ and exponential in $K, M$.

### *B. Approximation guarantees using polynomial time computable policies for arbitrary systems*

Saddle point strategies can be computed in polynomial time when either $n$ is a constant (using LP-CONTROLLER or LP-ACTOR) or $K, M$ are constants (using LP-CONTROLLER-CLASS or LP-ACTOR-CLASS). The computation however becomes intractable when two or more of these parameters are large. We now prove that simple linear ($O(n)$) or almost linear ($O(n\text{log}n) + K$) time computable policies can provably approximate the saddle point policies. Specifically, there exists a $O(n)$ time computable policy for the actor such that, irrespective of the policy of the controller, the utility of the actor with this policy is at least $1/(\min(k, M) + 1)$ times the max-min utility of the actor for arbitrary $k, M, n$ (Section V-C). Thus, the worst case approximation guarantee of this policy is $1/(k + 1)$ (attained for large $M$), and the approximation guarantee when all components are statistically identical ($M = 1$) is $1/2$. Also, the approximation improves with decrease in the number of classes and the number of components whose states can be concealed. We next show that there exists a $O(n\text{log}n + K)$ time computable policy for the controller such that, irrespective of the policy of the actor, the utility of the actor with this policy is at most $k + 1$ times the actor's min-max utility for arbitrary $K, M, n$, and at most $2$ times the actor's min-max utility for arbitrary $K, n$ and $M = 1$ (i.e., when all components are statistically identical) (Section V-D). Finally, we examine whether the above approximation guarantees are tight (Section V-E).

### *C. Approximation guarantees using a linear time computable policy for the actor*

Consider a symmetric sensible policy, denoted as $v^{\text{UNIFORM}} \in \mathcal{V}^s$, for the actor that (a) selects each concealed class and a revealed component with equal probabilities, i.e., $v^{s,\text{UNIFORM}}(\mathbf{m})_i = 1/\left(|\mathcal{C}(\mathbf{m})| + 1\right)$ for each $\mathbf{m} \in \mathcal{M}$, $i \in \mathcal{C}(\mathbf{m})$. Note that this uniquely describes any symmetric sensible policy since a symmetric policy selects uniformly among the concealed components in each class and a sensible policy selects only a revealed component with the highest state whenever it selects a revealed component. Clearly, the actor needs $O(n)$ time and memory to select a component using this policy.

We now prove the main result of this section.

*Theorem 5.4:* For any $\beta, k, n, K, M$,

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v^{\textsc{Uniform}}} \geq \frac{1}{\min(k,M)+1} \sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v}.$$

*Proof:* Consider an arbitrary sensible policy $v \in \mathcal{V}^s$. Let $T(\mathbf{l},\mathbf{m},v)$ be the utility of the actor if the system state vector is $\vec{x}$ such that $\mathbf{l}(\vec{x}) = \mathbf{l}$ and the actor's information is some $\vec{y}$ such that $\mathbf{m}(\vec{y}) = \mathbf{m}$ and the actor uses the policy $v$. Then,

$$\begin{aligned} T(\mathbf{l},\mathbf{m},v) &= (1 - \sum_{i\in\mathcal{C}(\mathbf{m})} v^s(\mathbf{m})_i) R_1(\mathbf{m}) + \sum_{i\in\mathcal{C}(\mathbf{m})} v^s(\mathbf{m})_i R_2(\mathbf{l},\mathbf{m},i) \\ &\leq \max\left(R_1(\mathbf{m}), \max_{i\in\mathcal{C}(\mathbf{m})} R_2(\mathbf{l},\mathbf{m},i)\right). \end{aligned} \tag{16}$$

Also,

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v} = \sum_{\mathbf{l}\in\mathcal{L}} \beta''(\mathbf{l}) \min_{\mathbf{m}\in\mathcal{M}_\mathbf{l}} T(\mathbf{l},\mathbf{m},v). \tag{17}$$

From (16) and (17),

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v^{\textsc{Uniform}}} = \sum_{\mathbf{l}\in\mathcal{L}} \beta(\mathbf{l}) \min_{\mathbf{m}\in\mathcal{M}_\mathbf{l}} T(\mathbf{l},\mathbf{m},v^{\textsc{Uniform}}), \quad \text{where,} \tag{18}$$

$$\begin{aligned} T(\mathbf{l},\mathbf{m},v^{\textsc{Uniform}}) &= \frac{R_1(\mathbf{m}) + \sum_{i\in\mathcal{C}(\mathbf{m})} R_2(\mathbf{l},\mathbf{m},i)}{|\mathcal{C}(\mathbf{m})|+1} \\ &\geq \frac{\max(R_1(\mathbf{m}), \max_{i\in\mathcal{C}(\mathbf{m})} R_2(\mathbf{l},\mathbf{m},i))}{\min(k,M)+1} \quad (\text{since } |\mathcal{C}(\mathbf{m})| \leq \min(k,M)) \end{aligned} \tag{19}$$

Now, let $v^*$ be the optimal solution of LP-ACTOR-CLASS. Then, from Theorem 5.3 and (17),

$$\sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v} = \sum_{\mathbf{l}\in\mathcal{L}} \beta''(\mathbf{l}) \min_{\mathbf{m}\in\mathcal{M}_\mathbf{l}} T(\mathbf{l},\mathbf{m},v^*).$$

Thus, from (18) it is sufficient to prove that $T(\mathbf{l},\mathbf{m},v^{\textsc{Uniform}}) \geq T(\mathbf{l},\mathbf{m},v^*)/\left(\min(k,M)+1\right)$ for each $\mathbf{l} \in \mathcal{L}, \mathbf{m} \in \mathcal{M}$.

Since $v^*$ is sensible, the result follows from (16) and (19). ■

For $K = 2$, the approximation ratio can be improved slightly. In this case, for a symmetric sensible saddle point strategy $v$ of the actor, $\sum_{i\in\mathcal{C}(\mathbf{m})} v^s(\mathbf{m})_i = 1$ if all revealed components are in state $0$ and $\sum_{i\in\mathcal{C}(\mathbf{m})} v^s(\mathbf{m})_i = 0$ otherwise. Using the above, it follows that the actor's policy that selects (a) a component in state $1$ if the state of at least one such component is revealed and (b) each concealed class with equal probability, otherwise, attains a $1/\min(k,M)$ approximation ratio.

### *D. Approximation guarantees using an almost linear time computable policy for the controller*

Consider the *Greedy for controller* policy of the controller where it conceals the components with $k$ highest states and breaks ties randomly and uniformly. We denote this policy as $u^{\mathrm{GC}}$.

Clearly, $u^{\text{GC}} \in \mathcal{U}^s$. Note that the controller needs $O(n\text{log}n + K)$ time and $O(n)$ memory to decide its actions using this policy.

*Theorem 5.5:* For any $\beta, k, n, K, M$,

$$\sup_{v\in\mathcal{V}} R_\beta^{u^{\text{GC}},v} \leq (k+1) \inf_{u\in\mathcal{U}} \sup_{v\in\mathcal{V}} R_\beta^{u,v}.$$

For any $\beta, k, n, K$ such that $M = 1$, that is, all components are statistically identical,

$$\sup_{v\in\mathcal{V}} R_\beta^{u^{\text{GC}},v} \leq 2 \inf_{u\in\mathcal{U}} \sup_{v\in\mathcal{V}} R_\beta^{u,v}.$$

*Proof:* We first provide a general framework for proving that

$$\sup_{v\in\mathcal{V}} R_\beta^{u^{\text{GC}},v} \leq \kappa \inf_{u\in\mathcal{U}} \sup_{v\in\mathcal{V}} R_\beta^{u,v}$$

for an arbitrary $\kappa$ and arbitrary $\beta, k, n, K, M$. We prove that

$$\sup_{v\in\mathcal{V}} R_\beta^{u^{\text{GC}},v} \leq \kappa \inf_{u\in\mathcal{U}} R_\beta^{u,v'} \text{ for some } v \in \mathcal{V}. \tag{20}$$

Now, the result follows since $\inf_{u\in\mathcal{U}} R_\beta^{u,v'} \leq \sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v} = \inf_{u\in\mathcal{U}} \sup_{v\in\mathcal{V}} R_\beta^{u,v}$.

Now, (20) can be proved as follows. Clearly,

$$\sup_{v\in\mathcal{V}} R_\beta^{u^{\text{GC}},v} = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x})\theta(\vec{x}) \tag{21}$$

for some real-valued function $\theta$ on $\mathcal{K}^n$ which depends on $\beta, k, n, K, M$. Let $T'(\vec{x}, \vec{y}, v')$ be the utility of the actor if the system state vector is $\vec{x}$ and the actor's information is $\vec{y}$ and the actor uses the policy $v'$. Then,

$$\inf_{u\in\mathcal{U}} R_\beta^{u,v'} = \sum_{\vec{x}\in\mathcal{K}^n} \beta(\vec{x}) \min_{\vec{y}\in\mathcal{A}_c(\vec{x})} T'(\vec{x}, \vec{y}, v). \tag{22}$$

Thus, from (21) and (22), (20) follows if we can prove that for each $\vec{x} \in \mathcal{K}^n$,

$$\theta(\vec{x}) \leq \kappa \min_{\vec{y}\in\mathcal{A}_c(\vec{x})} T'(\vec{x}, \vec{y}, v).$$

We introduce some terminologies first. Consider an arbitrary $\vec{x} \in \mathcal{K}^n$ and $\vec{y} \in \mathcal{A}_c(\vec{x})$. Let GC$(\vec{x})$ be the set of components whose states have been concealed by $u^{\text{GC}}$ when the system state vector is $\vec{x}$, $\mathcal{D}_1(\vec{x}, \vec{y}) = \text{GC}(\vec{x}) \setminus a(\vec{y})$ , and $\mathcal{D}_2(\vec{x}, \vec{y}) = a(\vec{y}) \setminus \text{GC}(\vec{x})$. Let $\vec{x}^{\text{GC}}$ be the actor's information under GC when $\vec{x}$ is the system state vector.

Note that the actor's best response to $u^{\text{GC}}$ is to select components whose states have been concealed since the state of any such component is at least as high as that of a component whose state has been revealed. Thus, $\theta(\vec{x}) = \sum_{i\in\text{GC}(\vec{x})} \gamma(\vec{x}^{\text{GC}})_i r(x_i)$ where $\gamma(\vec{x}^{\text{GC}})$ is a probability distribution on GC$(\vec{x})$ which depends on $\vec{x}^{\text{GC}}, \beta, k, n, K, M$.

We now consider the general case, that is, arbitrary $\beta, k, n, K, M$ and therefore need $\kappa = k + 1$. We consider $v'$ that selects each concealed component w.p. $1/(|a(\vec{y})| + 1)$ and the revealed component with the highest state w.p. $1/(|a(\vec{y})| + 1)$. Then, $T'(\vec{x}, \vec{y}, v') = (1/(|a(\vec{y})| +$

$1))\left(\max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i)+\sum_{i\in a(\vec{y})} r(x_i)\right)$. Since $|a(\vec{y})|\leq k$ as $\vec{y}\in\mathcal{A}_c(\vec{x})$, the result follows if we can show that

$$\theta(\vec{x})\leq \max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i)+\sum_{i\in a(\vec{y})} r(x_i)\ \forall\ \vec{x}\in\mathcal{K}^n,\vec{y}\in\mathcal{A}_c(\vec{x}). \tag{23}$$

$$\begin{aligned}
\theta(\vec{x})-\sum_{i\in a(\vec{y})} r(x_i) &= \sum_{i\in \text{GC}(\vec{x})}\gamma(\vec{x}^{\text{GC}})_i r(x_i)-\sum_{i\in a(\vec{y})} r(x_i)\\
&\leq \sum_{i\in\mathcal{D}_1(\vec{x},\vec{y})}\gamma(\vec{x}^{\text{GC}})_i r(x_i)-\sum_{i\in\mathcal{D}_2(\vec{x},\vec{y})} r(x_i)\ \text{(since } 0\leq\gamma(\vec{x}^{\text{GC}})_i\leq 1\ \forall\ i\in\text{GC}(\vec{x}))\\
&\leq \sum_{i\in\mathcal{D}_1(\vec{x},\vec{y})}\gamma(\vec{x}^{\text{GC}})_i r(x_i)\\
&\leq \max_{i\in\mathcal{D}_1(\vec{x},\vec{y})} r(x_i)\\
&\qquad\left(\text{since } \sum_{i\in\mathcal{D}_1(\vec{x},\vec{y})}\gamma(\vec{x}^{\text{GC}})_i\leq 1 \text{ and } 0\leq\gamma(\vec{x}^{\text{GC}})_i\leq 1\ \forall\ i\right)\\
&\leq \max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i)\ \text{(since } \mathcal{D}_1(\vec{x},\vec{y})=\text{GC}(\vec{x})\setminus a(\vec{y})\subset\mathcal{N}\setminus a(\vec{y}))
\end{aligned}$$

Thus, (23) follows.

We now consider the special case in which $M=1$. Thus, all components are statistically identical. In this case, from symmetry, $\gamma(\vec{x}^{\text{GC}(\vec{x})})_i=1/k$, for each $i\in\text{GC}(\vec{x})$, that is, the actor's best response is to select each concealed component w.p. $1/k$. Thus,

$$\theta(\vec{x})=\sum_{i\in\text{GC}(\vec{x})} r(x_i)/k. \tag{24}$$

We consider $v'$ that selects (a) each concealed component w.p. $1/(2|a(\vec{y})|)$ and the revealed component with the highest state w.p. $1/2$ if at least one component is concealed and (b) the revealed component with the highest state if no component is concealed. Then,

$$T'(\vec{x},\vec{y},v')=\left(\max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i)+\frac{\sum_{i\in a(\vec{y})} r(x_i)}{|a(\vec{y})|}\right)/2.$$

Here, we assume that the second term in the sum is $0$ if $a(\vec{y})=\phi$. Thus, from (24), the result follows if we can show that

$$\sum_{i\in\text{GC}(\vec{x})} r(x_i)/k\leq \max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i)+\frac{\sum_{i\in a(\vec{y})} r(x_i)}{|a(\vec{y})|}\ \forall\ \vec{x}\in\mathcal{K}^n,\vec{y}\in\mathcal{A}_c(\vec{x}). \tag{25}$$

If $a(\vec{y}) = \phi$, the result clearly holds as then the left hand side is $\max_{i\in\mathcal{N}} r(x_i)$, and since $|\text{GC}(\vec{x})| = k$, $\max_{i\in\mathcal{N}} r(x_i) \geq \sum_{i\in\text{GC}(\vec{x})} r(x_i)/k$. We therefore assume that $a(\vec{y}) \neq \phi$.

$$\begin{aligned}
\sum_{i\in\text{GC}(\vec{x})} \frac{r(x_i)}{k} - \frac{\sum_{i\in a(\vec{y})} r(x_i)}{|a(\vec{y})|} &\leq \sum_{i\in\text{GC}(\vec{x})} \frac{r(x_i)}{k} - \sum_{i\in a(\vec{y})} \frac{r(x_i)}{k} \text{ (since } |a(\vec{y})| \leq k \text{ as } \vec{y} \in \mathcal{A}_c(\vec{x})) \\
&\leq \sum_{i\in\mathcal{D}_1(\vec{x},\vec{y})} \frac{r(x_i)}{k} \\
&\leq \max_{i\in\mathcal{D}_1(\vec{x},\vec{y})} r(x_i) \text{ (since } |\mathcal{D}_1(\vec{x},\vec{y})| \leq k \text{ as } \mathcal{D}_1(\vec{x},\vec{y}) \subseteq \text{GC}(\vec{x})) \\
&\leq \max_{i\in\mathcal{N}\setminus a(\vec{y})} r(x_i).
\end{aligned}$$

Thus, (25) follows. ■

Note that when $K = 2$ the approximation factor turns out to be $k$ (instead of $k+1$) for arbitrary $\beta, k, n, M$. The proof is similar, but considers only states $\vec{y}$ in which all revealed components are in state $0$ and considers a policy $v'$ for the actor that differs from the corresponding one in the above proof in that it never selects any revealed component that is in state $0$.

### *E. Tightness of the approximation guarantees*

We now examine whether the approximation guarantees obtained so far are tight. We answer the relevant questions in part, and outline several important open problems.

We prove that the approximation bound obtained for for the uniform policy of the actor is tight. Specifically, given any $\epsilon > 0$, there exists a system with components whose state processes are mutually independent where the minimum utility obtained by the actor when it uses the uniform policy exceeds $1/(\min(k, M) + 1)$ times the max-min utility in the system by at most $\epsilon$. Consider a system where $M > 1$. Let the first class consist of only 1 component which is in state $K - 2$ w.p. $1 - \epsilon_1$ and in state $0$ w.p. $\epsilon_1$. The components in the other classes are either in states $0$ or $1$ (the probability distributions for the state processes for channels in different classes are different). The state processes of the components are mutually independent. Let $r(K-2) = 1 - \delta_1, r(1) = \delta_2$. Let $v_1 \in \mathcal{V}$ be the policy that always selects the component in the first class. Clearly $R_\beta^{u,v_1} = (1-\delta_1)(1-\epsilon_1)$ for any $u \in \mathcal{U}$. Thus, $\sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v} \geq (1-\delta_1)(1-\epsilon_1)$. Consider a $u_1 \in \mathcal{U}$ that conceals the component from the first class, and selects the rest of the components to be concealed in a round robin manner. Specifically, in the first round $u_1$ selects one component from classes $2, \ldots, M$ each, repeats the process in second, third rounds etc. until $k$ components have been selected. Thus, $\min(k, M)$ classes are concealed. Clearly, the state of the component that has the highest state among the revealed components is no more than $1$. Thus,

$$\begin{aligned}
R_\beta^{u,v^{\text{UNIFORM}}} &\leq (r(K-2) + \min(k, M)r(1))/(\min(k, M) + 1) \\
&\leq (1 + \min(k, M)\delta_2)/(\min(k, M) + 1) \\
&\leq \frac{(1-\delta_1)(1-\epsilon_1)}{(\min(k, M) + 1)} + \epsilon \text{ for sufficiently small } \delta_1, \delta_2, \epsilon_1 \\
&\leq \frac{\sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v}}{(\min(k, M) + 1)} + \epsilon.
\end{aligned}$$

Thus, $\inf_{u\in\mathcal{U}} R_\beta^{u,v^{\text{UNIFORM}}} \leq \left(\sup_{v\in\mathcal{V}} \inf_{u\in\mathcal{U}} R_\beta^{u,v}/(\min(k,M)+1)\right) + \epsilon$. The result follows. The scenario where this approximation factor turns out to be tight however rarely arises in practice, and as our numerical computations demonstrate, the minimum utility obtained by the uniform policy closely approximates the max-min utility of the actor in general.

## VI. Numerical example

Through a simple example we shall illustrate several points: (i) we show that GC is in general not a saddle-point strategy for the Controller, (ii) We illustrate the bound obtained with GC in Theorem 5.5, (iii) We show symmetrical statements for the Uniform policy for the Actor, (iv) Investigate tightness.

We consider $n = 3$ channels, $K = 3$ states per channel and $M = 1$ single class (i.e. all channels are symmetric). We have taken $k = 2$ (two channels are concealed). The probability distribution $\beta$ for each channel is given by the following vector: $\beta(\overrightarrow{x}) = (1/3, 1/3, 1/3)$. The reward function considered is the following vector: $r = (0, x, 1)$, where $x = r(1)$ is the reward when the channel is in state 1. We let $x$ vary and compare numerically in figure 1 the following:

- value of the game
- the performance when the Actor uses the uniform strategy and Controller plays optimally against it,
- the performance when the Controller uses the GC strategy and the Actor plays optimally against it,
- The bounds in Theorem 5.4
- The bounds in Theorem 5.5

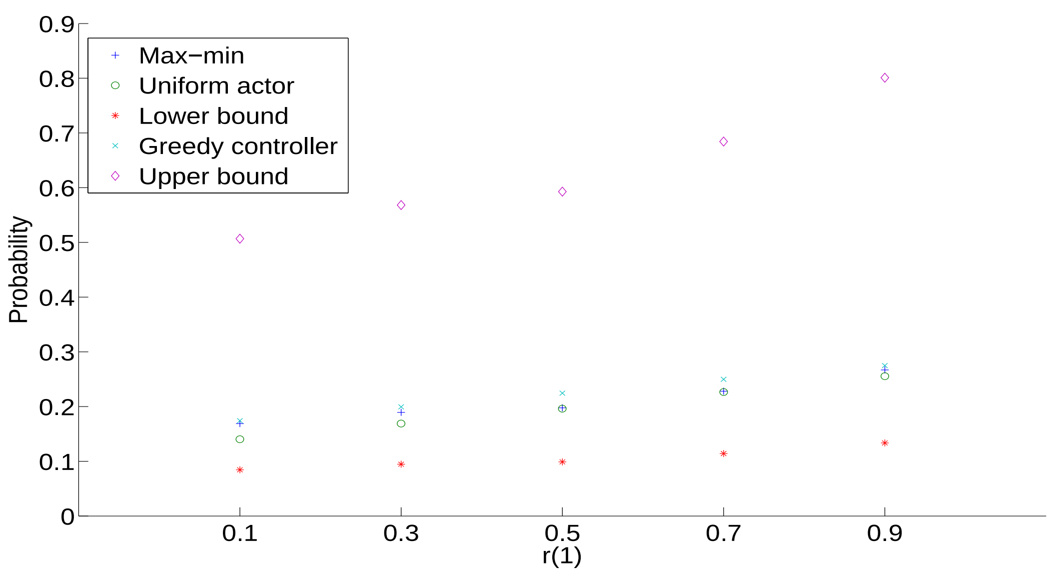


Fig. 1. Approximation performance of the uniform policy compared to the optimal.

The figure confirms that the GC is not a saddle-point point for the Controller and that the Uniform strategy is not a saddle-point for the Actor. However both these strategies are seen to perform well and to guarantee a performance close to the value of the ICG. The bounds given by Theorems 5.4 and 5.5 are indeed seen to hold. The optimal strategy of the sender turned out to choose the unconcealed channel with probability 1 for $x = 0.7$ or larger. It chooses with probability 1 a concealed channel for $x = 0.3$ or loss. For $x = 0.5$ the sender's policy was to randomize between the two (it chose the unconcealed channel with probabilitly 0.4).

## VII. Conclusions and open questions

We have studied a leader-follower game where the actions of the leader (Controller) determine the information available to the follower (Actor). By concealing information, the leader degrades the performance of the follower that attempts to choose one out of several resources with the best state among all. We have provided a rich body of computation and appproximation tools for that problem along with mathematical foundations and complexity analysis.

The question of tightness of the approximation guarantees is only partially solved. The approximation bound for the uniform policy of the Actor is indeed tight as we showed in Section V-E. The question regarding the tightness of the approximation ratio obtained for the Greedy policy of the Controller remains open. Other open problems include establishing that the computation of the saddle point policies is NP-hard, and determining whether the approximation factors can be substantially improved while using polynomial time computation. We plan to extend our study to the stochastic game framework in which the states can change in time according to some Markov structure.

## References


[1] R. J. Aumann and M. Maschler. Repeated games with incomplete information. avec la collaboration de R. Stearns. M.I.T. Press, 1995

[2] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, ”Infnormation Hiding — A Survey” *Proceedings of the IEEE, special Issue on protection of multimedia content*, **87**(7):1062–1078, July 1999.

[3] Harold W. Kuhn, ”Extensive games and the problem of information”, *Contributions to the to the Theory of Games*, eds. H.W.Kuhn and A.Tucker, Vol. 2, Princeton University Press, pp. 193-216, 1953.