Decentralized Leadership and Follower Deception in Stackelberg Games

Jiarui Gan

Jesus College

University of Oxford

Supervised by

Edith Elkind and Michael J. Wooldridge

Examined by

Paul W. Goldberg and David C. Parkes

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2020

Acknowledgements

I would like to express my heartfelt gratitude to my supervisors Profs. Edith Elkind and Michael Wooldridge. It was my great fortune to have them as my supervisors and their extraordinary supervision no doubt made my whole DPhil journey much easier and enjoyable. I am very grateful for the time and effort they have contributed to my work, their patience and trust, and the invaluable guidance and mentoring they have provided whenever needed. This thesis would simply not exist without them.

This thesis is also a fruit of collaboration and a collective wisdom belonging to all my collaborators. I want to thank my co-authors, especially those who contributed directly to this thesis: Bo An, Georgios Birmpas, Qingyu Guo, Alexandros Hollender, Sarit Kraus, Francisco Marmolejo-Cossío, Zinovi Rabinovich, Ninad Rajgopal, Long Tran-Thanh, Alexandros Voudouris, and Haifeng Xu. Special thanks go to Bo, who was my master advisor before I came to Oxford and introduced me to the multi-agent systems and AI communities. I am thankful for his support throughout the years.

I would like to thank Profs. Paul Goldberg and Elias Koutsoupias for generously serving as the examiners of my transfer and confirmation reports and providing extremely insightful feedback to my work. Paul was also an examiner of my final viva exam, along with Prof. David Parkes. This thesis benefited tremendously from their invaluable comments. I also want to thank many other leading researchers in the community, especially Profs. Haris Aziz, Vincent Conitzer, Nicola Gatti, Sarit Kraus, Zhongzhi Shi, Milind Tambe, and Yevgeniy Vorobeychik, for useful discussions with them and their generous help with my research and career development. I thank my college advisor Prof. Standa Zivny for giving me a lot of helpful advice.

My life in Oxford wouldn't be so memorable without those who kept me accompanied. They are: the members of Room 017 in the CS Department, Georgios Birmpas, Paul Harrenstein (honorary member), Ayumi Igarashi, Bo Li, Muhammad Najib, Dominik Peters, Warut Suksompong, and Alexandros Voudouris, who made this coolest (literally) office in the department even cooler; my other fellow graduate students Anna Gautier, Dongge Han, Lewis Hammond, James Paulin, Thomas Steeples, who were the best companions in the DPhil journey; Anne, Fred, and Bing Wang, whom I stayed with during the COVID-19 lockdown, when this thesis was compiled; and many other friends from the CS Department, Jesus College, and other random corners of Oxford. I wish to thank all of them.

I am deeply indebted to (and in debt to) my family, for their unconditional love and support for my decision to pursue a PhD degree. I am grateful to have Yue Sun (now Dr. Sun) in my life.

Last but not least, I would like to acknowledge the generous financial support I received towards my study from the Engineering and Physical Sciences Research Council (EPSRC).

Abstract

This thesis focuses on two aspects of Stackelberg games — decentralized leadership and follower deception — that stem from reasoning about strategic interactions at a higher level in Stackelberg games.

In the first part of the thesis, we study decentralized leadership in Stackelberg games. We focus on a variant of Stackelberg security games that involves multiple leaders, in which the leaders allocate their security resources to protect a set of targets against an attacker; the attacker acts as the follower in the game, surveying the leaders' strategies and responding rationally afterwards. We aim to understand the game under decentralized leadership and explore ways to coordinate the leaders. To this end, firstly, we propose a novel equilibrium concept to describe the outcome of the game, and we analyze the existence of an equilibrium and the complexity of computing it. Then, we take a mechanism design approach to coordinate the leaders, aiming to design a coordination mechanism that satisfies several natural properties concerning efficiency, incentives, strategyproofness, and stability of strategies it generates. We obtain impossibility results showing that certain combinations of these properties cannot be achieved by any mechanism. On top of that we also present mechanisms for property combinations that are not blocked by the impossibility results.

The second part of the thesis is motivated by the finding that, in a Stackelberg game, the follower can benefit from changing the leader's belief about the game. This can typically happen when the follower is able to tamper with the leader's attempt of information gathering: for instance, when the leader learns the follower's type by interacting with the follower, the follower can imitate the behavior of a different follower type, pretending that their payoffs are different than the actual ones. A leader who ignores such strategic manipulation will then be misled into playing optimally against the imitated follower type, which may correspond to a highly suboptimal strategy in the original game. We study the strategic interactions resulting from this deceptive behavior, both from the leader's and the follower's perspectives. From the leader's perspective, we propose a policy-based approach to mitigate potential losses due to follower deception and study the associated problem of computing the optimal policy. From the follower's perspective, we study the problem of computing the optimal strategy to deceive the leader. Our results provide an almost complete picture of the complexity landscape of these problems in various settings. It is shown that the problems facing the leader are hard in general, whereby we derive inapproximability bounds and also design algorithms that achieve matching approximation ratios; whereas the problems facing the follower are tractable, whereby we design polynomial-time algorithms to compute optimal (or near-optimal) deception strategies for the follower.

Our work sheds light on strategic interactions that may arise in Stackelberg games but are not captured by existing solution concepts. It provides frameworks to study these interactions and contributes to the understanding of them through the lens of computation.

Contents

1 Overview **1**
- 1.1 Thesis Outline and Main Contributions . . . 2
- 1.2 Publications . . . 3

2 Preliminaries **5**
- 2.1 Stackelberg Games . . . 5
- 2.2 Stackelberg Security Games . . . 8
- 2.3 Summary . . . 13

I Decentralized Leadership and Coordination **15**

3 Decentralized Leadership in Security Games **17**
- 3.1 Introduction . . . 17
- 3.2 Multi-defender SSG . . . 20
- 3.3 Characterizing an NSE . . . 22
- 3.4 Existence of an NSE . . . 28
- 3.5 Computing an NSE . . . 38
- 3.6 NSE in Other Settings . . . 40
- 3.7 Summary . . . 43
- Appendix 3.A Additional Examples of NSEs . . . 44
- Appendix 3.B Omitted Proofs . . . 46

4 Coordinating the Leaders: A Mechanism Design Approach **57**
- 4.1 Introduction . . . 57
- 4.2 Coordination Mechanism Design . . . 59
- 4.3 Impossibility Results . . . 60
- 4.4 Constrained Social Welfare Maximization Mechanisms . . . 64
- 4.5 Stability against Strategy Modification . . . 69
- 4.6 Summary . . . 77

II Follower Deception and Countermeasures **79**

5 Who are We Playing Against: Imitative Follower Deception **81**
- 5.1 Introduction . . . 81
- 5.2 A Motivating Example . . . 84

5.3 Stackelberg Games with Follower Deception . . . 86
5.4 Computing the Optimal Policy . . . 88
5.5 Generalization to Mixed Policy . . . 93
5.6 Empirical Evaluation . . . 97
5.7 Extensions — Beyond Bayesian Analysis . . . 99
5.8 Summary . . . 104

6 Deceiving the Leader Optimally **105**
6.1 Introduction . . . 105
6.2 Preliminaries . . . 106
6.3 Best Response Inducibility . . . 108
6.4 Payoff Inducibility . . . 111
6.5 Robustness with Respect to Equilibrium Selection . . . 113
6.6 Summary . . . 116
Appendix 6.A Omitted Proofs . . . 117

7 Follower Deception in SSGs and Ways to Counteract **123**
7.1 Introduction . . . 123
7.2 SSG with Attacker Deception . . . 125
7.3 Optimal Attacker Deception . . . 126
7.4 Commitment to a Policy . . . 131
7.5 Optimal Policy for a Finite Set of Attacker Types . . . 132
7.6 Beyond Finite Attacker Types . . . 134
7.7 Empirical Evaluation . . . 137
7.8 Summary . . . 137
Appendix 7.A Omitted Proofs . . . 138

III Conclusion **143**

8 Conclusion and Future Directions **145**

Bibliography **147**

List of Figures

2.1 The comb-sampling algorithm. 9

3.1 The tank-filling visualization. 22

5.1 Payoff matrices (proof of Theorem 5.5). 89
5.2 Payoff matrices (proof of Theorem 5.6). 91
5.3 The leader's utility obtained with different approaches. 97
5.4 Payoff matrices (proof of Theorem 5.17). 101
5.5 Payoff matrices (proof of Theorem 5.18). 103

6.1 No payoff matrix $\tilde{u}^{\mathsf{F}}$ realizes the polytopal BR correspondence $\widetilde{\mathsf{BR}}_{\mathrm{P}}$, such that $\ell \in \widetilde{\mathsf{BR}}_{\mathrm{P}}$ if and only if $\mathbf{y} \in R_\ell$. 111
6.2 A game where the optimal inducible utility is 1, but the optimal *strongly* inducible utility is 0. 113
6.3 A non-max-degenerate game for which the optimal inducible utility cannot be achieved by any strongly inducible profile. 114

7.1 Comparison of the EoP. 135
7.2 Comparison of the EoP (with the zero-sum attacker type). 136

Chapter 1

Overview

Stackelberg games are a simple yet powerful model for sequential interaction among two strategic players: a *leader*, who commits to a strategy first, and a *follower*, who responds upon knowing the leader's commitment. The equilibrium of the game captures rational behaviors of the two players, especially the optimal strategy commitment of the leader. The model dates back to 1934, when Heinrich von Stackelberg introduced it to analyze market competition between a large leader firm and a small follower firm [von Stackelberg, 1934, 2010]. Since then, it has been used to study a wide range of problems, such as principal-agent contract design, pricing and planning of transportation systems, influence maximization, or exam design [e.g. Page, 1991, Gan et al., 2013, Xiong et al., 2016, Carnes et al., 2007, Li and Conitzer, 2017], often from the leader's perspective and with the goal of identifying the optimal strategy for the leader to commit to. Recently, thanks primarily to the successful application of Stackelberg games to security resource allocation problems, known as the *Stackelberg security games* (SSGs) [Tambe, 2011], this classical game model becomes a subject of research interest again, and SSGs have become a major playground for research on Stackelberg games. This thesis focuses on two aspects of Stackelberg games — decentralized leadership and follower deception — that stem from reasoning about strategic interactions at a higher level in Stackelberg games and SSGs.

Our first focus is decentralized leadership. The canonical model of Stackelberg games has only one leader in it. Previous research has mostly focused on this model or similar single-leader variants, which cannot be used to study scenarios involving multiple leaders. Indeed, multi-leader scenarios abound in the real world. In security scenarios, examples such as multi-national patrol along international borders or co-located operations of different security agencies, fall right into this domain. Security agencies act as the leaders in the game and share a common interest in combating the same enemy — a potential attacker who poses a threat to the security of the patrolled region; the attacker acts as the follower in the game, surveying the leaders' strategies and responding rationally afterwards. Without proper coordination, defense is usually carried out inefficiently under decentralized leadership because of independent movement of the leaders. This motivates us to build a multi-leader SSG model to understand the interactions in the game, on top of which we study how to coordinate the leaders towards a more efficient defense. We are particularly interested in scenarios where the leaders' goals are not completely aligned: in a multi-national patrol, for instance, countries would naturally attach more importance to incidents involving their own properties, or that are in the vicinity of their own territories. Such inconsistency indicates that to coordinate the leaders means more than casting the game as a single-leader one, with a single

objective function that represents the benefit of all. Indeed, we face a collective decision-making problem, which calls for a mechanism to generate joint strategies for the leaders.

Our second focus is follower deception. Information uncertainty is one of the major challenges facing applications of game theory. In Stackelberg games, knowing the follower's payoffs is essential for the leader to compute their optimal strategy to commit to. When this information is uncertain or unknown, various approaches have been proposed to deal with uncertainties, typically by gathering information through interacting with the follower and observing the follower's responses. The second part of the thesis is motivated by the finding that, in the case of information asymmetry, the follower can tamper with the leader's information gathering attempt, with the aim of deceiving the leader into playing a strategy that benefits the follower. One particular approach the follower can use is to imitate the behavior of a different follower type, who has a different set of payoffs. A leader who ignores this strategic manipulation will be misled into playing the optimal strategy against the imitated follower type, which may be highly suboptimal in the original game. Essentially, the follower's deceptive behavior places an additional layer of strategic interaction between the leader and the follower on top of the standard Stackelberg game model. This brings us to a fundamental question: how to design a leader strategy in the presence of follower deception?

We aim at answering these questions in this thesis.

1.1 Thesis Outline and Main Contributions

In Chapter 2, we introduce the preliminaries and review the related literature. Then we present the two major technical parts: Part I consists of Chapters 3 and 4 and sets out to study decentralized leadership in SSGs and ways to coordinate the leaders from the perspective of mechanism design; and Part II consists of Chapters 5–7, which are developed around the topic of follower deception. We conclude the thesis in Chapter 8 with a discussion about potential directions for future work. In more detail, our two main technical parts are organized as follows.

Part I: Decentralized Leadership and Coordination

In Chapter 3, we present an SSG model with multiple leaders. We study the uncoordinated variant of this model, where the leaders move independently, and analyze the equilibrium of the game. Based on our understanding of the uncoordinated situation, we then study in Chapter 4 how to coordinate the leaders towards a better defense. The main contributions of Part I are as follows.

(a) We build a multi-leader SSG model and propose a novel equilibrium concept by combining the ideas of Nash and Stackelberg equilibria.

(b) We show that an equilibrium exists in every multi-leader SSG and one can be computed in polynomial time. The result is non-trivial: though our equilibrium concept resembles the Nash equilibrium, the leader-follower structure of the model makes the game fundamentally different from a normal-form game, so existing approaches to proving the existence of a Nash equilibrium by applying a fixed-point theorem does not apply to our equilibrium concept.

(c) With respect to mechanism design for coordination, we show that, unfortunately, it is not possible for any mechanism to achieve maximum efficiency if we also want the mechanism to be strategyproof and to guarantee every leader their equilibrium utility. Nevertheless, we design mechanisms that achieve a subset of the desired properties, or all of them in

special cases. We also consider coordination by way of strategy recommendation when implementation of joint strategies cannot be enforced; we show additional positive and negative results for this setting. Our results provide a complete picture of what can and cannot be achieved through mechanism design, with respect the properties we consider. Meanwhile, they also highlight the intrinsic difficulty of voluntary defense coordination in SSGs.

Part II: Follower Deception and Countermeasures

In Chapter 5, we identify the phenomenon of follower deception, study countermeasures to mitigate the leader's potential losses due to follower deception, and present algorithmic results for the problem of computing optimal countermeasures. In Chapter 6, we take the follower's perspective and study how the follower should optimally disguise their payoff matrix in a Stackelberg game when the leader is unaware of such deception. While we consider general Stackelberg games in these two chapters, we extend our work to SSGs in Chapter 7 and present results that are unique to SSGs. The main contributions of Part II are as follows.

(a) Conceptually, we are the first to identify the phenomenon of follower deception in Stackelberg games. We formalize the notion of *imitative follower deception* and demonstrate that it is an efficient tool to improve the follower's utility.

(b) We propose a *policy*-based approach for the leader to mitigate potential losses due to follower deception. We study the associated problem of computing the optimal policy in several natural settings and provide an almost complete picture of the complexity landscape: the problems facing the leader are hard even to approximate in general; on the other hand, we design efficient algorithms with approximation ratios that match our inapproximability results.

(c) From the follower's perspective, we present a polynomial-time algorithm to compute an optimal (or near-optimal) payoff matrix to imitate, i.e., a matrix that will maximize the follower's utility in the induced game if the leader is unaware of the deception.

(d) With respect to SSGs, we present a key finding that, it is often optimal for the follower to pretend that the game being played is zero-sum; consequently, the leader will be led to play a maximin strategy, which essentially invalidates the leader's information gathering attempt as to compute a maximin strategy requires no additional information about the opponent.

(e) We present a polynomial-time algorithm for the leader to compute an optimal policy in an SSG. This result contrasts the hardness results we obtain for general Stackelberg games. We also design a heuristic algorithm for the leader to use when the space of possible follower types is infinite or unknown; performance of the algorithm is evaluated via numerical experiments.

1.2 Publications

This thesis covers results from the following publications.

Part I (Chapters 3 and 4)

- Jiarui Gan, Edith Elkind, and Michael Wooldridge. Stackelberg security games with multiple uncoordinated defenders. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'18)*, pages 703–711, 2018. [Gan et al., 2018].

- Jiarui Gan, Edith Elkind, Sarit Kraus, and Michael Wooldridge. Mechanism design for defense coordination in security games. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'20)*, pages 402–410, 2020. [Gan et al., 2020].

Part II (Chapters 5–7)

- Jiarui Gan, Haifeng Xu, Qingyu Guo, Long Tran-Thanh, Zinovi Rabinovich, and Michael Wooldridge. Imitative follower deception in Stackelberg games. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC'19)*, pages 639–657, 2019b. [Gan et al., 2019b].

- Georgios Birmpas, Jiarui Gan, Alexandros Hollender, Francisco J. Marmolejo-Cossío, Ninad Rajgopal, and Alexandros A. Voudouris. Optimally deceiving a learning leader in Stackelberg games. In *Advances in Neural Information Processing Systems (NeurIPS'20)*, pages 20624–20635, 2020. [Birmpas et al., 2020].

- Jiarui Gan, Qingyu Guo, Long Tran-Thanh, Bo An, and Michael Wooldridge. Manipulating a learning defender and ways to counteract. In *Advances in Neural Information Processing Systems (NeurIPS'19)*, pages 8274–8283, 2019. [Gan et al., 2019a]

Chapter 2

Preliminaries

In this chapter, we introduce Stackelberg games.

2.1 Stackelberg Games

A Stackelberg game is a sequential game played between a *leader* and a *follower*. For ease of presentation, we will refer to the leader as a female and to the follower as a male throughout. The leader commits to a strategy, and the follower then reacts upon observing this commitment. For the entirety of this thesis, we focus on finite Stackelberg games, in which the leader and the follower have m and n *pure strategies* at their disposal, respectively, and their utilities for all possible combinations of pure strategies are given by two matrices $u^{\mathsf{L}}, u^{\mathsf{F}} \in \mathbb{R}^{m \times n}$. Games represented in this form are also called *bi-matrix games*. For any positive integer k, we write $[k] = \{1, \dots, k\}$. The entries $u^{\mathsf{L}}(i, j)$ and $u^{\mathsf{F}}(i, j)$ denote the utilities of the leader and the follower, under the *pure strategy profile* $(i, j) \in [m] \times [n]$, in which the leader plays strategy i and the follower plays j. The players are allowed to use *mixed* strategies whereby they flip a coin to decide which pure strategy to play. A mixed strategy of the leader is a probability distribution over $[m]$, denoted by a vector $\mathbf{x} = (x_1, \dots, x_m) \in \Delta^{m-1}$, where each x_i is the probability of pure strategy i being chosen and Δ^{m-1} denotes the standard $(m-1)$-simplex, i.e.,

$$\Delta^{m-1} = \left\{ \mathbf{x} \in \mathbb{R}^m : \sum_{i \in [m]} x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \in [m] \right\}.$$

By slightly abusing notation, we let

$$u^{\mathsf{L}}(\mathbf{x}, j) = \sum_{i \in [m]} x_i \cdot u^{\mathsf{L}}(i, j)$$

be the *expected utility* of the leader when she plays the mixed strategy $\mathbf{x}$ and the follower plays a pure strategy j. Similarly, we let $u^{\mathsf{F}}(\mathbf{x}, j) = \sum_{i \in [m]} x_i \cdot u^{\mathsf{F}}(i, j)$ denote the expected utility for the follower. For a given mixed strategy $\mathbf{x} \in \Delta^{m-1}$ of the leader, we say that $j \in [n]$ is a *follower best response* if $u^{\mathsf{F}}(\mathbf{x}, j) = \max_{\ell \in [n]} u^{\mathsf{F}}(\mathbf{x}, \ell)$; we denote the set of all follower best responses to $\mathbf{x}$ by $\mathsf{BR}(\mathbf{x})$, i.e.,

$$\mathsf{BR}(\mathbf{x}) = \arg\max_{j \in [n]} u^{\mathsf{F}}(\mathbf{x}, j).$$

2.1.1 Stackelberg Equilibrium

Stackelberg equilibrium is the basic solution concept for Stackelberg games. It captures the leader's optimal strategy commitment: in a Stackelberg equilibrium, the leader plays her optimal strategy $\mathbf{x}$, considering the follower's rational behavior upon observing this strategy (i.e., to best respond by playing $j \in \mathsf{BR}(\mathbf{x})$). Often, the leader faces the situation where there are multiple best responses in $\mathsf{BR}(\mathbf{x})$ that are by definition equally appealing to the follower. Hence, the assumption about the follower's tie-breaking behavior plays a role in determining an equilibrium.

Tie-breaking Assumptions Two types of Stackelberg equilibria can then be defined by using different tie-breaking assumptions. The strong form, know as *Strong Stackelberg equilibrium* (SSE) adopts the optimistic assumption that the follower always breaks ties in favor of the leader and chooses the best response that maximizes the leader's utility. In contrast, the weak counterpart of SSE, *weak Stackelberg equilibrium* (WSE), adopts the pessimistic assumption that the follower always seeks to minimize the leader's utility.

Definition 2.1 (Strong/weak Stackelberg equilibrium)**.** Suppose $\mathbf{x} \in \Delta^{m-1}$ and $j \in \mathsf{BR}(\mathbf{x})$. Then the strategy profile $(\mathbf{x}, j)$ is a *strong Stackelberg equilibrium* (SSE) if the following holds:

$$u^{\mathsf{L}}(\mathbf{x}, j) = \max_{\mathbf{y} \in \Delta^{m-1}} \max_{\ell \in \mathsf{BR}(\mathbf{y})} u^{\mathsf{L}}(\mathbf{y}, \ell).$$

It is a *weak Stackelberg equilibrium* (WSE) if the following holds:

$$u^{\mathsf{L}}(\mathbf{x}, j) = \max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in \mathsf{BR}(\mathbf{y})} u^{\mathsf{L}}(\mathbf{y}, \ell).$$

SSE has been used predominantly as the standard solution concept for Stackelberg games in the literature, even counterintuitively in adversarial scenarios such as security games. There are two reasons for this choice of assumption. First, the tie-breaking behavior of the follower in an SSE can often be induced if the leader makes an infinitesimal perturbation to her strategy. Second, an SSE always exists in a Stackeberg game, but a WSE may not [von Stengel and Zamir, 2004].

It is well known that the role as a first-mover gives the leader an advantage over the situation where both players move simultaneously. The leader's utility in an SSE is at least that in any Nash equilibrium, where each player plays a strategy that is optimal against the other player's strategy. Indeed, if the leader commits to her strategy in a Nash equilibrium that maximizes her utility amongst all Nash equilibria, the resulting strategy profile in a Stackelberg game will be the same as this optimal Nash equilibrium and provide the leader her utility in it. In addition to this utility guarantee, the computation of an SSE of a bi-matrix game is also tractable, which is in contrast to the PPAD hardness for computing a Nash equilibrium [Daskalakis et al., 2009]. Specifically, the following multiple-LP (linear program) approach computes an SSE of a game given by u^{L} and u^{F} [Conitzer and Sandholm, 2006]. It involves solving n LPs, each fixing the follower's best response to a pure strategy $j \in [n]$ and computing the optimal leader strategy that induces j as a best response of the follower; the solution of the LP that yields the maximum leader utility amongst all the LPs forms an SSE along with the corresponding follower best response j.

$$\begin{aligned} \max_{\mathbf{x}} \quad & u^{\mathsf{L}}(\mathbf{x}, j) \\ \text{subject to} \quad & u^{\mathsf{F}}(\mathbf{x}, j) \geq u^{\mathsf{F}}(\mathbf{x}, \ell) && \text{for all } \ell \in [n] \\ & \mathbf{x} \in \Delta^{m-1} \end{aligned}$$

As we can see, the constraints in this LP restrict the leader's strategy to be the ones that incentivize the follower to best respond by playing j. An LP can be solved in polynomial time [Bertsimas and Tsitsiklis, 1997] and there are many software tools with very good performance in solving LPs in practice (e.g., CPLEX[1]), which we also used to run experiments in this thesis.

2.1.2 Bayesian Stackelberg Games

Apparently, the follower's payoff information is necessary for the leader to compute her optimal commitment. When the leader only has a probabilistic belief about the follower's payoff information, *Bayesian Stackelberg game* (BSG) is a model proposed to study how to commit optimally in this case. In a BSG, the follower's type follows a distribution σ over a type space Θ. In the case of a finite Θ, each type $\theta \in \Theta$ corresponds to a payoff matrix u_θ^{F} and appears with probability σ_θ. This can also be thought of as the leader playing with a large population of followers, in which followers of each type θ constitute a portion σ_θ of the total population. It is assumed that σ is the common knowledge (and, in particular, is known to the leader), so we can define the *Bayesian Stackelberg equilibrium* (BSE) in a similar fashion to the setting with complete information; see Definition 2.2. In particular, we define the BSE using the strong form of tie-breaking assumption; the corresponding weak form will not be considered in this thesis. For simplicity, we also assume that the leader's payoffs are invariant with respect to the follower type she is playing against; hence, the leader's payoff matrix is always u^{L}. We extend the notation and let $\mathsf{BR}_\theta(\mathbf{x}) = \arg\max_{j\in[n]} u_\theta^{\mathsf{F}}(\mathbf{x}, j)$ denote the set of best responses of a type-θ follower against leader strategy $\mathbf{x}$.

Definition 2.2 (Bayesian Stackelberg equilibrium)**.** Suppose $\mathbf{x} \in \Delta^{m-1}$ and $j_\theta \in \mathsf{BR}_\theta(\mathbf{x})$ for each $\theta \in \Theta$. Then $\mathbf{x}$ and $(j_\theta)_{\theta\in\Theta}$ form a *(strong) Bayesian Stackelberg equilibrium* (BSE) if the following holds:

$$\mathbb{E}_{\theta\sim\sigma} u^{\mathsf{L}}(\mathbf{x}, j_\theta) = \max_{\mathbf{y}\in\Delta^{m-1}} \mathbb{E}_{\theta\sim\sigma} \max_{\ell\in\mathsf{BR}_\theta(\mathbf{y})} u^{\mathsf{L}}(\mathbf{y}, \ell).$$

A Bayesian game can be transformed into the bi-matrix form using the Harsanyi transformation [Harsanyi and Selten, 1972]. Nevertheless, since the transformation breaks the compact structure of the original game, the resulting game can be exponentially large in the size of the original one, rendering this approach unscalable in practice. In fact, it appears that there is no concise formulation for computing a BSE unless P = NP as Conitzer and Sandholm [2006] showed that the problem of computing a BSE is NP-hard. A naive way to compute a BSE, similar to the multiple-LP approach for computing an SSE, is to enumerate combinations of best responses of all the follower types; for each possible best response combination $(j_\theta)_{\theta\in\Theta}$, we solve the following LP:

$$\begin{aligned} \max_{\mathbf{x}} \quad & \sum_{\theta\in\Theta} \sigma_\theta \cdot u^{\mathsf{L}}(\mathbf{x}, j_\theta) \\ \text{subject to} \quad & u_\theta^{\mathsf{F}}(\mathbf{x}, j_\theta) \geq u_\theta^{\mathsf{F}}(\mathbf{x}, \ell) && \text{for all } \theta \in \Theta,\ \ell \in [n] \\ & \mathbf{x} \in \Delta^{m-1} \end{aligned}$$

The first constraint now restricts the leader's strategy to the ones that incentivize each follower type θ to take action j_θ.

[1] https://www.ibm.com/uk-en/analytics/cplex-optimizer

2.2 Stackelberg Security Games

One of the most prominent recent applications of Stackelberg games is its allocation in security resource allocations problems, i.e., *Stackelberg security games* (SSGs). An SSG models the scenario where: a defender, who acts as the leader, allocates their security resources (e.g., security guards, patrol canines, road blocks, etc.) to protect a set of targets (ports, boarding gates in an airport, transportation system, etc.); and a strategic attacker, who acts as the follower, conducts surveillance on the defender's allocation strategy and, then, launches an attack on the most attractive target, considering both the benefits of attacking this target and the probability of it being protected by the defender.

Let there be k (homogeneous) security resources and a set of targets $T = \{t_1, \ldots, t_m\}$. In the most basic setting, an attack on a target $t \in T$ will be unsuccessful if at least one resource is allocated to t, in which case the attacker receives a value $p^{\mathsf{F}}(t)$ as his utility, which we refer to as a *penalty*, and the defender receives a value $r^{\mathsf{L}}(t)$ referred to as a *reward*. Otherwise, i.e., when no resource is allocated to t, the attack will be successful, in which case the attacker receives a reward $r^{\mathsf{F}}(t)$ and the defender receives a penalty $p^{\mathsf{L}}(t)$. We say that a target is *protected*, or *covered*, if at least one resource is assigned to it; and *unprotected*, or *uncovered*, otherwise. It is assumed that $r^{\mathsf{F}}(t) > p^{\mathsf{F}}(t)$ and $r^{\mathsf{L}}(t) > p^{\mathsf{L}}(t)$ for all t, so the attacker always prefers a successful attack while the defender prefers it to be unsuccessful.

Since one resource is sufficient to fail an attack by the above payoff definition, the set of all rational ways of resource allocation corresponds to a subset of T consisting of at most k targets; each target in this subset is assigned a resource. We can represent an allocation as a vector $\mathbf{s} \in \{0,1\}^m$, such that $s_t = 1$ if one resource is allocated to target t, and $s_t = 0$, otherwise. We let $\mathcal{T}_k = \left\{\mathbf{s} \in \{0,1\}^m : \sum_{t \in T} s_t \leq k\right\}$ denote the set of feasible allocations when the defender has k resources. Since resources are often insufficient to cover all the targets, a good strategy for the defender is to further randomize the allocation and commit to a mixed strategy $\mathbf{x} \in \Delta(\mathcal{T}_k)$, where for any finite set S, $\Delta(S)$ denotes the set of all distributions over S, i.e.,

$$\Delta(S) = \left\{ \mathbf{x} \in \mathbb{R}^{|S|} : \sum_{\mathbf{s} \in S} x_{\mathbf{s}} = 1, \text{ and } x_{\mathbf{s}} \geq 0 \text{ for all } \mathbf{s} \in S \right\}.$$

The structure of SSGs allows a mixed strategy of the defender to be represented more compactly as a *coverage (vector)* $\mathbf{c} = (c_t)_{t \in T}$, with each c_t representing the probability that target t is protected. We will mostly stick to this representation and use the terms *coverage* and *defender mixed strategy* interchangeably. Sometimes, we also write a coverage vector as $\mathbf{c} = (c_1, \ldots, c_m)$, with c_i representing the coverage of target t_i. Under the constraint that the defender can use at most k resources, the set of feasible mixed strategies available to the defender is

$$\mathcal{C}_k = \left\{ \mathbf{c} \in \mathbb{R}^m : 0 \leq c_t \leq 1 \text{ for all } t \in T, \text{ and } \sum\nolimits_{t \in T} c_t \leq k \right\}. \tag{2.1}$$

Any coverage vector $\mathbf{c} \in \mathcal{C}_k$ can be implemented as a distribution $\mathbf{x} \in \Delta(\mathcal{T}_k)$, and any such distribution results in a coverage vector in $\mathcal{C}$: in one direction, a distribution $\mathbf{x}$ naturally results in each target being covered with probability $c_t = \sum_{\mathbf{s} \in \mathcal{T}_k} x_{\mathbf{s}} \cdot s_t$; whereas in the other, a distribution $\mathbf{x}$ implementing any $\mathbf{c} \in \mathcal{C}_k$ can be constructed efficiently, using e.g., the comb sampling algorithm [Tsai et al., 2010] (see Figure 2.1).

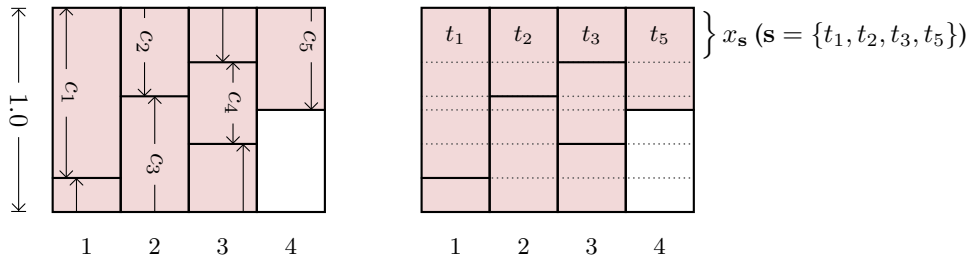


Figure 2.1: The comb-sampling algorithm: an example with $k = 4$ and $m = 5$. The algorithm creates k columns $1, \ldots, k$, which have height 1 and each correspond to a resource. Given a coverage $\mathbf{c}$, $\sum_{i=1}^{m} c_i \leq k$, we can arrange the coverage mass into the columns in the way shown on the left. The intersections of the coverage masses split the columns into at most $m + 1$ rows as shown on the right, each corresponding to a pure strategy; the height of the row corresponds to the probability that this pure strategy should chosen in the distribution $\mathbf{x}$ equivalent to $\mathbf{c}$ (e.g., $x_{\mathbf{s}}$ shown above is the probability of allocating four resources to protect targets t_1, t_2, t_3, and t_5).

When the defender plays a mixed strategy $\mathbf{c}$ and the attacker attacks a target $t \in T$, the expected utilities of the defender and the attacker, denoted $u^{\mathsf{L}}(\mathbf{c}, t)$ and $u^{\mathsf{F}}(\mathbf{c}, t)$, respectively, are given as follows.

$$u^{\mathsf{L}}(\mathbf{c}, t) = c_t \cdot r^{\mathsf{L}}(t) + (1 - c_t) \cdot p^{\mathsf{L}}(t)$$
$$u^{\mathsf{F}}(\mathbf{c}, t) = c_t \cdot p^{\mathsf{F}}(t) + (1 - c_t) \cdot r^{\mathsf{F}}(t)$$

Therefore, $u^{\mathsf{L}}(\mathbf{c}, t)$ is strictly increasing with respect to c_t, and $u^{\mathsf{F}}(\mathbf{c}, t)$ is strictly decreasing. We also use the notation $u^{\mathsf{L}}(c, t) = c \cdot r^{\mathsf{L}}(t) + (1-c) \cdot p^{\mathsf{L}}(t)$ and $u^{\mathsf{F}}(c, t) = c \cdot p^{\mathsf{F}}(t) + (1-c) \cdot r^{\mathsf{F}}(t)$ when the first variable c is a scalar representing the coverage of target t. As in a normal-form Stackelberg game, in an SSG the attacker can observe the defender's mixed strategy through surveillance before he launches an attack; however, the instantiated pure strategy is not observable.

According to Definition 2.1, a strategy profile $(\mathbf{c}, t)$ forms an SSE in a Stackelberg security game if the following holds:

$$(\mathbf{c}, t) \in \arg \max_{\mathbf{z} \in \mathcal{C},\, \ell \in \mathsf{BR}(\mathbf{z})} u^{\mathsf{L}}(\mathbf{z}, \ell).$$

It is worth mentioning that, in a security game, there is a straightforward way to induce the desired tie-breaking behavior of the attacker in an SSE: to make the attacker strictly prefer attacking some target in her best response set, the defender can simply reduce the coverage of that target by an infinitesimal amount. Although an SSG can be exponentially large when converted to the bi-matrix form (i.e., the size of the pure strategy set $\mathcal{T}_k$ grows exponentially with respect to k and n), the compact representation of mixed strategies allows us to use the same multiple-LP approach to compute an SSE. Specifically, the following is the LP to be solved when the attacker's best response is fixed to a target $t \in T$:

$$\begin{aligned} \max_{\mathbf{c}} \quad & u^{\mathsf{L}}(\mathbf{c}, t) \\ \text{subject to} \quad & u^{\mathsf{F}}(\mathbf{c}, t) \geq u^{\mathsf{F}}(\mathbf{c}, \ell) && \text{for all } \ell \in T \\ & \mathbf{c} \in \mathcal{C}_k \end{aligned}$$

2.2.1 An SSG Example

Let us look at a concrete example of an SSG. We consider a security agency (the defender) that wants to protect two conservation areas t_1 and t_2 from a poacher (the attacker). The defender has one security guard as her resource, who can patrol one of the two areas, while the poacher chooses one area to attack. The payoffs of the players are given as follows.

	t_1	t_2
reward r^{L}	1	0.99
penalty p^{L}	-1	-1

defender

	t_1	t_2
reward r^{F}	3	$1/3$
penalty p^{F}	-1	-1

attacker

The corresponding bi-matrix form of the game is given by the following payoff matrices u^{L} and u^{F} for the defender and the attacker, respectively. The defender is the row player and has three pure strategies, which correspond to the subsets $\{t_1\}$, $\{t_2\}$, and $\varnothing$ containing the targets the defender protects; the attacker is the column player and has two pure strategies, i.e., to attack t_1 or t_2.

$$u^{\mathsf{L}} = \begin{array}{c} \begin{matrix} t_1 & t_2 \end{matrix} \\ \begin{pmatrix} 1 & -1 \\ -1 & 0.99 \\ -1 & -1 \end{pmatrix} \end{array} \begin{matrix} \\ \{t_1\} \\ \{t_2\} \\ \varnothing \end{matrix} \qquad u^{\mathsf{F}} = \begin{array}{c} \begin{matrix} t_1 & t_2 \end{matrix} \\ \begin{pmatrix} -1 & 1/3 \\ 3 & -1 \\ 3 & 1/3 \end{pmatrix} \end{array} \begin{matrix} \\ \{t_1\} \\ \{t_2\} \\ \varnothing \end{matrix}$$

We can use the multiple-LP approach to compute an SSE of the game. In this game, there is a unique SSE in which the defender uses the coverage vector $\mathbf{c}^* = (3/4, 1/4)$, patrolling t_1 with probability $3/4$ and t_2 with $1/4$ (which is equivalent to mixed strategy $\mathbf{x}^* = (3/4,\ 1/4,\ 0)$ in the bi-matrix form of this game). When $\mathbf{c}^*$ is played, the attacker finds attacking t_1 and t_2 equally appealing choices since we have:

$$u^{\mathsf{F}}(\mathbf{c}^*, t_1) = (-1) \times \tfrac{3}{4} + 3 \times \tfrac{1}{4} = 0$$
$$u^{\mathsf{F}}(\mathbf{c}^*, t_2) = \tfrac{1}{3} \times \tfrac{3}{4} + (-1) \times \tfrac{1}{4} = 0$$

and hence $\mathsf{BR}(\mathbf{c}^*) = \{t_1, t_2\}$. By the optimistic tie-breaking assumption, the attacker breaks ties in favor of the defender and chooses to attack t_1 given that

$$u^{\mathsf{L}}(\mathbf{c}^*, t_1) = 1 \times \tfrac{3}{4} + (-1) \times \tfrac{1}{4} = \tfrac{1}{2}$$
$$u^{\mathsf{L}}(\mathbf{c}^*, t_2) = (-1) \times \tfrac{3}{4} + 0.99 \times \tfrac{1}{4} < \tfrac{1}{2}$$

Indeed, consider an arbitrary strategy $\mathbf{c}$ of the defender. We have $t_1 \in \mathsf{BR}(\mathbf{c})$ or $t_2 \in \mathsf{BR}(\mathbf{c})$ since at least one target has to be a best response of the attacker. The former means that $u^{\mathsf{F}}(\mathbf{c}, t_1) \geq u^{\mathsf{F}}(\mathbf{c}, t_2)$, which by the payoffs in this instance implies $c_1 \leq 3c_2$. Given that the defender has only one resource, we have $c_1 + c_2 \leq 1$ (i.e., $\mathbf{c} \in \mathcal{C}_1$) and it follows that:

$$c_1 + c_1/3 \leq c_1 + c_2 \leq 1, \quad \Longrightarrow \quad c_1 \leq 3/4, \ \text{ and } \ u^{\mathsf{L}}(\mathbf{c}, t_1) \leq u^{\mathsf{L}}(\mathbf{c}^*, t_1).$$

On the other hand, if it is the case that $t_2 \in \mathsf{BR}(\mathbf{c})$, then we have $c_1 \geq 3c_2$ and, again, by $c_1 + c_2 \leq 1$, it follows that

$$3c_2 + c_2 \leq c_1 + c_2 \leq 1, \quad \Longrightarrow \quad c_2 \leq 1/4, \ \text{ and } \ u^{\mathsf{L}}(\mathbf{c}, t_2) \leq u^{\mathsf{L}}(\mathbf{c}^*, t_2) < u^{\mathsf{L}}(\mathbf{c}^*, t_1).$$

In any case, the defender cannot obtain utility more than $u^{\mathsf{L}}(\mathbf{c}^*, t_1)$. Hence, $(\mathbf{c}^*, t_1)$ forms an SSE.

When the Attacker Type is Uncertain

Suppose that the defender is uncertain about the attacker's payoffs and now the attacker may be one of the following two types of poachers, who value the two targets differently due to their interests in animal species in these two areas. Suppose further that the attacker appears as types α or β with equal probability.

	t_1	t_2
r_α^{F}	3	$1/3$
p_α^{F}	-1	-1

poacher type α

	t_1	t_2
r_β^{F}	1	1
p_β^{F}	-1	-1

poacher type β

Now if the defender still commits to the same strategy $\mathbf{c}^*$ (which is optimal when the attacker's type is α with certainty), then the best responses of two poacher types are $\mathsf{BR}_\alpha(\mathbf{c}^*) = \{t_1, t_2\}$ and $\mathsf{BR}_\beta(\mathbf{c}^*) = \{t_2\}$. By the optimistic tie-breaking assumption, a type-α poacher will best respond by attacking t_1 and a type-β poacher will attack t_2. This results in the defender obtaining expected utility

$$0.5 \cdot u^{\mathsf{L}}(\mathbf{c}^*, t_1) + 0.5 \cdot u^{\mathsf{L}}(\mathbf{c}^*, t_2) = -1/800.$$

It turns out that the optimal strategy for the defender is to play $\mathbf{c}^{**} = (1/2, 1/2)$, by which we have $\mathsf{BR}_\alpha(\mathbf{c}^{**}) = \{t_1\}$ and $\mathsf{BR}_\beta(\mathbf{c}^*) = \{t_1, t_2\}$ and both poacher types will respond by attacking t_1; the defender's expected utility increases to

$$0.5 \cdot u^{\mathsf{L}}(\mathbf{c}^{**}, t_1) + 0.5 \cdot u^{\mathsf{L}}(\mathbf{c}^{**}, t_1) = 0.$$

The strategy profile $(\mathbf{c}^{**},\ t_1,\ t_1)$ is a BSE of this game.

2.2.2 Relevant Variants of SSGs

The study of the SSG originates from the pioneering work of Tambe *et al.*, which then led to a series of high-impact applications, including automated tools developed and put in use to assist the protection of important infrastructure, transportation systems, public events, and even wild life and natural resources [Tambe, 2011, Fang and Nguyen, 2016, Sinha et al., 2018]. Ever since, a multitude of variants of the standard SSG model have been proposed and studied in the literature. We provide a brief overview of the models that are most relevant to our work.

SSG with Scheduling Constraints

In the basic SSG model we introduced, the only constraint imposed on resource allocation is the budget constraint. Other realistic constraints may come into play in more complicated scenarios. For example, the Federal Air Marshal Service (FAMS) in the US places undercover police guards on flights to deter potential attacks. In the SSG model developed for studying the optimal placement of police guards, police guards are to be allocated to flights [Tsai et al., 2009]. Usually one police guard can be allocated to multiple flights in a day, but this is subject to the hard constraint that one police guard can only be allocated to a sequence of connected flights. As another example, the protection a resource provides may exhibit *externality*, when it protects not only the target it is allocated to, but also other associated ones, such as targets in the neighborhood on a graph or

targets near by in a Euclidean space [Gan et al., 2015b, 2017]. The presence of protection externality imposes constraints on the set of *effective* pure strategies, i.e., the coverage vectors resulting from deterministic resource allocation strategies, as it alters the landscape of feasible pure strategies.

In general, in models with additional allocation constraints, resources are not allocated directly to the targets but to subsets of the targets — be it a string of connected flights, or a group of targets in a neighborhood. These subsets are often referred to as *schedules* and the additional allocation constraints as *scheduling constraints*. With additional scheduling constraints, the problem of computing an SSE tends to be harder (e.g., NP-hard as in [Korzhyk et al., 2010, Gan et al., 2015b, 2017]). To deal with this computational obstacle, various approaches have been proposed. Amongst them, the *column generation* (CG) framework is perhaps the most effective and a widely applicable tool to tackle the complex strategy space resulting from the scheduling constraints. This idea was first introduced by Jain et al. (2010) to computing an SSE, and has been subsequently adopted for the purpose of solving a number of similar problems [e.g., Gan et al., 2015b, Guo et al., 2016a,b, Gan et al., 2017].

CG is inspired by the Carathéodory's theorem [Danninger-Uchida, 2001], which, in the context of SSG, is equivalent to saying that every defender mixed strategy, as a convex combination of coverage vectors resulting from feasible pure strategies, always admits a concise expression using as few as $m + 1$ pure strategies (the unused ones are chosen with probability 0 in this mixed strategy), where m is the number of targets. In other words, not all pure strategies in the exponentially large space are necessary for constructing an optimal strategy of the defender. In a nutshell, a CG-based algorithm starts by restricting the defender's choice of pure strategies to a small subset (often randomly chosen) of the entire pure strategy set. It repeatedly solves the small-scale problem constructed on top of this small strategy set and adds new strategies that are able to improve the solution quality into this set. When no such strategies can be found, the optimal solution to the small-scale problem is provably also optimal with respect to the original problem that involves the entire pure strategy set. In more complicated scenarios, where the attacker also faces hardness in computing a best response (e.g., when the attacker's pure strategy corresponds to a path in a road network, which amounts to exponentially many choices in the size of the network), a more sophisticated *double oracle* framework was often used to compute an SSE. This framework incorporates an additional but similar procedure that gradually expands the attacker's strategy space (see, e.g., [Jain et al., 2011, 2013, Wang et al., 2016, Yin and An, 2016]).

As we shall see in Section 3.6.2, Chapter 3, additional scheduling constraints indeed change results obtained with the standard model.

Bayesian SSG

As we have shown in Section 2.2.1, Bayesian Stackelberg games can be used to model uncertainty in SSGs. These models are known as Bayesian SSGs. Because of the complexity of computing a BSE, the main focus of the related literature on Bayesian SSGs is to develop algorithms that scale up to real applications. To the best of our knowledge, Paruchuri et al. [2007] first used the concept of Bayesian games to model an SSG with multiple types of followers. To bypass the computational complexity, they developed a heuristic approach called ASAP, in which they restrict the defender's mixed strategies to ones that have a special structure, in an attempt to alleviate the computational difficulty. In subsequent work [Paruchuri et al., 2008a], they further proposed an exact algorithm DOBSS based on mixed integer linear programs. These algorithms form the core of the ARMOR system that was later developed and deployed at the Los Angeles International Airport to automate scheduling of security personnel [Pita et al., 2008, Jain et al., 2008, Pita et al., 2009]. More recently,

Bayesian SSG models have also been used in a number of other applications across many fields, such as contagion control in a social network, routing selection in adversarial scenarios, maritime patrol, safe distribution of computation tasks, and protection of critical infrastructure [Tsai et al., 2013, Hochbaum et al., 2014, Oliva et al., 2018, Bai et al., 2019, Gu et al., 2019].

In Chapter 5, we will compare solutions of our policy-based framework with a BSE and show the improvement offered by our approach in tackling uncertainty.

Human Behavior Modeling in SSGs

Classic models in game theory assume that players are fully rational. In reality, however, human players hardly behaves like that. Human players do not use sophisticated algorithms to decide what to do from time to time, and may constantly deviate from utility maximizing strategies; limited or biased observations about the other players' strategies only add more randomness to their behavior. Without properly considering such behavioral features, the models we develop would be intrinsically inaccurate in describing or predicting outcomes of human interactions. Behavioral game theory sets out to remedy this inadequacy (see [Camerer, 2011]) and many models have been proposed along this research agenda [McKelvey and Palfrey, 1995, Costa-Gomes et al., 2001, Wright and Leyton-Brown, 2010]. Amongst these models, the *quantal response* (QR) model offers a simple, explainable framework McKelvey and Palfrey [1995], which has also been adopted in SSGs to model bounded rationality of an attacker [Pita et al., 2010, Yang et al., 2011, 2012, An et al., 2012, 2013, Shieh et al., 2013, Yang et al., 2013].

In the QR model, the players' actions are assumed to be possibly erroneous, instead of being always optimal. In a two-player game, a player would respond to the strategy $\mathbf{x}$ of the other player by playing each pure strategy i with probability

$$q_i(\mathbf{x}) = \frac{e^{\lambda U(\mathbf{x},j)}}{\sum_{i'} e^{\lambda U(\mathbf{x},j)}},$$

where e is the base of natural logarithm and λ is a constant representing the level of rationality: when $\lambda = 0$ the player behaves completely randomly, playing each of his pure strategies with equal probability, irrespective of utilities they yield; when $\lambda \to +\infty$, the player becomes fully rational, playing the pure strategies that maximizes his utility with certainty. For any $\lambda \in (0, +\infty)$, the probability that every strategy is selected is positively correlated to the utility the player gets by playing this strategy.

A player's behavior in the QR model increases the complexity of computing an optimal commitment in an SSG, making the associated optimization problem non-linear and non-convex (see, e.g., [Pita et al., 2010, Yang et al., 2012, 2013]). In Chapter 7, however, we will take advantage of the behavioral uncertainty introduced by this behavior as a countermeasure against attacker deception in SSGs.

2.3 Summary

While the SSG is the only concrete Stackelberg game model we will be focusing on in this thesis, we would like to point out that Stackelberg games have also been used to model strategic interactions in many other scenarios. For instance, this is the case for the principal-agent contract design problem [Page, 1991, Van Long and Sorger, 2010, Caldentey and Haugh, 2017], where a principal, the leader of the game, offers a contract to an agent who, upon receiving the contract, decides on

the amount of effort to spend towards fulfilling the contract to maximize their own payoff. The same paradigm fits into other similar applications, including exam design for large-scale tests, e.g., for MOOCs (massive open online courses) [Li and Conitzer, 2017], the competition among mining pools in the Bitcoin network [Marmolejo-Cossío et al., 2019, Sun et al., 2020], contagion control in social networks and influence maximization [Carnes et al., 2007, Tsai et al., 2013, Tanınmış et al., 2019], pricing and planning of transportation systems with strategic users [Gan et al., 2013, 2015a, Xiong et al., 2016, 2017, Gan and An, 2017], and even the analysis of a ballot manipulator's behavior in a district-based election [Elkind et al., 2019, 2021]. At a high level, most mechanism design problems can also be viewed as Stackelberg games, where the mechanism designer interacts with rational agents who best respond to the mechanism.

Part I

Decentralized Leadership and Coordination

Chapter 3

Decentralized Leadership in Security Games

We introduce a model of multi-defender security games, in which a group of defenders allocate security resources to protect some targets against an attacker. Our starting point is the observation that without coordination, the use of resources may be inefficient because of independent movement of the defenders, which motivates us to investigate the possibility of reducing this inefficiency through coordinating the defenders. Our task is twofold. First, we need to develop a good understanding of the uncoordinated situation as the baseline to improve. Second, the defenders' heterogeneous valuations for the targets give rise to a collective decision-making problem, which calls for a coordination mechanism to generate joint allocation strategies. In this chapter, we focus on the first task. We formulate a novel equilibrium concept suitable for the multi-defender scenario, which combines the ideas of Nash and Stackelberg equilibria. Our main result is that an equilibrium always exists and it can be computed efficiently. To prove this result requires non-standard techniques because the leader-follower structure prevents a fixed-point theorem from being applied (even if we view it as a game between the defenders). The results in this chapter serve as a foundation for us to tackle our second task, coordination mechanism design, in the next chapter.

3.1 Introduction

A group of defenders are patrolling the sea to combat their common enemy Captain Jack, a pirate who operates constantly in that area. Each defender acts independently, ordering their vessels to patrol the islands, which are the targets of Captain Jack. The choice of islands for day-to-day patrol is randomized. As a sophisticated pirate, Captain Jack plans carefully: he surveys the randomized patrolling patterns and chooses the best target to attack, taking into consideration both the likelihood that he runs into a patrol vessel on each target and the profits from attacking this target. Time after time, the defenders realize that independent movement may be rather inefficient. For example, when an island t is patrolled by vessels of two independent defenders, each with probability 0.5, the probability that t is patrolled by some vessel is $1-(1-0.5)^2=0.75$, and with probability $0.5\times 0.5=0.25$ two vessels appear at t simultaneously. Normally, one vessel is already capable of deterring an attack in confrontation with Captain Jack, so the latter scenario is a waste of resources. To increase efficiency, the defenders can reduce the unnecessary double protection by synchronising the defenders' movements, so that at any time exactly one vessel patrols at t. Then t would be covered with probability 1, while each defender would still contribute the same effort as

before. This example demonstrates that it can be very beneficial for a group of defenders who face a common adversary to collaborate.

Unfortunately, coordination may be much more challenging in real life because defenders may be heterogeneous and have different valuations for the targets. Each defender would hence prefer the group to adopt a different joint strategy, so as to better protect targets that are more important to them; the problem of choosing a joint strategy becomes very difficult. Our task is twofold. First, we need to develop a good understanding of the uncoordinated situation, as the baseline to improve. The scenario we want to study is essentially a Stackelberg security game (SSG). However, past research efforts on SSGs have been mostly put on models with only one defender; such models cannot be used to study scenarios involving multiple defenders, which are by nature very different. Second, when we want to coordinate the defenders, heterogeneity of the defenders gives rise to a collective decision-making problem, which calls for a coordination mechanism to generate joint allocation strategies. We seek a mechanism that improves the defenders' utility upon the uncoordinated situation, exploits efficiency gains, and incentivizes the defenders to report their true preferences and to execute the recommended strategies.

In this chapter, we focus on the first task and make the following contributions. We introduce a basic model of multi-defender SSG and formulate a novel equilibrium concept to characterize the outcome of the game when the defenders are uncoordinated. In contrast to existing multi-defender models — where the defenders can be assumed to have the same valuations over the targets [Jiang et al., 2013, Basilico et al., 2017], or each of them is associated with a disjoint target set [Smith et al., 2014, Lou and Vorobeychik, 2015, Lou et al., 2017] — we consider a more general setting where the defenders may value the targets differently, yet each of them can allocate security resources to any target. Our equilibrium concept, *Nash-Stackelberg equilibrium* (NSE), combines the ideas of Nash equilibrium and Stackelberg equilibrium to capture both the simultaneous movement among the defenders and the leader-follower action order between the defenders and the attacker. We prove that an NSE exists in every multi-leader SSG and can be computed in polynomial time. The proof of this result is non-trivial: though our equilibrium concept resembles the Nash equilibrium if we view the game as one between the defenders, the leader-follower structure in the model makes the utility functions of the defenders discontinuous, so existing approach to showing the existence of a Nash equilibrium by applying a fixed-point theorem does not apply to our model. The tractability of computing an equilibrium in our model is also in contrast to the PPAD complexity of computing a Nash equilibrium in normal-form games [Daskalakis et al., 2009].

Beyond hunting for Captain Jack, our model abstracts scenarios where heterogenous security agencies combat their common enemy in the same area, such as patrols along country borders or over international waters carried out by different countries, or co-located operations of police force and private security personnel. There are many such examples in the real world that motivate our interest in studying this model. To name but a few, it was reported that there were actions against oil-siphoning in the waters between Singapore, Indonesia, and Malaysia [Wang et al., 2018],[1] and against illegal fishing in the Palk Strait between Sri Lanka and India;[2] The US Coast Guard and local police departments are also known to patrol independently at some major ports in the USA [Jiang et al., 2013].

[1] https://www.economist.com/asia/2015/06/25/malacca-buccaneers

[2] http://thediplomat.com/2017/07/can-international-litigation-solve-the-india-sri-lanka-fishing-dispute/

3.1.1 Related Work

In the literature on Stackleberg games and SSGs, most models involve only one leader/defender, where the goal is normally to maximize the leader's expected utility in the game [Delle Fave et al., 2014, Paruchuri et al., 2008a, Pita et al., 2010]. Despite the dominance of single-leader models, there are a few exceptions. Jiang et al. [2013] and Basilico et al. [2017] studied SSGs with multiple *identical* defenders who have the same valuation for each target. Essentially, the defenders belong to the same organization but have to move independently because of technical constraints. These models can be studied within the single-defender framework since the goal is still to maximize a single utility function. When defenders are heterogeneous, however, the interaction between the defenders becomes much more complicated due to their unaligned interests. The outcome of the interaction can be predicted by means of equilibrium analysis. Specifically, Chan et al. [2012] considered games in which multiple players decide the amount of security investment. In contrast to our model, they assume that the attacker acts simultaneously with the defenders; besides that, there is no target in their model (each defender only decides to invest or not). In another series of papers [Smith et al., 2014, Lou and Vorobeychik, 2015, Lou et al., 2017], the model employed shares more similarities with ours: it is built on the SSG model and the defenders are heterogeneous. Nevertheless, it is assumed in these papers that the defenders protect disjoint subsets of the targets, which simplifies the problem considerably. In addition, these papers further depart from the standard model of SSGs by assuming that the attacker breaks ties uniformly at random. There are also some multi-defender SSG models intended for very specific applications, i.e., defense against spear-phishing attacks [Laszka et al., 2016], or traffic light control [Lou and Vorobeychik, 2016]). In contrast to the aforementioned work, our model is intended to be more general; it extends directly from the standard SSG, models heterogeneous defenders, and assumes that the defenders have access to and can protect all the targets. Our solution concept (the tie-breaking assumption, in particular) is also consistent with the standard definition of the Stackelberg equilibrium.

Beyond security domains, there are also research efforts toward multi-leader Stackelberg games [Leyffer and Munson, 2010, Kulkarni and Shanbhag, 2014], as well as multi-leader oligopoly models in which leaders choose the level of investment or production to maximize profits [Sherali, 1984, DeMiguel and Xu, 2009, Sinha et al., 2014]. However, these papers either employ a very different modeling approach or fall short of offering any algorithmic analysis. More recently, Castiglioni et al. [2019] studied the problem of computing correlated strategies for a group of leaders to commit to in generic Stackelberg games. Somewhat unusually, in their model a leader has the option to deviate from the correlated strategy so as to act as a follower in the game; such deviations would not be feasible in many security scenarios that we would like to model. Further, while the idea of correlating the leaders' strategies is conceptually similar to our approach, they did not consider strategyproofness and utility guarantees, so their results are very different from ours. Under their solution, players may be incentivized to misreport their preferences to the system and, hence, make other players worse off than even in the uncoordinated situation. In addition, their solution may, in a sense, force players to coordinate by threatening them with a worse outcome if they choose not to; we believe that this approach is not compatible with our agenda.

It is perhaps worth mentioning that, in contrast to the relatively sparse literature on multi-defender SSGs, games involving multiple attackers (or attacker types) have been studied quite extensively in various settings. These include Bayesian SSGs [Conitzer and Sandholm, 2006, Paruchuri et al., 2007, 2008b,a, Tsai et al., 2013], and coalitional SSGs with attacker networks [Guo

et al., 2016a, Wang et al., 2016]. A more recent paper also analyzed the worst equilibrium of a multi-follower Stackelberg game [Coniglio et al., 2017]. The difference in the structure of multi-leader games makes our analysis fundamentally different to that of games with multiple followers.

3.1.2 Chapter Overview

This chapter proceeds as follows. We introduce the multi-defender SSG model and present our equilibirum concept for the game in Section 3.2. Then, in Section 3.3, we provide a characterization of the equilibria we define, which helps us demonstrate the existence of an NSE in Section 3.4 and the easiness of the problem of computing an NSE in Section 3.5, where we present a polynomial-time algorithm for this problem. In addition to these main results, we also analyze the existence of an NSE in two extensions of our model and present the results in Section 3.6.

3.2 Multi-defender SSG

Our model extends the standard SSG model that we introduced in Chapter 2. To be consistent with the literature on SSGs, we will hereafter use the superscripts "D" and "A" (instead of "L" and "F") to indicate parameters for the defenders and for the attacker, respectively. There are n *defenders* $1, \dots, n$, who want to protect a set $T = \{t_1, \dots, t_m\}$ of m *targets* against an *attacker*. Each defender $i \in [n]$ has k_i identical *security resources* that can be allocated to the targets to protect them. We let $r^{\mathsf{A}}(t)$ and $p^{\mathsf{A}}(t)$ denote the attacker's reward and penalty, respectively, for each target $t \in T$; and let $r_i^{\mathsf{D}}(t)$ and $p_i^{\mathsf{D}}(t)$ be those of each defender $i \in [n]$. It is assumed that $r_i^{\mathsf{D}}(t) > p_i^{\mathsf{D}}(t)$ and $r^{\mathsf{A}}(t) > p^{\mathsf{A}}(t)$ for each $i \in [n]$ and $t \in T$, so each defender prefers an attack to be unsuccessful, and the attacker prefers the opposite. Thus, although defenders are heterogeneous, they all prefer every target to be safe. Suppose the attacker attacks target t and the defenders' movement results in coverage $\mathbf{c} = (c_{t_1}, \dots, c_{t_m})$. The expected utilities $u_i^{\mathsf{D}}(\mathbf{c}, t)$ of each defender i and the expected utility $u^{\mathsf{A}}(\mathbf{c}, t)$ of the attacker are given below:

$$u_i^{\mathsf{D}}(\mathbf{c}, t) = c_t \cdot r_i^{\mathsf{D}}(t) + (1 - c_t) \cdot p_i^{\mathsf{D}}(t) \tag{3.1}$$

$$u^{\mathsf{A}}(\mathbf{c}, t) = c_t \cdot p^{\mathsf{A}}(t) + (1 - c_t) \cdot r^{\mathsf{A}}(t) \tag{3.2}$$

By our assumption about the payoff parameters, $u_i^{\mathsf{D}}(\mathbf{c}, t)$ is strictly increasing with respect to c_t, and $u^{\mathsf{A}}(\mathbf{c}, t)$ is strictly decreasing.

We have yet to define how exactly the coverage vector is obtained from the defenders' strategies. Indeed, this depends on whether the defenders are coordinated or not. In the uncoordinated situation, the defenders move independently. We let $\mathbf{x}^i$ denote the mixed strategy of each defender $i \in [n]$, and let $X = (\mathbf{x}^1, \dots, \mathbf{x}^n)$ denote the defender strategy profile. Same as in a standard SSG, each $\mathbf{x}^i$ results in a coverage vector $\mathbf{c}^i = \mathsf{cov}(\mathbf{x}^i) \in \mathcal{C}_{k_i}$, where as defined in Chapter 2, $\mathcal{C}_k$ is the set of coverage vectors that can be implemented by using no more than k resources (see (2.1)). The profile X results in overall coverage (with slight abuse of notation) $\mathsf{cov}(X) = (\mathsf{cov}_{t_1}(X), \dots, \mathsf{cov}_{t_m}(X))$, such that

$$\mathsf{cov}_t(X) = 1 - \prod_{i \in [n]} \left(1 - \mathsf{cov}_t(\mathbf{x}^i)\right) \tag{3.3}$$

for each $t \in T$. Namely, a target is covered as long as some defender is protecting it.

In this chapter, we will mostly use a coverage vector $\mathbf{c} = \mathsf{cov}(\mathbf{x})$ to represent a defender mixed strategy $\mathbf{x}$ and view the strategy profile as $X = (\mathbf{c}^1, \dots, \mathbf{c}^n)$. For simplicity, we also write

$f(\cdot) = f(\mathsf{cov}(\cdot))$ for any function f that is supposed to take as input a coverage vector, e.g., $u_i^{\mathsf{D}}(X,t) = u_i^{\mathsf{D}}(\mathsf{cov}(X),t)$ and $\mathsf{BR}(X) = \mathsf{BR}(\mathsf{cov}(X))$. As defined in (3.3), the relation between the defenders' individual contributions and the overall coverage resemble an additive relation, so we define the operation $\oplus$ for two numbers $x, y \in [0,1]$, such that

$$x \oplus y = 1 - (1-x)(1-y).$$

The operation is subadditive, commutative, and associative, i.e., we have $x \oplus y \leq x + y$, $x \oplus y = y \oplus x$, and $(x \oplus y) \oplus z = x \oplus (y \oplus z)$. In addition, the result of the operation increases with every operant. We can then write

$$\mathsf{cov}_t(X) = c_t^1 \oplus c_t^2 \oplus \cdots \oplus c_t^n = \bigoplus_{i \in [n]} c_t^i.$$

We also define the operation $\ominus$ for two numbers $x, y \in [0,1]$, $x \geq y$, such that

$$x \ominus y = \begin{cases} 0, & \text{if } y = 1 \\ (x-y)/(1-y), & \text{otherwise} \end{cases}$$

We have $z = x \ominus y \Leftrightarrow x = y \oplus z$, and $x \ominus y \ominus z = x \ominus (y \oplus z)$.

3.2.1 Nash Stackelberg Equilibrium

In a multi-defender SSG, the attacker best-responds to the defenders' joint strategy, so there is a leader-follower structure in the game. On top of that the defenders also interact with each other, bearing in mind the goal of maximizing their own utilities. To capture this additional layer of interaction between the defenders, we incorporate the idea of Nash equilibrium when defining our equilibrium concept for multi-defender SSGs. We call the equilibrium concept *Nash Stackelberg equilibrium* (NSE). As defined below, in an NSE, no player has any incentive to change their strategy, which resembles the situation in a Nash equilibrium; however, when a defender evaluates the benefit of making a strategy deviation, it is assumed that every other defender will stick to their current strategy but the attacker will respond optimally to the strategy deviation, which is similar to the situation in a Stackelberg equilibrium.

Definition 3.1 (Nash-Stackelberg Equilibrium)**.** A defender strategy profile $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$ and an attacker best response $t \in \mathsf{BR}(X)$ form an NSE if for every defender $i \in [n]$ and every strategy deviation $\mathbf{c}' \in \mathcal{C}_{k_i}$, it holds for the resulting strategy profile $X' = (\mathbf{c}^{-i}, \mathbf{c}')$ that

$$\min_{\ell \in \mathsf{BR}(X')} u_i^{\mathsf{D}}(X', \ell) \leq u_i^{\mathsf{D}}(X, t). \tag{3.4}$$

We make two remarks about the above definition. First, since there might be multiple attacker best responses to X in an NSE, it appears that there is a lack of reasoning about how the attacker, who is the follower in the game, ends up with this tie-breaking behavior. Recall that in the standard single-defender model, the optimistic assumption about the tie-breaking behavior is justified by the fact that it can be induced via a small perturbation in the defender's strategy. In an NSE, the same argument applies: the specific attacker best response in an NSE can be induced if some defender reduces her protection to that target by an infinitesimal amount.

Second, in (3.4), however, we assume that the defenders adopt the *pessimistic* tie-breaking assumption when evaluating the benefit of a strategy deviation. It may then appear that we are deviating from the optimistic tie-breaking assumption for SSE, but this is not the case: the optimistic

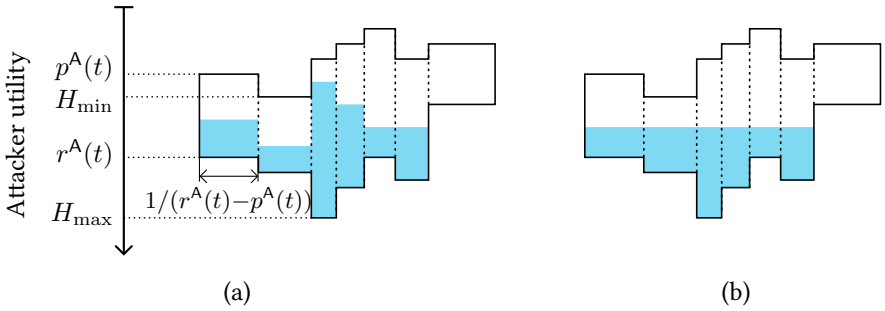


Figure 3.1: (a) The tank-filling visualization. (b) A level coverage.

assumption is about the attacker's response in an SSE, instead of his response resulting from a hypothetical deviation. In fact, it would be problematic if we adopted the optimistic tie-breaking assumption here. Consider the case when some defender makes a "pseudo" deviation that does not change the overall coverage vector (i.e., $\mathsf{cov}(X') = \mathsf{cov}(X)$ after the deviation). Essentially, nothing has changed from the attacker's perspective, but if the defender uses the optimistic tie-breaking assumption, she may assume that the attacker will change his response in her favor. This becomes especially an issue when there are multiple leaders in the game, as it may lead to an equilibrium in which the leaders hold different beliefs about which best response the follower will eventually choose, hence conflicting with the assumption that the follower will actually play some action. The pessimistic tie-breaking assumption we use avoids this issue as it requires a defender to explicitly induce a desired response of the attacker when making a strategy deviation. Proposition 3.2 also shows that Definition 3.1 is consistent with the definition of SSE, in the sense that in the degenerate case with only one defender, an NSE is equivalent to an SSE under a mild assumption.[3] The proof of the proposition can be found in the appendix of this chapter (Section 3.B.1).

Proposition 3.2. *Suppose that $X = (\mathbf{c})$ is a strategy profile (which contains only one defender's strategy) and $0 < c_\ell < 1$ for all $\ell \in \mathsf{BR}(X)$. Then (X, t) is an NSE if and only if $(\mathbf{c}, t)$ is an SSE.*

3.3 Characterizing an NSE

To better illustrate our ideas, we introduce a tank-filling model to visualize the game.

3.3.1 Tank-filling Visualization

As in Figure 3.1, we think of the coverage as water distributed in a set of connected tanks. Each tank corresponds to a target t, with its bottom area set to $\frac{1}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}$. Thus, the height of the water surface, as indicated on the vertical axis, is exactly the attacker's utility of attacking target t when the coverage of t is equal to the volume of water in the tank. Given the coverage of all the targets, the attacker's best responses then correspond to tanks with the lowest surface (e.g., the second tank in Figure 3.1 (a)). Therefore, in a single-defender SSG, if the lowest surface is strictly lower

[3] We need the mild assumption to cope with an inducibility issue that is ignored by the standard definition of SSE (e.g., when a best response of the attacker has coverage 0, we cannot actually induce it by using an infinitesimal reduction in its coverage). Unlike in an NSE, such inducibility issue rarely happens in an SSE and does not lead to inconsistent beliefs about the follower's tie-breaking behavior given that there is only one leader. We refer the interested reader to in-depth discussions about the inducibility issue of SSE [von Stengel and Zamir, 2004, Guo et al., 2019].

than some other surfaces, one can shift some water from the higher surfaces to the lowest one, to improve the coverage of the target that will be selected by the attacker and, hence, also the defender's utility. Moreover, if we connect the tanks and let gravity balance the water surfaces, the surfaces will end up being at the same height (see Figure 3.1 (b)); in which case, the height of the lowest surface is maximized, and the corresponding coverage is exactly the defender's strategy in an SSE as in a single-defender SSG. We call a coverage vector that corresponds to such balanced water distribution a *level coverage* and present Definition 3.3. We will hereafter let

$$H_{\min} = \max_{t \in T} p^{\mathsf{A}}(t) \quad \text{and} \quad H_{\max} = \max_{t \in T} r^{\mathsf{A}}(t),$$

so the interval $[H_{\min}, H_{\max}]$ contains the heights of all possible level coverages, as depicted in Figure 3.1 (a).

Definition 3.3 (Level coverage)**.** A coverage vector $\mathbf{c}$ is said to be *level* if there exists a value $h \in [H_{\min}, H_{\max}]$, such that

$$c_t = \overline{\mathsf{cov}}_t(h) := \max\left\{0, \frac{r^{\mathsf{A}}(t) - h}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)}\right\} \tag{3.5}$$

for all $t \in T$. A strategy profile X is level if $\mathsf{cov}(X)$ is level.

We write $\overline{\mathsf{cov}}(h) = (\overline{\mathsf{cov}}_1(h), \ldots, \overline{\mathsf{cov}}_m(h))$ and call the value h that characterizes a level coverage $\mathbf{c}$ in the above definition the *height* of the coverage (though it increases downwards in Figure 3.1). More generally, the height of a coverage vector $\mathbf{c}$ (not necessarily a level one) is defined by the following function:

$$\mathsf{height}(\mathbf{c}) = \max_{t \in T} u^{\mathsf{A}}(\mathbf{c}, t). \tag{3.6}$$

It is exactly the inverse function of $\overline{\mathsf{cov}}(h)$ for level coverage vectors according to Lemma 3.4 below. An alternative characterization of a level coverage, which is based on the attacker's best response, is presented in Lemma 3.5.

Lemma 3.4. *For any level coverage vector* $\mathbf{c} = \overline{\mathsf{cov}}(h)$*, it holds that* $\mathsf{height}(\mathbf{c}) = h$.

Proof. By definition, we have $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\}$. Substituting this into the utility function defined in (3.2) gives

$$u^{\mathsf{A}}(\mathbf{c}, t) = \min\{r^{\mathsf{A}}(t), h\}.$$

Since $h \leq H_{\max} = \max_{t \in T} r^{\mathsf{A}}(t)$, we have $\min\{r^{\mathsf{A}}(t), h\} = h$ for at least one t, while for all the others $\min\{r^{\mathsf{A}}(t), h\} \leq h$. Thus,

$$\mathsf{height}(\mathbf{c}) = \max_{t \in T} u^{\mathsf{A}}(\mathbf{c}, t) = \max_{t \in T} \min\{r^{\mathsf{A}}(t), h\} = h. \qquad \square$$

Lemma 3.5. *A coverage vector* $\mathbf{c}$ *is level if and only if* $c_t = 0$ *for all* $t \notin \mathsf{BR}(\mathbf{c})$.

Proof. First, suppose $\mathbf{c}$ is level. By definition, we can assume that $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\}$ for some $h \in [H_{\min}, H_{\max}]$. According to Lemma 3.4, if $t \notin \mathsf{BR}(\mathbf{c})$, we have

$$u^{\mathsf{A}}(\mathbf{c}, t) < \max_{t' \in T} u^{\mathsf{A}}(\mathbf{c}, t') = h. \tag{3.7}$$

Indeed, if $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\} > 0$, we would have $c_t = \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}$ and hence, $u^{\mathsf{A}}(\mathbf{c}, t) = h$, which contradicts (3.7). Thus, $c_t = 0$.

Conversely, suppose $c_t = 0$ for all $t \notin \mathsf{BR}(\mathbf{c})$. Let $h = \max_{t\in T} u^{\mathsf{A}}(\mathbf{c}, t)$. Trivially, $h \in [H_{\min}, H_{\max}]$ since $u^{\mathsf{A}}(\mathbf{c}, t) \in [H_{\min}, H_{\max}]$ for all $t \in T$. We show that $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\}$ for all t, which by definition will imply that $\mathbf{c}$ is level.

- For any $t \notin \mathsf{BR}(\mathbf{c})$, we have $c_t = 0$ and, since t is not a best response, we have

$$u^{\mathsf{A}}(\mathbf{c}, t) < \max_{t'} u^{\mathsf{A}}(\mathbf{c}, t') = h.$$

  Expanding $u^{\mathsf{A}}(\mathbf{c}, t)$ gives $\frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)} < c_t = 0$, so indeed, $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\}$.

- For any $t \in \mathsf{BR}(\mathbf{c})$, we have

$$u^{\mathsf{A}}(\mathbf{c}, t) = \max_{t'} u^{\mathsf{A}}(\mathbf{c}, t') = h,$$

  so $\frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)} = c_t \geq 0$ and $c_t = \max\left\{0, \frac{r^{\mathsf{A}}(t)-h}{r^{\mathsf{A}}(t)-p^{\mathsf{A}}(t)}\right\}$ holds, too.

This completes the proof. □

3.3.2 Necessary Conditions for an NSE

In a single-defender game, an SSE corresponds to a level coverage in the tank-filling model. By the same intuition, in the multi-defender case, we also expect that the overall coverage to be level in an equilibrium: if a target is overly protected, the defenders would have an incentive to balance the coverage, shifting some coverage of this target to the other targets. Formally, we present the following necessary conditions for an NSE.

Proposition 3.6. *Suppose that (X, t) forms an NSE, $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$, and $\mathsf{height}(X) > H_{\min}$. Then the following conditions hold:*

- **C1** *(levelness). X is level;*
- **C2** *(resource exhaustion). $\sum_{\ell\in T} c^i_\ell = k_i$ for all $i \in [n]$;*
- **C3** *(allocation prioritization). $c^i_\ell = 0$ if $U^{\mathsf{D}}_i(X, \ell) > U^{\mathsf{D}}_i(X, t)$.*

Intuitively, while **C1** balances the resource distribution as in an SSE, **C2** requires that every defender has used up their resource budget, so that no defender can leverage extra resource to further improve their utility, whereas **C3** prevents a defender from inducing a better attacker response ℓ by reducing the coverage of that target.

Proof of Proposition 3.6. We show that **C1**–**C3** must hold.

(i) Levelness. We first show that X must be level. Suppose for the sake of contradiction that X is not level. By Lemma 3.5, this means that $\mathsf{cov}_{t^*}(X) > 0$ for some $t^* \notin \mathsf{BR}(X)$. Hence, $c^i_{t^*} > 0$ for some $i \in [n]$ as $\mathsf{cov}_{t^*}(X) = \bigoplus_{i\in[n]} c^i_{t^*}$ by definition. The assumption that $\mathsf{height}(X) > H_{\min}$ implies that for every $\ell \in \mathsf{BR}(X)$ we have

$$u^{\mathsf{A}}(X, \ell) = \max_{\ell' \in T} u^{\mathsf{A}}(X, \ell') = \mathsf{height}(X) > H_{\min}.$$

Expanding $u^{\mathsf{A}}(X, \ell)$ gives

$$\mathsf{cov}_\ell(X) < \frac{r^{\mathsf{A}}(\ell) - H_{\min}}{r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)} \leq \frac{r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)}{r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)} = 1.$$

Hence, $c^\ell_i < 1$ for all $\ell \in \mathsf{BR}(X)$ as otherwise we would have $\mathsf{cov}_\ell(X) = \bigoplus_{i\in[n]} c^i_{t^*} = 1$.

Consider a strategy $\mathbf{c}'$ such that:

- $c'_{t^*} = c^i_{t^*} - \delta_{t^*}$,
- $c'_\ell = c^i_\ell + \delta_\ell$ for all $\ell \in \mathsf{BR}(X)$, and
- $c'_\ell = c^i_\ell$ for all other ℓ,

where $\delta_{t^*}, \delta_\ell > 0$. Namely, we shift some coverage from t^* to every $\ell \in \mathsf{BR}(X)$. Since $c^i_{t^*} > 0$ and $c^\ell_i < 1$ for all $\ell \in \mathsf{BR}(X)$, when δ_{t^*} and δ_ℓ are sufficiently small, we can ensure that $c'_\ell \in [0, 1]$ for all $\ell \in T$. In addition, when the δ_ℓ's are sufficiently smaller than δ_{t^*}, we can ensure $\sum_{\ell\in T} c'_\ell \leq \sum_{\ell\in T} c^i_\ell \leq k_i$; hence, $\mathbf{c}' \in \mathcal{C}_{k_i}$, so it is a feasible strategy of defender i.

Now suppose defender i deviates to playing $\mathbf{c}'$. Let $X' = (\mathbf{c}^{-i}, \mathbf{c}^i)$ be the resulting strategy profile. Before this deviation, we have $u^{\mathsf{A}}(X, \ell) \geq u^{\mathsf{A}}(X, t)$ for all $\ell \neq t$. Since the players' utilities for X' change continuously with δ_ℓ, when every δ_ℓ is sufficiently small and δ_t is sufficiently smaller than the other δ_ℓ's, we can ensure that $u^{\mathsf{A}}(X', \ell) > u^{\mathsf{A}}(X', t)$, so that t becomes the only best response of the attacker. As a result,

$$\min_{\ell\in\mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) = u^{\mathsf{D}}_i(X', t) > u^{\mathsf{D}}_i(X, t), \tag{3.8}$$

where $u^{\mathsf{D}}_i(X', t) > u^{\mathsf{D}}_i(X, t)$ is due to the fact that $c'_t = c^i_t + \delta_t > c^i_t$ and hence, $\mathsf{cov}_t(X') > \mathsf{cov}_t(X)$ as X' and X differ only in the strategy of defender i. This contradicts the assumption that (X, t) is an NSE. Therefore, X must be level.

(ii) Resource exhaustion. To see that $\sum_{\ell\in T} c^i_\ell = k_i$ for all $i \in [n]$, suppose that this is not the case, i.e., $\sum_{\ell\in T} c^i_\ell < k_i$ for some i. This means that defender i can distribute more resources to increase the coverage. We can construct a strategy $\mathbf{c}'$ such that $c'_\ell = c^i_\ell + \delta_\ell$ for all $\ell \in \mathsf{BR}(X)$, where $\delta_\ell > 0$. When all δ_ℓ's are sufficiently small, $\sum_{\ell\in T} c'_\ell$ can be made arbitrarily close to $\sum_{\ell\in T} c^i_\ell$, so we can ensure that $\sum_{\ell\in T} c'_\ell < k_i$ and $\mathbf{c}'$ is a feasible strategy for defender i. Same as above, when δ_t is sufficiently smaller than the other δ_ℓ's we can make t the only best response of the attacker while its coverage increases strictly. Hence, (3.8) holds, which contradicts the assumption that (X, t) is an NSE.

(iii) Allocation prioritization. Suppose that $c_\ell^i > 0$ and $u_i^{\mathsf{D}}(X, \ell) > u_i^{\mathsf{D}}(X, t)$ for some ℓ. Thus, $\mathsf{cov}_\ell(X) > 0$. Since we have shown that X is level, by Lemma 3.5, $\ell \in \mathsf{BR}(X)$, so if defender i reduces c_ℓ^i by a small amount δ, target ℓ will become the only best response in $\mathsf{BR}(X')$, where X' is the resulting strategy profile. In addition, by picking a δ sufficiently small, we can make $u_i^{\mathsf{D}}(X', \ell)$ arbitrarily close to $u_i^{\mathsf{D}}(X, \ell)$, so we can ensure that $u_i^{\mathsf{D}}(X', \ell) > u_i^{\mathsf{D}}(X, t)$. It follows that

$$\min_{\ell' \in \mathsf{BR}(X')} u_i^{\mathsf{D}}(X', \ell') = u_i^{\mathsf{D}}(X', \ell) > u_i^{\mathsf{D}}(X, t),$$

which contradicts the assumption that (X, t) is an NSE. □

As a useful corollary of the above lemma, if the defenders' resources are clearly insufficient for the height of their joint strategy to reach $H_{\min}$, then **C1**, **C2**, and **C3** must hold for an NSE.

Corollary 3.7. *Suppose (X, t) forms an NSE. If $\sum_{\ell \in T} \overline{\mathsf{cov}}_\ell(H_{\min}) > \sum_{i \in [n]} k_i$, then* **C1**, **C2**, *and* **C3** *hold.*

Proof. Let $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$. By Proposition 3.6, it suffices to show that $\mathsf{height}(X) > H_{\min}$. Suppose that $\mathsf{height}(X) \leq H_{\min}$. Then the following inequality holds for all $\ell \in T$:

$$u^{\mathsf{A}}(X, \ell) \leq \max_{\ell' \in T} u^{\mathsf{A}}(X, \ell') \leq H_{\min}.$$

Expanding $u^{\mathsf{A}}(X, \ell)$ gives $\mathsf{cov}_\ell(X) \geq \frac{r^{\mathsf{A}}(\ell) - H_{\min}}{r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)}$. Hence,

$$\overline{\mathsf{cov}}_\ell(H_{\min}) = \max\left\{0, \frac{r^{\mathsf{A}}(\ell) - H_{\min}}{r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)}\right\} \leq \mathsf{cov}_\ell(X).$$

Summing up the inequalities over all $\ell \in T$ and using the assumption that $\sum_{\ell \in T} \overline{\mathsf{cov}}_\ell(H_{\min}) > \sum_{i \in [n]} k_i$ give

$$\sum_{\ell \in T} \mathsf{cov}_\ell(X) \geq \sum_{\ell \in T} \overline{\mathsf{cov}}_\ell(H_{\min}) > \sum_{i \in [n]} k_i.$$

However, this contradicts the following fact:

$$\sum_{\ell \in T} \mathsf{cov}_\ell(X) = \sum_{\ell \in T} \left(c_t^1 \oplus c_t^2 \oplus \cdots \oplus c_t^n\right) \leq \sum_{t \in T} \sum_{i \in [n]} c_t^i = \sum_{i \in [n]} \sum_{t \in T} c_t^i = \sum_{i \in [n]} k_i.$$ □

3.3.3 Sufficient Conditions for an NSE

From **C1**, **C2**, and **C3** we can also derive a set of sufficient conditions for a strategy profile to form an NSE. We demonstrate this in Proposition 3.9, where we modify the resource exhaustion condition **C2** to incorporate the case when $\mathsf{height}(X) = H_{\min}$. The new condition **C2*** describes a type of *de facto* resource exhaustion, in which even though some defender may have not used up their resources, they cannot hope to use the remaining resources to further improve their utility.

Lemma 3.8. *Suppose that* $\mathbf{c}$ *and* $\mathbf{c}'$ *are two coverage vectors and* $\mathbf{c}$ *is a level coverage. If $c_t > c'_t$ for a target $t \in T$, then $c_\ell > c'_\ell$ for every $\ell \in \mathsf{BR}(\mathbf{c}')$.*

Proof. Since $c_t > c'_t$, by monotonicity we have $u^{\mathsf{A}}(\mathbf{c}, t) < u^{\mathsf{A}}(\mathbf{c}', t)$. In addition, now that $c_t > c'_t \geq 0$ and $\mathbf{c}$ is level, we have $t \in \mathsf{BR}(\mathbf{c})$ by Lemma 3.5, which means that $u^{\mathsf{A}}(\mathbf{c}, t) = \max_{j \in T} u^{\mathsf{A}}(\mathbf{c}, j)$. Hence,

$$u^{\mathsf{A}}(\mathbf{c}', t) > \max_{j \in T} u^{\mathsf{A}}(\mathbf{c}, j). \tag{3.9}$$

Suppose for the sake of contradiction that $\ell \in \mathsf{BR}(\mathbf{c}')$ but $c_\ell \leq c'_\ell$. Thus,

$$u^{\mathsf{A}}(\mathbf{c}', \ell) = \max_{j \in T} u^{\mathsf{A}}(\mathbf{c}', j) \geq u^{\mathsf{A}}(\mathbf{c}', t), \tag{3.10}$$

and by monotonicity,

$$u^{\mathsf{A}}(\mathbf{c}, \ell) \geq u^{\mathsf{A}}(\mathbf{c}', \ell). \tag{3.11}$$

Combining (3.9)–(3.11) gives the following contradiction:

$$u^{\mathsf{A}}(\mathbf{c}, \ell) > \max_{j \in T} u^{\mathsf{A}}(\mathbf{c}, j) \geq u^{\mathsf{A}}(\mathbf{c}, \ell).$$

Therefore, the assumption does not hold; we have $c_\ell > c'_\ell$ for every $\ell \in \mathsf{BR}(\mathbf{c}')$. □

Proposition 3.9. *A defender strategy profile $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$ and an attacker best response $t \in \mathsf{BR}(X)$ form an NSE if* **C1**, **C3**, *and the following condition hold.*

- **C2*** *(de facto resource exhaustion). For each $i \in [n]$, if $\sum_{\ell \in T} c^i_\ell < k_i$ then*

$$\min_{\ell \in \mathsf{BR}(\mathbf{z})} u^{\mathsf{D}}_i(\mathbf{z}, \ell) \leq u^{\mathsf{D}}_i(X, t)$$

for every coverage vector $\mathbf{z}$ that weakly dominates $\mathsf{cov}(X)$, i.e., $z_\ell \geq \mathsf{cov}_\ell(X)$ for all $\ell \in T$.

Proof. Suppose defender i deviates to playing an arbitrary strategy $\mathbf{c}' \in \mathcal{C}_{k_i}$, and let $X' = (\mathbf{c}^{-i}, \mathbf{c}')$ be the resulting strategy profile. We will argue that $\min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) \leq u^{\mathsf{D}}_i(X, t)$, so (X, t) forms an NSE by definition.

First, if it is the case that $\mathsf{cov}_\ell(X') < \mathsf{cov}_\ell(X)$ for some $\ell \in T$. By Lemma 3.8, we have $\mathsf{cov}_\ell(X') < \mathsf{cov}_\ell(X)$ for all $\ell \in \mathsf{BR}(X')$. It follows by monotonicity that

$$u^{\mathsf{D}}_i(X', \ell) < u^{\mathsf{D}}_i(X, \ell).$$

By **C3**, if $c^i_\ell > 0$ then we have $u^{\mathsf{D}}_i(X, \ell) \leq u^{\mathsf{D}}_i(X, t)$. Indeed, now that $\mathsf{cov}_\ell(X') < \mathsf{cov}_\ell(X)$, it must be that $c^i_\ell > 0$ as otherwise defender i cannot reduce the coverage of ℓ by deviating to X'. Hence, it holds for all $\ell \in \mathsf{BR}(X')$ that

$$u^{\mathsf{D}}_i(X', \ell) < u^{\mathsf{D}}_i(X, \ell) \leq u^{\mathsf{D}}_i(X, t),$$

which implies that $\min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) \leq u^{\mathsf{D}}_i(X, t)$, as desired.

It remains to consider the case when $\mathsf{cov}_\ell(X') \geq \mathsf{cov}_\ell(X)$ for all $\ell \in T$. We can further assume that $\mathsf{cov}_\ell(X') > \mathsf{cov}_\ell(X)$ for some ℓ: otherwise, $\mathsf{cov}(X') = \mathsf{cov}(X)$, and $\min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) \leq u^{\mathsf{D}}_i(X, t)$ holds trivially. Therefore, we have $c'_\ell \geq c^i_\ell$ for all ℓ and it is strict for at least one ℓ. Given that $\mathbf{c}' \in \mathcal{C}_{k_i}$, this implies that

$$\sum_{\ell \in T} c^i_\ell < \sum_{\ell \in T} c'_\ell \leq k_i.$$

It then follows immediately by **C2*** that $\min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) \leq u^{\mathsf{D}}_i(X, t)$ since $\mathsf{cov}(X')$ weakly dominates $\mathsf{cov}(X)$. □

Note that the sufficient conditions in the above proposition suggest that only the defenders' utilities in the given strategy profile X are relevant in characterizing an NSE. This contrasts Definition 3.1 where we also take into account the utilities in X' (as in (3.4)). This motivates us to define a special class of multi-defender games with simplified payoff structure in the next section, as a starting point for proving the existence of an NSE. Before that, we present an example of an NSE.

Example 3.10. There are two defenders α and β and a set $T = \{t_1, t_2, t_3\}$ of three targets. Each defender has one resource, i.e., $k_\alpha = k_\beta = 1$. The payoffs of the players are given in the tables below.

	t_1	t_2	t_3
r_α^{D}	3	7	9
p_α^{D}	2	4	8

defender α

	t_1	t_2	t_3
r_β^{D}	7	5	5
p_β^{D}	0	3	2

defender β

	t_1	t_2	t_3
r^{A}	6	5	5
p^{A}	-4	0	-5

attacker

Suppose defender α plays a strategy $\mathbf{c}^\alpha = (0.6, 0.2, 0.2)$ and defender β plays $\mathbf{c}^\beta = (0, 0.5, 0.5)$. By (3.3), the strategy profile $X = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$ results in $\mathsf{cov}(X) = (0.6, 0.6, 0.6)$ and hence,

$$u^{\mathsf{A}}(X, t_1) = 0,\ u^{\mathsf{A}}(X, t_2) = 2,\ \text{and } u^{\mathsf{A}}(X, t_3) = -1,$$

which implies $\mathsf{BR}(X) = \{t_2\}$ and, by Lemma 3.5, X is not level as $\mathsf{cov}_{t_1}(X) > 0$ but $t_1 \notin \mathsf{BR}(X)$. Since now we have $\mathsf{height}(X) = 2 > 0 = H_{\min}$, by Proposition 3.6, X cannot form an NSE with any $t \in T$.

Indeed, when X is played, the attacker will only attack t_2 (the only best response now), defender α will obtain $u_\alpha^{\mathsf{D}}(X, t_2) = 5.8$. If α deviates to playing, e.g., $\mathbf{c}' = (0.5, 0.3, 0.2)$, the resulting strategy profile X' gives $\mathsf{cov}(X') = (0.5, 0.65, 0.6)$ and $u^{\mathsf{A}}(X') = (1, 1.75, -1)$. Thus, t_2 is still the unique best response of the attacker, which, if attacked, gives defender α utility $u_\alpha^{\mathsf{D}}(X', t_2) = 5.95 > u_\alpha^{\mathsf{D}}(X, t_2)$.

On the other hand, if defender α plays $\mathbf{c}^\alpha = (0.5, 0.5, 0)$ and defender β plays $\mathbf{c}^\beta = (0, 0.6, 0.4)$, the strategy profile $Y = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$ results in $\mathsf{cov}(Y) = (0.5, 0.8, 0.4)$ and $u^{\mathsf{A}}(Y) = (1, 1, 1)$. Hence, $\mathsf{BR}(Y) = \{t_1, t_2, t_3\}$ and Y is level by Lemma 3.5. It can be verified that (Y, t_2) satisfies **C1**, **C2***, and **C3**, so it forms an NSE. □

3.4 Existence of an NSE

We take a two-step approach to prove the existence of an NSE, first proving the existence in a special class of games in which the *defenders* have *non-overlapping payoffs* defined below, and then extending the result to the general setting.

Definition 3.11 (Non-overlapping payoffs)**.** A defender i has *non-overlapping payoffs* if for any $t, \ell \in T$, one of the following holds:

(i) $r_i^{\mathsf{D}}(t) < p_i^{\mathsf{D}}(\ell)$, in which case we write $t \prec_i \ell$.

(ii) $r_i^{\mathsf{D}}(\ell) < p_i^{\mathsf{D}}(t)$, in which case we write $\ell \prec_i t$.

A game is said to be *payoff non-overlapping* if every defender has non-overlapping payoffs.

We also write $t \succeq_i \ell$ if $t \prec_i \ell$ does not hold; and $t \preceq_i \ell$ if $t \succ_i \ell$ does not hold. As we present in Proposition 3.12, when a defender has non-overlapping payoffs, her preference concerning the attacker's response is exactly the preference order $\prec_i$, which is irrespective of the coverage of the targets. For instance, defender α in Example 3.10 has non-overlapping payoffs that are captured by the preference order $t_1 \prec_\alpha t_2 \prec_\alpha t_3$, so this defender always gets a lower utility when t_1 is attacked than when t_2 is attacked (she even prefers a successful attack on target t_2 to a successful defense on t_1). If a game is payoff non-overlapping, the condition **C3** defined in Proposition 3.6 has the following simplified form.

- **C3**. $c_\ell^i = 0$ *if* $t \prec_i \ell$.

Proposition 3.12. *Suppose defender i has non-overlapping payoffs captured by the preference order $\prec_i$ over T and $t \prec_i \ell$. Then $u_i^{\mathsf{D}}(\mathbf{c}', \ell) < u_i^{\mathsf{D}}(\mathbf{c}, t)$ for any $\mathbf{c}$ and $\mathbf{c}'$.*

Proof. By assumption, every defender strictly prefers an unsuccessful attack to a successful one on the same target, i.e., $r_i^{\mathsf{D}}(j) > p_i^{\mathsf{D}}(j)$ for all $j \in T$, so it follows by the utility definition (see (3.1)) that

$$\begin{aligned} u_i^{\mathsf{D}}(\mathbf{c}, t) &= c_t \cdot r_i^{\mathsf{D}}(t) + (1 - c_t) \cdot p_i^{\mathsf{D}}(t) \\ &\leq c_t \cdot r_i^{\mathsf{D}}(t) + (1 - c_t) \cdot r_i^{\mathsf{D}}(t) = r_i^{\mathsf{D}}(t). \end{aligned}$$

Similarly,

$$u_i^{\mathsf{D}}(\mathbf{c}', \ell) \geq p_i^{\mathsf{D}}(\ell).$$

Now that $t \prec_i \ell$, we have $r_i^{\mathsf{D}}(t) < p_i^{\mathsf{D}}(\ell)$ by Definition 3.11, which along with the above two inequalities gives $u_i^{\mathsf{D}}(\mathbf{c}, t) < u_i^{\mathsf{D}}(\mathbf{c}', \ell)$. □

3.4.1 Existence of an NSE with Non-overlapping Payoffs

We prove the existence of an NSE by construction. We first describe our approach informally by way of the tank-filling analogy. Motivated by Proposition 3.9, we will aim at constructing a level strategy profile and we start by guessing the height of this profile.

1. We randomly pick a value $h \in [H_{\min}, H_{\max}]$ as the height of the level strategy profile we will construct next.

2. Starting from $i = 1$, we invite each defender i to distribute water (coverage) to the tanks. Each defender i has k_i units of water to distribute. Now that the defenders have non-overlapping payoffs, suppose that the preference of defender i is $t_{\ell_1} \prec_i t_{\ell_2} \prec_i \cdots \prec_i t_{\ell_m}$. We let defender i start from target t_{ℓ_1} and gradually add water into the tank corresponding to this target, until the water surface in the tank reaches the height h we selected in Step 1. Then we let defender i move on to the next target in her preference order, add water into the corresponding tank, and repeat, until there is no more target left or defender i has distributed all k_i units of water.

   Every time defender i distributes a positive amount of water into a tank, we also update t^*, letting it point to the target corresponding to this tank.

3. We move on to the next defender, let $i \leftarrow i + 1$, and repeat Steps 2 and 3 until $i > n$, when we finish the process with all the defenders. Note that when a defender adds c units of water to a tank where there are already z units of water contributed by other defenders, the volume of water in the tank increases to $c \oplus z$ as defined in (3.3).

At the end of the process, the water each defender i distributed in the tanks defines a strategy for this defender. The whole process will end in one of the following cases:

- **Deficit:** when every defender has used up their budget k_i but some tank is still underfilled, i.e., the water surface is still below h.
- **Surplus:** when all the tanks are filled up to h but some defender has not yet used up their budget.
- **Clearance:** when all the defenders have used up their budgets and all the tanks are filled up to h.

In the case of a clearance, we obtain a strategy profile, say X, that forms an NSE with t^*. Intuitively, since all the tanks are filled up to h, X is level, so **C1** is satisfied; since all the defenders have used up their budgets, **C2*** is satisfied; finally, the way t^* is updated at the end of Step 2 ensures that $\ell \succeq_i t^*$ if $c^i_\ell > 0$, so **C3** is satisfied. By Proposition 3.9, (X, t^*) is an NSE.

When there is a deficit, we do not immediately obtain an NSE, but we can lower the surface of the level coverage we aim to construct (i.e., increase the value of h) to reduce the deficit. Likewise, we can raise the surface to reduce a surplus. The hope is to find a clearance point. Indeed, if we start from $h = H_{\max}$, there will always a surplus as no water needs to be distributed to reach this height; when we decrease h, the surplus decreases continuously. The tricky issue is that it may be a surplus even when h drops to $H_{\min}$, in which case no $h \in [H_{\min}, H_{\max}]$ corresponds to a clearance point. To deal with this issue, we modify the coverage distributing procedure and this leads to a more sophisticated version of our algorithm, which we formally present in Algorithm 3.1.

Everything else being the same as the above described tank-filling procedure, Algorithm 3.1 distributes the coverage in two phases. It identifies a target τ_i (Line 5) for each defender i, which is a pessimistic belief about the attacker's response when every target is fully protected. In the first phase, each defender only distributes coverage to targets $t \prec_i \tau_i$, while in the second, they continue with the remaining targets. We make several observations about Algorithm 3.1 in Lemmas 3.13–3.16, with which we prove the existence of an NSE in every payoff non-overlapping game in Theorem 3.17.

Lemma 3.13. *When Algorithm 3.1 terminates, the following conditions hold (where* surp*,* $\mathcal{D}_t$ *and* $\mathcal{R}_i$ *are values of the corresponding variables in the algorithm):*

- *If* $\mathsf{surp} \geq 0$*, then* $\mathcal{D}_t = 0$ *for all* $t \in T$*,* $\mathsf{cov}(X) = \overline{\mathsf{cov}}(h)$*, and* $t^* \in \mathsf{BR}(X)$*;*
- *If* $\mathsf{surp} \leq 0$*, then* $\mathcal{R}_i = 0$ *for all* $i \in [n]$*.*

Proof. Suppose for the sake of contradiction that when Algorithm 3.1 terminates, we have $\mathcal{D}_t > 0$ for some t and also $\mathcal{R}_i > 0$ for some i. We check the iteration in which c^i_t is assigned a value at Line 9. If $\mathcal{D}_t \leq \mathcal{R}_i$ in that iteration, then c^i_t will be set to $\mathcal{D}_t$. Consequently, the subsequent Line 11 will set $\mathcal{D}_t \leftarrow \mathcal{D}_t \ominus \mathcal{D}_t = 0$, which contradicts the assumption that $\mathcal{D}_t > 0$ in the end.

Algorithm 3.1: Construct an NSE in a payoff non-overlapping game.

Input : a payoff non-overlapping game with defender preferences $\prec_1, \ldots, \prec_n$, and a height $h \in [H_{\min}, H_{\max}]$.

Output : $X := (\mathbf{c}^1, \ldots, \mathbf{c}^n), t^*$.

/* Initialization */

1 **for** *all* $i \in [n]$ *and* $t \in T$ **do**

2 $\quad c_t^i \leftarrow 0$; // $\mathbf{c}^i = (c^i_{t_1}, \ldots, c^i_{t_m})$

3 $\quad \mathcal{R}_i \leftarrow k_i$; // resource surplus of defender i

4 $\quad \mathcal{D}_t \leftarrow \overline{\mathsf{cov}}_t(h)$; // coverage demanded by t

5 $\quad \tau_i \leftarrow \arg\min_{j \in \mathsf{BR}(\mathbf{1}^m)} u_i^{\mathsf{D}}(\mathbf{1}^m, j)$; // $\mathbf{1}^m$ denotes coverage vector $(1, \ldots, 1)$

/* Phase 1 */

6 **for** *each defender* $i = 1, \ldots, n$ **do**

7 $\quad$ Suppose $t_{\ell_1} \prec_i t_{\ell_2} \prec_i \cdots \prec_i t_{\ell_m}$;

8 $\quad$ **for** *each target* $t = t_{\ell_1}, \ldots, t_{\ell_m}$ *such that* $t \prec_i \tau_i$ **do**

9 $\quad\quad c_t^i \leftarrow \min\{\mathcal{D}_t, \mathcal{R}_i\}$;

10 $\quad\quad \mathcal{R}_i \leftarrow \mathcal{R}_i - c_t^i$; // so that $\mathcal{R}_i = k_i - \sum_{t \in T} c_t^i$ throughout

11 $\quad\quad \mathcal{D}_t \leftarrow \mathcal{D}_t \ominus c_t^i$; // so that $\mathcal{D}_t = \overline{\mathsf{cov}}_t(h) \ominus \mathsf{cov}_t(X)$ throughout

12 $\quad\quad$ **if** $c_t^i > 0$ **then** $t^* \leftarrow t$;

/* Phase 2 */

13 Repeat Lines 6–12, with the inner loop being executed for each $t = \tau_i, \ldots, t_{\ell_m}$ (Line 8);

14 $\mathsf{surp} \leftarrow \sum_{i \in [n]} \mathcal{R}_i - \sum_{t \in T} \mathcal{D}_t$.

Thus, it must be that $\mathcal{D}_t > \mathcal{R}_i$, in which case, however, c_t^i will be set to $\mathcal{R}_i$, which then results in $\mathcal{R}_i \leftarrow \mathcal{R}_i - \mathcal{R}_i = 0$ in Line 10, contradicting the assumption that $\mathcal{R}_i > 0$ in the end.

Thus, $\sum_{t \in T} \mathcal{D}_t$ and $\sum_{i \in [n]} \mathcal{R}_i$ cannot both be positive. Since they are non-negative and $\mathsf{surp} = \sum_{i \in [n]} \mathcal{R}_i - \sum_{t \in T} \mathcal{D}_t$, this implies that $\mathcal{D}_t = 0$ for all t if $\mathsf{surp} \geq 0$, and $\mathcal{R}_i = 0$ for all i if $\mathsf{surp} \leq 0$.

In particular, since we have $\mathcal{D}_t = \overline{\mathsf{cov}}_t(h) \ominus \mathsf{cov}_t(X)$ throughout the algorithm, when $\mathcal{D}_t = 0$, we have $\overline{\mathsf{cov}}_t(h) \ominus \mathsf{cov}_t(X) = 0$ and hence, $\mathsf{cov}_t(X) = \overline{\mathsf{cov}}_t(h)$. Finally, the way t^* is set ensures that $c^i_{t^*} > 0$ and hence, $\mathsf{cov}_{t^*}(X) > 0$. Since $\mathsf{cov}(X) = \overline{\mathsf{cov}}(h)$ is level, by Lemma 3.5, $t^* \in \mathsf{BR}(X)$. □

Lemma 3.14. *In the output of Algorithm 3.1, it holds for all* $i \in [n]$ *that:* $t \preceq_i t^*$ *if* $c_t^i > 0$.

Proof. For an arbitrary defender i, suppose that ℓ^* is the last target t such that c_t^i is set to a positive value at Line 9. Thus, $t \preceq_i \ell^*$ if $c_t^i > 0$, and in the same iteration t^* is updated to ℓ^*.

It must be that $\mathcal{D}_t = 0$ for all $t \prec_i \ell^*$ before this iteration, and it remains to be so till the algorithm terminates. Indeed, if $\mathcal{D}_t > 0$, that means $c_t^i < \mathcal{D}_t$ in the same iteration when defender i distributes coverage to target t at Line 9, so we have $c_t^i = \mathcal{R}_i$ and in the subsequent line $\mathcal{R}_i \leftarrow \mathcal{R}_i - \mathcal{R}_i = 0$. Hence, in the iteration for target ℓ^* (which is preceded by the iteration for ℓ^* as $t \prec_i \ell^*$), we will have $c^i_{\ell^*} \leftarrow \min\{\mathcal{D}_{\ell^*}, 0\} = 0$, which contradicts our assumption that $c^i_{\ell^*}$ is set to a positive value at Line 9.

Now that $\mathcal{D}_t = 0$, it follows that $c_t^j = 0$ for the subsequent defenders $j = i+1, \ldots, n$. Hence, t^* will not be updated to any t such that $t \prec_i \ell^*$, which implies that $\ell^* \preceq_i t^*$. It follows that, if $c_t^i > 0$, we have $t \preceq_i \ell^* \preceq_i t^*$. □

Lemma 3.15. *In the output of Algorithm 3.1, it holds for all $i \in [n]$ that: $\tau_i \preceq_i t^*$ if $\mathsf{cov}_{\tau_i}(X) > 0$ and $\mathcal{R}_i > 0$.*

Proof. By definition (Line 5), we have $u_i^{\mathsf{D}}(\mathbf{1}^m, \tau_i) \leq u_i^{\mathsf{D}}(\mathbf{1}^m, t)$ for all $t \in \mathsf{BR}(\mathbf{1}^m)$; hence, $\tau_i \preceq_i t$. Therefore, no coverage will be distributed to these targets in Phase 1, we have $\mathsf{cov}_t(X) = 0$ for all $t \in \mathsf{BR}(\mathbf{1}^m)$ at the end of Phase 1, in particular, $\mathsf{cov}_{\tau_i}(X) = 0$.

If $\mathsf{cov}_{\tau_i}(X) > 0$ at the end of the algorithm, then some coverage must have been distributed to τ_i in Phase 2 and t^* must have been updated at least once. We argue that t^* must have not been updated to any $t \prec_i \tau_i$ to complete the proof.

Indeed, if t^* were updated to some $t \prec_i \tau_i$, then we have $\mathcal{D}_t > 0$ at the end of Phase 1. This implies that in the iteration when defender i distributes coverage to t, we have $c_t^i < \mathcal{D}_t$ (otherwise, $\mathcal{D}_t$ would have been updated to 0 in the subsequent lines). Hence, $c_t^i = \mathcal{R}_i$, so in the subsequent line we would have $\mathcal{R}_i \leftarrow \mathcal{R}_i - \mathcal{R}_i = 0$, which contradicts the assumption that $\mathcal{R}_i > 0$. □

Lemma 3.16. *Suppose that we fix an arbitrary payoff non-overlapping game in the input of Algorithm 3.1, and let $\mathsf{surp}(h)$ be the output of the algorithm on input h. Then $\mathsf{surp}(h)$ is a continuous function of h and it is strictly increasing in $[H_{\min}, H_{\max}]$.*

Proof. We also view the value of each variable in Algorithm 3.1 at the end of the algorithm as a function of h.

(i) Monotonicity. Let h_1 and h_2 be two arbitrary values such that $H_{\min} \leq h_1 < h_2 \leq H_{\max}$. It is easy to see that the following conditions hold right after the initialization stage:

- $\mathsf{surp}(h_1) < \mathsf{surp}(h_2)$,
- $\mathcal{D}_t(h_1) > \mathcal{D}(h_2)$, or $\mathcal{D}_t(h_1) = \mathcal{D}(h_2) = 0$, and
- $\mathcal{R}_i(h_1) \leq \mathcal{R}_i(h_2)$.

In addition, it can be verified that if these conditions hold after defender ℓ finishes distributing her coverage in Phase 1, they will also hold after the iteration of defender $\ell + 1$, and the same applies to the iterations in Phase 2. Therefore, we have $\mathsf{surp}(h_1) < \mathsf{surp}(h_2)$ when the algorithm terminates, so $\mathsf{surp}(h)$ is strictly increasing.

(ii) Continuity. It is obvious that the values of all the variables after the initialization stage changes continuously with h. In the subsequent lines, all variables except $\mathcal{D}_t$ are continuous functions of other variables (or their values in the previous iterations), so they remain continuous functions of h as long as the variables they depend on are continuous functions of h. In terms of $\mathcal{D}_t$, the $\ominus$ operation in Line 11 is not continuous at the point $c_t^i = 1$. We will argue that this in fact does not affect $\mathcal{D}_t(h)$ being continuous.

Observe that we have

$$c_t^i(h) \leq \mathcal{D}'_t(h) \leq \overline{\mathsf{cov}}_t(h) \leq \overline{\mathsf{cov}}_t(H_{\min}) \leq 1,$$

where $\mathcal{D}'_t(h)$ denotes the value of $\mathcal{D}_t$ before c^i_t is assigned a value at Line 9. In other words, the coverage distributed is at most the coverage demanded, which is at most the level coverage $\overline{\mathsf{cov}}_t(h)$ we aim to construct. This means that $c^i_t(h) = 1$ only when $h = H_{\min}$. To be more specific, for a variable v, we let $v^{(\ell)}(h)$ denote its value at the end of the ℓ-th iteration in Phase 1, i.e., the iteration defender ℓ distributes resources (let $v^{(0)}(h)$ denote the value right after initialization). Consider the following cases.

- $\mathcal{D}^{(\ell)}_t(H_{\min}) > 0$. In this case, it must be that $c^{i\,(\ell)}_t(H_{\min}) < 1$. Since we also have $c^i_t(h) \le \overline{\mathsf{cov}}_t(h) < 1$ for all $h \in (H_{\min}, H_{\max}]$ (see (3.5)), now $c^i_t(h) < 1$ for all $h \in [H_{\min}, H_{\max}]$. By definition of the $\ominus$ operation we have $\mathcal{D}^{(\ell-1)}_t(h) \ominus c^{i\,(\ell)}_t(h) = \frac{\mathcal{D}^{(\ell-1)}_t(h) - c^{i\,(\ell)}_t(h)}{1 - c^{i\,(\ell)}_t(h)}$, so effectively Line 11 in this iteration is equivalent to $\mathcal{D}_t \leftarrow \frac{\mathcal{D}_t - c^i_t}{1 - c^i_t}$, the right side of which is a continuous function of the variables involved.
- $\mathcal{D}^{(\ell)}_t(H_{\min}) = 0$. Since we have shown that $\mathcal{D}^{(\ell)}_t(h)$ is monotonic, we have $\mathcal{D}^{(\ell)}_t(h) \le \mathcal{D}^{(\ell)}_t(H_{\min}) = 0$, which means $\mathcal{D}^{(\ell)}_t(h) = 0$ for all $h \in [H_{\min}, H_{\max}]$. Thus, $\mathcal{D}^{(\ell)}_t(h)$ is, too, a continuous function of h.

The same analysis applies to Phase 2, so $\mathcal{D}_t(h)$ is a continuous function of h, and the same holds with $\mathsf{surp}(h)$. □

Theorem 3.17. *There exists an NSE in every payoff non-overlapping game.*

Proof. The proof relies on Algorithm 3.1. By Lemma 3.16, surp is a continuous function of h, and $\mathsf{surp}(H_{\max}) = \sum_{i\in[n]} k_i \ge 0$ as $\overline{\mathsf{cov}}_t(H_{\max}) = 0$ for all $t \in T$ (see (3.5)). Consider the following possibilities in terms of the value of $\mathsf{surp}(H_{\min})$.

(i) $\mathsf{surp}(H_{\min}) \le 0$. In this case, there must be a zero $h^* \in [H_{\min}, H_{\max}]$ of surp at which $\mathsf{surp}(h^*) = 0$. We let $X = (\mathbf{c}^1, \dots, \mathbf{c}^n)$ and t^* be the output of Algorithm 3.1 on input $h = h^*$.

(ii) $\mathsf{surp}(H_{\min}) > 0$. In this case, we let $X = (\mathbf{c}^1, \dots, \mathbf{c}^n)$ and t^* be the output of Algorithm 3.1 on input $h = h^*$.

We will argue that in both cases $t^* \in \mathsf{BR}(X)$ and (X, t^*) satisfies all three conditions in Proposition 3.9, thus forming an NSE.

In both cases $\mathsf{surp}(h^*) \ge 0$, so by Lemma 3.13, $t^* \in \mathsf{BR}(X)$, X is level; hence, **C1** holds. In addition, by Lemma 3.14, **C3** holds. It remains to show that **C2*** holds.

Indeed, since $\mathsf{surp}(h^*) = 0$ in Case (i), by Lemma 3.13 we also have $\mathcal{R}_i = 0$ and hence, $k_i - \sum_{t\in T} c^i_t = \mathcal{R}_i = 0$ for all $i \in [n]$, so **C2*** holds. In Case (ii), however, this is not the case: it may happen that $\mathcal{R}_i > 0$ for some i and hence $\sum_{t\in T} c^i_t < k_i$. To argue that **C2*** is still satisfied, note that we have $\mathsf{height}(X) = H_{\min}$ in Case (ii), so it must be that $\tau_i \in \mathsf{BR}(X)$ and $\mathsf{cov}_{\tau_i}(X) = 1$. Indeed, $\tau_i \in \mathsf{BR}(\mathbf{1}^m)$ by definition (Line 5), so we have $u^{\mathsf{A}}(\mathbf{1}^m, \tau_i) = \max_{t\in T} u^{\mathsf{A}}(\mathbf{1}^m, t)$, which implies the following:

$$\begin{aligned}\max_{t\in T} u^{\mathsf{A}}(X,t) \ge u^{\mathsf{A}}(X,\tau_i) &\ge u^{\mathsf{A}}(\mathbf{1}^m,\tau_i)\\ &= \max_{t\in T} u^{\mathsf{A}}(\mathbf{1}^m,t)\\ &= \max_{t\in T} p^{\mathsf{A}}(t) = H_{\min} = \mathsf{height}(X) = \max_{t\in T} u^{\mathsf{A}}(X,t),\end{aligned}$$

where $u^{\mathsf{A}}(X, \tau_i) \geq u^{\mathsf{A}}(\mathbf{1}^m, \tau_i)$ is due to monotonicity of the utility function and the fact that $c_t^i \leq 1$; and $\mathsf{height}(X) = \max_{t \in T} u^{\mathsf{A}}(X, t)$ by (3.6). Therefore, the inequality is in fact an equality. We have $u^{\mathsf{A}}(X, \tau_i) = \max_{t \in T} u^{\mathsf{A}}(X, t)$, which implies

$$\tau_i \in \mathsf{BR}(X);$$

in addition, $u^{\mathsf{A}}(X, \tau_i) = u^{\mathsf{A}}(\mathbf{1}^m, \tau_i)$, which implies

$$\mathsf{cov}_{\tau_i}(X) = 1.$$

For any coverage vector $\mathbf{z}$ such that $z_t \geq \mathsf{cov}_t(X)$ for all t, we have $u^{\mathsf{A}}(\mathbf{z}, t) \leq u^{\mathsf{A}}(X, t)$ by monotonicity. Meanwhile, since $\mathsf{cov}_{\tau_i}(X) = 1$, we have $z_{\tau_i} = 1$ as the coverage value must be within $[0, 1]$; hence, $u^{\mathsf{A}}(\mathbf{z}, \tau_i) = u^{\mathsf{A}}(X, \tau_i)$ and $u_i^{\mathsf{D}}(\mathbf{z}, \tau_i) = u_i^{\mathsf{D}}(X, \tau_i)$. The former further implies that, for all $t \in T$,

$$u^{\mathsf{A}}(\mathbf{z}, \tau_i) = u^{\mathsf{A}}(X, \tau_i) \geq u^{\mathsf{A}}(X, t) \geq u^{\mathsf{A}}(\mathbf{z}, t),$$

where the second transition is due to the fact that $\tau_i \in \mathsf{BR}(X)$ as we argued above. Hence, τ_i is also a best response against $\mathbf{z}$, i.e., $\tau_i \in \mathsf{BR}(\mathbf{z})$. We have

$$\min_{t \in \mathsf{BR}(\mathbf{z})} u_i^{\mathsf{D}}(\mathbf{z}, t) \leq u_i^{\mathsf{D}}(\mathbf{z}, \tau_i) = u_i^{\mathsf{D}}(X, \tau_i) \leq u_i^{\mathsf{D}}(X, t^*),$$

where $u_i^{\mathsf{D}}(X, \tau_i) \leq u_i^{\mathsf{D}}(X, t^*)$ holds as $\tau_i \preceq_i t^*$ (if $\mathcal{R}_i > 0$) by Lemma 3.15. Therefore, **C2*** holds. □

3.4.2 Existence of an NSE in Games with General Payoffs

In contrast to games with non-overlapping payoffs, in the general payoff setting, a defender's preference over the attacker's responses depends on the actual coverage vector. Thus, if we take a set of preference orders $\prec_1, \ldots, \prec_n$ and attempt to construct an NSE using our approach for non-overlapping payoffs, the defender preferences in the constructed strategy profile may end up being different from $\prec_1, \ldots, \prec_n$. If that is the case, then $\ell \prec_i t$ does not imply $u_i^{\mathsf{D}}(X, \ell) \leq u_i^{\mathsf{D}}(X, t)$ anymore, so **C3** may break and our approach may fail to generate an NSE. Nevertheless, we will show that we can find a "stationary point" where the preferences used to construct the NSE coincide with the actual preferences of the defenders in the NSE constructed. We call the set of defender preferences $(\prec_1, \ldots, \prec_n)$ a *preference profile* and define the *induced preference profile* of a strategy profile X to be the defenders' preferences at X (Definition 3.18).

Definition 3.18 (Induced preference profile)**.** A preference profile *induced by a coverage* $\mathbf{c}$ is a list of binary relations $(\preceq_1^{\mathbf{c}}, \ldots, \preceq_n^{\mathbf{c}})$ such that $t \preceq_i^{\mathbf{c}} \ell$ if and only if $u_i^{\mathsf{D}}(\mathbf{c}, t) \leq u_i^{\mathsf{D}}(\mathbf{c}, \ell)$. A preference profile induced by a strategy profile X, denoted $(\preceq_1^X, \ldots, \preceq_n^X)$, is the preference profile induced by $\mathsf{cov}(X)$.

Note that in contrast to the previously defined preference order, a defender may be indifferent between different targets in an induced preference order. For any $t, \ell \in T$, we will write $t \prec_i^X \ell$, if $t \preceq_i^X \ell$ but not $\ell \preceq_i^X t$. With the above definition, **C3** can be described equivalently as follows, in a form similar to the one we obtain with non-overlapping payoffs; we will use **C3** in this form in the remainder of this section.

- **C3**. $c_\ell^i = 0$ *if* $t \prec_i^X \ell$.

Existence of a Stationary Point

Imagine that we increase h from $H_{\min}$ to $H_{\max}$. The corresponding level coverage changes from $\overline{\mathsf{cov}}(H_{\min})$ to $\overline{\mathsf{cov}}(H_{\max})$, and in the meantime the preference profile induced by $\overline{\mathsf{cov}}(h)$ changes when some defender's utilities under different attacker responses become equal, i.e., when $u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t) = u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t')$ for some $t \neq t'$. Let

$$\mathcal{H} = \left\{ h \in [H_{\min}, H_{\max}] : u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t) = u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t'), i \in [n], t \neq t' \right\} \tag{3.12}$$

be the values of h at these points. Note that if two utility functions are the same in an interval, there are infinitely many intersection points. In this case, we only include the endpoints of that interval in $\mathcal{H}$. This suffices for our purpose because our goal is to identify distinct induced preference profiles, and the defender's preferences do not change in this interval.

Let $h^{(\ell)}$ be the ℓ-th element in $\mathcal{H} \cup \{H_{\min}, H_{\max}\}$ in ascending order. For each ℓ, all the level coverage vectors $\overline{\mathsf{cov}}(h)$ with $h \in (h^{(\ell)}, h^{(\ell+1)})$ thus induce the same preference profile; denote it by $P^{(\ell)} = (\preceq_1^{(\ell)}, \ldots, \preceq_n^{(\ell)})$. Further, we map each $P^{(\ell)}$ to an NSE $(X^{(\ell)}, t^{(\ell)})$ in the following way: Fix the defender's preferences to $P^{(\ell)}$, find the zero of $\mathsf{surp}(h)$ using Algorithm 3.1, let $H^{(\ell)}$ be the zero (or let it be $H_{\min}$ if there is no zero in $[H_{\min}, H_{\max}]$), and let $X^{(\ell)}$ be the level strategy profile generated by Algorithm 3.1 on input $h = H^{(\ell)}$. Note that Algorithm 3.1 only accepts strict preference orders in its input, so we convert the orders in $P^{(\ell)}$ to strict ones using the lexicographic tie-breaking rule.

Now each interval $(h^{(\ell)}, h^{(\ell+1)})$ is associated with a value $H^{(\ell)}$. Intuitively, if some $H^{(\ell)}$ happens to fall back into the interval $(h^{(\ell)}, h^{(\ell+1)})$, then $H^{(\ell)}$ is a stationary point we desire. Lemma 3.19 establishes the existence of two types of stationary points, with which we demonstrate in Theorem 3.20 the existence of a strategy profile satisfying **C1**, **C2***, and **C3**, and hence the existence of an NSE.

Lemma 3.19. *Let $h^{(1)}, \ldots, h^{(L)}$ and $H^{(1)}, \ldots, H^{(L-1)}$ be two sequences of real numbers, such that $h^{(1)} < h^{(2)} < \cdots < h^{(L)}$ and $h^{(1)} \leq H^{(\ell)} \leq h^{(L)}$ for all ℓ. Then there exists either $H^{(\ell)} \in [h^{(\ell)}, h^{(\ell+1)}]$ or $h^{(\ell)} \in (\!|H^{(\ell-1)}, H^{(\ell)}|\!)$, where $(\!|a, b|\!) = \{x \in \mathbb{R} : a < x < b \text{ or } b < x < a\}$ denotes the open interval between points $a, b \in \mathbb{R}$.*

Proof. Suppose for the sake of contradiction that neither $h^{(\ell)}$ nor $H^{(\ell)}$ exists. Then for all ℓ it holds that (i) $H^{(\ell)} \notin [h^{(\ell)}, h^{(\ell+1)}]$ and (ii) $h^{(\ell)} \notin (\!|H^{(\ell-1)}, H^{(\ell)}|\!)$. In particular, $H^{(1)} \notin [h^{(1)}, h^{(2)}]$. Since $h^{(1)} \leq H^{(\ell)} \leq h^{(L)}$ for all ℓ, this implies $H^{(1)} > h^{(2)}$.

Now, suppose that $H^{(t)} > h^{(t+1)}$ for all $t < \ell$. Consider the case $t = \ell + 1$, if $H^{(\ell+1)} \leq h^{(\ell+2)}$ then either $h^{(\ell+1)} \leq H^{(\ell+1)} \leq h^{(\ell+2)}$, which contradicts (i), or $H^{(\ell+1)} < h^{(\ell+1)}$, which implies that $H^{(\ell+1)} < h^{(\ell+1)} < H^{(\ell)}$ and contradicts (ii). Thus, we must have $H^{(\ell+1)} > h^{(\ell+2)}$, and, by induction, $H^{(\ell)} > h^{(\ell+1)}$ for all $\ell \leq L$. However, this contradicts the assumption that $H^{(L-1)} \leq h^{(L)}$. □

Theorem 3.20. *There exists an NSE in every multi-defender SSG.*

Proof. By Lemma 3.19, there exists either (i) $H^{(\ell)} \in [h^{(\ell)}, h^{(\ell+1)}]$, or (ii) $h^{(\ell)} \in (\!|H^{(\ell-1)}, H^{(\ell)}|\!)$. In Case (i), $H^{(\ell)}$ is a stationary point we desire, so the arguments for proving Theorem 3.17 can be applied to show that $t^{(\ell)} \in \mathsf{BR}(X^{(\ell)})$ and $(X^{(\ell)}, t^{(\ell)})$ satisfies **C1**, **C2***, and **C3**; thus $(X^{(\ell)}, t^{(\ell)})$ is an NSE by Proposition 3.9. In Case (ii), however, we do not obtain a stationary point immediately; neither $(X^{(\ell-1)}, t^{(\ell-1)})$ nor $(X^{(\ell)}, t^{(\ell)})$ satisfies the conditions. We will show that we can 'merge'

these two profiles, $(X^{(\ell-1)}, t^{(\ell-1)})$ and $(X^{(\ell)}t^{(\ell)})$, into a new one (X, t^*), $t^* \in \mathsf{BR}(X)$, that satisfies all the three conditions. To do this, we will use Algorithm 3.2. Without loss of generality, we can assume that $H^{(\ell-1)} < H^{(\ell)}$, so $h^{(\ell)} \in (H^{(\ell-1)}, H^{(\ell)})$; indeed, by switching the preference profiles associated with $\ell - 1$ and ℓ, we can replicate the arguments for the case $H^{(\ell)} < H^{(\ell-1)}$.

Algorithm 3.2 is again a two-phase procedure; each phase calls Algorithm 3.1 once: Phase 1 fixes the preference to $\prec_i^{(\ell-1)}$ and the resource budget to $\gamma \cdot k_i$ for each defender i; Phase 2 changes the preference to $\prec_i^{(\ell)}$ and continues the resource distributing process with the remaining budget $(1-\gamma)\cdot k_i$.

Note that the way the coverage is distributed (Lines 11–15) is slightly different from that in Algorithm 3.1: In Phase 1, the procedures are effectively the same as in Algorithm 3.1 because $c_t^i = 0$ before each defender i distributes resouces; In Phase 2, however, we need to take into account the amount of coverage already distributed by the same defender in Phase 1, which makes the formulation slightly different from Algorithm 3.1. The demand is set to $(\mathcal{D}_t \oplus c_t^i) - c_t^i$ at Line 11; indeed, if defender i increases the coverage by $(\mathcal{D}_t \oplus c_t^i) - c_t^i$ on top of the coverage c_t^i she already distributed, her total contribution will be $\mathcal{D}_t \oplus c_t^i$ and the overall coverage of target t will become

$$\begin{aligned}(\mathcal{D}_t \oplus c_t^i) \oplus \bigoplus_{j \neq i} c_t^j &= \overline{\mathsf{cov}}_t(h^{(\ell)}) \ominus \mathsf{cov}_t(X) \oplus \bigoplus_{j \in [n]} c_t^j \\ &= \overline{\mathsf{cov}}_t(h^{(\ell)}) \ominus \mathsf{cov}_t(X) \oplus \mathsf{cov}_t(X) \\ &= \overline{\mathsf{cov}}_t(h^{(\ell)}),\end{aligned}$$

which is the level coverage we aim to construct.

Fixing all the other input parameters, we view surp^+ as a function of the control parameter γ. We make two observations about surp^+.

First, surp^+ is a continuous function of γ: all the operations at Lines 11 and 14 are continuous with respect to the variables involved. Unlike in Algorithm 3.1, here at Line 14, we do not need to worry about the discontinuity that may arise at $\mathsf{cov}_t(X) = 1$ when we perform the $\ominus$ operation; indeed, we always have $h^{(\ell)} > H_{\min}$ and hence, $\mathsf{cov}_t(X) \leq \overline{\mathsf{cov}}_t(h^{(\ell)}) < 1$.

Second, when $\gamma = 1$, Phase 2 is actually skipped and Algorithm 3.2 is equivalent to Algorithm 3.1 with input $\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}$ and $h = H^{(\ell)}$; namely, we have

$$\mathsf{surp}^+(1) = \mathsf{surp}\left(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}; H^{(\ell)}\right),$$

where we view surp defined in Algorithm 3.1 as a function of the input preference profile and the height. By definition, $H^{(\ell-1)}$ is the zero of $\mathsf{surp}(h)$ when the input preferences are $\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}$. Since $\mathsf{surp}(h)$ is an increasing function when we fix a preference profile (see Lemma 3.16) and now we have $H^{(\ell-1)} < H^{(\ell)}$, it follows that

$$\mathsf{surp}\left(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}; H^{(\ell)}\right) > \mathsf{surp}\left(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}; H^{(\ell-1)}\right) = 0.$$

Thus,

$$\mathsf{surp}^+(1) = \mathsf{surp}\left(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}; H^{(\ell)}\right) > \mathsf{surp}\left(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)}; H^{(\ell-1)}\right) = 0.$$

Similarly, when $\gamma = 0$, Phase 1 is skipped and we have

$$\mathsf{surp}^+(0) = \mathsf{surp}\left(\prec_1^{(\ell)}, \ldots, \prec_n^{(\ell)}; H^{(\ell)}\right) < \mathsf{surp}\left(\prec_1^{(\ell)}, \ldots, \prec_n^{(\ell)}; H^{(\ell)}\right) = 0.$$

Algorithm 3.2: Construct an NSE using two preference profiles.

Input : two preference profiles $(\prec_1^{(\ell-1)}, \ldots, \prec_n^{(\ell-1)})$ and $(\prec_1^{(\ell)}, \ldots, \prec_n^{(\ell)})$; $h^{(\ell)}$; $\gamma \in [0,1]$.

Output : $X := (\mathbf{c}^1, \ldots, \mathbf{c}^n), t^*$.

`/* Initialization */`

1 **for** *all* $i \in [n]$ *and* $t \in T$ **do**

2 $\quad c_t^i \leftarrow 0$;

3 $\quad \mathcal{R}_i^1 \leftarrow \gamma \cdot k_i$; `// resource surplus of Phase 1`

4 $\quad \mathcal{R}_i^2 \leftarrow (1-\gamma) \cdot k_i$; `// resource surplus of Phase 2`

5 $\quad \mathcal{D}_t \leftarrow \overline{\mathsf{cov}}_t(h^{(\ell)})$;

6 $\quad \tau_i \leftarrow \arg\min_{j \in \mathsf{BR}(\mathbf{1}^m)} u_i^{\mathsf{D}}(\mathbf{1}^m, j)$,

`/* Phase 1 */`

7 Let $\lambda \leftarrow 1$ and $\prec_i \leftarrow \prec_i^{(\ell-1)}$ for all $i \in [n]$;

8 **for** *defender* $i = 1, \ldots, n$ **do**

9 $\quad$ Suppose $t_{j_1} \prec_i t_{j_2} \prec_i \cdots \prec_i t_{j_m}$;

10 $\quad$ **for** *each target* $t = t_{j_1}, \ldots, t_{j_m}$ *such that* $t \prec_i \tau_i$ **do**

11 $\quad\quad \delta \leftarrow \min\left\{(\mathcal{D}_t \oplus c_t^i) - c_t^i,\ \mathcal{R}_i^\lambda\right\}$;

12 $\quad\quad c_t^i \leftarrow c_t^i + \delta$;

13 $\quad\quad \mathcal{R}_i^\lambda \leftarrow \mathcal{R}_i^\lambda - \delta$;

14 $\quad\quad \mathcal{D}_t \leftarrow \overline{\mathsf{cov}}_t(h^{(\ell)}) \ominus \mathsf{cov}_t(X)$;

15 $\quad\quad$ **if** $\delta > 0$ **then** $t^* \leftarrow t$;

16 Repeat Lines 8–15, with the inner loop being executed for each $t = \tau_i, \ldots, t_{j_m}$ (Lines 10);

`/* Phase 2 */`

17 Let $\lambda \leftarrow 2$ and $\prec_i \leftarrow \prec_i^{(\ell)}$ for all $i \in [n]$;

18 Repeat Lines 8–15;

19 $\mathsf{surp}^+ \leftarrow \sum_{i \in [n]} \left(\mathcal{R}_i^1 + \mathcal{R}_i^2\right) - \sum_{t \in T} \mathcal{D}_t$.

Therefore, there must be a zero point $\gamma^0 \in (0,1)$ such that $\mathsf{surp}^+(\gamma^0) = 0$. Let (X, t^*) be the output of Algorithm 3.2 on input $\gamma = \gamma^0$, where $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$. Applying the same arguments for proving Lemma 3.13, from $\mathsf{surp}^+(\gamma^0) = 0$, we know that $\mathsf{cov}(X) = \overline{\mathsf{cov}}(h^{(\ell)})$ and $k_i - \sum_{t \in T} c_t^i = 0$ for all i, so indeed $t^* \in \mathsf{BR}(X)$, and **C1** and **C2*** are satisfied by X.

To see that **C3** is satisfied, observe that, similar to Lemma 3.14, we have $t \preceq_i^{(\ell-1)} t^*$ for each target t that receives a positive coverage from defender i in Phase 1, and we have $t \preceq_i^{(\ell)} t^*$ for each target t that receives a positive coverage from defender i in Phase 2. Therefore, if $c_t^i > 0$, defender i must have distributed some coverage to t in Phase 1 or Phase 2, so we have $t \preceq_i^{(\ell-1)} t^*$ or $t \preceq_i^{(\ell)} t^*$. Recall that $\preceq_i^{(\ell-1)}$ is induced by any $\overline{\mathsf{cov}}(h)$ with $h \in (h^{(\ell-1)}, h^{(\ell)})$. By Definition 3.18, $t \preceq_i^{(\ell-1)} t^*$ then implies, for all $h \in (h^{(\ell-1)}, h^{(\ell)})$, that

$$u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h), t\right) - u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h), t^*\right) \leq 0.$$

By continuity, we then have

$$u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h^{(\ell)}), j\right) - u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h^{(\ell)}), j^*\right) = \lim_{h \underset{<}{\longrightarrow} h^{(\ell)}} u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h), j\right) - u_i^{\mathsf{D}}\left(\overline{\mathsf{cov}}(h), j^*\right) \leq 0.$$

Thus, $t \succeq_i^{\overline{\mathsf{cov}}(h^{(\ell)})} t^*$ by Definition 3.18, or equivalently, $t \succeq_i^{X} t^*$ as we have $\mathsf{cov}(X) = \overline{\mathsf{cov}}(h^{(\ell)})$. The same can be shown when $t \preceq_i^{(\ell)} t^*$. Thus, **C3** is satisfied.

To summarize our proof, in Case (i), the strategy profile $(X^{(\ell)}, t^{(\ell)})$ satisfies **C1**, **C2***, and **C3** and forms an NSE; in Case (ii), we can merge $X^{(\ell-1)}$ and $X^{(\ell)}$ into a strategy profile (X, t^*) using Algorithm 3.2, which satisfies the same conditions. By Proposition 3.9, these strategy profiles form NSEs. Therefore, an NSE always exists. □

While we have shown that an NSE always exists in a game, we remark that in some games there might be more than one NSE. The defenders' strategy profile may also correspond to different level coverages in different NSEs. We present an example with multiple NSEs in the appendix (see Section 3.A.2).

3.5 Computing an NSE

We have demonstrated the existence of an NSE. Now we consider the problem of computing an NSE. Indeed, our constructive proof for the existence of an NSE already gives us a general idea about how to find an NSE. An NSE may involve irrational numbers as the probabilities in the players' strategies, and sometimes it is the only NSE in a game (we present an example in the appendix; see Section 3.A.1). To deal with this issue, we take the approach that is used to deal with the same issue with the computation of a Nash equilibrium in a normal form game, by seeking only approximate solutions [Daskalakis et al., 2009]. We say that a strategy profile (X, t) is an ϵ-NSE, if no player can improve their utility by more than ϵ by playing a different strategy; the defenders take into account the attacker's approximately best responses when evaluating the value of a strategy deviation. We denote by

$$\mathsf{BR}^{\epsilon}(\mathbf{c}) = \left\{t \in T : u^{\mathsf{A}}(\mathbf{c}, t) \geq u^{\mathsf{A}}(\mathbf{c}, \ell) - \epsilon \text{ for all } \ell \in T\right\}$$

the set of approximately best responses of the attacker. The ϵ-NSE is defined below.

Definition 3.21 (ϵ-NSE)**.** A strategy profile (X, t) form an ϵ-NSE if $t \in \mathsf{BR}^\epsilon(X)$ and, for every defender $i \in [n]$ and every strategy deviation $\mathbf{c}' \in \mathcal{C}_{k_i}$, it holds for the resulting strategy profile $X' = (\mathbf{c}^{-i}, \mathbf{c}')$ that

$$\min_{\ell \in \mathsf{BR}^\epsilon(X')} u_i^{\mathsf{D}}(X', \ell) \leq u_i^{\mathsf{D}}(X, t) + \epsilon. \tag{3.13}$$

In what follows, we normalize the payoff parameters to be in $[0, 1]$. Thus, the players' utilities also lie in this range and we only consider ϵ-NSEs with $\epsilon < 1$ (trivially, any strategy profile forms an ϵ-NSE if $\epsilon > 1$). We let ξ denote the precision of the payoff parameters, i.e., $|a - b| \geq \xi$ for any two parameters $a \neq b$. Thus, by our assumption that reward values are greater than the corresponding penalties, we have $r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t) \geq \xi$ and $r_i^{\mathsf{D}}(t) - p_i^{\mathsf{D}}(t) \geq \xi$ for all $i \in [n]$ and $t \in T$.

Theorem 3.22. *For any $\epsilon > 0$, an ϵ-NSE can be computed in time* $\mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, m, n)$.

Proof sketch. The full proof can be found in the appendix of this chapter (see Section 3.B.2). The following steps find an ϵ-NSE in polynomial time.

1. Compute the set $\mathcal{H}$ of intersection points as defined in (3.12). This can be done in time $\mathsf{poly}(n, m)$ by enumerating all the triples $(i, t, t') \in [n] \times T^2$. Then sort the elements in $\mathcal{H}$ and let $h^{(1)}, \ldots, h^{(L)}$ be the elements in ascending order.

2. For each interval $(h^{(\ell)}, h^{(\ell+1)})$, pick an arbitrary value h in this interval and compute the preference profile $P^{(\ell)} = (\preceq_1^{(\ell)}, \ldots, \preceq_n^{(\ell)})$ induced by $\overline{\mathsf{cov}}(h)$. Fix the preferences to $P^{(\ell)}$, we run Algorithm 3.1 on both $h = h^{(\ell)}$ and $h = h^{(\ell+1)}$ to determine if there is a zero of surp in $[h^{(\ell)}, h^{(\ell+1)}]$. Note that we can compute the exact value of $\mathsf{surp}(h)$ for any given rational number h as Algorithm 3.1 involves only basic arithmetic operations, and the size of the value in binary representation is polynomial in that of h.

   If it is the case that $\mathsf{surp}(h^{(\ell)}) \cdot \mathsf{surp}(h^{(\ell+1)}) \leq 0$, then there is indeed a zero in this interval. Since $\mathsf{surp}(h)$ is an increasing function, we can use binary search to approximate the zero and obtain an ϵ-NSE. The key to this approach is to argue that binary search will indeed converge to the desired precision level in time $\mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, m, n)$, which is not a trivial task as the output of Algorithm 3.1 changes very fast with respect to h when h approaches $H_{\min}$; indeed, the proof is technically involved and we leave the details to the full proof in the appendix of this chapter (Appendix 3.B).

   If it is the case that $\mathsf{surp}(h^{(\ell)}) \cdot \mathsf{surp}(h^{(\ell+1)}) > 0$ for all $\ell = 1, \ldots, L-1$, we continue with the next step.

3. By Lemma 3.19, now there must be an $h^{(\ell)} \in (\!|H^{(\ell-1)}, H^{(\ell)}|\!)$, where $H^{(\ell-1)}$ and $H^{(\ell)}$ are heights of the defender strategy profiles in NSEs obtained at the zeros of surp, when the preference profile is fixed to $P^{(\ell-1)}$ and $P^{(\ell)}$, respectively. Note that, to identify such $h^{(\ell)}$, we do not actually need to know $H^{(\ell-1)}$ or $H^{(\ell)}$. Instead, we check if $\mathsf{surp}(h^{(\ell)})$ has different signs when we fix the input preference profile to $P^{(\ell-1)}$ and $P^{(\ell)}$, respectively (we have already computed the profiles in Step 2). The zeros $H^{(\ell-1)}$ and $H^{(\ell)}$ must lie at different sides of $h^{(\ell)}$ on the axis if the signs are different. Using binary search, we can obtain an ϵ-NSE in time $\mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, m, n)$ in this case, too. We leave the details to the full proof. □

3.6 NSE in Other Settings

To conclude this chapter, we present two extended variants of our model in which NSE may fail to exist.

3.6.1 Games with Malicious Defenders

It is assumed that $r_i^{\mathsf{D}}(t) > p_i^{\mathsf{D}}(t)$ in our model, in the sense that every defender prefers the attack to be unsuccessful no matter which target is attacked. We show that without this assumption an NSE may not exist. Consider the following game.

Example 3.23. There are two defenders α and β and a set $T = \{t_1, t_2\}$ of two targets. Each defender has one resource, i.e., $k_\alpha = k_\beta = 1$. The players' payoffs are given below.

	t_1	t_2
r_α^{D}	0	1
p_α^{D}	2	3

defender α

	t_1	t_2
r_β^{D}	3	1
p_β^{D}	2	0

defender β

	t_1	t_2
r^{A}	1	1
p^{A}	0	0

attacker

Thus, defender β has non-overlapping payoffs captured by the preference order $t_2 \prec_\beta t_1$. In addition, defender α prefers maliciously both targets to be attacked successfully. The attacker's payoffs are identical for the two targets, so his best response is the target with the lower coverage. □

We show that the above game does not admit an NSE. Suppose for the sake of contradiction that (X, t^*) forms an NSE in this game. Let $X = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$. Observe that X must be level for the same reason as Proposition 3.6: if X was not level, defender β would have an incentive to rebalance the coverage to improve the coverage of t^* and, hence, her utility. Indeed, in this example, it is feasible for defender β to rebalance X till it is level, no matter what strategy defender α plays in X. In this example, X being level implies that $\mathsf{cov}_{t_1}(X) = \mathsf{cov}_{t_2}(X)$.

Consider the following two cases.

Case 1. $t^* = t_1$. Now that $\mathsf{cov}_{t_1}(X) = \mathsf{cov}_{t_2}(X)$, according to the payoffs of defender α, we have $u_\alpha^{\mathsf{D}}(X, t_1) < u_\alpha^{\mathsf{D}}(X, t_2)$. As a result, it must be that $c_{t_2}^\alpha = 0$; otherwise, α can reduce $c_{t_2}^\alpha$ by a small amount to induce the attacker to attack t_2 and improve her utility. Therefore, $\mathsf{cov}_{t_2}(X) = c_{t_2}^\beta$. If $\mathsf{cov}_{t_2}(X) < 1$, then α would have an incentive to deviate to playing $\mathbf{c}' = (1, 0)$: this will result in the strategy profile to change to X' such that $\mathsf{cov}_{t_1}(X') = 1$ and $\mathsf{cov}_{t_2}(X') = c_{t_2}^\beta < 1$, so $\mathsf{BR}(X') = \{t_2\}$; it follows that

$$\min_{t \in \mathsf{BR}(X')} u_\alpha^{\mathsf{D}}(X', t) = u_\alpha^{\mathsf{D}}(X', t_2) > u_\alpha^{\mathsf{D}}(X, t^*),$$

which contradicts the assumption that (X, t^*) is an NSE.

Thus, the only possibility is that $\mathsf{cov}_{t_2}(X) = 1$; hence, $\mathsf{cov}_{t_1}(X) = 1$ as well and $u_\alpha^{\mathsf{D}}(X, t^*) = 0$. Since t_2 is protected by β only, we have $c_{t_2}^\beta = 1$ and hence, $c_{t_1}^\beta = 0$ given that $k_\beta = 1$. However, now defender α would be better off deviating to playing $\mathbf{c}' = (0, 0)$: the resulting strategy profile X' is such that $\mathsf{cov}_{t_1}(X') = c_{t_1}^\beta < c_{t_2}^\beta = \mathsf{cov}_{t_2}(X')$, so we have $\mathsf{BR}(X') = \{t_1\}$; it follows that

$$\min_{t \in \mathsf{BR}(X')} u_\alpha^{\mathsf{D}}(X', t) = u_\alpha^{\mathsf{D}}(X', t_1) = u_\alpha^{\mathsf{D}}(0, t_1) = 2 > u_\alpha^{\mathsf{D}}(X, t^*),$$

so (X, t^*) is not an NSE — a contradiction.

Case 2. $t^* = t_2$. Similar to (i), in this case, it must be that $c^\beta_{t_1} = 0$ since otherwise β can reduce the coverage of t_1 to induce the attacker to choose t_1, which benefits β since $t_2 \prec_\beta t_1$. Thus, $\mathsf{cov}_{t_1}(X) = c^\alpha_{t_1}$. It must be that $\mathsf{cov}_{t_1}(X) = c^\alpha_{t_1} = 1$: if $\mathsf{cov}_{t_1}(X) < 1$, then β can play $(0,1)$ to induce the attacker to choose t_1. Therefore, $\mathsf{cov}_{t_2}(X) = \mathsf{cov}_{t_1}(X) = 1, \implies \mathbf{c}^\alpha = (1,0)$, $\mathbf{c}^\beta = (0,1)$, and $u^{\mathsf{D}}_\alpha(X, t^*) = 1$. Again, defender α now has an incentive to deviate to playing $(0,0)$, which will result in $\mathsf{BR}(X') = \{t_1\}$ and

$$\min_{t \in \mathsf{BR}(X')} u^{\mathsf{D}}_\alpha(X', t) = u^{\mathsf{D}}_\alpha(X', t_1) = u^{\mathsf{D}}_\alpha(0, t_1) = 2 > u^{\mathsf{D}}_\alpha(X, t^*),$$

contradicting the assumption that (X, t^*) is an NSE.

Therefore, there is no NSE in Example 3.23.

3.6.2 Games with Scheduling Constraints

In games we have considered so far, other than the defender's resource budget, no other constraints are imposed. We show that in games with scheduling constraints, an NSE may not exist. In the presence of scheduling constraints, defenders allocate their resources to *schedules*, which are subsets of targets. A resource allocated to a schedule protects all the targets in this schedule (see Section 2.2.2 in Chapter 2 for a discussion about SSGs with scheduling constraints). The NSE definition for this setting remains the same as Definition 3.1, except that the feasible strategy space of defender i is not $\mathcal{C}_{k_i}$. For this reason, in what follows we will describe a defender (mixed) strategy as a probability distribution over the pure strategies, instead of a coverage vector. Consider the following game.

Example 3.24. There are two defenders α and β and a set $T = \{t_1, t_2, t'_1, t'_2\}$ of four targets. Each defender has one resource, i.e., $k_\alpha = k_\beta = 1$. In addition to the empty set, there are two schedules available for defender α: $\{t_1, t_2\}$ and $\{t'_1, t'_2\}$; and two available for defender β: $\{t_2\}$ and $\{t'_2\}$. Let $S^\alpha = \{\varnothing, \{t_1, t_2\}, \{t'_1, t'_2\}\}$ and $S^\beta = \{\varnothing, \{t_2\}, \{t'_2\}\}$. Now that each defender $i \in \{\alpha, \beta\}$ has only one resource, we can represent their mixed strategy as $\mathbf{x}^i \in \Delta(S^i)$ and denote by x^i_s the probability they allocate their resource to schedule s. The players' payoffs are given below, where t_i and t'_i $(i = 1, 2)$ are identical for every player.

	t_1, t'_1	t_2, t'_2
r^{D}_α	3	1
p^{D}_α	2	0

defender α

	t_1, t'_1	t_2, t'_2
r^{D}_β	1	3
p^{D}_β	0	2

defender β

	t_1, t'_1	t_2, t'_2
r^{A}	3	4
p^{A}	0	1

attacker

The payoffs are partially non-overlapping: $t_2 \prec_\alpha t_1$ and $t'_2 \prec_\alpha t'_1$ for α; and $t_1 \prec_\beta t_2$ and $t'_1 \prec_\beta t'_2$ for β. □

Suppose for the sake of contradiction that (X, t^*) is an NSE in the above game, $X = (\mathbf{x}^\alpha, \mathbf{x}^\beta)$. According to the defenders' feasible schedules, we can establish the following equations.

$$\begin{aligned}
\mathsf{cov}_{t_1}(X) &= x^\alpha_{\{t_1,t_2\}} \\
\mathsf{cov}_{t_2}(X) &= x^\alpha_{\{t_1,t_2\}} \oplus x^\beta_{\{t_2\}} \\
\mathsf{cov}_{t'_1}(X) &= x^\alpha_{\{t'_1,t'_2\}} \\
\mathsf{cov}_{t'_2}(X) &= x^\alpha_{\{t'_1,t'_2\}} \oplus x^\beta_{\{t'_2\}}
\end{aligned}$$

We make two observations.

Claim 1. $t^* \notin \{t_1, t_1'\}$.

Proof. Indeed, defender β always prefers t^* to be t_2 or t_2' rather than t_1 or t_1', and she can prevent t_1 and t_1' from being the attacker's best responses by assigning her resource to $\varnothing$ with probability 1: in the resulting strategy profile X', the coverage is contributed solely by defender α, so we have $\mathsf{cov}_{t_1}(X') = \mathsf{cov}_{t_2}(X')$ and $\mathsf{cov}_{t_1'}(X') = \mathsf{cov}_{t_2'}(X')$ given that t_1 and t_2 (respectively, t_1' and t_2') always appear in the same schedule; hence, according to the attacker's payoffs, we have $u^{\mathsf{A}}(X', t_1) < u^{\mathsf{A}}(X', t_2)$ and $u^{\mathsf{A}}(X', t_1') < u^{\mathsf{A}}(X', t_2')$, which implies that $\{t_1, t_1'\} \cap \mathsf{BR}(X') = \varnothing$. □

Without loss of generality, we can assume that $t^* = t_2$ (in the case $t^* = t_2'$, the same argument applies by switching t_2 and t_2').

Claim 2. $x^{\beta}_{\{t_2\}} \leq 1/3$ *and* $x^{\beta}_{\{t_2'\}} \leq 1/3$.

Proof. Indeed, if $x^{\beta}_{\{t_2\}} > 1/3$, defender α would have an incentive to deviate. She can reduce $x^{\alpha}_{\{t_1,t_2\}}$ to 0. This will make t_1 the only best response of the attacker and induce the attacker to attack it, which benefits α given that $t^* = t_2 \prec_{\alpha} t_1$. Specifically, let $\mathbf{y}$ be the strategy α deviates to playing, i.e.,

$$y_{\{t_1,t_2\}} = 0, \quad y_{\varnothing} = x_{\varnothing} + x^{\alpha}_{\{t_1,t_2\}}, \quad \text{and} \quad y^{\alpha}_{\{t_1',t_2'\}} = x^{\alpha}_{\{t_1',t_2'\}};$$

let $X' = (\mathbf{y}, \mathbf{x}^{\beta})$ be the resulting strategy profile. We have $\mathsf{cov}_{t_1}(X') = y_{\{t_1,t_2\}} = 0$ and hence,

$$\mathsf{cov}_{t_1}(X') + \frac{1}{3} < x^{\beta}_{\{t_2\}} \oplus y_{\{t_1,t_2\}} = \mathsf{cov}_{t_2}(X'),$$

which implies that

$$u^{\mathsf{A}}(X', t_1) > u^{\mathsf{A}}(X', t_2).$$

In addition,

$$\mathsf{cov}_{t_2}(X') = y_{\{t_1,t_2\}} \oplus x^{\beta}_{\{t_2\}} \leq x^{\alpha}_{\{t_1,t_2\}} \oplus x^{\beta}_{\{t_2\}} = \mathsf{cov}_{t_2}(X),$$

and since $y^{\alpha}_{\{t_1',t_2'\}} = x^{\alpha}_{\{t_1',t_2'\}}$ for the other two targets $t \in \{t_1', t_2'\}$ we have

$$\mathsf{cov}_{t}(X') = \mathsf{cov}_{t}(X).$$

Hence, by monotonicity $u^{\mathsf{A}}(X', t_2) \geq u^{\mathsf{A}}(X, t_2)$ and $u^{\mathsf{A}}(X, t) = u^{\mathsf{A}}(X', t)$. It follows that

$$u^{\mathsf{A}}(X', t_1) > u^{\mathsf{A}}(X', t_2) \geq u^{\mathsf{A}}(X, t_2) \geq u^{\mathsf{A}}(X, t) = u^{\mathsf{A}}(X', t),$$

which implies that, indeed, $\mathsf{BR}(X') = \{t_1\}$ after the deviation. The same argument applies if we assume $x^{\beta}_{\{t_2'\}} \leq 1/3$, when defender α would have an incentive to reduce $x^{\alpha}_{\{t_1',t_2'\}}$ to 0. □

The fact that $x^{\beta}_{\{t_2\}} \leq 1/3$ further implies that

$$\mathsf{cov}_{t_1}(X) + \frac{1}{3} = x^{\alpha}_{\{t_1,t_2\}} + \frac{1}{3} > x^{\alpha}_{\{t_1,t_2\}} \oplus \frac{1}{3} \geq x^{\alpha}_{\{t_1,t_2\}} \oplus x^{\beta}_{\{t_2\}} = \mathsf{cov}_{t_2}(X),$$

so according to the attacker's payoffs, we have $u^{\mathsf{A}}(X, t_1) < u^{\mathsf{A}}(X, t_2)$. Similarly, we also have $u^{\mathsf{A}}(X, t_1') < u^{\mathsf{A}}(X, t_2')$.

Now if it is the case that $\mathsf{cov}_{t_2}(X) < 1$, since $x^{\beta}_{\{t_2\}} \leq 1/3$ and $x^{\beta}_{\{t_2'\}} \leq 1/3$, defender β can deviate to playing a strategy $\mathbf{y}$ that increases the coverage of t_2 while keeping it the only attacker best response. Specifically, we let

$$y_{\{t_2\}} = x_{\{t_2\}} + \delta, \quad y_{\{t_2'\}} = x_{\{t_2'\}} + \delta', \quad \text{and} \quad y_{\varnothing} = x_{\varnothing} - \delta - \delta'.$$

When $\delta > 0$, the coverage of t_2 increases, i.e.,

$$\mathsf{cov}_{t_2}(X) = y_{\{t_2\}} \oplus x^{\alpha}_{\{t_1,t_2\}} > x_{\{t_2\}} \oplus x^{\alpha}_{\{t_1,t_2\}} = \mathsf{cov}_{t_2}(X),$$

so by monotonicity we have $u^{\mathsf{D}}_{\delta}(X', t_2) > u^{\mathsf{D}}_{\delta}(X, t_2)$. When $\delta, \delta' > 0$ are sufficiently small, we can keep $u^{\mathsf{A}}(X', t_1) < u^{\mathsf{A}}(X', t_2)$ and $u^{\mathsf{A}}(X', t_1') < u^{\mathsf{A}}(X', t_2')$ satisfied with the strategy profile X' after the deviation; and when δ is sufficiently smaller than δ', we can have $u^{\mathsf{A}}(X', t_2) > u^{\mathsf{A}}(X', t_2')$ (we have $u^{\mathsf{A}}(X, t_2) \geq u^{\mathsf{A}}(X, t_2')$ as $t_2 = t^* \in \mathsf{BR}(X)$). Thus, indeed, $\mathsf{BR}(X') = \{t_2\}$. It follows that,

$$\min_{t \in \mathsf{BR}(X')} u^{\mathsf{D}}_{\delta}(X', t) = u^{\mathsf{D}}_{\delta}(X', t_2) > u^{\mathsf{D}}_{\delta}(X, t_2),$$

which contradicts the assumption that (X, t^*) is an NSE.

It remains to check the case $\mathsf{cov}_{t_2}(X) = 1$. Indeed, by Claim 2, $x^{\beta}_{\{t_2\}} < 1$, so $\mathsf{cov}_{t_2}(X) = x^{\alpha}_{\{t_1,t_2\}} \oplus x^{\beta}_{\{t_2\}} = 1$ implies that $x^{\alpha}_{\{t_1,t_2\}}$. Hence, $x^{\alpha}_{\{t_1',t_2'\}} = x^{\alpha}_{\varnothing} = 0$, which means that t_1' and t_2' are protected by defender β only. Therefore,

$$\mathsf{cov}_{t_2'}(X) = x^{\beta}_{\{t_2'\}} < 1/3 < \mathsf{cov}_{t_2}(X),$$

which implies that

$$u^{\mathsf{A}}(X, t_2) < u^{\mathsf{A}}(X, t_2').$$

This contradicts the assumption that $t_2 \in \mathsf{BR}(X)$.

Therefore, (X, t^*) cannot be an NSE; there does not exist an NSE in Example 3.24.

3.7 Summary

We have presented a model of multi-defender SSG, proposed the NSE as the solution concept for the uncoordinated situation, and demonstrated the existence of an NSE in every game, as well as the tractability of computing an NSE. In the next chapter, we will use the NSE to define one of the basic properties to be achieved through defense coordination and proceed with our task of coordination mechanism design.

Appendix 3.A Additional Examples of NSEs

3.A.1 A Game with No Rational NSE

We show that the following game instance does not have an NSE in which the players' strategies can be described using rational numbers only.

Example 3.25. There are two defenders α and β and three targets t_1, t_2, and t_3. The defenders each have one resource and have non-overlapping payoffs that are captured by the preference orders $t_1 \prec_\alpha t_2 \prec_\alpha t_3$ and $t_3 \prec_\beta t_2 \prec_\beta t_1$. For the attacker, the three targets are identical: $r^{\mathsf{A}}(t) = 0$ and $p^{\mathsf{A}}(t) = -1$ for all t, so his utility from attacking a target is exactly the negative of the coverage of this target, and his best response set contains targets with the lowest coverage. □

Suppose (X, t^*) forms an NSE; $X = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$ with $\mathbf{c}^\alpha$ and $\mathbf{c}^\beta$ being coverage vectors. We claim that $\mathbf{c}^\alpha = \left(\frac{\sqrt{5}-1}{2}, \frac{3-\sqrt{5}}{2}, 0\right)$ and $\mathbf{c}^\beta = \left(0, \frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}-1}{2}\right)$, which involves irrational numbers. Indeed, we can verify this by noting the following facts.

First, the attacker's response in the NSE must be $t^* = t_2$. Indeed, suppose for the sake of contradiction that $t^* = t_1$. Since t_1 is the least preferred attacker response of defender α, this defender can deviate to protecting target t_1 only, i.e., playing $\mathbf{c}' = (1, 0, 0)$. This will improve the coverage of t_1 to 1 and prevent t_1 to be a best response of the attacker (at least one of t_2 and t_3 must have coverage less than 1 no matter what strategy β plays). As a result, α is better off making this deviation, so (X, t^*) cannot be an NSE. Likewise, if we assume that $t^* = t_3$, defender β can make a beneficial strategy deviation. Thus, $t^* = t_2$.

According to the attacker's payoffs, we have $H_{\min} = \max_{t \in T} p^{\mathsf{A}}(t) = -1$; hence, $\overline{\mathsf{cov}}_t(H_{\min}) = 1$ and $\sum_{t \in T} \overline{\mathsf{cov}}_t(H_{\min}) = 3 > k_\alpha + k_\beta$. By Corollary 3.7, we have: (i) X is level. (ii) $c_1^\alpha + c_2^\alpha = 1$ and $c_2^\beta + c_3^\beta = 1$. (iii) $c_1^\beta = c_3^\alpha = 0$ (since $t_2 \prec_\alpha t_3$ and $t_2 \prec_\beta t_1$). Observe that a target is surely an attacker best response in this game if its coverage is zero, i.e., $t \in \mathsf{BR}(X)$ if $\mathsf{cov}_t(X) = 0$. On the other hand, the fact that X is level implies that $t \in \mathsf{BR}(X)$ if $\mathsf{cov}_t(X) > 0$ (Lemma 3.5). Thus, every $t \in T$ is in $\mathsf{BR}(X)$; we have $\mathsf{cov}_{t_1}(X) = \mathsf{cov}_{t_2}(X) = \mathsf{cov}_{t_3}(X)$, i.e., $c_1^\alpha \oplus c_1^\beta = c_2^\alpha \oplus c_2^\beta = c_3^\alpha \oplus c_3^\beta$.

We obtain the following system of equations:

$$
\begin{aligned}
&c_1^\alpha \oplus c_1^\beta = c_2^\alpha \oplus c_2^\beta = c_3^\alpha \oplus c_3^\beta \\
&c_1^\beta = c_3^\alpha = 0 \\
&c_1^\alpha + c_2^\alpha = 1 \\
&c_2^\beta + c_3^\beta = 1
\end{aligned}
$$

The only solution of the above system with all the coverage values within $[0, 1]$ is $\mathbf{c}^\alpha = \left(\frac{\sqrt{5}-1}{2}, \frac{3-\sqrt{5}}{2}, 0\right)$ and $\mathbf{c}^\beta = \left(0, \frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}-1}{2}\right)$.

3.A.2 A Game with Multiple NSEs

On the other hand, there are games that admit more than one NSEs and even more than one NSEs with the same attacker response. For NSEs with the same attacker response, the defenders' strategy profiles are not necessarily in a connected set.

Example 3.26. There are two defenders α and β and a set $T = \{t_1, t_2, t_3\}$ of three targets. The targets are identical for the attacker: $r^{\mathsf{A}}(t) = 1$ and $p^{\mathsf{A}}(t) = 0$ for all $t \in T$. The defenders' payoffs are given below.

	t_1	t_2	t_3
r^{D}_α	0.8	0.9	1
p^{D}_α	1/4	1/8	0

defender α

	t_1	t_2	t_3
r^{D}_β	1	0.9	0.8
p^{D}_β	0	1/8	1/4

defender β

Each defender has one resource, i.e., $k_\alpha = k_\beta = 1$. □

There are two types of NSEs in the example above. The first type (X, t^*) is the same as the only NSE in Example 3.25, where $t^* = t_2$ and $X = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$: $\mathbf{c}^\alpha = \left(\frac{\sqrt{5}-1}{2}, \frac{3-\sqrt{5}}{2}, 0\right)$ and $\mathbf{c}^\beta = \left(0, \frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}-1}{2}\right)$. Indeed, X corresponds to the level coverage $\left(\frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}-1}{2}\right)$, at which the defenders' preferences are $t_1 \prec_\alpha t_2 \prec_\alpha t_3$ and $t_3 \prec_\beta t_2 \prec_\beta t_1$, which are exactly the same as in Example 3.25.

The second type of NSEs (X, t^*) is such that $\mathsf{cov}_1(X) = \mathsf{cov}_2(X) = \mathsf{cov}_3(X) = 5/9$ with t^* being any $t \in T$. This can happen when each defender distributes their resources uniformly over the three targets, which results in overall coverage $1 - (1 - 1/3)^2 = 5/9$ for each target. In this case, the overall coverage results in both defenders getting the same utility no matter which target the attacker chooses to attack, while the attacker also finds it equally beneficial to attack each target. Hence, (X, t^*) form an NSE.

These are the only NSEs in this game. Indeed, we expect an NSE X to be level as we have demonstrated in this chapter, so in this particular example, we must have $\mathsf{cov}_1(X) = \mathsf{cov}_2(X) = \mathsf{cov}_3(X) = h$ for some $h \in [0, 1]$. Now if $h \neq 5/9$, the defenders preferences are either $t_1 \prec_\alpha t_2 \prec_\alpha t_3$ and $t_3 \prec_\beta t_2 \prec_\beta t_1$, or the opposite, by which we can only obtain an NSE of the first type. If $h = 5/9$, then we obtain NSEs of the second type. The defenders' strategy profiles are not in a connected set in this case, even for NSEs in which the attacker attacks t_2.

Appendix 3.B Omitted Proofs

3.B.1 Proof of Proposition 3.2

We prove in two parts.

(i) NSE implies SSE. Suppose that $(\mathbf{c}, t)$ forms an SSE. We show that (X, t) forms an NSE. By definition $t \in \mathsf{BR}(\mathbf{c})$, so we only need to argue that (3.4) holds for any deviation $\mathbf{c}' \in \mathcal{C}_k$, where k is the number of resources the defender has. Indeed, suppose for the sake of contradiction that (3.4) is violated. We would have

$$\min_{\ell \in \mathsf{BR}(\mathbf{c}')} u^{\mathsf{D}}(\mathbf{c}', \ell) > u^{\mathsf{D}}(X, t) = u^{\mathsf{D}}(\mathbf{c}, t).$$

Since $(\mathbf{c}, t)$ forms an SSE, by definition, we have

$$u^{\mathsf{D}}(\mathbf{c}, t) = \max_{\mathbf{z} \in \mathcal{C}_k, \ell \in \mathsf{BR}(\mathbf{z})} u^{\mathsf{D}}(\mathbf{z}, \ell).$$

This leads to the following contradiction:

$$\max_{\ell \in \mathsf{BR}(\mathbf{c}')} u^{\mathsf{D}}(\mathbf{c}', \ell) > u^{\mathsf{D}}(\mathbf{c}, t) = \max_{\mathbf{z} \in \mathcal{C}_k, \ell \in \mathsf{BR}(\mathbf{z})} u^{\mathsf{D}}(\mathbf{z}, \ell) \geq \max_{\ell \in \mathsf{BR}(\mathbf{c}')} u^{\mathsf{D}}(\mathbf{c}', \ell).$$

Thus, (X, t) is an NSE.

(ii) SSE implies NSE. Suppose that (X, t) forms an NSE and, for the sake of contradiction, that $(\mathbf{c}, t)$ is not an SSE. Suppose that $(\mathbf{c}', t')$ is an SSE. By definition, we have

$$u^{\mathsf{D}}(\mathbf{c}', t') > u^{\mathsf{D}}(\mathbf{c}, t) = u^{\mathsf{D}}(X, t).$$

By assumption, $c_\ell < 1$ for all $\ell \in \mathsf{BR}(X)$, so

$$\mathsf{height}(X) = \max_{\ell' \in T} u^{\mathsf{A}}(\mathbf{c}, \ell') = u^{\mathsf{A}}(\mathbf{c}, \ell) > p^{\mathsf{A}}(\ell);$$

in addition, for all $\ell \notin \mathsf{BR}(X)$, we have

$$\mathsf{height}(X) = \max_{\ell' \in T} u^{\mathsf{A}}(\mathbf{c}, \ell') > u^{\mathsf{A}}(\mathbf{c}, \ell) \geq p^{\mathsf{A}}(\ell).$$

Hence, $\mathsf{height}(X) > \max_{\ell \in T} p^{\mathsf{A}}(\ell) = H_{\min}$, which by Proposition 3.6 implies that X is level and $\sum_{\ell \in T} c_\ell = k$.

Consider the following two cases.

(i) $u^{\mathsf{A}}(\mathbf{c}', t') < \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{c}, \ell)$. Since $(\mathbf{c}', t')$ is an SSE, we have $t' \in \mathsf{BR}(\mathbf{c}')$, or $u^{\mathsf{A}}(\mathbf{c}', \ell) \leq u^{\mathsf{A}}(\mathbf{c}', t')$ for all $\ell \in T$. Thus, for all $\ell \in \mathsf{BR}(\mathbf{c})$, we have

$$u^{\mathsf{A}}(\mathbf{c}', \ell) \leq u^{\mathsf{A}}(\mathbf{c}', t') < \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{c}, \ell) = u^{\mathsf{A}}(\mathbf{c}, \ell),$$

which implies that $c'_\ell > c_\ell$ by monotonicity of the utility function. For those $\ell \notin \mathsf{BR}(\mathbf{c})$, since $\mathbf{c}$ is level, by Lemma 3.5, we have $c_\ell = 0$. It follows that

$$\sum_{\ell \in T} c'_\ell > \sum_{\ell \in T} c_\ell = k, \tag{3.14}$$

which means that $\mathbf{c}'$ violates the budget constraint and cannot be a feasible strategy in $\mathcal{C}_k$. This contradicts the assumption that $(\mathbf{c}', t')$ is an SSE.

(ii) $u^{\mathsf{A}}(\mathbf{c}', t') \geq \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{c}, \ell)$. In this case, we have $u^{\mathsf{A}}(\mathbf{c}', t') \geq \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{c}, \ell) \geq u^{\mathsf{A}}(\mathbf{c}, t')$, which means that $c'_{t'} \leq c_{t'}$ by monotonicity.

If $t' \notin \mathsf{BR}(\mathbf{c})$, then since $\mathbf{c}$ is level, by Lemma 3.5, $c_{t'} = 0 \leq c'_{t'}$. Thus, $u^{\mathsf{A}}(\mathbf{c}', t') \leq u^{\mathsf{A}}(\mathbf{c}, t')$ and the following inequality holds for all $\ell \in \mathsf{BR}(\mathbf{c})$:

$$u^{\mathsf{A}}(\mathbf{c}', \ell) \leq u^{\mathsf{A}}(\mathbf{c}', t') \leq u^{\mathsf{A}}(\mathbf{c}, t') < u^{\mathsf{A}}(\mathbf{c}, \ell),$$

where the first transition is due to $t' \in \mathsf{BR}(\mathbf{c}')$ and the last is due to the fact that $t' \notin \mathsf{BR}(\mathbf{c})$ but $\ell \in \mathsf{BR}(\mathbf{c})$. Hence, $c'_\ell > c_\ell$ by monotonicity. For those $\ell \notin \mathsf{BR}(\mathbf{c})$, we have $c_\ell = 0$ by Lemma 3.5. Thus, similar to (3.14), the budget constraint would be violated, contradicting the assumption that $(\mathbf{c}', t')$ is an SSE.

Therefore, the only possibility is that $t' \in \mathsf{BR}(\mathbf{c})$. By assumption $c_{t'} > 0$, so by reducing $c_{t'}$ by a small amount $\delta > 0$ we can make target t' the only best response of the attacker against the coverage $\tilde{\mathbf{c}}$ resulting from this coverage reduction. Moreover, since $c'_{t'} \leq c_{t'}$, we have $\tilde{c}_{t'} = c_{t'} - \delta \geq c'_{t'} - \delta$, and since $u^{\mathsf{D}}(\cdot, t')$ is a continuous function, when δ is sufficiently small, we can ensure that $u^{\mathsf{D}}(\tilde{\mathbf{c}}, t') \geq u^{\mathsf{D}}(\mathbf{c}', t') - \epsilon$ for any $\epsilon > 0$. Choosing an arbitrary $\epsilon \in (0,\ u^{\mathsf{D}}(\mathbf{c}', t') - u^{\mathsf{D}}(\mathbf{c}, t))$ then gives:

$$\begin{aligned}\min_{\ell \in \mathsf{BR}(\tilde{\mathbf{c}})} u^{\mathsf{D}}(\tilde{\mathbf{c}}, \ell) = u^{\mathsf{D}}(\tilde{\mathbf{c}}, t') &\geq u^{\mathsf{D}}(\mathbf{c}', t') - \epsilon \\ &> u^{\mathsf{D}}(\mathbf{c}', t') - \Big(u^{\mathsf{D}}(\mathbf{c}', t') - u^{\mathsf{D}}(\mathbf{c}, t)\Big) = u^{\mathsf{D}}(\mathbf{c}, t),\end{aligned}$$

so the defender would be better-off deviating from (X, t) to playing $\tilde{\mathbf{c}}$, contradicting the assumption that (X, t) is an NSE.

3.B.2 Proof of Theorem 3.22

As we have demonstrated in the proof sketch in Section 3.5, the key to the proof is to argue that, in the second and third steps of the algorithm we presented in the proof sketch, binary search can indeed find an ϵ-NSE in polynomial time. We prove this using the following two lemmas.

Lemma 3.27. *Given two numbers a and b, such that $H_{\min} \leq a < b \leq H_{\max}$, $\mathsf{surp}(a) \leq 0$ and $\mathsf{surp}(b) \geq 0$ when the input preference profile is fixed to P, and $\overline{\mathsf{cov}}(h)$ induces P for all $h \in (a, b)$. For any given $\epsilon > 0$, an ϵ-NSE can be computed in time $\mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, m, n)$.*

Lemma 3.28. *Given a number $h^* \in (H_{\min}, H_{\max})$ and two induced preference profiles P^a and P^b, such that $\overline{\mathsf{cov}}(h)$ induces P^a for all $h \in (a, h^*)$, and $\overline{\mathsf{cov}}(h)$ induces P^b for all $h \in (h^*, b)$, $a < h^* < b$. Let $\mathsf{surp}(P, h)$ denote the value of surp on input preference profile P and height h. If $\mathsf{surp}(P^a, h^*) \cdot \mathsf{surp}(P^b, h^*) < 0$, then for any given $\epsilon > 0$, an ϵ-NSE can be computed in time $\mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, m, n)$.*

To prove the above two lemmas, we need Proposition 3.29.

Proposition 3.29. *Let $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$ and $Y = (\mathbf{z}^1, \ldots, \mathbf{z}^n)$ be two profiles of defender strategies (in the form of coverage vectors). Suppose that (X, t^*) forms an NSE, X is level, and $\sum_{\ell \in T} c^i_\ell = k_i$ for all $i \in [n]$. Then (Y, ℓ^*) forms an ϵ-NSE if all of the following hold:*

(i) $\ell^* \in \mathsf{BR}^{\epsilon/3}(Y)$.

(ii) $|c^i_\ell - z^i_\ell| \leq \sigma = \frac{\epsilon}{3^{n+1}}$ *for all* $i \in [n]$ *and* $\ell \in T$.

(iii) $z_\ell^i = 0$ *if* $\ell \succ_i^X \ell^*$.

Proof. Since we already have $\ell^* \in \mathsf{BR}^{\epsilon/3}(Y) \subseteq \mathsf{BR}^{\epsilon}(Y)$, it remains to show that no defender can improve their utility by more than ϵ by deviating from Y.

Observe that the following holds if $y_1, y_2 \in [0,1]$ and $|\delta_1|, |\delta_2| \le \delta \le 1$ for some δ:

$$|(y_1+\delta_1)(y_2+\delta_2) - y_1 \cdot y_2| = |\delta_1 \cdot \delta_2 + \delta_2 \cdot y_1 + \delta_1 \cdot y_2| \le 3\delta. \tag{3.15}$$

If $y_1, y_2, y_3 \in [0,1]$ and $|\delta_1|, |\delta_2|, |\delta_3| \le \delta \le 1/3$, we have:

$$\begin{aligned}|(y_1+\delta_1)(y_2+\delta_2)(y_3+\delta_3) - y_1 \cdot y_2 \cdot y_3| &\le |(y_1 \cdot y_2 + 3\delta)(y_3+\delta_3) - y_1 \cdot y_2 \cdot y_3| \\ &\le |(y_{12}+\delta_{12})(y_3+\delta_3) - y_{12} \cdot y_3| \\ &\le 3^2\delta,\end{aligned}$$

where we let $y_{12} = y_1 \cdot y_2$ and $\delta_{12} = 3\delta$ and apply (3.15). Likewise, if $y_1, \dots, y_n \in [0,1]$ and $|\delta_1|, \dots, |\delta_n| \le \delta \le \frac{1}{3^n}$, we will have $\left|\prod_{i\in[n]}(y_i+\delta_i) - \prod_{i\in[n]} y_i\right| \le 3^n\delta$.

Now let $\delta_\ell = z_\ell^i - c_\ell^i$ for each $\ell \in T$, so $|\delta_\ell| \le \sigma \le \frac{1}{3^n}$ by the assumption (ii) of this lemma. We have

$$\begin{aligned}|\mathsf{cov}_\ell(Y) - \mathsf{cov}_\ell(X)| &= \left|\prod_{j\in[n]}\left(1-z_\ell^j\right) - \prod_{j\in[n]}\left(1-c_\ell^j\right)\right| \\ &= \left|\prod_{j\in[n]}\left(1-c_\ell^j-\delta_\ell\right) - \prod_{j\in[n]}\left(1-c_\ell^j\right)\right| \le 3^n\sigma \le \frac{\epsilon}{3}.\end{aligned} \tag{3.16}$$

Hence, the gap between the attacker's utilities under Y and X can be bounded from above as follows.

$$|u^{\mathsf{A}}(Y,\ell) - u^{\mathsf{A}}(X,\ell)| \le (r^{\mathsf{A}}(\ell) - p^{\mathsf{A}}(\ell)) \cdot |\mathsf{cov}_\ell(Y) - \mathsf{cov}_\ell(X)| \le \frac{\epsilon}{3}. \tag{3.17}$$

Similarly, for the defenders' utilities, we have

$$|u_i^{\mathsf{D}}(Y,\ell) - u_i^{\mathsf{D}}(X,\ell)| \le (r_i^{\mathsf{D}}(\ell) - p_i^{\mathsf{D}}(\ell)) \cdot |\mathsf{cov}_\ell(Y) - \mathsf{cov}_\ell(X)| \le \frac{\epsilon}{3}. \tag{3.18}$$

Suppose now defender i makes a strategy deviation, playing $\mathbf{c}' \in \mathcal{C}_{k_i}$, instead. Let $X' = (\mathbf{c}^{-i}, \mathbf{c}')$ and $Y' = (\mathbf{z}^{-i}, \mathbf{c}')$ be the resulting strategy profiles when the other defenders play according to X and Y, respectively. We can bound the gap between the coverages similarly as in (3.16), hence obtaining the following.

$$|u^{\mathsf{A}}(Y',\ell) - u^{\mathsf{A}}(X',\ell)| \le \frac{\epsilon}{3} \tag{3.19}$$

$$|u_i^{\mathsf{D}}(Y',\ell) - u_i^{\mathsf{D}}(X',\ell)| \le \frac{\epsilon}{3} \tag{3.20}$$

(3.19) implies that $\ell \in \mathsf{BR}^{\epsilon}(Y')$ if $\ell \in \mathsf{BR}(X')$, i.e.,

$$u^{\mathsf{A}}(X,\ell) \ge u^{\mathsf{A}}(X,\ell') \text{ for all } \ell' \in T \quad \Longrightarrow \quad u^{\mathsf{A}}(Y,\ell) \ge u^{\mathsf{A}}(Y,\ell') - \epsilon \text{ for all } \ell' \in T.$$

In other words,

$$\mathsf{BR}(X') \subseteq \mathsf{BR}^{\epsilon}(Y'). \tag{3.21}$$

We consider the following cases and show that $\min_{\ell\in\mathsf{BR}^{\epsilon}(Y')} u_i^{\mathsf{D}}(Y',\ell) \le u_i^{\mathsf{D}}(Y,\ell^*) + \epsilon$ in all the cases to finish the proof (which then implies that (Y,ℓ^*) forms an ϵ-NSE by Definition 3.21).

Case 1. When $\mathsf{cov}_\ell(X) \le \mathsf{cov}_\ell(X')$ for all $\ell \in T$. This means that defender i does not increase her contribution to the coverage of every target ℓ after deviating, so $c^i_\ell \le c'_\ell$. Since $\mathbf{c}' \in \mathcal{C}_{k_i}$, it follows that $k_i \ge \sum_{\ell \in T} c'_\ell \ge \sum_{\ell \in T} c^i_\ell = k_i$, so the inequalities cannot be strict. Hence, $c'_\ell = c^i_\ell$ — essentially, the deviation results in defender i contributing the same coverage $\mathbf{c}^i$, so $\mathsf{cov}_\ell(X') = \mathsf{cov}_\ell(X)$. It follows that

$$\begin{aligned}|u^{\mathsf{A}}(Y', \ell) - u^{\mathsf{A}}(Y, \ell)| &\le |u^{\mathsf{A}}(Y', \ell) - u^{\mathsf{A}}(X', \ell)| + |u^{\mathsf{A}}(X', \ell) - u^{\mathsf{A}}(Y, \ell)| \\ &= |u^{\mathsf{A}}(Y', \ell) - u^{\mathsf{A}}(X', \ell)| + |u^{\mathsf{A}}(X, \ell) - u^{\mathsf{A}}(Y, \ell)| \\ &\le \frac{2\epsilon}{3},\end{aligned}$$

which implies that $\ell \in \mathsf{BR}^\epsilon(Y')$ if $\ell \in \mathsf{BR}^{\epsilon/3}(Y)$. Indeed, $\ell^* \in \mathsf{BR}^{\epsilon/3}(Y)$ by assumption (i), so $\ell^* \in \mathsf{BR}^\epsilon(Y')$. We have

$$\min_{\ell \in \mathsf{BR}^\epsilon(Y')} u^{\mathsf{D}}_i(Y', \ell) \le u^{\mathsf{D}}_i(Y', \ell^*) \le u^{\mathsf{D}}_i(Y, \ell^*) + \epsilon,$$

as desired.

Now we assume that $\mathsf{cov}_\ell(X) > \mathsf{cov}_\ell(X')$ for some ℓ in the remaining cases. By Lemma 3.8, this implies a coverage decrease for all $t \in \mathsf{BR}(X')$:

$$\mathsf{cov}_t(X) > \mathsf{cov}_t(X'). \tag{3.22}$$

Case 2. When there exists $t \in \mathsf{BR}(X')$, such that $t \preceq^X_i \ell^*$. Thus, $u^{\mathsf{D}}_i(X, t) \le u^{\mathsf{D}}_i(X, \ell^*)$ (see Definition 3.18). We can establish the following inequalities:

$$\begin{aligned}\min_{\ell \in \mathsf{BR}^\epsilon(Y')} u^{\mathsf{D}}_i(Y', \ell) &\le \min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(Y', \ell) && (\text{as } \mathsf{BR}(X') \subseteq \mathsf{BR}^\epsilon(Y') \text{ by (3.21)}) \\ &\le \min_{\ell \in \mathsf{BR}(X')} u^{\mathsf{D}}_i(X', \ell) + \epsilon/3 && (\text{by (3.20)}) \\ &\le u^{\mathsf{D}}_i(X', t) + \epsilon/3 && (\text{as } t \in \mathsf{BR}(X') \text{ in this case}) \\ &< u^{\mathsf{D}}_i(X, t) + \epsilon/3 && (\text{by (3.22)}) \\ &\le u^{\mathsf{D}}_i(X, \ell^*) + \epsilon/3 && (\text{as } t \preceq^X_i \ell^*) \\ &\le u^{\mathsf{D}}_i(Y, \ell^*) + 2\epsilon/3 && (\text{by (3.18)}) \\ &\le u^{\mathsf{D}}_i(Y, \ell^*) + \epsilon.\end{aligned}$$

Case 3. When $\ell^* \prec^X_i \ell$ for all $\ell \in \mathsf{BR}(X')$. We will argue that in this case, there exists $t \in \mathsf{BR}^\epsilon(Y')$ such that $t \preceq^X_i \ell^*$ and $\mathsf{cov}_t(X') \le \mathsf{cov}_t(X) + \frac{\epsilon}{3}$; this will imply

$$u^{\mathsf{D}}_i(X', t) \le u^{\mathsf{D}}_i\left(\mathsf{cov}_t(X) + \frac{\epsilon}{3},\ t\right) \le u^{\mathsf{D}}_i(X, t) + \frac{\epsilon}{3}$$

and establish the following.

$$\begin{aligned}\min_{\ell \in \mathsf{BR}^\epsilon(Y')} u^{\mathsf{D}}_i(Y', \ell) &\le u^{\mathsf{D}}_i(Y', t) \\ &\le u^{\mathsf{D}}_i(X', t) + \epsilon/3 && (\text{by (3.20)}) \\ &\le u^{\mathsf{D}}_i(X, t) + 2\epsilon/3 \\ &< u^{\mathsf{D}}_i(X, \ell^*) + 2\epsilon/3 && (\text{as } t \preceq^X_i \ell^*) \\ &\le u^{\mathsf{D}}_i(Y, \ell^*) + \epsilon. && (\text{by (3.18)})\end{aligned}$$

Claim. *In Case 3, there exists* $t \in \mathsf{BR}^\epsilon(Y')$ *such that* $t \preceq_i^X \ell^*$ *and* $\mathsf{cov}_t(X') \le \mathsf{cov}_t(X) + \frac{\epsilon}{3}$.

Pick some $t \preceq_i^X \ell^*$. We first argue that defender does not increase her coverage to t significantly after deviating to $\mathbf{c}'$, i.e., $c'_t \le c^i_t + \frac{\epsilon}{3}$, which results in $\mathsf{cov}_t(X') \le \mathsf{cov}_t(X) + \frac{\epsilon}{3}$.

Suppose for the sake of contradiction that $c'_t > c^i_t + \frac{\epsilon}{3}$. Since $c^i_\ell \le z^i_\ell + \sigma$ and $z^i_\ell = 0$ for all $\ell \succ_i^X \ell^*$ by the assumptions (ii) and (iii) of the lemma, we have $c^i_\ell \le z^i_\ell + \sigma \le \sigma$. Hence, $c'_\ell \ge 0 \ge c^i_\ell - \sigma$ for all $\ell \succ_i^X \ell^*$. It follows that,

$$\begin{aligned}
\sum_{\ell \in T} c'_\ell &= \sum_{\ell \preceq_i^X \ell^*} c'_\ell + \sum_{\ell \succ_i^X \ell^*} c'_\ell \\
&> \sum_{\ell \preceq_i^X \ell^*} \left(c^i_\ell + \epsilon/3\right) + \sum_{\ell \succ_i^X \ell^*} \left(c^i_\ell - \sigma\right) \\
&\ge \sum_{\ell \in T} c^i_\ell + \epsilon/3 - (n-1)\sigma \\
&> \sum_{\ell \in T} c^i_\ell \\
&= k_i,
\end{aligned}$$

which contradicts the fact that $\mathbf{c}' \in \mathcal{C}_{k_i}$. Therefore, $c'_t \le c^i_t + \frac{\epsilon}{3}$. The increment of the overall coverage of target t is at most that of the contribution of defender i, so we have

$$\mathsf{cov}_t(X') \le \mathsf{cov}_t(X) + \frac{\epsilon}{3}. \tag{3.23}$$

Let $j \in \arg\max_{\ell \in T} u^{\mathsf{A}}(Y', \ell)$; thus, $j \in \mathsf{BR}^\epsilon(Y')$ and $j \succ_i^X \ell^*$ by the assumption of Case 3. This further implies $z^i_j = 0$ by the assumption (iii) of the lemma. Thus, no matter which strategy the defender deviates to, the coverage of target j will not decrease; $\mathsf{cov}_j(Y') \ge \mathsf{cov}_j(Y)$ and hence

$$u^{\mathsf{A}}(Y', j) \le u^{\mathsf{A}}(Y, j). \tag{3.24}$$

If $\mathsf{cov}_t(X) > 0$, we have $t \in \mathsf{BR}(X)$ given that X is level, so $u^{\mathsf{A}}(X, t) \ge u^{\mathsf{A}}(X, j)$, in which case

$$\begin{aligned}
u^{\mathsf{A}}(Y', t) &\ge u^{\mathsf{A}}(X', t) - \epsilon/3 && \text{(by (3.19))} \\
&= u^{\mathsf{A}}(X, t) - (r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t))(\mathsf{cov}_t(X) - \mathsf{cov}_t(X')) - \epsilon/3 \\
&\ge u^{\mathsf{A}}(X, t) - 2\epsilon/3 && \text{(by (3.23))} \\
&\ge u^{\mathsf{A}}(X, j) - 2\epsilon/3 \\
&\ge u^{\mathsf{A}}(Y, j) - \epsilon && \text{(by (3.17))} \\
&\ge u^{\mathsf{A}}(Y', j) - \epsilon, && \text{(by (3.24))}
\end{aligned}$$

which implies $t \in \mathsf{BR}^\epsilon(Y')$.

Indeed, there must be a $t \preceq_i^X \ell^*$ such that $\mathsf{cov}_t(X) > 0$: otherwise, $c^i_\ell \le \mathsf{cov}_\ell(X) = 0$ for all $\ell \preceq_i^X \ell^*$; we would have $c^i_\ell \le \sigma$ for all $\ell \in T$ as for those $\ell \succ_i^X \ell^*$ we have $c^i_\ell \le z^i_\ell + \sigma = \sigma$ by the assumptions (ii) and (iii) of the lemma. It then follows that $k_i = \sum_{\ell \in T} c^i_\ell \le n \cdot \sigma < 1$, so $k_i = 0$ given that it is an integer, in which case defender i cannot make a strategy deviation in the first place. □

Now we are ready to present the proof of Lemma 3.27.

Proof of Lemma 3.27. We suppose that the input preference profile of Algorithm 3.1 is fixed to P throughout this proof. Let $\varepsilon = \epsilon \cdot \xi$. We consider the following two cases:

- **Case 1.** $\mathsf{surp}(H_{\min} + \varepsilon) > 0$;
- **Case 2.** $\mathsf{surp}(H_{\min} + \varepsilon) \leq 0$.

Since $\mathsf{surp}(h)$ is an increasing function of h, and by assumption $\mathsf{surp}(a) \leq 0$, it must be that $a < H_{\min} + \varepsilon$ in the first case, and $b \geq H_{\min} + \varepsilon$ in the second.

Case 1. Let $\hbar = \min\{b, H_{\min} + \varepsilon\}$; thus, $\hbar \in (a, b]$, and we have $\mathsf{surp}(\hbar) \geq 0$ given that $\mathsf{surp}(b) \geq 0$ and $\mathsf{surp}(H_{\min} + \varepsilon) > 0$. We will argue that in this case, on input $h = \hbar$, the output (X, t^*) of Algorithm 3.1 forms an ϵ-NSE. Since $\mathsf{surp}(\hbar) \geq 0$, by Lemma 3.13, X is level and $t^* \in \mathsf{BR}(X) \subseteq \mathsf{BR}^\epsilon(X)$.

Suppose now defender i makes a strategy deviation, playing $\mathbf{c}' \in \mathcal{C}_{k_i}$ instead. We show that (3.13) holds. Let $X' = (\mathbf{c}^{-i}, \mathbf{c}')$ be the resulting strategy profile. We have either (i) $\mathsf{cov}_t(X) > \mathsf{cov}_t(X')$ for some $t \in T$, or (ii) $\mathsf{cov}_\ell(X) \leq \mathsf{cov}_\ell(X')$ for all $\ell \in T$.

If it is (i), by Lemma 3.8, some target t with $\mathsf{cov}_t(X) > 0$ will appear in $\mathsf{BR}(X') \subseteq \mathsf{BR}^\epsilon(X')$, and after the deviation it holds that $\mathsf{cov}_t(X') < \mathsf{cov}_t(X)$, which implies

$$\min_{\ell \in \mathsf{BR}^\epsilon(X')} u_i^{\mathsf{D}}(X', \ell) \leq u_i^{\mathsf{D}}(X', t) < u_i^{\mathsf{D}}(X, t),$$

where the second inequality follows by monotonicity.

Suppose $P = (\preceq_1, \ldots, \preceq_n)$. Since now we have $\mathsf{cov}_t(X) > \mathsf{cov}_t(X')$, it must be that $c_t^i > 0$. Moreover, since P is the input preference profile, we have $t \preceq_i t^*$ by Lemma 3.14. By assumption, $\overline{\mathsf{cov}}(h)$ induces P for all $h \in (a, b)$, so we have

$$u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t) \leq u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t^*),$$

which by continuity must also hold for all $h \in [a, b]$. Now that $\mathsf{height}(X) = \hbar \in (a, b]$, we have

$$u_i^{\mathsf{D}}(X, t) \leq u_i^{\mathsf{D}}(X, t^*).$$

It follows that

$$\min_{\ell \in \mathsf{BR}^\epsilon(X')} u_i^{\mathsf{D}}(X', \ell) \leq u_i^{\mathsf{D}}(X, t^*),$$

so (3.13) holds.

If it is (ii), we have $c_\ell^i \leq c'_\ell$ for all $\ell \in T$; hence, $\sum_{\ell \in T} c_\ell^i \leq \sum_{\ell \in T} c'_\ell$. We can further assume that $\sum_{\ell \in T} c_\ell^i < k_i$. Indeed, if $\sum_{\ell \in T} c_\ell^i = k_i$, we must also have $\sum_{\ell \in T} c'_\ell = k_i$ as $\mathbf{c}'$ is also subject to the budget constraint; hence, $c_\ell^i = c'_\ell$ for all ℓ, which means nothing changes after the deviation and the defender cannot be better off.

Let τ_i be an arbitrary target in $\arg\min_{\ell \in \mathsf{BR}(\mathbf{1}^m)} u_i^{\mathsf{D}}(\mathbf{1}^m, \ell)$. Since $\tau_i \in \mathsf{BR}(\mathbf{1}^m)$, we have $u^{\mathsf{A}}(\mathbf{1}^m, \tau_i) = \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{1}^m, \ell)$. We have $u^{\mathsf{A}}(\mathbf{1}^m, \ell) = p^{\mathsf{A}}(\ell)$, i.e., the attacker gets a penalty when he attacks a target that has coverage is 1. It follows that

$$p^{\mathsf{A}}(\tau_i) = u^{\mathsf{A}}(\mathbf{1}^m, \tau_i) = \max_{\ell \in T} u^{\mathsf{A}}(\mathbf{1}^m, \ell) = \max_{\ell \in T} p^{\mathsf{A}}(\ell) \geq p^{\mathsf{A}}(\tau_i),$$

which means

$$p^{\mathsf{A}}(\tau_i) = \max_{\ell \in T} p^{\mathsf{A}}(\ell) = H_{\min}. \tag{3.25}$$

Thus, it must be that $\mathsf{cov}_{\tau_i}(X) > 0$. Indeed, if it was the case that $\mathsf{cov}_{\tau_i}(X) = 0$, we would have

$$u^{\mathsf{A}}(X, \tau_i) = u^{\mathsf{A}}(0, \tau_i) = r^{\mathsf{A}}(\tau_i) \geq p^{\mathsf{A}}(\tau_i) + \xi = H_{\min} + \xi > H_{\min} + \varepsilon,$$

so by monotonicity and (3.5), we have

$$\mathsf{cov}_{\tau_i}(X) < \frac{r^{\mathsf{A}}(\tau_i) - (H_{\min} + \varepsilon)}{r^{\mathsf{A}}(\tau_i) - p^{\mathsf{A}}(\tau_i)} \leq \overline{\mathsf{cov}}_{\tau_i}(H_{\min} + \varepsilon),$$

which contradicts Lemma 3.13 that the defender strategy profile produced by the algorithm should match the goal $\overline{\mathsf{cov}}(H_{\min} + \varepsilon)$ when there is a surplus.

Since X is level, $\mathsf{cov}_{\tau_i}(X) > 0$ then implies $\tau_i \in \mathsf{BR}(X)$ by Lemma 3.5. Again, by Lemma 3.13, $\mathsf{cov}(X) = \overline{\mathsf{cov}}(H_{\min} + \varepsilon)$, so $\tau_i \in \mathsf{BR}(X)$ implies that

$$\begin{aligned} u^{\mathsf{A}}(X, \tau_i) &= \max_{\ell \in T} u^{\mathsf{A}}(X, \ell) \\ &= \max_{\ell \in T} u^{\mathsf{A}}(\overline{\mathsf{cov}}(H_{\min} + \varepsilon), \ell) \\ &= H_{\min} + \varepsilon. && \text{(by Lemma 3.4)} \end{aligned} \tag{3.26}$$

Expanding the utility function and rearranging the above equality gives

$$\begin{aligned} \mathsf{cov}_{\tau_i}(X) &= \frac{r^{\mathsf{A}}(\tau_i) - (H_{\min} + \varepsilon)}{r^{\mathsf{A}}(\tau_i) - p^{\mathsf{A}}(\tau_i)} \\ &= \frac{r^{\mathsf{A}}(\tau_i) - p^{\mathsf{A}}(\tau_i) - \varepsilon}{r^{\mathsf{A}}(\tau_i) - p^{\mathsf{A}}(\tau_i)} && \text{(by (3.25))} \\ &\geq 1 - \varepsilon/\xi \\ &= 1 - \epsilon. \end{aligned} \tag{3.27}$$

Now back to Condition (b) $\mathsf{cov}_\ell(X) \leq \mathsf{cov}_\ell(X')$ for all $\ell \in T$, we have $u^{\mathsf{A}}(X, \ell) \geq u^{\mathsf{A}}(X', \ell)$ by monotonicity. Hence, for all $\ell \in T$,

$$\begin{aligned} u^{\mathsf{A}}(X', \tau_i) &\geq u^{\mathsf{A}}(\mathbf{1}^m, \tau_i) && \text{(as } \mathsf{cov}_{\tau_i}(X') \leq 1\text{)} \\ &= (H_{\min} + \varepsilon) - \varepsilon && \text{(by (3.25))} \\ &= u^{\mathsf{A}}(X, \tau_i) - \varepsilon && \text{(by (3.26))} \\ &\geq u^{\mathsf{A}}(X, \ell) - \varepsilon && \text{(as } \tau_i \in \mathsf{BR}(X)\text{)} \\ &\geq u^{\mathsf{A}}(X', \ell) - \varepsilon \\ &\geq u^{\mathsf{A}}(X', \ell) - \epsilon, \end{aligned}$$

which means that $\tau_i \in \mathsf{BR}^{\epsilon}(X')$.

Since $\sum_{\ell \in T} c_\ell^i < k_i$ and $\mathsf{cov}_{\tau_i}(X) > 0$ as we have argued above, by Lemma 3.15 we have $\tau_i \preceq_i^X t^*$; hence, $u_i^{\mathsf{D}}(X, \tau_i) \le u_i^{\mathsf{D}}(X, t^*)$. It follows that

$$\begin{aligned}
\min_{\ell \in \mathsf{BR}^\epsilon(X')} u_i^{\mathsf{D}}(X', \ell) &\le u_i^{\mathsf{D}}(X', \tau_i) \\
&\le u_i^{\mathsf{D}}(\mathbf{1}^m, \tau_i) \\
&= u_i^{\mathsf{D}}(X, \tau_i) + (r^{\mathsf{A}}(\tau_i) - p^{\mathsf{A}}(\tau_i))(1 - \mathsf{cov}_{\tau_i}(X)) \\
&\le u_i^{\mathsf{D}}(X, \tau_i) + \epsilon \qquad \text{(by (3.27))} \\
&\le u_i^{\mathsf{D}}(X, t^*) + \epsilon,
\end{aligned}$$

so (3.13) holds as well.

Case 2. We let $\overline{a} = \max\{a, H_{\min} + \varepsilon\}$. Since it holds that $b \ge H_{\min} + \varepsilon$ in this case, we have $a \le \overline{a} \le b$, and since $\mathsf{surp}(H_{\min} + \varepsilon) \le 0$, there exists a zero $h^* \in [\overline{a}, b]$ of $\mathsf{surp}(h)$. The output (X, t^*) of Algorithm 3.1 on input $h = h^*$ satisfies $t^* \in \mathsf{BR}(X)$ and conditions **C1** and **C2**. In addition, for all $\ell \in T$ we have $\ell \preceq_i t^*$ if $c_\ell^i > 0$ by Lemma 3.14, where $\mathbf{c}$ is the strategy of defender i in X and $\preceq_i$ is the preference of defender i in P. Since $\overline{\mathsf{cov}}(h)$ induces P for all $h \in (a, b)$, we have the following if $c_\ell^i > 0$:

$$u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), \ell) \le u_i^{\mathsf{D}}(\overline{\mathsf{cov}}(h), t^*),$$

which by continuity must also hold for all $h \in [a, b]$. Now that $\mathsf{height}(X) = h^* \in [\overline{a}, b] \subseteq [a, b]$, we have

$$u_i^{\mathsf{D}}(X, \ell) \le u_i^{\mathsf{D}}(X, t^*).$$

Thus, **C3** holds, and (X, t^*) is an NSE.

It remains to show that we can approximate (X, t^*) in polynomial time. We use binary search to find an approximation of h^*. We show that in polynomial time we can find a point $\tilde{h}$ that is sufficiently close to h^*, such that the output, say $(\tilde{X}, \tilde{t}^*)$, of Algorithm 3.1 on input $h = \tilde{h}$ satisfies all the conditions in Proposition 3.29 and forms an ϵ-NSE. Trivially, any output of Algorithm 3.1 satisfies the third condition according to Lemma 3.14, so we will argue for the first two conditions below. We start from the second condition.

Condition (ii). We show that after polynomially many iterations of binary search the approximation $\tilde{h}$ it generates satisfies Condition (ii). To this end, we show that for two values h_1 and h_2, the gap $|c_t^i(h_1) - c_t^i(h_2)|$ is at most exponentially large with respect to the gap $|h_1 - h_2|$; hence, binary search — that closes the gap $|\tilde{h} - h^*|$ at an exponential rate — is able to find a sufficiently good approximation in polynomial time.

Suppose $H_{\min} + \varepsilon \le h_1 \le h_2 \le H_{\max}$. By (3.5), for any $t \in T$, if $h_1 \ge r^{\mathsf{A}}(t)$, we have

$$|\overline{\mathsf{cov}}_t(h_1) - \overline{\mathsf{cov}}_t(h_2)| = 0 \le \frac{1}{\varepsilon} \cdot |h_1 - h_2|.$$

If $h_1 < r^{\mathsf{A}}(t)$, we have

$$\begin{aligned}
& r^{\mathsf{A}}(t) > h_1 \ge H_{\min} + \varepsilon \ge p^{\mathsf{A}}(t) + \varepsilon \\
\implies \quad & |\overline{\mathsf{cov}}_t(h_1) - \overline{\mathsf{cov}}_t(h_2)| \le \frac{r^{\mathsf{A}}(t)}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)} \cdot |h_1 - h_2| \le \frac{1}{\varepsilon} \cdot |h_1 - h_2|.
\end{aligned}$$

In both cases, $|\overline{\mathsf{cov}}_t(h_1) - \overline{\mathsf{cov}}_t(h_2)|$ is bounded by $\frac{1}{\varepsilon} \cdot |h_1 - h_2|$.

Now we view the variables in Algorithm 3.1 as functions of the input parameter h. For simplicity, we also let $z_t(h) = \mathsf{cov}_t(X(h))$ be the overall coverage of each target t; we will also derive a bound for z_t, which will be useful for our subsequent proof. Observe that z_t can be computed incrementally using $z_t \leftarrow z_t \oplus c_t^i$ in each iteration.

Before the first defender's iteration in Phase 1 of the algorithm, we have

$$|c_t^i(h_1) - c_t^i(h_2)| = 0$$

for all i and t, which also holds for z_t and $\mathcal{R}_i$. For $\mathcal{D}_t$, the above bounds on $\overline{\mathsf{cov}}_t(h)$ gives

$$|\mathcal{D}_t(h_1) - \mathcal{D}_t(h_2)| = |\overline{\mathsf{cov}}_t(h_1) - \overline{\mathsf{cov}}_t(h_2)| \leq \frac{1}{\varepsilon} \cdot |h_1 - h_2|.$$

Suppose that before the iteration of defender i in Phase 1, it holds for all the variables f in the algorithm that $|f(h_1) - f(h_2)| \leq \rho \cdot |h_1 - h_2|$ for some $\rho > 0$. We will use the following observation to bound the gap of the variables after this iteration. Suppose function $\Phi(x_1, \ldots, x_\lambda)$ is differentiable in the domain $\mathcal{X}$. Then it holds that

$$|\Phi(\mathbf{a}) - \Phi(\mathbf{b})| \leq \sum_{l \in [\lambda]} |a_l - b_l| \cdot \sup_{\mathbf{x} \in \mathcal{X}} \frac{\partial \Phi}{\partial f_l}(\mathbf{x}).$$

Applying the above observation, we obtain the following bounds:

$$|c_t^i(h_1) - c_t^i(h_2)| \leq |\mathcal{D}_t^*(h_1) - \mathcal{D}_t^*(h_2)| + |\mathcal{R}_i^*(h_1) - \mathcal{R}_i^*(h_2)| \leq 2\rho \cdot |h_1 - h_2|.$$

where f^* denotes the value of f before this iteration. Moreover,

$$|\mathcal{R}_i(h_1) - \mathcal{R}_i(h_2)| \leq |\mathcal{R}_i^*(h_1) - \mathcal{R}_i^*(h_2)| + |c_t^i(h_1) - c_t^i(h_2)| \leq 3\rho \cdot |h_1 - h_2|.$$

For $\mathcal{D}_t$, note that now that $h \geq H_{\min} + \varepsilon > H_{\min}$, we have $c_t^i \leq \overline{\mathsf{cov}}(h) < \overline{\mathsf{cov}}(H_{\min}) = 1$ throughout the algorithm, in which case $\mathcal{D}_t$ is effectively updated as $\mathcal{D}_t \leftarrow \mathcal{D}_t \ominus c_t^i = \frac{\mathcal{D}_t - c_t^i}{1 - c_t^i}$. Thus,

$$\begin{aligned}
|\mathcal{D}_t(h_1) - \mathcal{D}_t(h_2)| &\leq |\mathcal{D}_t^*(h_1) - \mathcal{D}_t^*(h_2)| \cdot \sup_{x \in [H_{\min}+\epsilon, H_{\max}]} \frac{\partial}{\partial \mathcal{D}_t^*} \left(\frac{\mathcal{D}_t^* - c_t^i}{1 - c_t^i} \right) \Bigg|_{h=x} + \\
&\qquad |c_t^i(h_1) - c_t^i(h_2)| \cdot \sup_{x \in [H_{\min}+\epsilon, H_{\max}]} \frac{\partial}{\partial c_t^i} \left(\frac{\mathcal{D}_t^* - c_t^i}{1 - c_t^i} \right) \Bigg|_{h=x} \\
&\leq \rho \cdot |h_1 - h_2| \cdot \sup_{h \in [H_{\min}+\epsilon, H_{\max}]} \left(\left| \frac{1}{1 - c_t^i(h)} \right| + 2 \left| \frac{\mathcal{D}_t(h) - 1}{(1 - c_t^i(h))^2} \right| \right) \\
&\leq \rho \cdot |h_1 - h_2| \cdot \left(\left| \frac{1}{\varepsilon} \right| + 2 \left| \frac{1}{\varepsilon^2} \right| \right) \\
&\leq \frac{3}{\varepsilon^2} \cdot \rho \cdot |h_1 - h_2|.
\end{aligned}$$

where to derive the third inequality we use the fact that $1 - c_t^i(h) \geq \epsilon$ which is due to the following that holds when $h \in [H_{\min} + \varepsilon, H_{\max}]$:

$$1 - c_t^i(h) \geq 1 - \overline{\mathsf{cov}}_t(h) \geq 1 - \overline{\mathsf{cov}}_t(H_{\min} + \varepsilon) \geq \frac{H_{\min} + \varepsilon - p^{\mathsf{A}}(t)}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)} \geq \frac{\varepsilon}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)} \geq \varepsilon.$$

Finally, for z_t, applying the updating rule $z_t \leftarrow z_t \oplus c_t^i$ we can obtain the following bound:

$$\begin{aligned}
|z_t(h_1) - z_t(h_2)| \le |z_t(h_1) - z_t(h_2)| \cdot \sup_{h\in[H_{\min}+\epsilon,H_{\max}]} \frac{\partial}{\partial z_t^*}\left(z_t^* \oplus c_t^i\right)\big|_{h=x} + \\
|c_t^i(h_1) - c_t^i(h_2)| \cdot \sup_{h\in[H_{\min}+\epsilon,H_{\max}]} \frac{\partial}{\partial c_t^i}\left(z_t^* \oplus c_t^i\right)\big|_{h=x} \\
\le \rho \cdot |h_1 - h_2| \cdot \left(\sup_{h\in[H_{\min}+\epsilon,H_{\max}]} (1 - c_t^i(h)) + \sup_{h\in[H_{\min}+\epsilon,H_{\max}]} (1 - z_t^*(h))\right) \\
\le 2\rho \cdot |h_1 - h_2|,
\end{aligned}$$

Overall, the bound of every variable blows up by a factor of at most $\frac{3}{\varepsilon^2}$.

The result can be demonstrated for each iteration of Phase 2. Thus, the value f of every variable at the end of the algorithm is such that $|f(h_1) - f(h_2)| \le \left(\frac{3}{\varepsilon^2}\right)^{2n+1} \cdot |h_1 - h_2|$. This suggests that, whenever $|\tilde{h} - h^*| \le \frac{\epsilon}{3^{n+1}} \cdot \left(\frac{\varepsilon^2}{3}\right)^{2n+1}$, we can ensure that $|f(\tilde{h}) - f(h^*)| \le \frac{\epsilon}{3^{n+1}}$; in particular, we have

$$|\tilde{c}_t^i - c_t^{*i}| = |c_t^i(\tilde{h}) - c_t^i(h^*)| \le \frac{\epsilon}{3^{n+1}}.$$

Indeed, using binary search and starting from $|\tilde{h} - h^*| \le |a - b| \le 1$, we need at most $\log \frac{1}{\frac{\epsilon}{3^{n+1}} \cdot \left(\frac{\varepsilon^2}{3}\right)^{2n+1}} = \mathsf{poly}(\log \frac{1}{\epsilon}, \log \frac{1}{\xi}, n)$ iterations to reach the point when $|f(\tilde{h}) - f(h^*)| \le \frac{\epsilon}{3^{n+1}}$, which is as desired.

Condition (i). It remains to show that the first condition in Proposition 3.29 is also satisfied, in particular, when $\tilde{h}$ is sufficiently close to h^* as above. Since in Algorithm 3.1 t^* is updated when a positive amount of coverage is distributed to a target, we have $\mathsf{cov}_{\tilde{t}^*}(\tilde{X}) > 0$. Hence, $\overline{\mathsf{cov}}_{\tilde{t}^*}(\tilde{h}) \ge \mathsf{cov}_{\tilde{t}^*}(\tilde{X}) > 0$, so by (3.5) we have $\overline{\mathsf{cov}}_{\tilde{t}^*}(\tilde{h}) = \frac{r^{\mathsf{A}}(\tilde{t}^*) - \tilde{h}}{r^{\mathsf{A}}(\tilde{t}^*) - p^{\mathsf{A}}(\tilde{t}^*)}$ and in turn $\mathsf{cov}_{\tilde{t}^*}(\tilde{X}) \le \frac{r^{\mathsf{A}}(\tilde{t}^*) - \tilde{h}}{r^{\mathsf{A}}(\tilde{t}^*) - p^{\mathsf{A}}(\tilde{t}^*)}$. Substituting this into the utility function gives

$$u^{\mathsf{A}}(\tilde{X}, \tilde{t}^*) \ge \tilde{h}.$$

Now suppose for the sake of contradiction that Condition (i) is not satisfied, so $\tilde{t}^* \notin \mathsf{BR}^{\epsilon/3}(\tilde{X})$, which means $u^{\mathsf{A}}(\tilde{X}, \tilde{t}^*) < \max_{\ell\in T} u^{\mathsf{A}}(\tilde{X}, \ell) - \frac{\epsilon}{3}$, and hence $\tilde{h} < \max_{\ell\in T} u^{\mathsf{A}}(\tilde{X}, \ell) - \frac{\epsilon}{3}$. Pick an arbitrary $j \in \arg\max_{\ell\in T} u^{\mathsf{A}}(\tilde{X}, \ell)$; we have,

$$u^{\mathsf{A}}(\tilde{X}, j) - \frac{\epsilon}{3} > \tilde{h}.$$

Recall that we have shown above that $|\mathsf{cov}_t(\tilde{X}) - \mathsf{cov}_t(X)| = |z_t(\tilde{h}) - z_t(h^*)| \le \frac{\epsilon}{3^{n+1}}$ for all t, so we have

$$|u^{\mathsf{A}}(\tilde{X}, j) - u^{\mathsf{A}}(X, j)| \le |\mathsf{cov}_t(\tilde{X}) - \mathsf{cov}_t(X)| \le \frac{\epsilon}{3^{n+1}}.$$

Combining these facts and the condition $|\tilde{h} - h^*| \le \frac{\epsilon}{3^{n+1}} \cdot \left(\frac{\varepsilon^2}{3}\right)^{2n+1}$, we obtain the following inequality.

$$\begin{aligned}
u^{\mathsf{A}}(X, j) &\ge u^{\mathsf{A}}(\tilde{X}, j) - \frac{\epsilon}{3^{n+1}} \\
&> \tilde{h} + \frac{\epsilon}{3} - \frac{\epsilon}{3^{n+1}} \\
&\ge h^* - \frac{\epsilon}{3^{n+1}} \cdot \left(\frac{\varepsilon^2}{3}\right)^{2n+1} + \frac{\epsilon}{3} - \frac{\epsilon}{3^{n+1}} \\
&> h^*.
\end{aligned}$$

Expanding $u^{\mathsf{A}}(X, j)$ and rearranging the terms then gives

$$\mathsf{cov}_j(X) < \frac{r^{\mathsf{A}}(j) - h^*}{r^{\mathsf{A}}(j) - p^{\mathsf{A}}(j)} \leq \overline{\mathsf{cov}}_j(h^*).$$

This contradicts the assumption that $\mathsf{surp}(h^*) = 0$, when we expect $\mathsf{cov}(X) = \overline{\mathsf{cov}}(h^*)$ according to Lemma 3.13. Therefore, the first condition must also be satisfied. □

With respect to the proof of Lemma 3.28, we omit the details and remark that the similar arguments as we use in the proof of Lemma 3.27 can be applied, and in polynomially many iterations binary search will lead us to an ϵ-NSE.

Chapter 4

Coordinating the Leaders: A Mechanism Design Approach

In this chapter, we explore the possibility of reducing inefficiency of uncoordinated defense by coordinating the defenders — specifically, by pooling the defenders' resources and allocating them jointly. We aim at designing a coordination mechanism to generate the joint allocation strategy for the defenders, that takes into account their individual preferences about how the targets should be protected. It is desired that the mechanism improves every defender's utility upon the uncoordinated situation we analyzed in the previous chapter, exploits efficiency gains, and incentivizes the defenders to report their true preferences and to execute the recommended strategies. Our analysis reveals that, unfortunately, even these basic properties clash with each other and no mechanism can achieve them simultaneously; we present several impossibility results. On the positive side, we design mechanisms that fulfill subsets of these properties, which match exactly our impossibility results and thereby depict a complete picture of what can and cannot be achieved through mechanism design with respect to the properties we consider.

4.1 Introduction

As we mentioned in the fictional story in Chapter 3, defense against Captain Jack could be enhanced through coordination. Now that we know how to analyze equilibria of the uncoordinated situation, let us take a look at some more concrete examples. Consider Example 3.10, if we coordinate the defenders and allocate their resources as a single defender, then the coverage $\mathbf{c} = (0.6, 0.9, 0.5)$ can be implemented since it satisfies the resource budget constraint $\mathbf{c} \in \mathcal{C}_2$ for this more powerful defender, who has two resources. This coverage dominates the coverage $\mathsf{cov}(Y) = (0.5, 0.8, 0.4)$ in the NSE (Y, t_2) we derived for Example 3.10. Hence, if we further induce the attacker to attack t_2, the outcome $(\mathbf{c}, t_2)$ improves both defenders' utilities upon what they obtain in an NSE. While the utility improvement is not significant in this example, in some other cases the improvement can be unbounded as we demonstrate in the following example.

Example 4.1 (Efficiency improvement via coordination)**.** Consider a variant of the game in Example 3.10, that has an additional target t_4. There are two defenders α and β and a set $T = \{t_1, t_2, t_3, t_4\}$ of four targets. Each defender has one resource, i.e., $k_\alpha = k_\beta = 1$. The payoffs of the players are given in the tables below, where M is a large number.

	t_1	t_2	t_3	t_4
r^{D}_α	3	7	9	M
p^{D}_α	2	4	8	0

defender α

	t_1	t_2	t_3	t_4
r^{D}_β	7	5	5	M
p^{D}_β	0	3	2	0

defender β

	t_1	t_2	t_3	t_4
r^{A}	6	5	5	0.9
p^{A}	-4	0	-5	0

attacker

The strategy profile (Y, t_2) we derive in Example 3.10 also forms an NSE in this game if we augment the defenders' strategies by letting $c^\alpha_{t_4} = c^\beta_{t_4} = 0$. In this NSE, defender α plays $\mathbf{c}^\alpha = (0.5, 0.5, 0, 0)$ and defender β plays $\mathbf{c}^\beta = (0, 0.6, 0.4, 0)$, resulting in overall coverage $\mathsf{cov}(Y) = (0.5, 0.8, 0.4, 0)$ and $u^{\mathsf{A}}(Y) = (1, 1, 1, M)$. It can be verified that it satisfies **C1**, **C2** (and hence, **C2***), and **C3**.

Now consider the outcome $(\mathbf{c}, t_4)$, where $\mathbf{c} = (0.6, 0.9, 0.49, 0.01)$ is a feasible coverage in $\mathcal{C}_K$ ($K = 2$). By calculating the attacker's utilities, we can also verify that $t_4 \in \mathsf{BR}(\mathbf{c})$, so $(\mathbf{c}, t_4)$ is indeed a valid outcome. When $(\mathbf{c}, t_4)$ is implemented, both defenders obtain utility $u^{\mathsf{D}}_\alpha(\mathbf{c}, t_4) = u^{\mathsf{D}}_\beta(\mathbf{c}, t_4) = 0.01M$. When $M \to \infty$, we have $\frac{u^{\mathsf{D}}_\alpha(\mathbf{c}, t_4)}{u^{\mathsf{D}}_\alpha(Y, t_2)} \to \infty$ (which holds even when we normalize the defenders' payoffs to be in $[0, 1]$).[1] □

In this chapter, we aim to design a coordination mechanism to generate joint defense strategies, that has the following basic properties: **(i) Efficiency**, that the allocation strategies generated should be *Pareto efficient*, so that no other allocation can improve the social welfare while not making any defender worse off; **(ii) Utility preservation**, that the mechanism should guarantee every defender at least their equilibrium utility in the uncoordinated situation, so that they indeed benefit from coordination; **(iii) Strategyproofness**, that the mechanism should incentivize the defenders to report their true preferences (valuations to the targets), which are held private. In some circumstances the mechanism can only recommend strategies for the defenders, but cannot force the defenders to actually use the recommended strategies. We consider an additional property: **(iv) Correlated stability**, that the recommended strategies form a correlated equilibrium in which no defender has an incentive to modify their strategy.

We show in this chapter that, unfortunately, it is impossible for any mechanism to achieve the above basic properties simultaneously. This theoretical barrier drives us to explore mechanisms that fulfill a subset of these properties or special problem settings that allow more of them to coexist. We present several mechanisms, including: a mechanism that achieves Pareto efficiency and utility guarantee with respect to the worst-case equilibrium utility in the uncoordinated situation; and a mechanism that achieves the first three of our properties when we are given a status-quo equilibrium and the defenders' preferences do not depend on how the targets are protected. To achieve correlated stability, we define a broader class of mechanisms that only recommend strategies for the defenders to take, without delegation of resources. We show that this additional property is incompatible with strategyproofness; moreover, even when strategyproofness is set aside, it cannot be achieved together with efficiency and utility guarantee.

Our results depicts a complete picture of what can and cannot be achieved, with respect to the properties we consider: for every combination of the properties, we show that they are either achieved by some mechanism we design, or cannot be achieved by any mechanism through

[1] In other words, the *Price of Anarchy (PoA)* of a multi-defender security game, as from the defenders' perspective, is unbounded. The PoA is defined as the ratio between the efficiency of a system under the optimal centralized solution and that in the worst equilibrium. It measures the degradation of the system efficiency due to agents' selfish behavior [Koutsoupias and Papadimitriou, 2009].

our impossibility results. On the other hand, the impossibility results indicate that voluntary coordination among defenders may be insufficient to achieve efficient resource utilization in security domains, hence calling for other innovative ways to promote defense coordination beyond the domain of mechanism design.

4.1.1 Related Work

The mechanism design approach we take to coordinate the defenders is related to a substantial literature on collective decision-making and mechanism design, which are among the key considerations in computational social choice [Brandt et al., 2016] and algorithmic game theory [Nisan et al., 2007]. There is a plethora of applications across various domains, to name but a few, auction and market design, elections, facility location, and fair division [e.g., Myerson, 1981, Roth, 2015, Bartholdi et al., 1989b,a, Miyagawa, 2001, Conitzer et al., 2007, Procaccia and Tennenholtz, 2013, Procaccia, 2013].

Conceptually, our work relates in paticular to collective decision-making for public resource allocation, see, e.g., [Bogomolnaia et al., 2005, Duddy, 2015, Aziz et al., 2019, Airiau et al., 2019]. We distribute security resources which form a divisible probability mass in the mixed strategy setting. Other than the specific utility structure that is realized through the SSG model, one feature of our problem that distinguishes it from prior work is that resources are contributed by the players, who need additional incentives to agree to cooperate in the first place. A similar phenomenon can be observed in the donor coordination problem studied by Brandl et al. [2020], in which philanthropists with different preferences pool their resources and jointly distribute them to charity projects. However, the underlying problem is very different and they do not use equilibria of the uncoordinated situation as benchmarks for designing mechanisms. A prominent feature of our model as a result of the SSG is that there is a strategic player, i.e., the attacker, who is not a participant of collective decision making.

Our approach to proving our impossibility results relies on the Gibbard–Satterthwaite theorem [Gibbard, 1973, Satterthwaite, 1975], which rules out the hope of designing incentive-compatible social-choice functions — a very important implication on social choice theory and mechanism design that drives social choice function or mechanism designers to focus on restricted classes of preferences or desiderata. We follow the same roadmap in designing the coordination mechanism.

4.1.2 Chapter Overview

We introduce the coordination framework and formalizing our desiderata in Section 4.2. In Section 4.3, we present our key impossibility results and then, in Section 4.4, we present two mechanisms that achieve a subset of our desired properties or all of them in a special circumstance. We consider the correlated stability property in Section 4.5 and present additional impossibility results and mechanisms with respect to this additional property. Finally, we summarize our results in Section 4.6.

4.2 Coordination Mechanism Design

To coordinate the resource allocation, the defenders delegate their resources to a coordination authority, which then chooses an allocation strategy to implement according to a *coordination mechanism* π, as if it is a single defender. Let $K = \sum_{i\in[n]} k_i$ be the total number of delegated resources. We call the tuple $\theta_i = (r_i^{\mathsf{D}}, p_i^{\mathsf{D}})$ of payoffs the *type* of defender i, and $\theta^{\mathsf{A}} = (r^{\mathsf{A}}, p^{\mathsf{A}})$

the type of the attacker. Let $\Theta = \left\{(r,p) \in \mathbb{R}^{2m} : r(t) > p(t) \text{ for all } t \in T\right\}$ denote the set of valid player types that are consistent with our assumption above. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ and $\mathbf{k} = (k_1, \ldots, k_n)$, so they encode the information of the defenders; we call $(\boldsymbol{\theta}, \mathbf{k})$ a *defender profile*.

The mechanism takes as input a defender profile $\boldsymbol{\theta}$ and an attacker type θ^{A}, and outputs an allocation strategy $\mathbf{c} \in \mathcal{C}_K$ of K resources. To deal with the tie-breaking issue, we assume that the mechanism explicitly specifies an attacker best response $t \in \mathsf{BR}(\mathbf{c})$ as part of the allocation strategy; such a t can be induced via a small perturbation in the allocation strategy. Thus, we require a coordination mechanism to output a tuple $(\mathbf{c}, t)$, called an *outcome*. An outcome is feasible if and only if $\mathbf{c} \in \mathcal{C}_K$ and $t \in \mathsf{BR}(\mathbf{c})$.

Our goal is to find a mechanism $\pi : (\boldsymbol{\theta}, \mathbf{k}, \theta^{\mathsf{A}}) \mapsto (\mathbf{c}, t)$ that satisfies the previously mentioned properties which we will formally define next.

4.2.1 Desired Properties

We formally define the desired properties we introduced in Section 3.1. First, we want the mechanism to exploit all possibilities of efficiency improvement and the outcome it produces to be *Pareto efficient*, so that no defender's utility can be further improved without making any other defender worse off.

Property 1 (Pareto Efficiency (PE)). An outcome $(\mathbf{c}, t)$ *Pareto dominates* another outcome $(\mathbf{c}', t')$ if $u_i^{\mathsf{D}}(\mathbf{c}, t) \geq u_i^{\mathsf{D}}(\mathbf{c}', t')$ for all $i \in [n]$ and $u_i^{\mathsf{D}}(\mathbf{c}, t) > u_i^{\mathsf{D}}(\mathbf{c}', t')$ for some $i \in [n]$. A mechanism π is *Pareto efficient* if for any $\mathbf{k} \in \mathbb{N}^n$, $\boldsymbol{\theta} \in \Theta^n$, and $\theta^{\mathsf{A}} \in \Theta$ no feasible outcome Pareto dominates $\pi(\boldsymbol{\theta}, \mathbf{k}, \theta^{\mathsf{A}})$.

For coordination to be meaningful, it is desired that every defender is better off than when they are uncoordinated. We use the NSE that we have analyzed in the previous sections as a benchmark and the next property *utility preservation* requires the mechanism to offer every defender at least their utility in the NSE. We define utility preservation in the strong and weak senses, which use the best and the worst NSE utilities as the benchmarks, respectively.

Property 2 (Utility Preservation (UP)). A mechanism π is *strongly* (resp. *weakly*) *utility preserving* if for any $\mathbf{k} \in \mathbb{N}^n$, $\boldsymbol{\theta} \in \Theta^n$, and $\theta^{\mathsf{A}} \in \Theta$ the utility $u_i^{\mathsf{D}}(\pi(\boldsymbol{\theta}, \mathbf{k}, \theta^{\mathsf{A}}))$ of each defender $i \in [n]$ is at least their utility in the best (resp. worst) NSE.

Finally, we would like our mechanism to be strategyproof and incentivize every defender to report their true payoffs.

Property 3 (Strategyproofness (SP)). A mechanism π is *not strategyproof* if for some $\mathbf{k} \in \mathbb{N}^n$, $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta^n$, $\theta^{\mathsf{A}} \in \Theta$, and $i \in [n]$, we have $\theta_j' = \theta_j$ for all $j \in [n] \setminus \{i\}$ and $u_i^{\mathsf{D}}(\pi(\boldsymbol{\theta}, \mathbf{k}, \theta^{\mathsf{A}})) < u_i^{\mathsf{D}}(\pi(\boldsymbol{\theta}', \mathbf{k}, \theta^{\mathsf{A}}))$, where each u_i^{D} is the utility function of defender type θ_i. Otherwise, π is strategyproof.

4.3 Impossibility Results

Unfortunately, it turns out that the basic properties defined so far already conflict with each other. We present a couple of impossibility results.

Strong UP is Impossible

We first show that strong UP cannot be satisfied on its own.

Theorem 4.2. *No mechanism is strongly UP.*

Proof. It suffices to present a game with no strongly UP outcome. Consider a game with two defenders α and β, and a target set $T = \{t_1, t_2\}$. Each defender has one resource, and both of them have non-overlapping payoffs captured by preference orders $t_1 \prec_\alpha t_2$ and $t_2 \prec_\beta t_1$. The targets are identical to the attacker, so the attacker's best response is always the target(s) with the lowest coverage.

Suppose that defender α always protects t_1 and defender β always protects t_2. The coverage contributed by the defenders is $\mathbf{c}^\alpha = (c^\alpha_{t_1}, c^\alpha_{t_2}) = (1, 0)$ and $\mathbf{c}^\beta = (c^\beta_{t_1}, c^\beta_{t_2}) = (0, 1)$, respectively. Overall, both targets are fully protected and are the attacker's best responses. Let the joint strategy be $X = (\mathbf{c}^\alpha, \mathbf{c}^\beta)$. It is easy to see that (X, t_2) is an NSE. Indeed, for α this is already the best outcome. Defender β prefers the attacker to choose t_1, but she cannot induce this attacker action: she can neither increase the coverage of t_2 to make it less attractive to the attacker nor decrease the coverage of t_1 to make it more attractive. Similarly, (X, t_1) is also an NSE, and it is the best for β.

Now suppose some outcome $(\mathbf{z}, t)$ is strongly UP. If $t = t_1$, we have $u_1^{\mathsf{D}}(\mathbf{z}, t) < u_1^{\mathsf{D}}(\mathbf{z}, t_2) \leq u_1^{\mathsf{D}}(X, t_2)$, so α would be worse off than in her most preferred NSE. By symmetry, if $t = t_2$, β would be worse off than in her most preferred NSE. Thus, no outcome of this game is strongly UP. □

SP and Weak UP are Incompatible

As strong UP turns out to be too demanding, we focus on weak UP. Trivially, every NSE (X, t) can be turned into an outcome $(\mathsf{cov}(X), t)$ (i.e., by simulating independent defender movement specified by X) that is a weakly UP by definition, so a weakly UP outcome always exists. Nevertheless, the weak UP property clashes with SP.

Theorem 4.3. *There exists no mechanism that is weakly UP and SP, even when the defenders' payoffs are restricted to non-overlapping payoffs.*

To prove Theorem 4.3, we show that for $|T| \geq 3$ a strategyproof coordination mechanism has to be a *dictatorship* — loosely speaking, it has to always output the favorite outcome of a distinguished defender (the dictator). While a dictatorship is a valid coordination mechanism, its output will not be acceptable to a defender whose preference conflicts with that of the dictator; indeed, under this mechanism her utility may be lower than in her worst NSE. Hereafter, we let L be the set of linear orders on T. Our argument proceeds by transforming a coordination mechanism into a function $f : L^n \to T$, i.e., a *social choice function*, and invoking the classic Gibbard–Satterthwaite theorem (Theorem 4.6), which states that, for more than two alternatives, any onto strategyproof social choice function is a dictatorship (also see Definitions 4.4 and 4.5).

Definition 4.4 (Strategyproof social choice function)**.** A social choice function $f : L^n \to T$ is *not strategyproof* if for some $\prec_1, \ldots, \prec_n \in L$, some $\prec'_i \in L$, and $i \in [n]$, we have $\tau \prec_i \tau'$ where $\tau = f(\prec_1, \ldots, \prec_i, \ldots, \prec_n)$ and $\tau' = f(\prec_1, \ldots, \prec'_i, \ldots, \prec_n)$. Otherwise, f is strategyproof.

Definition 4.5 (Dictatorship)**.** A voter $i \in [n]$ is a *dictator* in social choice function $f : L^n \to T$ if for all $\prec_1, \ldots, \prec_n \in L$, it holds that: $f(\prec_1, \ldots, \prec_n) = \tau^*$ if $\tau \prec_i \tau^*$ for all $\tau \in T \setminus \{\tau^*\}$. The function f is called a *dictatorship* if some $i \in [n]$ is a dictator in it.

Theorem 4.6 (Gibbard–Satterthwaite theorem [Gibbard, 1973, Satterthwaite, 1975]). *Let $f : L^n \to T$ be a strategyproof social choice function onto T. If $|T| \geq 3$, then f is a dictatorship.*

Let g be a function that maps a linear order $\prec \in L$ to a defender type $(r^{\mathsf{D}}, p^{\mathsf{D}}) \in \Theta$, where for each $t \in T$,

$$r^{\mathsf{D}}(t) = |\{\tau \in T : \tau \prec t\}| + 1/2 \quad \text{and} \quad p^{\mathsf{D}}(t) = |\{\tau \in T : \tau \prec t\}|.$$

Hence, a defender i of type $g(\prec)$ has non-overlapping payoffs captured by $\prec$: for any coverage $\mathbf{c}$ and $\mathbf{c}'$ it holds that $u_i^{\mathsf{D}}(\mathbf{c}, t) \leq r^{\mathsf{D}}(t) < p^{\mathsf{D}}(t') \leq u_i^{\mathsf{D}}(\mathbf{c}', t')$ if $t \prec t'$. Overloading this notation, given a preference profile $(\prec_1, \dots, \prec_n)$, we write $g(\prec_1, \dots, \prec_n) = (g(\prec_1), \dots, g(\prec_n))$.

Given a mechanism π, we construct a social choice function f_π: for any preference profile $(\prec_1, \dots, \prec_n) \in L^n$, we let $f_\pi(\prec_1, \dots, \prec_n) = t$ if $\pi(g(\prec_1, \dots, \prec_n), \mathbf{k}, \theta_*^{\mathsf{A}}) = (\mathbf{c}, t)$ for some $\mathbf{c}$, where

- $\mathbf{k}$ is a defender resource profile such that $k_1 = k_2 = 1$ and $k_i = 0$ for all $i = 3, \dots, n$ (hence, $K = 2$).
- $\theta_*^{\mathsf{A}} = (r^{\mathsf{A}}, p^{\mathsf{A}})$ is an attacker type such that $r^{\mathsf{A}}(t) = 0$ and $p^{\mathsf{A}}(t) = -1$ for all $t \in T$ (so all the targets are identical to θ_*^{A} and $u^{\mathsf{A}}(\mathbf{c}, t) = -c_t$ for any $\mathbf{c}$).

For the Gibbard–Satterthwaite theorem to be applicable, we need to argue that f_π is strategyproof, and it is *onto* T (i.e., for every $t \in T$ there exists some $(\prec_1, \dots, \prec_n) \in L^n$ such that $f(\prec_1, \dots, \prec_n) = t$).

Lemma 4.7. *If π is weakly UP then f_π is onto T.*

Proof. We will show that f_π satisfies a weak form of unanimity, i.e., for any $\tau \in T$ and any linear order $\prec$ such that $\ell \prec \tau$ for all $\ell \in T \setminus \{\tau\}$ it holds that $f_\pi(\prec, \dots, \prec) = \tau$. To this end, we consider a game where the defender profile is given by $(g(\prec, \dots, \prec), \mathbf{k})$ and the attacker type is θ_*^{A}.

We first argue that $t = \tau$ for every NSE (X, t) in this game. Suppose for the sake of contradiction that (X, t) is an NSE, $X = (\mathbf{c}^1, \dots, \mathbf{c}^n)$, but $t \neq \tau$. Since the attacker has type θ_*^{A}, by his payoffs we have $H_{\min} = \max_{t \in T} p^{\mathsf{A}}(t) = -1$; hence, $\overline{\mathsf{cov}}_t(H_{\min}) = \frac{r^{\mathsf{A}}(t) - H_{\min}}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)} = 1$ and $\sum_{t \in T} \overline{\mathsf{cov}}_t(H_{\min}) = m > K$. Using Corollary 3.7, we can find that (i) X is level, (ii) $\sum_{\ell \in T} c_\ell^i = k_i$ for all $i \in [n]$, and (iii) $c_\tau^i = 0$ for all $i \in [n]$ (we have $u_i^{\mathsf{D}}(X, t) < u_i^{\mathsf{D}}(X, \tau)$ as $t \prec_i \tau$ by our assumption).

The fact that $c_\tau^i = 0$ for all $i \in [n]$ implies that $\mathsf{cov}_\tau(X) = 0$. We can further argue that $\mathsf{cov}_\ell(X) = 0$ for all $\ell \in T$. Indeed, if $\mathsf{cov}_\ell(X) > 0$, the fact that X is level will suggest $\ell \in \mathsf{BR}(X)$ by Lemma 3.5, so $u^{\mathsf{A}}(X, \ell) \geq u^{\mathsf{A}}(X, \tau) = 0$ (as $\mathsf{cov}_\tau(X) = 0$); expanding $u^{\mathsf{A}}(X, \ell)$ and rearranging the terms gives $\mathsf{cov}_\ell(X) = 0$, which is a contradiction. However, if $\mathsf{cov}_\ell(X) = 0$ for all $\ell \in T$, we will have a contradiction with the fact (ii): we have $\sum_{\ell \in T} c_\ell^1 = k_1 > 0$, so some target must have a positive coverage.

Therefore, $t = \tau$. The assumption that π is weakly UP implies that

$$\pi\left(g(\prec, \dots, \prec), \mathbf{k}, \theta_*^{\mathsf{A}}\right) = (\mathbf{c}, \tau)$$

for some $\mathbf{c}$; hence, $f_\pi(\prec, \dots, \prec) = \tau$ by definition, which finishes the proof. Indeed, if the outcome is some tuple $(\mathbf{c}', \tau')$ and $\tau' \neq \tau$, we have $\tau' \prec \tau$ by our assumption, so it follows that $u_i^{\mathsf{D}}(\mathbf{c}', \tau') < u_i^{\mathsf{D}}(X, \tau) = u_i^{\mathsf{D}}(X, t)$, even when (X, t) is the worst NSE for defender i. This contradicts the assumption that π is weakly UP. □

Lemma 4.8. *If π is strategyproof then f_π is strategyproof.*

Proof. Suppose that f_π is not strategyproof. Then there exist linear orders $\succ_1, \ldots, \succ_n$ and $\succ_i'$ over T and alternatives $\tau, \tau' \in T$ such that $f(\succ_1, \ldots, \succ_i, \ldots, \succ_n) = \tau$, $f(\succ_1, \ldots, \succ_i', \ldots, \succ_n) = \tau'$, and $\tau \succ_i \tau'$, i.e., voter i benefits from reporting $\succ_i'$ instead of $\succ_i$.

Let $\boldsymbol{\theta} = g(\succ_1, \ldots, \succ_i, \ldots, \succ_n)$ and $\boldsymbol{\theta}' = g(\succ_1, \ldots, \succ_i', \ldots, \succ_n)$. Hence, $\theta_j = \theta_j'$ for all $j \in [n] \setminus \{i\}$. By the definition of f_π, we have $\pi\left(\boldsymbol{\theta}, \mathbf{k}, \theta_*^{\mathsf{A}}\right) = (\mathbf{c}, \tau)$ and $\pi\left(\boldsymbol{\theta}', \mathbf{k}, \theta_*^{\mathsf{A}}\right) = (\mathbf{c}', \tau')$ for some $\mathbf{c}$ and $\mathbf{c}'$. Consider a defender i of type θ_i. Since $\tau \succ_i \tau'$, we have

$$u_i^{\mathsf{D}}\left(\pi(\boldsymbol{\theta}, \mathbf{k}, \theta_*^{\mathsf{A}})\right) = u_i^{\mathsf{D}}(\mathbf{c}, \tau) < u_i^{\mathsf{D}}(\mathbf{c}', \tau') = u_i^{\mathsf{D}}\left(\pi(\boldsymbol{\theta}', \mathbf{k}, \theta_*^{\mathsf{A}})\right).$$

Thus, π is not strategyproof. □

We are now ready to prove Theorem 4.3.

Proof of Theorem 4.3. Suppose for the sake of contradiction that a mechanism π is SP and weakly UP. By Lemmas 4.7 and 4.8, the social choice function f_π is onto T and strategyproof. Thus, according to the Gibbard–Satterthwaite theorem, when $|T| \geq 3$, function f_π is a dictatorship. Assume that voter λ is the dictator and, without loss of generality, that $\lambda \neq 1$ (if $\lambda = 1$, we exchange voters 1 and 2 in the arguments below).

Now, consider linear orders $\succ_1', \ldots, \succ_n'$, in which

$$t_1 \succ_\lambda' t_2 \succ_\lambda' \cdots \succ_\lambda' t_m, \quad \text{and} \quad t_m \succ_1' t_{m-1} \succ_1' \cdots \succ_1' t_1.$$

Since λ is the dictator in f_π, we have $f_\pi(\succ_1', \ldots, \succ_n') = t_m$. By the definition of f_π, this implies that $\pi(g(\succ_1', \ldots, \succ_n'), \mathbf{k}, \theta_*^{\mathsf{A}}) = (\mathbf{c}, t_m)$ for some $\mathbf{c}$. Let $\boldsymbol{\theta} = g(\succ_1', \ldots, \succ_n')$. Thus, in the game given by defender profile $(\boldsymbol{\theta}, \mathbf{k})$ and attacker type θ_*^{A}, the mechanism offers defender 1 utility

$$u_1^{\mathsf{D}}(\pi(\boldsymbol{\theta}, \mathbf{k}, \theta_*^{\mathsf{A}})) = u_1^{\mathsf{D}}(\mathbf{c}, t_m) \leq r_1^{\mathsf{D}}(t_m) = 1/2.$$

However, we will show next that defender 1 can obtain utility at least 1 even in the worst NSE, which contradicts the assumption that π is weakly UP.

Observe that defender 1 can prevent t_m (her least preferred response of the attacker) from being adopted by the attacker by always allocating her resource to t_m: the resulting strategy profile $\tilde{X}$ gives $\mathsf{cov}_{t_m}(\tilde{X}) = 1$ and $\mathsf{cov}_t(\tilde{X}) < 1$ for at least one target $t \in T$. Thus, $u^{\mathsf{A}}(\tilde{X}, t_m) > u^{\mathsf{A}}(\tilde{X}, t)$, so $t_m \notin \mathsf{BR}(\tilde{X})$. Since $t_m \succ_1' t$ for all $t \in T \setminus \{t_m\}$, we have

$$\min_{t \in \mathsf{BR}(\tilde{X})} u_1^{\mathsf{D}}(\tilde{X}, t) \geq u_1^{\mathsf{D}}(\tilde{X}, t_{m-1}) \geq p_1^{\mathsf{D}}(t_{m-1}) = 1.$$

Therefore, a strategy profile cannot be an NSE if it does not offer defender 1 utility 1, as otherwise the defender can always allocate her resource to t_m to increase her utility. The worst NSE should offer defender 1 utility 1, which completes the proof. □

Since our desirable properties cannot be satisfied simultaneously, we can ask what is a maximal subset of them that can be satisfied at the same time. SP alone is easy to achieve, by appointing one defender (e.g., defender 1 or the defender who contributes the most resources) as a dictator and maximizing her utility, which also gives us PE if we fix a tie-breaking order; see Mechanism 4.1 and Theorem 4.9. Nevertheless, since some defenders may end up even less happy than in the worst NSE, this defeats the purpose of coordination. In the next section, we explore a class of mechanisms that can achieve PE and weak UP, as well as SP in a special circumstance.

Mechanism 4.1: DICT — a dictatorship.

Compute an outcome that maximizes the utility of defender 1. This is equivalent to computing an SSE in a game, where defender 1 is the only defender and has K resources. More specifically:

1. Solve the following linear program for each $t \in T$, where $\mathbf{c}$ are the variables.

$$\begin{aligned}
&\text{maximize} && u_1^{\mathsf{D}}(\mathbf{c}, t) && && (4.1)\\
&\text{subject to} && u^{\mathsf{A}}(\mathbf{c}, j) \le u^{\mathsf{A}}(\mathbf{c}, t) && \text{for all } j \in T && (4.1a)\\
& && \sum_{j \in T} c_j \le K && && (4.1b)\\
& && 0 \le c_j \le 1 && \text{for all } j \in T && (4.1c)
\end{aligned}$$

2. Pick $t^* \in T$ that leads to the maximum optimal value of the above LP amongst all $t \in T$; tie breaking in favor of defender 2 (and the next defender if it is still a tie). Let the corresponding optimal solution be $\mathbf{c}^*$. Output $(\mathbf{c}^*, t^*)$.

Theorem 4.9. DICT *is SP and PE.*

Proof. Suppose no tie is encountered throughout the implementation of Mechanism 4.1, then it is straightforward that it is SP and PE. It is SP since defender 1 has no incentive to misreport her payoffs as she already obtains the best possible outcome from the mechanism by truthful reporting; whereas the other defenders cannot change the outcome by misreporting as the mechanism does not use their payoff information. It is PE as all other outcomes gives defender 1 a lower utility.

When ties are encountered but only with respect to defender 1, the tie-breaking rule employed in the mechanism means that the arguments above now applies to defender 2, and inductively, they apply to all defenders. Hence, Mechanism 4.1 is SP and PE. □

4.4 Constrained Social Welfare Maximization Mechanisms

We propose a Constrained Social Welfare maXimization mechanism (CSWX) and show that it achieves PE and weak UP (but not SP). For the special case where all defenders have non-overlapping payoffs we show that when an NSE is provided, a variant of CSWX mechanism can achieve PE, Weak UP, and SP simultaneously. Both CSWX and its variant run in polynomial time.

4.4.1 A PE and Weakly UP Mechanism

CSWX is presented as Mechanism 4.2. In LP-t, Constraint (4.2a) requires each defender to obtain at least their utility in the NSE (X, τ); (4.1a) requires that the attacker is incentivized to attack target t; the remaining two constraints ensure that $\mathbf{c} \in \mathcal{C}_K$ is a feasible strategy. We will now argue that LP-t is feasible for at least one $t \in T$, so the mechanism will always output some outcome. Moreover, this outcome is PE and weakly UP.

Lemma 4.10. *LP-t is feasible for at least one $t \in T$.*

Mechanism 4.2: CSWX — constrained social welfare maximization.

1. Compute an NSE (X, τ), which can be done in polynomial time for any desired precision using our approach in Section 3.5.

2. Solve the following linear program LP-t for each $t \in T$, where $\mathbf{c}$ are the variables.

$$\text{LP-}t\text{:} \quad \text{maximize} \quad \sum_{i\in[n]} u_i^{\mathsf{D}}(\mathbf{c}, t) \tag{4.2}$$

$$\text{subject to} \quad u_i^{\mathsf{D}}(\mathbf{c}, t) \geq u_i^{\mathsf{D}}(X, \tau) \qquad \text{for all } i \in [n] \tag{4.2a}$$

$$\text{Constraints (4.1a)–(4.1c)}$$

3. Pick $t^* \in T$ such that the optimal value of LP-t^* is maximized amongst all LP-t (let ties be broken lexicographically).

   Output $(\mathbf{c}^*, t^*)$, where $\mathbf{c}^*$ is the optimal solution of LP-t^*.

Proof. In fact, LP-τ is always feasible as $\mathbf{z} = \mathsf{cov}(X)$ satisfies all the constraints. Specifically, (4.2a) is satisfied trivially when $\mathbf{z} = \mathsf{cov}(X)$. (4.1a) is satisfied because (X, τ) is an NSE, so by definition $\tau \in \mathsf{BR}(X) = \mathsf{BR}(\mathbf{z})$. Let $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$. We have

$$z_j = c_j^1 \oplus c_j^2 \oplus \cdots \oplus c_j^n \leq \sum_{i=1}^{n} c_j^i$$

for each $j \in T$, so summing up the inequalities over $j \in T$ gives

$$\sum_{j\in T} z_j \leq \sum_{j\in T} \sum_{i\in[n]} c_j^i \leq \sum_{i\in[n]} k_i = K,$$

so Constraint (4.1b) is satisfied. Finally, Constraint (4.1c) is satisfied as $0 \leq \mathsf{cov}_j(X) \leq 1$ for all $j \in T$ by definition. □

Theorem 4.11. CSWX *is PE and weakly UP.*

Proof. Let $(\mathbf{c}^*, t^*)$ be the outcome of CSWX. It is weakly UP since by Constraint (4.2a) every defender gets at least their utility in the NSE (X, τ).

Suppose for the sake of contradiction that $(\mathbf{c}^*, t^*)$ is not PE, and it is Pareto dominated by another outcome $(\mathbf{z}, \ell)$. We have $\sum_{i\in[n]} u_i^{\mathsf{D}}(\mathbf{z}, \ell) > \sum_{i\in[n]} u_i^{\mathsf{D}}(\mathbf{c}^*, t^*)$. Since $(\mathbf{c}^*, t^*)$ maximizes the optimal values of LP-t's that admit feasible solutions, $\mathbf{z}$ is not a feasible solution of LP-ℓ. However, since $(\mathbf{z}, \ell)$ is a feasible outcome, it satisfies (4.1a)–(4.1c), so the only constraint it can violate is (4.2a); we have $u_j^{\mathsf{D}}(\mathbf{z}, \ell) < u_j^{\mathsf{D}}(X, \tau)$ for some $j \in [n]$. Further, since $(\mathbf{c}^*, t^*)$ is a feasible solution of LP-t^*, we have $u_i^{\mathsf{D}}(\mathbf{c}^*, t^*) \geq u_i^{\mathsf{D}}(X, \tau)$ for all $i \in [n]$. Thus, $u_j^{\mathsf{D}}(\mathbf{z}, \ell) < u_j^{\mathsf{D}}(X, \tau) \leq u_j^{\mathsf{D}}(\mathbf{c}^*, t^*)$, which contradicts the assumption that $(\mathbf{z}, \ell)$ Pareto dominates $(\mathbf{c}^*, t^*)$. □

CSWX is nevertheless not strategyproof; we provide an example below (Example 4.12) to illustrate how a defender can manipulate this mechanism by misreporting her payoffs. To obtain even higher social welfare, we may relax the right-hand side of Constraint (4.2a) to the utility in the worst NSE for each defender i. However, this requires us to compute the worst NSE, which turns out to be an NP-hard problem (Proposition 4.13).

Example 4.12. There are two defenders α and β and four targets t_1, t_2, t_3, and t_4. The defenders each have one resource and have the following non-overlapping payoffs:

$$t_1 \prec_\alpha t_2 \prec_\alpha t_3 \prec_\alpha t_4$$
$$t_1 \prec_\beta t_2 \prec_\beta t_4 \prec_\beta t_3$$

For the attacker, these four targets are identical: $r^{\mathsf{A}}(t) = 0$ and $p^{\mathsf{A}}(t) = -1$ for all t, so his utility from attacking a target is exactly the negative of the coverage of this target, and his best response set contains targets with the lowest coverage.

Suppose CSWX uses the algorithm we present in Theorem 3.22 to generate a benchmark NSE. In this case, the algorithm will generate an NSE (X, t_3), where in X defender α plays $\mathbf{c}^\alpha = (0, 0, 1/2, 1/2)$ and β plays $\mathbf{c}^\beta = (1/2, 1/2, 0, 0)$. Specifically, now that the game has non-overlapping payoffs, there is no intersection point of the defenders' utility functions, so $\mathcal{H} = \{H_{\min}, H_{\max}\}$. Hence, we expect a zero h of surp in the interval $[H_{\min}, H_{\max}]$, which turns out to be $h = 1/2$. It can be verified that on input $h = 1/2$, Algorithm 3.1 will output exactly the strategy profile (X, t_3) as we describe above: during the execution of the algorithm, α fills the coverage of t_1 and then t_2 to $1/2$; then β fills t_4 and then t_3, so t^* (as in Line 12 of Algorithm 3.1) will be updated to t_3 in the end. Now that (X, t_3) is the benchmark in Constraint (4.2a), CSWX will eventually output an outcome $(\mathbf{c}^*, t_3)$ for some $\mathbf{c}^*$. Indeed, if it outputs $(\mathbf{c}^*, t)$ and $t \neq t_3$, (4.2a) cannot be satisfied for $i = \alpha$.

It follows that now defender α has an incentive to misreport her preference order as

$$t_3 \prec_\alpha t_1 \prec_\alpha t_2 \prec_\alpha t_4$$

This changes the order the defenders distribute coverage in Algorithm 3.1: now α first fills t_3 and then t_1; then β fills t_2 and then t_4, so the algorithm will output t_4 as the attacker's response in the NSE. Consequently, CSWX will output an outcome $(\mathbf{z}^*, t_4)$ for some $\mathbf{z}^*$ as otherwise (4.2a) cannot be satisfied for $i = \beta$. Defender α is better off because $u_\alpha^{\mathsf{D}}(\mathbf{z}^*, t_4) > u_\alpha^{\mathsf{D}}(\mathbf{c}^*, t_3)$ given that $t_3 \prec_\alpha t_4$ according to her true preference. □

Proposition 4.13. *Given a number μ, it is NP-hard to decide whether there exists an NSE that offers a particular defender at most utility μ.*

Proof. We provide a reduction from the classic NP-complete problem *exact cover by 3-sets* (X3C). Given a set $Q = \{e_1, \ldots, e_{3l}\}$ of $3l$ elements and a collection of subsets $\mathcal{S} = \{S_i\}_{i\in[\lambda]}$ of Q, $\lambda \geq l$, each of size 3, X3C asks if there exists an *exact cover*, i.e., a subcollection $\mathcal{S}'$ of $\mathcal{S}$ such that each element in Q is contained in exactly one subset in $\mathcal{S}'$.

Given an instance of X3C, we construct a game with $n = \lambda + 1$ defenders, and a set of targets $T = H \cup H' \cup \{t_*\}$ with $H = \{t_1, \cdots, t_{3l}\}$ and $H' = \{t'_1, \ldots, t'_{\lambda-l}\}$. Each defender has one resource except defender $\lambda + 1$ who has zero resource. Define $H_i = \{t_j \in H : e_j \in S_i\}$ and $H_{-i} = \{t_j \in H : e_j \notin S_i\}$ for each $i \in [\lambda]$. The payoffs are given in the tables below; the value in each entry applies to the corresponding target or to all targets in the corresponding subset. We will now argue that Q admits an exact cover if and only if there is an NSE in which defender $\lambda + 1$ obtains utility at most $\mu = 1$.

	H_i	H_{-i}	H'	t_*
r_i^{D}	6	10	2	8
p_i^{D}	0	9	1	7

each defender $i \in [\lambda]$

	$T \setminus \{t_*\}$	t_*
r_i^{D}	10	1
p_i^{D}	9	0

defender $\lambda + 1$

	H	H'	t_*
r^{A}	3	3	2
p^{A}	0	2	1

attacker

First, suppose there exists an exact cover. Let $I \subseteq [\lambda]$ be the indices of subsets in this exact cover, i.e., $\bigcup_{i \in I} S_i = S$. We have $|I| = l$. One can verify that the following strategy profile X and the attacker's response t_* form an NSE; it will give defender $\lambda + 1$ utility at most $r^{\mathsf{D}}_{\lambda+1}(t_*) = 1$. In X, each of the l defenders $i \in I$ protects the three targets in H_i, each with probability $1/3$, so $\mathsf{cov}_t(X) = \frac{1}{3}$ for all $t \in H$, as I is an exact cover; and each of the remaining $\lambda - l$ defenders protects a unique target in H' with probability 1, so $\mathsf{cov}_t(X) = 1$ for all $t \in H'$. No defender protects t_*, so $\mathsf{cov}_{t_*}(X) = 0$. Indeed, (X, t_*) satisfies the conditions in Proposition 3.9. Specifically, when X is played, the attacker gets utility 2 no matter which target he attacks. Hence, $\mathsf{BR}(X) = T$ and X is level; **C1** holds. In addition, **C2*** holds as the coverage every defender contributes equals their resource budget. **C3** holds because: (i) for each defender $i \in [\lambda]$, $c_t^i > 0$ only for targets $t \in H_i$, and we have $u_i^{\mathsf{D}}(X, t) = 2 \leq u_i^{\mathsf{D}}(X, \ell)$ for all $\ell \in T$; (ii) for defender $\lambda + 1$, $c_t^i = 0$ for all $t \in T$.

Conversely, suppose there does not exist an exact cover, we argue that the attacker's strategy must be t_* in every NSE; hence, defender $\lambda + 1$ obtains at least $\min_{t \in T \setminus \{t_*\}} p^{\mathsf{D}}_{\lambda+1}(t) = 9$ in every NSE. Suppose for the sake of contradiction that (X, t_*) is an NSE. By definition, $t_* \in \mathsf{BR}(X)$, so $u^{\mathsf{A}}(X, t_*) \geq u^{\mathsf{A}}(X, t)$ for all $t \in T$. According to the attacker's payoffs, this implies:

$$\mathsf{cov}_t(X) = 1 \quad \text{for all } t \in H'$$

and

$$\mathsf{cov}_t(X) \geq 1/3 \quad \text{for all } t \in H.$$

Thus,

$$\sum_{t \in T} \mathsf{cov}_t(X) \geq |H'| + \frac{1}{3}|H| = \lambda.$$

Let $\mathbf{c}^i$ be the coverage vector defender i plays in X. We have $\mathbf{c}^i \in \mathcal{C}_{k_i}$, so $\sum_{t \in T} c_t^i \leq k_i$ and summing this up over $i \in [n]$ gives $\sum_{i \in [n]} \sum_{t \in T} c_t^i \leq \sum_{i \in [n]} k_i = \lambda$. Hence,

$$\sum_{i \in [n]} \sum_{t \in T} c_t^i \leq \sum_{t \in T} \mathsf{cov}_t(X) = \sum_{t \in T} \left(c_t^1 \oplus \cdots \oplus c_t^n \right).$$

For this inequality to hold, it must be that for each $t \in T$: $c_t^i > 0$ for at most one $i \in [n]$; otherwise, the left side will be strictly larger than the right side. In other words, every target is protected by at most one defender.

Now that $\mathsf{cov}_t(X) = 1$ for each $t \in H'$, there are $\lambda - l$ defenders each protect a distinct target in H' with probability 1. The remaining l defenders who do have resources protect targets in H; given that $\mathsf{cov}_t(X) = 1/3$ for all $t \in H$, each of them protects three targets. If each defender i among these l defenders only protects targets in H_i, that would make an exact cover, which is not possible by our assumption. Thus, some defender i must be protecting a target $t \in H_{-i}$. In this case, however, defender i would be better off not protecting this target: this will induce the attacker to attack t and increase the utility of this defender. Therefore, (X, t_*) cannot be an NSE. The proof is completed. □

4.4.2 NSE-induced SP Mechanism

Sometimes a (truthful) NSE is known before the defenders decide to collaborate, e.g., when the status quo is an NSE (indeed, since an NSE in our model can be computed efficiently, it is plausible to assume that one may arise from interaction between the players). If in this NSE every target is protected with some probability but not probability 1 and every defender has non-overlapping payoffs, we can achieve SP, PE, and weak UP simultaneously. Intuitively, the known NSE spares us the trouble of dealing with non-truthful input information when we compute the NSE in Step 1 of CSWX. Given also the defenders' consistent preferences with non-overlapping payoffs, we are able to generate a PE and weakly UP outcome without any additional information.

Mechanism 4.3: CSWX$^+$ — constrained social welfare maximization with a given NSE.

1. Given an NSE (X, τ), solve LP-τ defined in (4.2). Let the optimal solution obtained be $\mathbf{c}^*$.
2. Output $(\mathbf{c}^*, \tau)$.

Let (X, τ) be a known NSE such that $0 < \mathsf{cov}_t(X) < 1$ for all $t \in T$. Mechanism 4.3 is a simple variant of CSWX. The argument in Lemma 4.10 shows that LP-τ is always feasible. We will show that CSWX$^+$ is SP, PE, and weakly UP.

Theorem 4.14. CSWX$^+$ *is SP, PE, and weakly UP when the defenders have non-overlapping payoffs.*

Proof. Observe that $u_i^{\mathsf{D}}(\mathbf{c}, \tau)$ depends only on c_τ and is increasing with respect to c_τ. Thus, LP-τ is equivalent to the following LP with variables $\mathbf{c}$:

$$\begin{array}{lll} \text{maximize} & c_\tau & \\ \text{subject to} & c_\tau \geq \mathsf{cov}_\tau(X) & \text{for all } i \in [n] \\ & \text{Constraints (4.1a)–(4.1c)} & \end{array}$$

The formulation does not rely on the defenders' payoffs. Thus, misreporting does not change the output and CSWX$^+$ is SP.

To show that it is PE, suppose for the sake of contradiction that $(\mathbf{c}^*, \tau)$ is Pareto dominated by another outcome $(\mathbf{z}, \ell)$. If $\ell = \tau$, we have $u_i^{\mathsf{D}}(\mathbf{z}, \tau) \geq u_i^{\mathsf{D}}(\mathbf{c}^*, \tau)$ for all $i \in [n]$, so $\mathbf{z}$ is a feasible solution to LP-τ; and $\sum_{i\in[n]} u_i^{\mathsf{D}}(\mathbf{z}, \tau) > \sum_{i\in[n]} u_i^{\mathsf{D}}(\mathbf{c}^*, \tau)$, so $\mathbf{z}$ is actually a better solution than $\mathbf{c}^*$, contradicting the fact that $\mathbf{c}^*$ is an optimal solution of LP-τ.

Therefore, $\ell \neq \tau$. Since $\mathsf{cov}_t(X) < 1$ for all $t \in T$, we have $u^{\mathsf{A}}(X, t) > p^{\mathsf{A}}(t)$; hence, $\mathsf{height}(X) = \max_{t\in T} u^{\mathsf{A}}(X, t) > \max_{t\in T} p^{\mathsf{A}}(t) = H_{\min}$. Thus, by Corollary 3.7 (iii) $c_t^i > 0$ only if $t \preceq_i \tau$. Indeed, now that $\mathsf{cov}_\ell(X) > 0$ by assumption, we have $c_\ell^i > 0$ for some defender i, so $\ell \prec_i \tau$ as now $\ell \neq \tau$. This implies $u_i^{\mathsf{D}}(\mathbf{z}, \ell) < u_i^{\mathsf{D}}(\mathbf{c}^*, \tau)$ and contradicts the assumption that $(\mathbf{z}, \ell)$ Pareto dominates $(\mathbf{c}^*, \tau)$.

Finally, weak UP is guaranteed by Constraint (4.2a) of LP-τ. □

4.5 Stability against Strategy Modification

So far, we assumed that the mechanism was in charge of the defenders' resources. In some scenarios the mechanism does not have direct access to the resources; it can only *recommend* allocation strategies to the defenders and cannot enforce that the defenders will indeed follow the recommendations. In this section, we consider an additional property: *correlated stability* (CS). This property requires that no defender has an incentive to deviate from the strategy recommended by the coordination mechanism.

4.5.1 Recommendation Strategies and CS

In allocation strategies considered so far, resources are treated as anonymous; we do not care about the ownership of resources in the allocation. We will now extend the definition of allocation strategies to incorporate the ownership information, and call the newly defined strategies *recommendation strategies*. In this new model, a joint pure strategy $S = (s^1, \dots, s^n)$ contains a pure allocation strategy $s^i \in \mathcal{T}_{k_i}$ for each defender i. We have $\mathsf{cov}_t(S) = 1$ if and only if $\mathsf{cov}_t(s^1) = 1$ for some $i \in [n]$. A recommendation strategy $\mathbf{x}$ is then a distribution over joint pure strategies (Definition 4.15). A mechanism generates an outcome $(\mathbf{x}, t)$ conditioned on the input player profile as previously defined, and to implement the outcome, the mechanism samples a joint pure strategy $(s^1, \dots, s^n)$ from $\mathbf{x}$ and recommends each defender i to take strategy s^i. The recommendation is privately sent to each defender, and only the distribution $\mathbf{x}$ is publicly known to everyone. We define *strategy modification* to formalize how a defender can deviate from the recommended strategy.

Definition 4.15 (Recommendation strategy)**.** A *recommendation strategy* $\mathbf{x}$ is a distribution over joint pure strategies, i.e., $\mathbf{x} \in \Delta(\mathcal{T}_{\mathbf{k}})$, where $\mathcal{T}_{\mathbf{k}} = \prod_{i \in [n]} \mathcal{T}_{k_i}$.

Definition 4.16 (Strategy modification)**.** A strategy modification of a defender $i \in [n]$ is a function $\phi : \mathcal{T}_{k_i} \to \Delta(\mathcal{T}_{k_i})$, under which the defender plays a mixed strategy $\phi(s)$ whenever she is instructed to play a pure strategy $s \in \mathcal{T}_{k_i}$ by the coordination mechanism.

For a distribution $\mathbf{x} \in \Delta(\mathcal{T}_{\mathbf{k}})$, we abuse the notation and let $\phi(\mathbf{x})$ denote the distribution that arises when defender i adopts ϕ. By $\mathbf{x}' = \phi(\mathbf{x})$, the probability that each joint pure strategy $S = (s^1, \dots, s^n) \in \mathcal{T}_{\mathbf{k}}$ is chosen is

$$x'_S = \sum_{Q=(q^1,\dots,q^n) \in \mathcal{T}_{\mathbf{k}}} x_Q \cdot \phi_{s^i}(q^i),$$

where $\phi_{s^i}(q^i)$ is the probability s^i is chosen by the mixed strategy $\phi(q^i)$. We are now ready to define correlated stability.

Property 4 (Correlated Stability (CS)). An outcome $(\mathbf{x}, t)$ is *correlated stable* if for every defender $i \in [n]$ it holds that

$$u_i^{\mathsf{D}}(\mathbf{x}, t) \geq \min_{t' \in \mathsf{BR}(\phi(\mathbf{x}))} u_i^{\mathsf{D}}(\phi(\mathbf{x}), t')$$

for every strategy modification $\phi : \mathcal{T}_{k_i} \to \Delta(\mathcal{T}_{k_i})$. A coordination mechanism is CS if only outputs CS outcomes.

Put differently, CS requires that the outcome to form a correlated equilibrium among the defenders and, indeed, every NSE is by definition also an CS outcome. Given our results in the previous sections, we could hope to find a PE and weakly UP mechanism that is also CS, or a mechanism that is SP, PE, and CS. Unfortunately, as we will show next, such mechanisms do not exist.

4.5.2 Impossibility Results with CS

We present two impossibility results with CS.

PE, Weak UP, and CS are Incompatible

First, we show that PE, weak UP, and CS are incompatible by describing a game in which no outcome satisfies these properties.

Example 4.17. There are three defenders and five targets $T = \{a, b, c, d, e\}$. Each defender has one resource. The defenders have non-overlapping payoffs captured by the following preference orders:

$$a \prec_1 b \prec_1 c \prec_1 d \prec_1 e;$$
$$c \prec_2 b \prec_2 a \prec_2 d \prec_2 e;$$
$$e \prec_3 d \prec_3 c \prec_3 b \prec_3 a.$$

The attacker's payoffs are as follows.

	a	b	c	d	e
r^{A}	1	1	1	0.7	0.7
p^{A}	0	0	0	0	0

□

We make two observations about the above exmaple in Lemmas 4.18 and 4.19, with which we prove Theorem 4.20 by showing that no outcome is PE, weakly UP, and CS in this example.

Lemma 4.18. *In Example 4.17 the attacker's best response is to attack target b in every NSE.*

Proof. First, since a, c, and e are the least favored responses of the attacker for some defender, the attacker's response cannot be any of these targets in an NSE: otherwise the defender for whom this target is the least favored response of the attacker can deviate to always allocating her resource to this target; the target will have coverage 1 and cannot be an attacker best response in this example.

We will now argue that target d cannot be the attacker's best response either. We show that defender 3 can always avoid targets d and e to be the attacker's best responses by playing the strategy $\mathbf{c}^3 = (0, 0, 0, \frac{1}{2}, \frac{1}{2})$, irrespective of the other defenders' strategies. When $\mathbf{c}^3$ is played, the coverage of both targets d and e will be at least 0.5. The attacker obtains utility at most $0.7 \times 0.5 + 0 \times 0.5 = 0.35$ by attacking any of them. Thus, for either of them to be a best response of the attacker, each of a, b, and c needs to receive coverage at least 0.65 (for the attacker to get utility at least 0.35 by attacking them); this is impossible. Essentially, let $\mathbf{c}^1$ and $\mathbf{c}^2$ be the first two

defenders' strategy, so each target $t \in \{a, b, c\}$ has coverage $1 - (1 - c_t^1)(1 - c_t^2)$; we show that the following constraints cannot be satisfied simultaneously to complete the proof.

$$\begin{cases} 1 - (1 - c_t^1)(1 - c_t^2) \geq 0.65, & \text{for all } t \in \{a, b, c\} \\ 0 \leq c_t^i \leq 1, & \text{for all } i = 1, 2, \text{ and } t \in \{a, b, c\} \\ c_a^i + c_b^i + c_c^i \leq k_i = 1, & \text{for all } i = 1, 2 \end{cases}$$

Indeed, suppose that the above constraints are satisfied by some c_t^i's. For simplicity, let

$$(x_1, x_2, x_3) = (1 - c_a^1, 1 - c_b^1, 1 - c_c^1)$$
$$(y_1, y_2, y_3) = (1 - c_a^2, 1 - c_b^2, 1 - c_c^2)$$

Then the following constraints should be satisfied.

$$\begin{cases} x_j \cdot y_j \leq 0.35, & \text{for all } j = 1, 2, 3 \\ 0 \leq x_j \leq 1, \text{ and } 0 \leq y_j \leq 1, & \text{for all } j = 1, 2, 3 \\ x_1 + x_2 + x_3 \geq 2, \text{ and } y_1 + y_2 + y_3 \geq 2 \end{cases}$$

Assume without loss of generality that $x_1 \leq x_2 \leq x_3$. In fact, if the inequalities are satisfiable, there is a satisfying solution such that $y_1 \geq y_2 \geq y_3$ because if, e.g., $y_1 \leq y_2$, we can exchange the values of y_1 and y_2 while the inequalities will still be satisfied. Further, assume without loss of generality that $x_1 \cdot y_3 \leq x_3 \cdot y_1$. Let $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) = (x_1 - \delta,\ x_2,\ x_3 + \delta)$ and $(\tilde{y}_1, \tilde{y}_2, \tilde{y}_3) = (y_1 + \frac{\delta \cdot y_3}{x_3 + \delta},\ y_2,\ y_3 - \frac{\delta \cdot y_3}{x_3 + \delta})$. We have

$$\begin{aligned} &\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 = x_1 + x_2 + x_3 \geq 2, \\ &\tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 = y_1 + y_2 + y_3 \geq 2, \\ &\tilde{x}_3 \cdot \tilde{y}_3 = (x_3 + \delta) \cdot (y_3 - \frac{\delta}{x_3 + \delta} \cdot y_3) = x_3 \cdot y_3 \leq 0.35, \\ &\tilde{x}_2 \cdot \tilde{y}_2 = x_2 \cdot y_2 \leq 0.35, \\ &\tilde{x}_1 \cdot \tilde{y}_1 = (x_1 - \delta) \cdot (y_1 + \frac{\delta}{x_3 + \delta} \cdot y_3) \\ &\qquad = x_1 y_1 + \frac{\delta(x_1 \cdot y_3 - y_1 \cdot x_3 - y_1 \cdot \delta - y_3 \cdot \delta)}{x_3 + \delta} \leq x_1 y_1 \leq 0.35. \end{aligned}$$

However, if we gradually increase δ, $0 \leq x_j \leq 1$ and $0 \leq y_j \leq 1$ will be violated. Consider the first constraint that becomes tight.

- Suppose that $\tilde{x}_1$ drops to 0 first. We have $\tilde{x}_1 = 0$ while all the constraints are satisfied. By $\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 \geq 2$ and $\tilde{x}_2, \tilde{x}_3 \in [0, 1]$ we have $\tilde{x}_2 = \tilde{x}_3 = 1$. By $\tilde{x}_j \cdot \tilde{y}_j \leq 0.35$ we have $\tilde{y}_2 \leq 0.35$ and $\tilde{y}_3 \leq 0.35$. It follows that, $\tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 \leq 1 + 0.35 + 0.35 < 2$, which is a contradiction.
- Suppose that $\tilde{x}_3$ reaches 1 first. We have $\tilde{x}_3 = 1$ while all the constraints are satisfied. It follows by $\tilde{x}_3 \cdot \tilde{y}_3 \leq 0.35$ that $\tilde{y}_3 \leq 0.35$ and, by $\tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3 = 2$ and $\tilde{y}_1 \geq \tilde{y}_2$, that $\tilde{y}_1 \geq 0.825$. Applying these constraints repeatedly, we have

$$\begin{aligned} & \tilde{x}_1 \leq \tfrac{0.35}{\tilde{y}_1} \leq \tfrac{0.35}{0.825} < 0.425 \\ \Rightarrow\quad & \tilde{x}_2 \geq 2 - \tilde{x}_1 - \tilde{x}_3 \geq 2 - 0.425 - 1 = 0.575 \\ \Rightarrow\quad & \tilde{y}_2 \leq \tfrac{0.35}{\tilde{x}_2} \leq \tfrac{0.35}{0.575} < 0.61 \\ \Rightarrow\quad & \tilde{y}_1 = 2 - \tilde{y}_2 - \tilde{y}_3 > 2 - 0.61 - 0.35 > 1, \end{aligned}$$

which is a contradiction.

- When $\tilde{y}_1$ reaches 1 first or $\tilde{y}_3$ drops to 0 first, the same analysis can be applied. Both cases lead to contradictions as above.

Thus, no x_j's and y_j's can satisfy the constraints. □

Lemma 4.19. *Suppose that $(\mathbf{x}, \tau)$ is a PE outcome in Example 4.17, and $\mathbf{c} = \mathsf{cov}(\mathbf{x})$. Then $c_a = c_b = c_c = \frac{24}{31}$ and $c_d = c_e = \frac{21}{31}$.*

Proof. We first argue that $\mathbf{c}$ indeed corresponds to a feasible outcome. Note that in this example, $K = 3 = \sum_{t \in T} c_t$ by the above valuation of $\mathbf{c}$, so $\mathbf{c} \in \mathcal{C}_K$ is a feasible strategy. Trivially, any coverage vector that can be implemented by some distribution $\mathbf{x} \in \Delta(\mathcal{T}_K)$ can also be implemented by a distribution in $\mathbf{x} \in \Delta(\mathcal{T}_{\mathbf{k}})$ over the non-anonymous joint pure strategy $\mathcal{T}_{\mathbf{k}}$, so $\mathbf{c}$ is feasible.

Suppose for the sake of contradiction that $(\mathbf{x}', \tau')$, $\tau' \in \mathsf{BR}(\mathbf{c}')$, is a PE outcome, but $\mathbf{c}' = \mathsf{cov}(\mathbf{x}') \neq \mathbf{c}$. We have $\sum_{t \in T} c'_t \leq K = \sum_{t \in T} c_t$, which implies that $c'_\ell < c_\ell$ for some $\ell \in T$ as otherwise $\mathbf{c}' = \mathbf{c}$. Taking $\mathbf{c}$ to the utility function, we find that $u^{\mathsf{A}}(\mathbf{c}, t) = \frac{7}{31}$ for all $t \in T$, so $\mathsf{BR}(\mathbf{c}) = T$. Since $u^{\mathsf{A}}(\mathbf{c}, t)$ decreases with c_t, we have

$$\max_{t \in T} u^{\mathsf{A}}(\mathbf{c}', t) \geq u^{\mathsf{A}}(\mathbf{c}', \ell) > u^{\mathsf{A}}(\mathbf{c}, \ell) = \frac{7}{31}.$$

By assumption $\tau' \in \mathsf{BR}(\mathbf{c}')$, so

$$u^{\mathsf{A}}(\mathbf{c}', \tau') = \max_{t \in T} u^{\mathsf{A}}(\mathbf{c}', t) > \frac{7}{31} = u^{\mathsf{A}}(\mathbf{c}, \tau').$$

On the other hand, the utility $u_i^{\mathsf{D}}(\mathbf{c}, t)$ of each defender i increase with c_t, so $u^{\mathsf{A}}(\mathbf{c}', \tau') > u^{\mathsf{A}}(\mathbf{c}, \tau')$ implies $u_i^{\mathsf{D}}(\mathbf{c}', \tau') < u_i^{\mathsf{D}}(\mathbf{c}, \tau')$ for all $i \in [n]$, which means that now there is another outcome $(\mathbf{x}, \tau')$ that Pareto dominates $(\mathbf{x}', \tau')$, contradicting the assumption that $(\mathbf{x}', \tau')$ is PE. Indeed, here $(\mathbf{x}, \tau')$ is a feasible outcome since $\tau' \in \mathsf{BR}(\mathbf{c}) = T$. □

Theorem 4.20. *There exists no mechanism that is PE, weakly UP, and CS, even when the defenders' payoffs are non-overlapping.*

Proof. It suffices to show that Example 4.17 does not admit a PE, weakly UP, and CS outcome.

First, by Lemma 4.18 and the fact that the defenders have non-overlapping utilities, if an outcome $(\mathbf{x}, \tau)$ is weakly UP then we must have $\tau = b$. By Lemma 4.19, $\mathsf{BR}(\mathbf{c}) = T$. This implies that no defender i will receive a recommendation to protect any target $t \succ_i b$ with positive probability, i.e., for all $S = (s^1, s^2, s^3) \in \mathcal{T}_{\mathbf{k}}$ such that $x_S > 0$, $\mathsf{cov}_t(s^i) = 0$ if $t \succ_i b$ (otherwise, the defender can remove the resource allocated to this target to induce a better response of the attacker). Hence, targets d and e will only be protected by defender 3 who has only one resource. We have $\mathsf{cov}_e(S) + \mathsf{cov}_d(S) \leq 1$ in every pure joint strategy in the support set of $\mathbf{x}$; it follows that

$$\mathsf{cov}_e(\mathbf{x}) + \mathsf{cov}_d(\mathbf{x}) = \sum_{S \in \mathcal{T}_{\mathbf{k}} : x_S > 0} x_S \cdot (\mathsf{cov}_e(S) + \mathsf{cov}_d(S)) \leq \sum_{S \in \mathcal{T}_{\mathbf{k}}} x_S = 1.$$

This contradicts Lemma 4.19 where we have $c_d + c_e = \frac{42}{31} > 1$. □

Mechanism 4.4: NSE — a mechanism that outputs an NSE.

1. Compute an NSE (X, τ), where $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$.
2. Convert each $\mathbf{c}^i$ to a distribution $\mathbf{x}^i$ over $\mathcal{T}_{k_i}$ using a standard approach (See Section 2.2), whereby the support set $\mathcal{S}^i = \left\{s \in \mathcal{T}_{k_i} : x^i_s > 0\right\}$ is of size $O(m)$.
3. Output $(\mathbf{x}, \tau)$, where $\mathbf{x}$ is such that $x_S = \prod_{i\in[n]} x^i_{s^i}$ for each joint pure strategy $S = (s^1, \ldots, s^n) \in \prod_{i\in[n]} \mathcal{S}^i$.[2]

SP and CS are Incompatible

If we drop weakly UP and consider SP, PE, and CS, the next result, similar to Theorem 4.3, show that even SP and CS are incompatible.

Theorem 4.21. *There exists no mechanism that is SP and CS, even when the defenders' payoffs are non-overlapping.*

Proof. Suppose for the sake of contradiction that a mechanism π is SP and CS. We can construct the same social choice function f_π and replicate the arguments for proving Lemmas 4.7 and 4.8 to show that f_π is onto T and strategyproof. Hence, f_π is a dictatorship when $|T| \geq 3$.

Specifically, to see that f_π is onto T, we argue that any CS outcome for the game with attacker type θ^{A}_* and defender profile $(g(\prec, \ldots, \prec), \mathbf{k})$ must be of the form $(\mathbf{x}, \tau)$, where $\ell \prec \tau$ for all $\ell \in T \setminus \{\tau\}$ (see the definitions of θ^{A}_*, g, and $\mathbf{k}$ in Section 4.3). We can show that $\mathbf{x}$ must result in a level coverage $\mathbf{c}$. Indeed, if this was not the case, defender 1 or 2 would be able to increase the coverage of a least protected target t by switching her resource from a better protected target t' to t with some probability (the fact that t' is better protected implies that there are pure strategies in the support of $\mathbf{x}$, in which t' is protected but t is not). Since $\mathbf{c}$ is level (and in this game this means $c_t < 1$ for all $t \in T$), if a CS outcome was $(\mathbf{x}, t)$, $t \neq \tau$, then unless $c_t = 0$ for all t some defender would be incentivized to reduce c_τ to make τ the only best response of the attacker (e.g., by doing nothing whenever she is recommended to allocated a resource to τ); whereas in the other case when $c_t = 0$ for all t, defender 1 can trivially distribute some resources to all targets but τ to achieve the same. Thus, all CS outcomes must be in the form $(\mathbf{x}, \tau)$.

Next, we can use the argument in the proof of Theorem 4.3 to finish the proof: Suppose voter $\lambda \neq 1$ is the dictator in f_π; An SP mechanism would generate outcome of the form $(\mathbf{x}, t_m)$, which would only incentivize defender 1 to modify her strategy to, e.g., always protecting t_m to prevent t_m from being the best response of the attacker. Thus, $(\mathbf{x}, t_m)$ cannot be CS, and no mechanism can be SP and CS at the same time. □

4.5.3 CS Mechanisms

The above impossibility results force us to select a strict subset of the conflicting properties again. If we drop SP and PE, essentially we will be looking at mechanisms that generate correlated equilibria. Trivially, every NSE, as a special correlated equilibrium, is weakly UP and CS, and we know that

[2] Note that the support set of x_S may be exponentially large with the explicit representation above. Alternatively, we can just use the set of distributions $\mathbf{x}^1, \ldots, \mathbf{x}^n$ to represent $\mathbf{x}$, by which a joint pure strategy can be sampled efficiently by independently sampling pure strategies from each $\mathbf{x}^i$.

Mechanism 4.5: ACD — additive coverage distribution.

1. Compute an SSE $(\mathbf{z}^*, \ell^*)$ in a game where defender 1 is the only defender and has K resources (same as Mechanism 4.1). Let $\overline{\mathbf{z}} = \overline{\mathsf{cov}}(u^{\mathsf{A}}(\mathbf{z}^*, \ell^*))$.[3]

   Let $(\prec_1, \ldots, \prec_n)$ be the preference profile induced by $\overline{\mathbf{z}}$, i.e., for all $i \in [n]$, we have $t \prec_i \ell$ if $u_i^{\mathsf{D}}(\overline{\mathbf{z}}, t) < u_i^{\mathsf{D}}(\overline{\mathbf{z}}, \ell)$ (or $u_i^{\mathsf{D}}(\overline{\mathbf{z}}, t) = u_i^{\mathsf{D}}(\overline{\mathbf{z}}, \ell)$ and $t < \ell$).

2. Run Algorithm 3.1 on input $\prec_1, \ldots, \prec_n$ and $h = \mathsf{height}(\overline{\mathbf{z}})$, where:

   - We change the operations $\oplus$ and $\ominus$ to $+$ and $-$, respectively.
   - Let the output of the algorithm be $X = (\mathbf{c}^1, \ldots, \mathbf{c}^n)$ and t^*. If there exists $\ell \in T$, such that $(\overline{\mathbf{z}}, \ell)$ Pareto dominates $(\overline{\mathbf{z}}, t^*)$, then update t^* to ℓ; repeat until no other $(\overline{\mathbf{z}}, \ell)$ Pareto dominates $(\overline{\mathbf{z}}, t^*)$.

3. Convert $(\mathbf{c}^1, \ldots, \mathbf{c}^n)$ to a recommendation strategy $\mathbf{x} \in \Delta(\mathcal{T}_{\mathbf{k}})$, such that the following equations hold (see Lemma 4.22).

$$\sum_{S=(s^1,\ldots,s^n)\in\mathcal{T}_{\mathbf{k}}} x_S \cdot \mathsf{cov}_t(s^i) = c_t^i \qquad \text{for all } i \in [n] \tag{4.3}$$

$$\sum_{S\in\mathcal{T}_{\mathbf{k}}} x_S \cdot \mathsf{cov}_t(S) = \overline{z}_t \qquad \text{for all } t \in T \tag{4.4}$$

   (In other words, each defender i protects t with probability c_t^i overall, and each t is protected with probability $\overline{z}_t$ overall.)

   Output $(\mathbf{x}, t^*)$.

they always exist and can be computed efficiently; see Mechanism 4.4. However, in this way we do not obtain any benefits from coordination, which violates our original motivation in some sense.

If we drop SP and UP, the following mechanism ACD (Additive Coverage Distribution; Mechanism 4.5) is PE and CS. The idea is to mimic the way we generate an NSE by Algorithm 3.1, where we let each defender distribute coverage according to the importance order of the targets. To achieve PE, we need the coverage to reach maximum as in DICT (Mechanism 4.1), which means that we need to avoid any waste of resources, which occurs whenever more than one of them are allocated to the same target in a pure strategy. Consequently, the additive sum of the coverage every defender contributes to a target must be exactly the overall coverage this target gets, so that when we distribute coverage according to Algorithm 3.1, the subadditive operation $\oplus$ operation we use to calculate the overall coverage needs to be replaced with the addition operation. The key to this mechanism is to convert the coverage we obtained from additive coverage distribution (which represents each defender's marginal contribution) to an equivalent recommendation strategy. We show that this is possible via Lemma 4.22 and, finally, via Theorem 4.23 we show that ACD is PE and CS.

[3] This operation is used to ensure that when there are multiple SSEs that correspond to different coverage vectors (which can happen when the coverage of some target reaches 1), we always pick an SSE with a level coverage. There is only one such level coverage. This further ensures that the preference profile generated next results in a CS outcome in the subsequent steps.

Lemma 4.22. *A recommendation strategy* $\mathbf{x}$ *satisfying* (4.3) *and* (4.4) *always exists and can be computed in polynomial time.*

Proof. In Step 2 of Mechanism 4.5, the defenders distribute resources with the aim of constructing $\overline{\mathbf{z}}$. We must have $\overline{z}_t \leq z_t^*$ for all $t \in T$. Specifically, this holds trivially if $\overline{z}_t = 0$. If $\overline{z}_t > 0$, then we have $\overline{z}_t = \frac{r^{\mathsf{A}}(t) - u^{\mathsf{A}}(\mathbf{z}^*, \ell^*)}{r^{\mathsf{A}}(t) - p^{\mathsf{A}}(t)}$ by (3.5), which means $u^{\mathsf{A}}(\overline{\mathbf{z}}, t) = u^{\mathsf{A}}(\mathbf{z}^*, \ell^*) \geq u^{\mathsf{A}}(\mathbf{z}^*, t)$, where the last inequality holds since $(\mathbf{z}^*, \ell^*)$ is an SSE; hence, $\overline{z}_t \leq z_t^*$ by monotonicity of the utility function.

It follows that $\sum_{t \in T} \overline{z}_t \leq \sum_{t \in T} z_t^* \leq K$. Now that the defenders' contributions to the coverage is additive, this means that they have enough resources to construct $\overline{\mathbf{z}}$; hence, $\mathsf{surp} \geq 0$, and using arguments similar to Lemma 3.13, we have $\sum_{i \in [n]} \mathbf{c}^i = \mathsf{cov}(X) = \overline{\mathsf{cov}}(h) = \overline{\mathbf{z}}$, where $\mathsf{cov}(X) = \sum_{i \in [n]} \mathbf{c}^i$ as now the $\oplus$ operation is replaced with $+$.

Next, we show that we can compute $\mathbf{x}$ in polynomial time and that it satisfies (4.3) and (4.4). Let $\widetilde{T} = T \cup \{0\}$ be the set of targets including an additional dummy target 0. Let $c_0^i = k_i - \sum_{t \in T} c_t^i$ and $z_0^* = \sum_{i \in [n]} c_0^i$. Consider a flow network with a source node v^+ and a sink node v^-. In addition, there is a node v_i for each $i \in [n]$, and a node v_t for each $t \in \tilde{T}$. The nodes are connected in the following way:

- There is an edge (v^+, v_i) with capacity k_i for each $i \in [n]$.
- For each $i \in [n]$ and $t \in T$, there is an edge (v_i, v_t) with capacity 1 if and only if $c_t^i > 0$. In addition, there is an edge (v_i, v_0) with capacity k_i for every $i \in [n]$.
- There is an edge (v_t, v^-) with capacity 1 for every $t \in T$, and an edge (v_0, v^-) with capacity K.

It is easy to see that flow $\mathbf{f}$ is a maximum flow between v^+ and v^-, where

- $f_{(v^+, v_i)} = k_i$ for each $i \in [n]$.
- $f_{(v_i, v_t)} = c_t^i$ for each $i \in [n]$ and $t \in \widetilde{T}$.
- $f_{(v_t, v^-)} = z_t^*$ for each $t \in \widetilde{T}$.

Since the network has integral capacities, a useful result (see e.g., [Khandekar et al., 2012, Lemma 2.5]) is that flow $\mathbf{f}$ can be decomposed to a convex combination $\mathbf{f} = \sum_{j=1}^{\lambda} y_j \cdot \mathbf{p}^j$ of a set of integral maximum flows $\mathbf{p}^1, \ldots, \mathbf{p}^\lambda$ between v^+ and v^-; moreover, the decomposition can be computed in polynomial time, with the obtained λ being polynomial in the size of the network. We show next that we can obtain a recommendation strategy $\mathbf{x}$ from the coefficients $\mathbf{y} = (y_1, \ldots, y_\lambda)$, which satisfies (4.3) and (4.4) .

First, for each of the integral flow $\mathbf{p}^j$, $j = 1, \ldots, \lambda$, we construct a joint pure strategy $S = (s^1, \ldots, s^n)$, such that $\mathsf{cov}_t(s^i) = p^j_{(v_i, v_t)}$ (we have $p^j_{(v_i, v_t)} \in \{0, 1\}$ as (v_i, v_t) has capacity 1). Indeed, S also satisfies the capacity constraints as we have

$$\sum_{t \in T} \mathsf{cov}_t(s^i) = \sum_{t \in T} p^j_{(v_i, v_t)} \leq p^j_{(v^+, v_i)} \leq k_i$$

for all $i \in [n]$, given that each edge (v^+, v_i) has capacity k_i. In addition, we also have

$$\mathsf{cov}_t(S) = \min\left\{1, \textstyle\sum_{i \in [n]} \mathsf{cov}_t(s^i)\right\} = \min\left\{1, \textstyle\sum_{i \in [n]} p^j_{(v_i, v_t)}\right\} = p^j_{(v_t, v^-)},$$

where the last equality holds as (v_t, v^-) has capacity 1 and $\mathbf{p}^j$ is integral.

We assign probability $x_S = y_j$ to each strategy S constructed above. For all other strategies $S' \in \mathcal{T}_\mathbf{k}$ that are not constructed with any flow $\mathbf{p}^j$, we let $x_{S'} = 0$. Hence, $\mathbf{x} = (x_S)_{S \in \mathcal{T}_\mathbf{k}}$ is a distribution over $\mathcal{T}_\mathbf{k}$. We now have

$$\sum_{S \in \mathcal{T}_\mathbf{k}} x_S \cdot \mathsf{cov}_t(s^i) = \sum_{j=1}^{\lambda} x_j \cdot p^j_{(v_i, v_t)} = f_{(v_i, v_t)} = c^i_t$$

for each $i \in [n]$ and $t \in T$; and

$$\sum_{S \in \mathcal{T}_\mathbf{k}} x_S \cdot \mathsf{cov}_t(S) = \sum_{j=1}^{\lambda} x_j \cdot p^j_{(v_t, v^-)} = f_{(v_t, v^-)} = \overline{z}_t$$

for each $t \in T$. Hence, (4.3) and (4.4) hold, which completes the proof. □

Theorem 4.23. ACD *is PE and CS.*

Proof. We show that the outcome $(\mathbf{x}, t^*)$ ACD outputs is PE and CS. The reason that it is CS is similar to the reason that an outcome generated by Algorithm 3.1 is an NSE: if any defender deviates, that would only make a target they protect a best response of the attacker, hence making the defender worse off.

To see that $(\mathbf{x}, t^*)$ is PE, suppose for the sake of contradiction that it is Pareto dominated by another outcome $(\mathbf{y}, \tau)$, where we assume that $\mathbf{y}$ is represented as a coverage vector. Note that $\mathbf{x}$ results in coverage $\overline{\mathbf{z}}$ because of (4.4), so $(\mathbf{x}, t^*)$ is equivalent to $(\overline{\mathbf{z}}, t^*)$ in terms of the utilities they yield. If $y_\tau < \overline{z}_\tau$, then for all defenders $i \in [n]$, we have $u^\mathsf{D}_i(\mathbf{y}, \tau) < u^\mathsf{D}_i(\mathbf{x}, \tau)$. In particular, for those i with $c^i_\tau > 0$, we also have $\tau \preceq_i t^*$ (similar to Lemma 3.14), and hence, $u^\mathsf{D}_i(\mathbf{y}, \tau) < u^\mathsf{D}_i(\mathbf{x}, \tau) \leq u^\mathsf{D}_i(\mathbf{x}, t^*)$, contradicting the assumption that $(\mathbf{y}, \tau)$ Pareto dominates $(\mathbf{x}, t^*)$.

Thus, the only possibility is that $y_\tau \geq \overline{z}_\tau$. If $y_\tau = \overline{z}_\tau$, then we have $u^\mathsf{D}_i(\mathbf{y}, \tau) = u^\mathsf{D}_i(\overline{\mathbf{z}}, \tau)$ for all $i \in [n]$, so $(\overline{\mathbf{z}}, \tau)$ must also Pareto dominate $(\overline{\mathbf{z}}, t^*)$, which contradicts the operation in Step 2 of Mechanism 4.5. If $y_\tau > \overline{z}_\tau$, we then have

$$u^\mathsf{A}(\mathbf{y}, \ell^*) \leq u^\mathsf{A}(\mathbf{y}, \tau) < u^\mathsf{A}(\overline{\mathbf{z}}, \tau) \leq \max_{t \in T} u^\mathsf{A}(\overline{\mathbf{z}}, t) = \mathsf{height}(\overline{\mathbf{z}}) = u^\mathsf{A}(\mathbf{z}^*, \ell^*),$$

where $u^\mathsf{A}(\mathbf{y}, \ell^*) \leq u^\mathsf{A}(\mathbf{y}, \tau)$ since $\tau \in \mathsf{BR}(\mathbf{y})$ for $(\mathbf{y}, \tau)$ to be a feasible outcome; $u^\mathsf{A}(\overline{\mathbf{z}}, t) = \mathsf{height}(\overline{\mathbf{z}})$ by definition (see (3.6)); and $\mathsf{height}(\overline{\mathbf{z}}) = u^\mathsf{A}(\mathbf{z}^*, \ell^*)$ by Lemma 3.4 and the way $\overline{\mathbf{z}}$ is defined in Step 1 of the mechanism.

It then follows that $u^\mathsf{A}(\mathbf{y}, \ell^*) \leq u^\mathsf{A}(\mathbf{y}, \tau) < u^\mathsf{A}(\mathbf{z}^*, \ell^*)$, which implies that $y_{\ell^*} > z^*_{\ell^*}$ by monotonicity of the utility function. This means that if we reduce the coverage of ℓ^* in $\mathbf{y}$ by an appropriate amount, we would obtain a coverage $\tilde{\mathbf{y}}$, such that $u^\mathsf{A}(\mathbf{y}, \tau) < u^\mathsf{A}(\tilde{\mathbf{y}}, \ell^*) < u^\mathsf{A}(\mathbf{z}^*, \ell^*)$. Since the coverage of the other targets is not changed and τ is the attacker's best response against $\mathbf{y}$, now ℓ^* would become the only best response of the attacker, while $\tilde{y}_{\ell^*} > z^*_{\ell^*}$. It follows that for defender 1 we have

$$\max_{t \in \mathsf{BR}(\tilde{\mathbf{y}})} u^\mathsf{D}_1(\tilde{\mathbf{y}}, t) = u^\mathsf{D}_1(\tilde{\mathbf{y}}, \ell^*) > u^\mathsf{D}_1(\mathbf{z}^*, \ell^*),$$

which contradicts that fact that $(\mathbf{z}^*, \ell^*)$ is an SSE as defined in Step 1 of the mechanism. □

	DICT	CSWX	CSWX$^+$	NSE	ACD	Thm 4.2	Thm 4.3	Thm 4.20	Thm 4.21
PE	✓	✓	✓		✓			✗	
SP	✓		✓*				✗		✗
weak-UP		✓	✓	✓			✗	✗	
strong-UP						✗			
CS				✓	✓			✗	✗

Table 4.1: Summary of mechanisms, properties they achieve, and impossibility results: a "✓" mark indicates that the mechanism achieves the corresponding property (CSWX$^+$ achieves SP only with a given NSE and non-overlapping payoffs); the "✗" marks at each column indicate that the corresponding properties cannot be achieved together by any mechanism.

Now we have designed mechanisms that achieve weak UP and CS, or PE and CS; meanwhile, weak UP, PE, and CS cannot be achieved by any mechanism. Given the importance of UP to the original purpose of coordination, an interesting question is whether we can relax PE and make these three properties coexists. Namely, we could aim to find a weak UP and CS outcome that is PE among all such outcomes (i.e., one with the maximum social welfare). We leave this question open for future work and only highlight the challenges here. Unlike in normal-form games or many other succinctly representable multiplayer games where CS outcomes (i.e., correlated equilibria) can be computed efficiently [Papadimitriou and Roughgarden, 2008], our model features a very different utility structure, where a player's utility from playing a mixed strategy is not a linear combination of their utilities from playing pure strategies, because of different responses of the attacker induced by these strategies. This means that we are unable to use existing approaches to compute a correlated equilibrium in our setting.

4.6 Summary

In this chapter, we explored ways to coordinate defenders in a multi-defender game from the perspective of mechanism design. We identified several basic properties to be achieved. On the one hand, our impossibility results preclude these properties to coexist; on the other, we design mechanisms that fulfill subsets of the properties. Our results are summarized in Table 4.1, where the mechanisms we design match our impossibility results exactly: every combination of the properties that is not blocked by the impossibility results is achieved by some mechanism. Thus, our results provide a complete picture of what can and cannot be achieved through mechanism design, with respect to the considered properties. Our results demonstrate an intriguing phenomenon: even when the defenders face a common enemy, it is often not possible for them to reach a consensus about how to coordinate. The underlying cause of this phenomenon, as highlighted by our analysis, is a competition among the defenders for incentivizing the attacker to attack a target that they consider less important.

Part II

Follower Deception and Countermeasures

Chapter 5

Who are We Playing Against: Imitative Follower Deception

> *"If you know the enemy and know yourself, you need not fear the result of a hundred battles."* (知彼知己，百战不殆。)
>
> — Sun Tzu, *The Art of War*

Information uncertainty is one of the major challenges facing applications of game theory. In Stackelberg games, various approaches have been proposed to deal with the leader's incomplete knowledge about the follower's payoffs, typically by gathering information from the leader's interaction with the follower. In this and the following chapters, we examine a weakness of such approaches that results from a key assumption they rely on — that is, the follower behaves truthfully when interacting with the leader. As we will show in this chapter, the follower can easily manipulate the leader's information gathering by imitating the behavior of a different follower type, which benefits the follower but may result in the leader choosing a highly suboptimal strategy. This raises a fundamental question: how to design the leader's strategy in the presence of such deceptive behavior?

To answer this question, in this chapter, we present a basic model of Stackelberg games with *imitative follower deception* and show that the leader is indeed able to mitigate potential losses due to follower deception by using carefully designed *policies*. We provide a systematic study on the problem of computing the optimal leader policy and draw an almost complete picture of the complexity landscape, providing matching positive and negative complexity results for natural variants of the model.

5.1 Introduction

As we have introduced in Chapter 2, to compute the optimal strategy commitment in a Stackelberg game, the leader needs to know the follower's payoffs. In reality, however, this information is not always available. To address this issue, various approaches were used. When the leader can estimate the follower's payoffs to within certain intervals, strategies that are robust against small interval uncertainties are computed using worst-case analysis [Letchford et al., 2009, Kiekintveld

et al., 2013, Nguyen et al., 2014]; When the leader knows a distribution of the follower's payoffs, Bayesian Stackelberg equilibria are computed to maximize the leader's average utility [Conitzer and Sandholm, 2006, Paruchuri et al., 2008b, Jain et al., 2008]; When the leader can interact with the follower, algorithms are designed to learn optimal strategies to commit to based on the follower's best responses [Letchford et al., 2009, Blum et al., 2014, Haghtalab et al., 2016, Roth et al., 2016, Peng et al., 2019].

Despite their differences, the above approaches share a common key step — obtaining payoff-relevant information about the follower. However, knowing that the leader is making an attempt at information gathering, a strategic follower would have an incentive to tamper with the leader's information gathering and distort information the leader acquires. One natural way for the follower to do this is to imitate the behavior of a different follower type — a phenomenon we term *imitative follower deception*. Consider perhaps the most basic learning setting where the follower has an uncertain type $\theta \in \Theta$. Knowing the set Θ of all possible types but not the actual θ, the leader wants to learn the follower's type by interacting with the follower. If the follower is truthful, the leader learns the true type θ and, hence, the optimal strategy to commit to against θ. However, if the follower behaves completely according to the payoffs of another type $\theta' \in \Theta$, the leader can only learn the optimal strategy against the false type θ', which may be highly suboptimal with respect to θ.

Unfortunately, all aforementioned approaches to tackling information uncertainty ignore the possibility of a deceitful follower and adopt a simplistic assumption that any follower information the algorithm gathers is truthful. This raises a fundamental question: how to design the leader's strategy in the presence of imitative follower deception?

We put forward a basic model in an attempt to formalize this question and aim to understand how imitative follower deception affects the leader's choice of strategies. We note that, even though there might be other forms of follower deception as well, imitative deception is arguably the simplest and the most natural deception approach, which we expect to happen the most frequently in practice. Our basic model in this chapter is Bayesian by nature: the type of the follower is drawn from a finite set Θ; the follower knows his true type, but the leader only has a probabilistic belief about it. To play against a deceitful follower, we propose that the leader should commit to a *policy* — a "menu" that specifies a mixed strategy to play for each type the follower behaves like. The strategic interaction between the two players under commitment to a policy then results in a Stackelberg game at a higher level. We study the problem of computing the optimal policy of the leader in this game and several natural variants of this problem that are concerned with incentive compatibility and *mixed policy* — a concept that generalizes the policy by using randomization. Our results provide an almost complete picture of the complexity landscape: the problem is hard to approximate in general, while we also present fixed-parameter tractability results and design efficient algorithms with approximation ratios that match the inapproximability results.

5.1.1 Related Work

The problems we address in this and the next two chapters are related to an emerging line of work at the intersection of machine learning and algorithmic game theory, dealing with scenarios where the samples used for training learning algorithms are controlled by strategic agents, who aim to optimize their personal benefit. In particular, there has been recent interest in the analysis of the effect of strategic behavior on the efficiency of existing algorithms, as well as the design of algorithms resilient to strategic manipulation for linear regression [Ben-Porat and Tennenholtz, 2019, Chen et al., 2018, Dekel et al., 2010, Hossain and Shah, 2020, Perote and Perote-Peña, 2004,

Waggoner et al., 2015] and classification [Chen et al., 2019, Dong et al., 2018, Meir et al., 2012, Zhang et al., 2019].

More directly, our work is motivated by a line of work that aims at designing algorithms to learn the optimal strategy for a leader to commit to in a Stackelberg game. The algorithms learn by actively querying the follower's best responses. Letchford et al. [2009] first proposed an algorithm to accomplish this task. Their approach is, in a nutshell, to reconstruct the hyperplanes that divide the leader's strategy space into regions that induce different responses of the follower. They showed that, with high probability, their algorithm is able to reconstruct the hyperplanes by using only polynomially many queries and thereby learn a near-optimal strategy for the leader. In follow-up work, Blum et al. [2014] then designed an efficient algorithm for learning a near-optimal defender strategy in an SSG; Roth et al. [2016] studied the same problem in several other variants of Stackelberg games, such as the profit maximization problem and the optimal toll problem in nonatomic congestion games. Even more recently, Peng et al. [2019] improved the result of Letchford et al. [2009] by getting rid of the assumed constant lower bound on the volume of the smallest region in the leader's strategy space. [Haghtalab et al., 2016] analyzed the learnability of attacker behavior in an SSG that is hypothesized to follow the QR model.

The policy-based countermeasure we design to deal with follower deception is related to Bayesian persuasion, where an agent strategically reveals information to persuade another agent to change her action [Kamenica and Gentzkow, 2011]. In our game, the follower has this information advantage, whereas the leader leverages the power of commitment (to a policy) to play against strategic information revelation of the follower. The game resulting from the leader's commitment to a policy can be viewed as a Bayesian Stackelberg game (see Section 2.1.2, Chapter 2) on top of the original Stackelberg game played between the leader and the follower. Meanwhile, it can also be viewed as a generalization of a BSG, where the leader is allowed to play different strategies depending on the type information revealed by the follower.

Beyond the strategic considerations above, our work is also related to the study of query protocols for learning game-theoretic equilibria. Similarly to our work, algorithms for computing equilibria via utility and best response queries are a natural starting point. For utility queries, there has been much work in proving exponential lower bounds for randomized computation of exact, approximate and well-supported Nash equilibria [Babichenko and Rubinstein, 2017, Babichenko, 2016, Chen et al., 2017, Goldberg and Roth, 2016, Hart and Mansour, 2010, Hart and Nisan, 2016], as well as providing query-efficient protocols for approximate Nash equilibrium computation in bimatrix games, congestion games [Fearnley et al., 2015], anonymous games [Goldberg and Turchetta, 2017], and large games [Goldberg et al., 2019]. Best response queries are weaker than utility queries, but they arise naturally in practice, and are also expressive enough to implement fictitious play, a dynamic first proposed in [Brown, 1949], and proven to converge in [Robinson, 1951] for two-player zero-sum games to an approximate Nash equilibrium. In terms of equilibrium computation, Goldberg and Marmolejo-Cossío [2018] also provide query-efficient algorithms for computing approximate Nash equilibria for bimatrix games via best response queries provided one agent has a constant number of strategies.

Finally, learning via incentive queries in games is directly related to the theory of preference elicitation, where the goal is to mine information about the private parameters of the agents by interacting with them [Blum et al., 2004, Lahaie and Parkes, 2004, Zinkevich et al., 2003, Goldberg et al., 2020]. This has many applications, most notably combinatorial auctions, where access to the valuation functions of the agents is achieved via value or demand queries [Blumrosen and Nisan, 2007, Conen and Sandholm, 2001, Nisan and Segal, 2006].

5.1.2 Chapter Overview

We first provide an example of follower deception in Section 5.2, which illustrates the losses deception may bring to the leader, and how the leader can reduce potential losses by using carefully designed strategies. Then, in Section 5.3, we introduce our model and formalize the leader's countermeasures as the problem of computing an optimal policy. We present the complexity results for this problem in Sections 5.4 and 5.5, where the latter contains results obtained by considering mixed policies. We evaluate our approaches in Section 5.6 in randomly generated games. Finally, in Section 5.7 we consider an extension of our model where no prior information about the follower's type distribution is available and present complexity results for this model.

5.2 A Motivating Example

We illustrate how the follower's deceptive behavior may result in a highly suboptimal strategy for the leader and how we can overcome this issue by using carefully-designed strategies. For illustrative purpose, we use an SSG as an example, though we note that the underlying phenomenon also generalizes to other applications.

Recall the SSG we presented in Section 2.2.1. A defender wants to protect two conservation areas t_1 and t_2 from a poacher's attack. The defender has one security guard that can be sent to patrol one of these two areas, while the poacher chooses one area to attack. At different seasons of the year, the poacher's payoffs vary due to their interests in different animal products. We model this uncertainty using two possible types of poacher payoffs, types α and β. Assume that the defender's payoff is independent of the poacher's type. The poacher knows his true type, but the defender only has a prior belief that each type appears with probability 0.5 (which can be obtained via surveying past prices on black markets). The payoff matrix u^{L} of the defender (row player) and the matrices u^{F}_{α} and u^{F}_{β} of the two poacher types (column player) are given below.

$$u^{\mathsf{L}} = \begin{array}{c} \\ t_1 \\ t_2 \end{array} \begin{array}{c} \begin{array}{cc} t_1 & t_2 \end{array} \\ \begin{pmatrix} 1 & -1 \\ -1 & 0.99 \end{pmatrix} \end{array} \qquad u^{\mathsf{F}}_{\alpha} = \begin{array}{c} \\ t_1 \\ t_2 \end{array} \begin{array}{c} \begin{array}{cc} t_1 & t_2 \end{array} \\ \begin{pmatrix} -1 & 1/3 \\ 3 & -1 \end{pmatrix} \end{array} \qquad u^{\mathsf{F}}_{\beta} = \begin{array}{c} \\ t_1 \\ t_2 \end{array} \begin{array}{c} \begin{array}{cc} t_1 & t_2 \end{array} \\ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

Namely, the defender has two actions she can choose from: to patrol t_1 or t_2 (for simplicity we assume that the defender always sends the security guard to some area, never leaving it idle; this does not change the SSE and BSE in this game); and the poacher also has two actions: to attack t_1 or t_2. Each entry of the matrices represents the utility the player will get when the players take actions of the corresponding row and column.

The only difference between these poacher types lies in their values for successful attacks (i.e., attacks on an unprotected area). Standard calculation (see Section 2.2.1) reveals that if the poacher has type α, the optimal defender strategy is $\left(\frac{3}{4}, \frac{1}{4}\right)$, i.e., patrolling t_1 and t_2 with probability $\frac{3}{4}$ and $\frac{1}{4}$, respectively, resulting in poacher utility 0 at both areas. By the standard assumption, the poacher thus breaks the tie *in favor* of the defender and attacks t_1 (which can be explicitly induced by playing $\left(\frac{3}{4} - \epsilon, \frac{1}{4} + \epsilon\right)$ with $\epsilon \to 0$), which results in the defender obtaining utility 0.5. Similarly, the optimal defender strategy is $\left(\frac{1}{2}, \frac{1}{2}\right)$ against a type-β poacher, which induces the poacher to also attack t_1 and yields defender utility 0.

Ideally, the defender would like to play the optimal patrolling strategy against the poacher's true type at any time; however, this type is unknown to the defender. Nevertheless, if the poacher were to behave truthfully according to his type, the defender can easily learn the poacher's type, e.g., by observing their different best responses to strategy $\left(\frac{3}{5}, \frac{2}{5}\right)$. Indeed, previous approaches to learning an optimal strategy rely crucially on the assumption that the follower will truthfully respond (e.g., [Letchford et al., 2009, Blum et al., 2014, Haghtalab et al., 2016, Roth et al., 2016]). However, in this example, a type-α poacher has a strong incentive to be untruthful: by imitating a type-β poacher and reacting exactly according to u_β^{F}, a type-α poacher is able to induce the defender to play $\left(\frac{1}{2}, \frac{1}{2}\right)$ — the optimal strategy against a type-β poacher — which results this type-α poacher obtaining a utility of $(-1) \times \frac{1}{2} + 3 \times \frac{1}{2} = 1$. This strictly improves upon his previous utility 0 from reacting truthfully. Simple calculation shows that a type-β poacher does not benefit from deception, under the assumption that they have to behave constantly according to one type throughout the game; thus, a type-β poacher will behave truthfully. As a result, the defender will not be able to distinguish type α from β under the poacher's strategic deception.

A Countermeasure The issue we discovered above is due to the fact that the defender focuses solely on optimizing her utility without considering the poacher's strategic behavior. We now show how the defender can overcome this issue by taking into account the poacher's deception. It happens that in this example the optimal policy for the defender is to still play $\left(\frac{3}{4}, \frac{1}{4}\right)$ for type α and $\left(\frac{1}{2}, \frac{1}{2}\right)$ for type β, but induce a type-β poacher to *break ties "against" the defender* by attacking t_2 (i.e., by playing $\left(\frac{1}{2} + \epsilon,\ \frac{1}{2} - \epsilon\right)$ with $\epsilon \to 0$). This slightly decreases the defender's utility against a type-β poacher, from 0 to $(-1) \times \frac{1}{2} + 0.99 \times \frac{1}{2} = -0.005$, but it also eliminates the incentive of a type-α poacher to imitate type β: if the type-α poacher imitates β to attack t_2 under defender strategy $\left(\frac{1}{2}, \frac{1}{2}\right)$, his expected utility would become -1, which is worse than his utility 0 of behaving truthfully.

Remarks (i) This example illustrates an intriguing phenomenon: in the presence of follower deception, it is not always optimal to induce tie breaking in favor of the leader with respect to the type the follower appears to be; rather, we want to induce tie-breaking behavior that benefits the leader overall. What we are designing can be viewed as the best strategy the leader can learn in the presence of follower deception.

(ii) One might wonder why the defender does not ignore the poacher's deception by simply playing a single strategy against all types of poachers, which is precisely the optimal commitment in a Bayesian Stackelberg game. The reason is that our nuanced approach, through the additional information exchange with the follower, can achieve even better defender utility than in the Bayesian Stackelberg equilibrium. In this example, it can be computed that the defender would play $\left(\frac{1}{2}, \frac{1}{2}\right)$ in the Bayesian Stackelberg equilibrium, obtaining utility 0 (see Section 2.2.1), whereas the strategy we design above yields defender utility $\frac{1}{4} - \frac{1}{400}$. (More formally, we will present a result regarding the utility improvement in Proposition 5.3.)

(iii) The price of ignoring follower deception can be huge. Example 5.1 illustrates a game in which the price — the ratio between the leader's utilities when deception is ignored and when it is properly dealt with — is unbounded. (Payoffs are normalized to be in $[0, 1]$ for the ratio to be meaningful.) Nevertheless, perhaps counter-intuitively, follower deception is not always bad for the leader. One example is a Stackelberg game version of the prisoner's dilemma illustrated in Example 5.2.

Example 5.1. Suppose the SSG now has the following payoff matrices, where the payoffs are normalized to be in the interval $[0, 1]$.

$$u^{\mathsf{L}} = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix} \qquad u^{\mathsf{F}}_{\alpha} = \begin{pmatrix} 0 & 0.2 \\ 0.8 & 1 \end{pmatrix} \qquad u^{\mathsf{F}}_{\beta} = \begin{pmatrix} 0.2 & 0 \\ 1 & 0.8 \end{pmatrix}$$

The follower appears as each type with equal probability. When deception is ignored, the leader assumes the follower to be truthful and will play $(0, 1)$ against type α and $(1, 0)$ against type β. However, this gives a type-β follower an incentive to imitate type α, yielding utility ϵ for the leader overall. Now, if the leader properly deals with follower deception and plays, instead, $(1, 0)$ for both types α and β, a type-β follower will have no incentive to imitate type α; the leader obtains utility 0.5 as a result. The ratio $0.5/\epsilon$ is unbounded when ϵ is arbitrarily close to 0. □

Example 5.2. Consider a Stackelberg game where the players can choose to play actions C or D. The true payoffs of the leader and the follower are given by u^{L} and u^{F}, respectively.

$$u^{\mathsf{L}} = \begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} -1 & -3 \\ 0 & -2 \end{pmatrix} \end{array} \qquad u^{\mathsf{F}} = \begin{array}{c} 0 \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} -1 & 0 \\ -3 & -2 \end{pmatrix} \end{array}$$

Note that D is a strictly dominant strategy of the follower, so the optimal commitment of the leader is to play D with probability 1; the pure strategy profile (D, D) forms an SSE of the game, in which both players obtain utility -2. However, suppose that the follower (column player) imitates a type that has utility 0, instead of -1, for the action profile (C, C) as annotated above (suppose that this is indeed also a possible type of the follower). In this case, the leader's optimal commitment is to play C with probability 1, and the follower best responds with C (tie-breaking in favor of the leader), resulting in utility -1 for both players. □

5.3 Stackelberg Games with Follower Deception

We now describe a basic Stackelberg game model with (imitative) follower deception. The model builds on a standard $m \times n$ Stackelberg game which we presented in Chapter 2, where the leader and the follower have m and n pure strategies, respectively; the reader may also refer back to our motivating example as an instantiation of this model. In this game, the leader faces a follower with an uncertain type, which falls in a finite set Θ. Let $u^{\mathsf{F}}_{\theta} \in \mathbb{R}^{m \times n}$ be the payoff matrix of each follower type $\theta \in \Theta$. For simplicity, we assume that the leader's payoff does *not* depend on θ and let $u^{\mathsf{L}} \in \mathbb{R}^{m \times n}$ be the leader's payoff matrix, though we note that all our results easily generalize to the case with type-dependent leader payoffs. By slightly abusing notation, we let

$$u^{\mathsf{L}}(\mathbf{x}, j) = \sum_{i \in [m]} x_i \cdot u^{\mathsf{L}}(i, j)$$

be the *expected utility* of the leader when she plays the mixed strategy $\mathbf{x}$ and the follower plays a pure strategy j. Similarly, we let $u^{\mathsf{F}}(\mathbf{x}, j) = \sum_{i \in [m]} x_i \cdot u^{\mathsf{F}}(i, j)$ be the expected utility of the follower.

To model the leader's prior knowledge regarding the follower's type, we adopt the classic Bayesian perspective and assume that the leader has a prior distribution σ, i.e., each type θ appears with probability σ_{θ}. Note that σ, Θ and the payoff matrices are common knowledge.

Commitment to a Policy We assume that the leader can commit to a *policy* $\pi : \Theta \to \Delta^{m-1} \times [n]$, i.e., a "menu" that prescribes for each type $\theta \in \Theta$ an *outcome* $\pi(\theta)$ of the game. An outcome is a tuple $(\mathbf{x}, j)$ that consists of a mixed strategy $\mathbf{x}$ of the leader and a best response $j \in \mathsf{BR}_\theta(\mathbf{x})$ of the follower type θ that this outcome is prescribed for, where

$$\mathsf{BR}_\theta(\mathbf{x}) = \arg\max_j u_\theta^{\mathsf{F}}(\mathbf{x}, j)$$

denotes the best response set of a type-θ follower. Since Θ is a finite set, we can also represent π as a set $\{(\mathbf{x}^\theta, j^\theta)\}_{\theta \in \Theta}$, i.e., $\pi(\theta) = (\mathbf{x}^\theta, j^\theta)$.

The game then proceeds as follows:

1. *Policy commitment.* The leader commits to a policy $\pi = \{(\mathbf{x}^\theta, j^\theta)\}_{\theta \in \Theta}$ and announce it to the follower.

2. *Type reporting.* Knowing the leader's policy π, a follower of type θ, who appears with probability σ_θ, chooses a type $\hat{\theta}(\pi) \in \Theta$ that will maximize his utility in the realization step below and behaves according to it. The leader then learns that the follower is of type $\hat{\theta}$ by interacting with him. We omit the cost incurred for both players and view this step as a reporting step, where the follower simply reports his type to the leader.

3. *Realization.* The outcome $\left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)}\right)$ is realized: the leader plays the strategy $\mathbf{x}^{\hat{\theta}(\pi)}$ prescribed by her policy; the follower "best responds" to $\mathbf{x}^{\hat{\theta}(\pi)}$ and, in the case of a tie, the leader induces him to choose the particular response $j^{\hat{\theta}(\pi)}$. The leader and the follower obtain utilities $u^{\mathsf{L}}\left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)}\right)$ and $u^{\mathsf{F}}\left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)}\right)$, respectively.

   Thus, the fact that the follower is a rational player means that

   $$\hat{\theta}(\pi) \in \arg\max_{\beta \in \Theta} u_\theta^{\mathsf{F}}(\mathbf{x}^\beta, j^\beta)$$

   in the type reporting step; we assume that if there are multiple best types to report, the follower breaks ties in favor of the leader. Given the prior knowledge about the follower's type, the leader's overall utility, as the expectation over the distribution of the follower's true types, is given as follows:

   $$\mathrm{EU}(\pi) = \sum_{\theta \in \Theta} \sigma_\theta \cdot u^{\mathsf{L}}\left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)}\right). \tag{5.1}$$

Our task is to find the best policy that achieves $\max_\pi \mathrm{EU}(\pi)$. Note that we assume the follower to behave consistently according to the type he reports. This means that the follower may be playing a fake best response — hence, a suboptimal one — in the realized outcome and may exploit the leader for an even higher utility by switching back to his true best response. Nevertheless, since such a change in his behavior will inevitably make the leader aware of the deception and further complicates the interaction, we ignore the possibility of such a behavior change and adopt this simplified model to capture the essence of follower deception. Our model can be viewed as a generalization of the classic Bayesian Stackelberg game (BSG) model, in the senes that the leader's policy is restricted to prescribing the same mixed strategy for all follower types in the classic BSG model.

Incentive Compatibility (IC) We can further impose the IC constraints on π so that reporting truthfully is a (weakly) dominant strategy for every follower type θ, i.e., $u_\theta^{\mathsf{F}}(\mathbf{x}^\theta, j^\theta) \geq u_\theta^{\mathsf{F}}(\mathbf{x}^\beta, j^\beta)$ for all $\beta \in \Theta$. We will consider both settings, i.e., with and without IC constraints. Indeed, an IC policy always exists and the best IC policy always achieves at least what the leader obtains in a Bayesian Stackelberg equilibrium (same as what we demonstrate in the motivating example).

Proposition 5.3. *There always exists an IC policy and the optimal IC policy achieves at least the leader's utility in a Bayesian Stackelberg equilibrium.*

Proof. Let $\mathbf{x}$ be the leader's strategy in a Bayesian Stackelberg equilibrium (BSE) and let $j^\theta \in \arg\max_{j \in \mathsf{BR}_\theta(\mathbf{x})} u^{\mathsf{L}}(\mathbf{x}, j)$ be the best response of a type-θ follower in the BSE. A policy π with $\pi(\theta) = (\mathbf{x}, j^\theta)$ for each $\theta \in \Theta$ is trivially IC because no matter what type the follower reports, the leader will always play $\mathbf{x}$, while by reporting truthfully the follower will also be induced to play his true best response j^θ against $\mathbf{x}$. This policy offers the leader utility $\sum_{\theta \in \Theta} \sigma_\theta \cdot u^{\mathsf{L}}(\mathbf{x}, j^\theta)$ — exactly what she obtains in the BSE. □

Note that the optimal IC policy may perform strictly worse than the optimal non-IC policy; see the example below. This is due to the fact that different follower types often correspond to different sets of feasible outcomes, so we cannot apply a revelation-principle-style argument to show that considering only IC policies is without loss of generality.

Example 5.4. Consider a game with the same payoffs as Example 5.1, but with $\epsilon = 2$ instead of a small number. Suppose for simplicity that the follower appears as type α with probability 0, and as type β with probability 1. In this game, any IC policy will incentivize a type-β follower to best responds with action 1 since action 1 strictly dominates action 2 in u_β^{F}. The leader obtains expected utility at most 1 according to the first column of u^{L}. On the other hand, consider a policy that prescribes outcome $((0,1),2)$ for type α, and outcome $((1,0),1)$ for type β. This policy will incentivize a type-β follower to misreport that they are of type α, so it is not IC. Nevertheless, it yields expected utility 2 for the leader. □

5.4 Computing the Optimal Policy

In this section, we present complexity results and algorithms for computing the optimal leader policy. We will refer to the problem of computing the optimal policy as Opt-Policy, and as Opt-Policy-IC when IC constraints are imposed.

5.4.1 Hardness of Approximation

We first present in Theorems 5.5 and 5.6 that both Opt-Policy and Opt-Policy-IC are NP-hard, even when we only want to approximate the optimal policy to within a meaningful ratio and even when the follower only has a small number of pure strategies. Our approximation ratios are all multiplicative and are defined as the ratio between the leader's expected utility provided by the solution of the algorithm and that provided by the optimal policy; by convention, we normalize the leader's payoff parameters to be in $[0, 1]$ for the ratios to be meaningful. We will shortly show that these inapproximability bounds are essentially tight by providing efficient algorithms that guarantee the matching approximation ratio.

Theorem 5.5. *For any constant $\epsilon > 0$,* Opt-Policy *does not admit a polynomial-time $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximation algorithm unless P = NP, even when the follower has three actions only.*

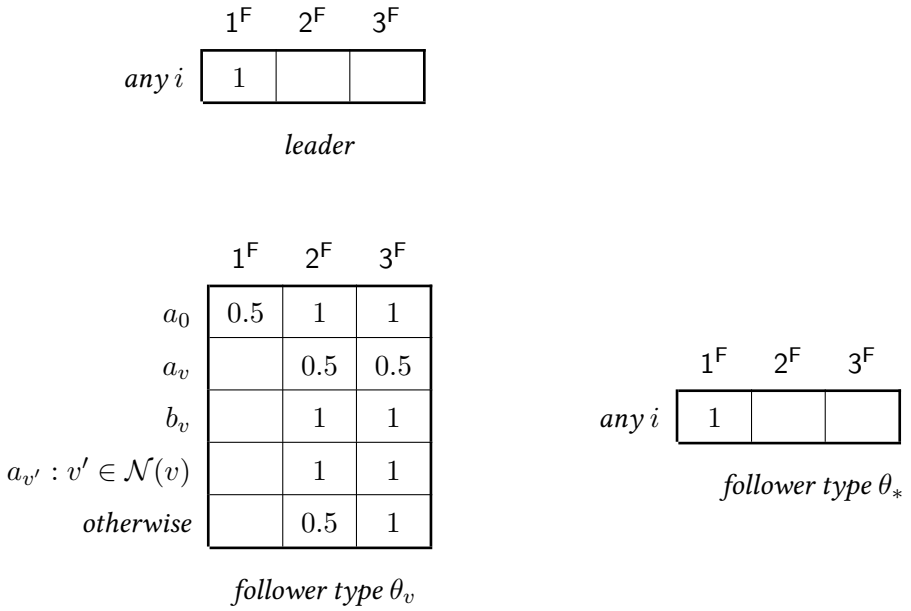


Figure 5.1: Payoffs matrices (blank entries are 0).

Proof. We show a reduction from the MAX-INDEPENDENT-SET problem, which asks for the size of the maximum independent set in a graph $G = (V, E)$. A set of nodes $V' \subseteq V$ is an *independent set* of G if no edge in E connects any pair of nodes in V'; a maximum independent set has the largest size among all independent sets. It is known that no polynomial-time $\frac{1}{|V|^{1-\epsilon}}$-approximation algorithm exists for MAX-INDEPENDENT-SET, unless P = NP [Zuckerman, 2006].

We construct a game with $|V| + 1$ follower types $\Theta = \{\theta_*\} \cup \{\theta_v : v \in V\}$, where each θ_v corresponds to node v. Let $\sigma_{\theta_*} = 0$ and $\sigma_\theta = \frac{1}{|\Theta|-1}$ for all other θ. The leader has $2|V| + 1$ actions $\{a_v : v \in V\} \cup \{b_v : v \in V\} \cup \{a_0\}$; and the follower has three actions $\{1^\mathsf{F}, 2^\mathsf{F}, 3^\mathsf{F}\}$. The payoffs are given in Figure 5.1, where $\mathcal{N}(v)$ denotes the set of neighboring nodes of v, and all the empty entries are 0.

The following are a few observations about the game:

- For follower type θ_*, $j = 1^\mathsf{F}$ strictly dominates all the other actions. A follower can only be induced to play $j = 1^\mathsf{F}$ when θ_* is reported, in which case the leader gets utility 1.
- For each θ_v, when they are reported, $j = 1^\mathsf{F}$ cannot be induced by any strategy of the leader because it is strictly dominated by $j = 2^\mathsf{F}$. As a result, the leader gets 0 once θ_v is reported.
- Given the above, the leader's overall utility is proportional to the number of follower types θ_v that are motivated to report θ_*.

Now for any independent set V' of size k in G, we can construct the following leader policy π with $\pi(\theta_*) = (a_0, 1^\mathsf{F})$, $\pi(\theta_v) = (a_v, 2^\mathsf{F})$ if $v \in V'$, and $\pi(\theta_v) = (b_v, 2^\mathsf{F})$ if $v \notin V'$. Given this policy, all types θ_v, $v \in V'$, will find reporting θ_* optimal, while all the others are better-off reporting truthfully. As a result, the leader receives overall utility $\frac{k}{|\Theta|-1}$.

Conversely, we claim that for any policy π, the set of nodes v such that the corresponding types θ_v are incentivized to report θ_* form an independent set, i.e., $\{v \in V : \hat{\theta}_v(\pi) = \theta_*\}$ is an

independent set. To see this, suppose for a contradiction that types θ_v and $\theta_{v'}$ are incentivized to report θ_* but $\{v, v'\} \in E$. By our observations above, both θ_v and $\theta_{v'}$ will be induced to play 1^{F} and obtain utility 0.5 each. Thus, we need the following conditions in order for them to have no incentive to report another type.

(i) $x^{\theta_{v'}}_{a_{v'}} = 0$, since otherwise the fourth row of the payoff matrix of θ_v would be chosen with some probability, and θ_v would be better-off reporting $\theta_{v'}$ to obtain utility strictly greater than 0.5.

(ii) $x^{\theta_{v'}}_{a_0} = x^{\theta_{v'}}_{b_{v'}} = 0$, and $x^{\theta_{v'}}_{a_{v''}} = 0$ for all $v'' \in \mathcal{N}(v')$, since otherwise one of the first, third and fourth rows of the payoff matrix of $\theta_{v'}$ would be chosen with some probability, and $\theta_{v'}$ would be better-off reporting truthfully to obtain utility strictly greater than 0.5.

It follows that, now $\mathbf{x}^{\theta_{v'}}$ will pick the last row of the payoff matrix with probability 1, so 3^{F} becomes the only best response of a type-$\theta_{v'}$ player against $\mathbf{x}^{\theta_{v'}}$. Thus, we have $\pi(\theta_{v'}) = (\mathbf{x}^{\theta_{v'}}, 3^{\mathsf{F}})$, in which case, however, $\theta_{v'}$ is able to obtain utility 1 by reporting truthfully, which contradicts the assumption that $\theta_{v'}$ is incentivized to report θ^*.

Therefore, the number of follower types that are motivated to report θ_* by the optimal policy is exactly to the size of the maximum independent set (and proportional to the leader's overall utility). Any $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximation algorithm would also provide a $\frac{1}{|V|^{1-\epsilon}}$-approximation to the Max-Independent-Set problem, which does not exists unless P = NP. This completes the proof. □

Next, we show that the same inapproximability result holds for Opt-Policy-IC. Our reduction will still be from Max-Independent-Set but the underlying idea and the instance constructed are different: the previous reduction maps follower types that *violate* IC to an independent set; this approach does not apply when IC constraints are imposed.

Theorem 5.6. *For any constant $\epsilon > 0$,* Opt-Policy-IC *does not admit a polynomial-time $\frac{1}{|\Theta|^{1-\epsilon}}$-approximation algorithm unless P = NP, even when the follower has three actions only.*

Proof. We show a reduction from the Max-Independent-Set problem. Given an instance of Max-Independent-Set specified by a graph $G = (V, E)$, we construct a game with $|V|$ follower types $\Theta = \{\theta_v : v \in V\}$; each θ_v corresponds to a node v and appears with probability $\frac{1}{|\Theta|}$. The leader has $2|V|$ actions $\{a_v : v \in V\} \cup \{b_v : v \in V\}$; and the follower has three actions $\{1^{\mathsf{F}}, 2^{\mathsf{F}}, 3^{\mathsf{F}}\}$. The payoffs are given in Figure 5.2, where $\mathcal{N}(v)$ denotes the set of neighboring nodes of v, and all the empty entries are 0.

Now for any independent set V' of size k in G, consider a leader policy π with: $\pi(\theta_v) = (a_v, 1^{\mathsf{F}})$ if $v \in V'$, and $\pi(\theta_v) = (b_v, 2^{\mathsf{F}})$ if $v \notin V'$. In this policy, all types find truthful reporting to be optimal. Hence, the policy is IC and offers the leader overall utility $\frac{k}{|\Theta|}$.

Conversely, for any IC policy π, we claim that the set of nodes v such that the corresponding follower types θ_v are motivated to respond with $j = 1^{\mathsf{F}}$ form an independent set, i.e., $\{v \in V : \pi(\theta_v) = (i, 1^{\mathsf{F}})$ for some $i\}$ is an independent set. Suppose for a contradiction that θ_v and $\theta_{v'}$ both respond with $j = 1^{\mathsf{F}}$ while $(v, v') \in E$. Observe that, when θ_v is reported, $j = 1^{\mathsf{F}}$ can be induced only by the strategy a_v because it is strictly dominated by $j = 3^{\mathsf{F}}$ in all other cases. Thus, $\pi(\theta_v) = (a_v, 1^{\mathsf{F}})$ and $\pi(\theta_{v'}) = (a_{v'}, 1^{\mathsf{F}})$, in which case both θ_v and $\theta_{v'}$ obtain utility 0, so $\theta_{v'}$ would be better-off reporting θ_v, which contradicts the assumption that π is IC.

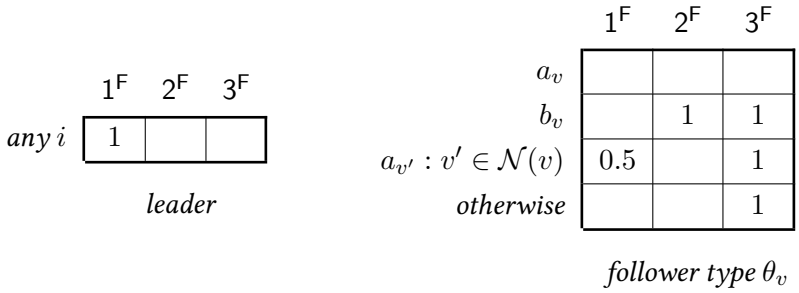


Figure 5.2: Payoffs matrices (blank entries are 0).

Therefore, under the optimal IC policy, the number of follower types that respond with $j = 1^{\mathsf{F}}$ (which is proportional to the leader's overall utility) is equal to the size of the maximum independent set in G. Any $\frac{1}{|\Theta|^{1-\epsilon}}$-approximation algorithm would also provide a $\frac{1}{|V|^{1-\epsilon}}$-approximation to the Max-Independent-Set problem. □

Tightness of the Inapproximability Bounds The bounds in Theorems 5.5 and 5.6 are essentially tight as suggested by a polynomial-time approximation algorithm in the next theorem, which guarantees a matching approximation ratio.

Theorem 5.7. *Both* Opt-Policy *and* Opt-Policy-IC *admit polynomial-time* $\frac{1}{|\Theta|}$*-approximation algorithms.*

Proof. We show an approximation algorithm for Opt-Policy. The algorithm for Opt-Policy-IC follows a similar procedure.

For each follower type $\theta \in \Theta$, we compute a policy π_θ that maximizes the actual utility the leader obtains on this specific type. We show that (i) π_θ can be computed in polynomial time, and (ii) the best π_θ over all $\theta \in \Theta$ achieves at least $\frac{1}{|\Theta|}$ of the leader's utility under the optimal policy.

(i) Polynomial-time computability. To compute π_θ, the idea is to enumerate all possible reporting strategies $\beta \in \Theta$ of type θ and, for each β, compute the best policy *under the constraint that it is indeed optimal for a type-θ follower to report β*. An observation that helps simplifying the computation is that we only need to guarantee reporting β to be (weakly) preferred by a type-θ follower to reporting θ, i.e., the truthful report, so it suffices to consider only the two associated outcomes in the leader's policy, i.e., $(\mathbf{x}^\theta, j^\theta)$ and $(\mathbf{x}^\beta, j^\beta)$. This is due to the fact that a policy that prescribes the same strategy $\mathbf{x}^\theta$ for all other reports $\beta' \notin \{\theta, \beta\}$ will trivially make the truthful report (weakly) preferred to all other reporting strategies $\beta' \notin \{\theta, \beta\}$: by reporting truthfully, the follower will already be induced to play his true best response to $\mathbf{x}^\theta$. Now, the problem reduces to solving the following LP (linear program), in which $\mathbf{x}^\theta$ and $\mathbf{x}^\beta$ are the variables, for every pair of possible values of j^θ and j^β. Since $j^\theta, j^\beta \in [n]$, there are n^2 LPs to be solved for each β; the polynomial-time computability follows immediately.

$$
\begin{aligned}
\text{maximize} \quad & u^{\mathsf{L}}(\mathbf{x}^\beta, j^\beta) && && (5.2)\\
\text{subject to} \quad & u^{\mathsf{F}}_\theta(\mathbf{x}^\theta, j^\theta) \geq u^{\mathsf{F}}_\theta(\mathbf{x}^\theta, j) && \text{for all } j \in [n] && (5.2a)\\
& u^{\mathsf{F}}_\beta(\mathbf{x}^\beta, j^\beta) \geq u^{\mathsf{F}}_\beta(\mathbf{x}^\beta, j) && \text{for all } j \in [n] && (5.2b)\\
& u^{\mathsf{F}}_\theta(\mathbf{x}^\beta, j^\beta) \geq u^{\mathsf{F}}_\theta(\mathbf{x}^\theta, j^\theta) && && (5.2c)\\
& \mathbf{x}^\theta, \mathbf{x}^\beta \in \Delta^{m-1} && && (5.2d)
\end{aligned}
$$

Namely, we maximize the leader's utility obtained on a type-θ follower under the condition that the follower is incentivized to report β. Here, (5.2a) and (5.2b) guarantee that $j^\theta \in \mathsf{BR}_\theta(\mathbf{x})$ and $j^\beta \in \mathsf{BR}_\beta(\mathbf{x})$ as required in the definition of leader policy; and (5.2c) ensures that reporting β is indeed a better choice for a type-θ follower than reporting θ. The policy π_θ is given by the LP with the best optimal solution among all the $n^2|\Theta|$ LPs.

(ii) Utility guarantee. It remains to show that the best π_θ over all $\theta \in \Theta$ offers the desired utility. For each θ, let μ_θ^{L} be the utility offered by π_θ. We have $\mathrm{EU}(\pi_\theta) \geq \sigma_\theta \cdot \mu_\theta^{\mathsf{L}}$ because $\sigma_\theta \cdot \mu_\theta^{\mathsf{L}}$ is only the part of the leader's utility obtained on type θ. Now consider an arbitrary feasible policy $\varpi = (\mathbf{z}^\theta, k^\theta)_{\theta\in\Theta}$. Let $\hat{\theta}(\varpi)$ denote a type-θ follower's best reporting strategy in response to policy ϖ. It holds that $\mu_\theta^{\mathsf{L}} \geq u^{\mathsf{L}}(\mathbf{z}^{\hat{\theta}(\varpi)}, k^{\hat{\theta}(\varpi)})$ because $\mathbf{x}^\theta = \mathbf{z}^\theta$ and $\mathbf{x}^\beta = \mathbf{z}^{\hat{\theta}(\varpi)}$ are feasible under (5.2a)–(5.2d) when j^θ and j^β are set to k^θ and $k^{\hat{\theta}(\varpi)}$, respectively. Let $\theta^* = \arg\max_{\theta\in\Theta} \sigma_\theta \cdot \mu_\theta^{\mathsf{L}}$. It follows that, for any feasible π,

$$\mathrm{EU}(\pi_{\theta^*}) \geq \sigma_{\theta^*} \cdot \mu_{\theta^*}^{\mathsf{L}} \geq \frac{1}{|\Theta|}\sum_\theta \sigma_\theta \cdot \mu_\theta^{\mathsf{L}} \geq \frac{1}{|\Theta|}\sum_\theta \sigma_\theta \cdot u^{\mathsf{L}}(\mathbf{x}^{\hat{\theta}(\varpi)}, j^{\hat{\theta}(\varpi)}) = \frac{1}{|\Theta|}\cdot \mathrm{EU}(\varpi).$$

Since the choice of ϖ is arbitrary, π_{θ^*} serves as a $\frac{1}{|\Theta|}$-approximation of the optimal policy, which completes the proof. □

5.4.2 An MILP Formulation and the Tractability with Small Θ

Given our hardness results, efficient algorithms for computing the optimal policy seem unlikely. We provide an algorithm based on an MILP as a practical approach to tackle the problem. The MILP formulation also forms the basis of a polynomial-time algorithm for Opt-Policy and Opt-Policy-IC when the size of the follower type set Θ is fixed.

First, when IC constraints are not imposed, the following program (not linear yet) computes the optimal policy of the leader, where μ_θ^{L}, $\mathbf{x}^\theta$, y_β^θ, and p_j^θ are the variables ($\theta, \beta \in \Theta$, and $j \in [n]$).

$$\begin{aligned}
&\text{maximize} && \textstyle\sum_\theta \sigma_\theta \cdot \mu_\theta^{\mathsf{L}} && && (5.3)\\
&\text{subject to} && \mu_\theta^{\mathsf{L}} \leq \textstyle\sum_j p_j^\beta \cdot u^{\mathsf{L}}(\mathbf{x}^\beta, j) + (1 - y_\beta^\theta)\cdot M && \text{for all } \theta, \beta \in \Theta && (5.3a)\\
& && p_j^\theta \cdot \left[u_\theta^{\mathsf{F}}(\mathbf{x}^\theta, j) - u_\theta^{\mathsf{F}}(\mathbf{x}^\theta, k)\right] \geq 0 && \text{for all } \theta \in \Theta,\ j,k \in [n] && (5.3b)\\
& && \textstyle\sum_j p_j^\beta \cdot u_\theta^{\mathsf{F}}(\mathbf{x}^\beta, j) - \sum_j p_j^\gamma \cdot u_\theta^{\mathsf{F}}(\mathbf{x}^\gamma, j) \geq -(1 - y_\beta^\theta)\cdot M && && \\
& && && \text{for all } \theta, \beta, \gamma \in \Theta && (5.3c)\\
& && \textstyle\sum_i x_i^\theta = \sum_j p_j^\theta = \sum_\beta y_\beta^\theta = 1 && \text{for all } \theta \in \Theta && (5.3d)\\
& && \mathbf{x}^\theta \in [0,1]^m,\ \ y_\beta^\theta \in \{0,1\},\ \ p_j^\theta \in \{0,1\} && \text{for all } \theta, \beta \in \Theta,\ j \in [n] && (5.3e)
\end{aligned}$$

Here, the leader's policy is represented by variables $\mathbf{x}^\theta$ and p_j^θ, where $\mathbf{x}^\theta$ is the leader's mixed strategy prescribed for report θ, and $p_j^\theta \in \{0,1\}$ captures the induced follower action. Through (5.3b), it is guaranteed that $p_j^\theta = 1$ only if j is a best response of type θ to $\mathbf{x}^\theta$, and $p_j^\theta = 0$, otherwise. Similarly, y_β^θ captures the optimal reporting strategy of type θ, i.e., $y_\beta^\theta = 1$ only if reporting β is optimal for θ, and this is guaranteed by (5.3c), in which M is a sufficiently large constant. Finally, each μ_θ^{L} captures the utility the leader obtains from true type θ by (5.3a).

This program is not yet linear because of the quadratic terms $p_j^\theta \cdot \mathbf{x}^\theta$. To linearize the program, we replace these terms with a set of new variables $\tilde{x}_{ji}^\theta$, subject to $0 \leq \tilde{x}_{ji}^\theta \leq x_i^\theta$ (for all $\theta \in \Theta$,

$j \in [n]$, and $i \in [m]$) and $\sum_i \tilde{x}^\theta_{ji} = p^\theta_j$ (for all $\theta \in \Theta$ and $i \in [m]$). This guarantees $\tilde{x}^\theta_{ji} = p^\theta_j \cdot x^\theta_i$ given that p^θ_j is restricted to be in $\{0, 1\}$.

In the situation with IC constraints, we force $y^\theta_\theta = 1$ for all $\theta \in \Theta$ in the above MILP. The windfall of the MILP formulation is Theorem 5.8, the tractability of Opt-Policy and Opt-Policy-IC with a small Θ.

Theorem 5.8. *For a fixed number of follower types, both* Opt-Policy *and* Opt-Policy-IC *can be solved in polynomial time.*

Proof. Following Program (5.3), when $|\Theta|$ is a constant, the size of the joint space of feasible $\mathbf{y}$ and $\mathbf{p}$ is bounded by a polynomial in n. We can enumerate every combination in this space and solve Program (5.3) with $\mathbf{y}$ and $\mathbf{p}$ fixed to values in the combination; the program becomes a linear program as the remaining variables are continuous. □

5.5 Generalization to Mixed Policy

The leader policy we have considered so far consists of a single outcome $(\mathbf{x}^\theta, j^\theta)$ for each θ. In this section, we extend the definition of leader policy, and allow a policy to prescribe a distribution of outcomes for each follower type. We call an outcome distribution a *mixture* and refer to a leader policy consisting of mixtures a *mixed* policy, as opposed to *pure* policies in the previous sections.

Now the game proceeds as follows. The leader first commits to a mixed policy. The follower observes the mixed policy and reports a type θ. The leader then samples an outcome $(\mathbf{x}, j)$ from the mixture prescribed for θ and plays $\mathbf{x}$; the follower responds with action j.

Obviously, the utility the leader obtains by playing an optimal mixed policy is at least what she can obtain by playing a pure policy, as all pure policies are also mixed policies by definition. An immediate question is why further randomization over the outcomes will benefit the leader since the leader is already playing a mixed strategy in every outcome, and randomizing over mixed strategies normally will not bring any extra power to a player. It turns out that this additional randomization may be beneficial for the leader when there is follower deception. Intuitively, such randomization can help us tame the incentive constraints for the follower's type reporting, while the randomization in mixed strategies is responsible for inducing desired follower action responses. The following example provides a more concrete illustration.

Example 5.9 (Utility improvement with mixed policies). Consider a security game with two targets t_1 and t_2. The defender can patrol exactly one of these areas and the poacher chooses one of them to attack. The poacher has three types α, β, and γ, which all appear with probability $1/3$. The payoff matrices of the players are given in the tables below (similar to the matrices in Section 5.2, where the row actions are to patrol t_1 and t_2, and the column actions are to attack t_1 and t_2).

$$u^{\mathsf{L}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad u^{\mathsf{F}}_\alpha = \begin{pmatrix} 0 & 0 \\ 1 & -2 \end{pmatrix} \qquad u^{\mathsf{F}}_\beta = \begin{pmatrix} -2 & 1 \\ 0 & 0 \end{pmatrix} \qquad u^{\mathsf{F}}_\gamma = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

Note that the two areas are of equal value to both the defender and poacher type γ. Thus, γ will always attack the less frequently patrolled area, so for any outcome $(\mathbf{x}, j)$ prescribed for report γ, it always holds that $x_j \leq 0.5$. Suppose first that $j = t_1$. Then a type-α poacher is able to obtain utility at least 0.5 by reporting γ, in which case the actual utility the defender obtains on type α, using a pure policy, cannot be more than 0, because for a type-α poacher to obtain utility at least

0.5, the bottom left entries of the payoff matrices need to be chosen with probability at least 0.5. The same argument applies to type β if we suppose $j = t_2$. As a result, any pure policy can obtain a positive utility on at most one of α and β. Further, observe that u_γ^{F} make the game zero-sum. Thus, the actual utility the defender obtains on poacher type γ is at most 0 because a type-γ poacher is guaranteed utility 0 by reporting truthfully, and when playing against a type-γ poacher, the defender always gets the negative of the poacher's actual utility. As a result, a pure policy obtains positive utility on at most one poacher type; the defender's overall utility is at most $1/3$.

Now we consider a mixed policy as follows:

- If type α is reported, choose outcome $((1, 0), t_1)$ with probability 1;
- If β is reported, choose outcome $((0, 1), t_2)$ with probability 1;
- If type γ is reported, choose outcomes $\big((\frac{1}{2}, \frac{1}{2}), t_1\big)$ and $\big((\frac{1}{2}, \frac{1}{2}), t_2\big)$, each with probability 0.5.

As such, a poacher of types α or β can only obtain utility -0.5 in expectation if they report γ. This incentivizes these two poacher types to report truthfully, yielding defender utility 1 in each case, and overall an expected utility of $2/3$ for the defender given that each poacher type appears with probability $1/3$ — a utility improvement of at least $1/3$ compared to the utility yielded by the optimal pure policy. □

5.5.1 Mixture with a Finite Support

To compute the optimal mixed policy, one challenge is that a mixture may be supported on an infinite set since there are infinitely many mixed strategies and, hence, as many outcomes. Fortunately, we can prove that it suffices to consider mixtures supported on at most n outcomes; each outcome induces the follower to choose a distinct action in $[n]$. This is shown in Proposition 5.10 below.

Proposition 5.10. *For any feasible mixture ϕ prescribed for a reported type $\theta \in \Theta$, there is a feasible mixture with support size $\tilde{n} \le n$ that yields the same utility as does ϕ, for both the leader and the follower.*

Proof. We will show that for any given mixture ϕ prescribed for report θ we can merge outcomes that induce the same response of the follower into one outcome, and this will not change both players' expected utility.

Formally, let $\mathcal{O}$ be the support set of ϕ; let $p_{(\mathbf{x},j)} > 0$ be the probability assigned to outcome $(\mathbf{x}, j) \in \mathcal{O}$. For each $j \in [n]$, let $\mathcal{O}_j = \{(\mathbf{x}, j') \in \mathcal{O} : j' = j\}$ be the set of outcomes that induce follower action j. Let $\tilde{p}_j = \sum_{(\mathbf{x},j)\in\mathcal{O}_j} p_{(\mathbf{x},j)}$ and $\tilde{\mathbf{x}}_j = \frac{1}{\tilde{p}_j} \sum_{(\mathbf{x},j)\in\mathcal{O}_j} p_{(\mathbf{x},j)} \cdot \mathbf{x}$. Let $\tilde{\phi}$ be a mixture supported on $\{(\tilde{\mathbf{x}}_1, 1), \ldots, (\tilde{\mathbf{x}}_n, n)\}$, with probability $\tilde{p}_j$ for each $(\tilde{\mathbf{x}}_j, j)$. This new mixture will not change the leader's utility when θ is reported because of the following:

$$\begin{aligned}
\sum_{(\mathbf{x},j)\in\mathcal{O}} p_{(\mathbf{x},j)} \cdot u^{\mathsf{L}}(\mathbf{x}, j) &= \sum_{j\in[n]} \sum_{(\mathbf{x},j)\in\mathcal{O}_j} p_{(\mathbf{x},j)} \cdot u^{\mathsf{L}}(\mathbf{x}, j) \\
&= \sum_{j\in[n]} \tilde{p}_j \cdot u^{\mathsf{L}}\Big(\frac{1}{\tilde{p}_j} \sum_{(\mathbf{x},j)\in\mathcal{O}_j} p_{(\mathbf{x},j)} \cdot \mathbf{x}\,,\ \ j \Big) \\
&= \sum_{j\in[n]} \tilde{p}_j \cdot u^{\mathsf{L}}(\tilde{\mathbf{x}}_j,\ j).
\end{aligned}$$

By changing labels of the utility function, we can obtain $\sum_{(\mathbf{x},j)\in\mathcal{O}} p_{(\mathbf{x},j)} \cdot u_\beta^{\mathsf{F}}(\mathbf{x}, j) = \sum_{j\in[n]} \tilde{p}_j \cdot u_\beta^{\mathsf{F}}(\tilde{\mathbf{x}}_j,\ j)$ for any follower type $\beta \in \Theta$, so $\tilde{\phi}$ will not change the follower's utility, either, irrespective of his true type.

It remains to show that $\tilde{\phi}$ is indeed a valid mixture. Clearly, $\tilde{p}_j \geq 0$ and $\sum_j \tilde{p}_j = \sum_j \sum_{(\mathbf{x},j)\in\mathcal{O}_j} p_{(\mathbf{x},j)} = \sum_{(\mathbf{x},j)\in\mathcal{O}} p_{(\mathbf{x},j)} = 1$, so $\tilde{p}_j$ is a valid probability distribution. In addition, we need each outcome $(\tilde{\mathbf{x}}_j, j)$ to be feasible, i.e., $j \in \mathsf{BR}_\theta(\tilde{\mathbf{x}}_j)$, so that j can indeed be induced by $\tilde{\mathbf{x}}_j$. This is true because the set $\mathcal{X}_j = \{\mathbf{x} \in \Delta^{m-1} : j \in \mathsf{BR}_\theta(\mathbf{x})\} = \{\mathbf{x} \in \Delta^{m-1} : u_\theta^{\mathsf{F}}(\mathbf{x}, j) \geq u_\theta^{\mathsf{F}}(\mathbf{x}, j') \text{ for all } j' \in [n]\}$ is a convex set, and by definition, all $\mathbf{x}$ such that $(\mathbf{x}, j) \in \mathcal{O}_j$ are in $\mathcal{X}_j$. Thus, $\tilde{\mathbf{x}}$, as a convex combination of $\mathbf{x} \in \mathcal{X}_j$, also lies in $\mathcal{X}_j$, so $j \in \mathsf{BR}_\theta(\tilde{\mathbf{x}}_j)$. Therefore, $(\tilde{\mathbf{x}}_j, j)$ is a feasible outcome and $\tilde{\phi}$ a valid mixture. This completes the proof. □

Following this result, we can represent the mixture for each type θ as a vector $(p_1^\theta, \ldots, p_n^\theta) \in \Delta^{n-1}$, along with a support set $\{\mathbf{x}_1^\theta, \ldots, \mathbf{x}_n^\theta\}$ of n mixed strategies; the mixture samples each outcome $(\mathbf{x}_j^\theta, j)$ with probability p_j^θ. Note that this representation allows outcomes involved to be invalid because it is possible that some follower action j cannot be induced by any strategy of the leader, i.e., $j \notin \mathsf{BR}_\theta(\mathbf{x})$ for all $\mathbf{x} \in \Delta^{m-1}$. Thus, a mixture is valid only if $j \in \mathsf{BR}_\theta(\mathbf{x}_j^\theta)$ for all j such that $p_j^\theta > 0$.

5.5.2 Computing the Optimal Mixed Policy

Proposition 5.10 implies that there exists an optimal mixed policy of polynomial size. Nevertheless, our next theorem, Theorem 5.11, shows that the problem of computing the optimal mixed policy, referred to as Opt-XPolicy, remains inapproximable in the situation without IC constraints. Interestingly, after we impose IC constraints, the problem Opt-XPolicy-IC becomes tractable. We present these results next.

Theorem 5.11. *For any constant $\epsilon > 0$,* Opt-XPolicy *does not admit a polynomial-time $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximation algorithm unless P = NP, even when the follower has three actions only.*

Proof. We show a reduction from the same Max-Independent-Set problem as in the proof of Theorem 5.5, but we will need to adapt the correctness proof to mixed policies here.

For one direction of the reduction, if there is an independent set V' of size k in G, the policy we constructed in the proof of Theorem 5.5 achieves overall utility $\frac{k}{|\Theta|-1}$ by the same argument.

For the other direction, observe that for mixed strategies it still holds that the leader gets 1 when θ_* is reported, and 0, otherwise; hence, the leader's overall utility is proportional to the number of types θ_v, $v \in V$, who report θ_*. We show that, for any mixed policy π, $\{v \in V : \hat{\theta}_v(\pi) = \theta_*\}$ is an independent set to conclude the proof. Suppose for a contradiction that $\hat{\theta}_v(\pi) = \theta_*$ and $\hat{\theta}_{v'}(\pi) = \theta_*$ but $\{v, v'\} \in E$. We make the following observations:

(i) $p_{1^{\mathsf{F}}}^{\theta_v} = 0$ because 1^{F} is strictly dominated by 2^{F} for a follower of type θ_v and hence $1^{\mathsf{F}} \notin \mathsf{BR}_{\theta_v}(\mathbf{x})$ for any strategy $\mathbf{x}$ of the leader. It follows that we need $p_{2^{\mathsf{F}}}^{\theta_v} > 0$ or $p_{3^{\mathsf{F}}}^{\theta_v} > 0$, as $(p_{1^{\mathsf{F}}}^{\theta_v}, p_{2^{\mathsf{F}}}^{\theta_v}, p_{3^{\mathsf{F}}}^{\theta_v}) \in \Delta^2$.

(ii) Since a follower of type $\theta_{v'}$ gets utility 0.5 by reporting θ_*, we need $p_{2^{\mathsf{F}}}^{\theta_v} \cdot x_{2^{\mathsf{F}},a_v}^{\theta_v} = 0$ and $p_{3^{\mathsf{F}}}^{\theta_v} \cdot x_{3^{\mathsf{F}},a_v}^{\theta_v} = 0$ as otherwise the action profile $(a_v, 2^{\mathsf{F}})$ would be chosen with some positive probability and $\theta_{v'}$ would be better-off reporting θ_v to get more than 0.5.

Now, consider the two possible cases implied by (i).

- If $p^{\theta_v}_{3^{\mathsf{F}}} > 0$, we then have $x^{\theta_v}_{3^{\mathsf{F}},a_v} = 0$ by (ii), which means some action profile $(i, 3^{\mathsf{F}})$, $i \neq a_v$, would be chosen with some positive probability, in which case a follower of type θ_v would be better-off reporting truthfully instead of reporting θ_* — a contradiction.
- If $p^{\theta_v}_{2^{\mathsf{F}}} > 0$, we have $x^{\theta_v}_{2^{\mathsf{F}},a_v} = 0$ by (ii). Now in order for a type-θ_v follower to still find reporting θ_* optimal, we need $x^{\theta_v}_{2^{\mathsf{F}},i} = 0$ for all $i \in \{a_0, b_v\} \cup \mathcal{N}(v)$ (i.e., for all the actions corresponding to the first, third, and fourth rows of the payoff matrix shown in Figure 5.1) as otherwise truthful report would gain type θ_v a higher utility. Consequently, only the fifth row of the payoff matrix will be chosen by $\mathbf{x}^{\theta_v}_{2^{\mathsf{F}}}$ with a positive probability. Thus, we have $\mathsf{BR}_{\theta_v}(\mathbf{x}^{\theta_v}_{2^{\mathsf{F}}}) = \{3^{\mathsf{F}}\}$, which means that $(\mathbf{x}^{\theta_v}_{2^{\mathsf{F}}}, 2^{\mathsf{F}})$ cannot be a valid outcome — a contradiction to the assumption that $p^{\theta_v}_{2^{\mathsf{F}}} > 0$.

This completes the proof. □

The bound in the above inapproximability result is tight following an argument similar to that for Opt-Policy. Moreover, an MILP formulation for Opt-XPolicy can be obtained by modifying Program (5.3), and this yields a polynomial-time algorithm for Opt-XPolicy when the number of follower types is fixed. If we further impose IC constraints, the constraints remove the only integer variables y^{θ}_{β} in the MILP (variables p^{θ}_{j} are relaxed now), which results in a linear program; this establishes the tractability of Opt-XPolicy-IC. We sketch the results in Theorems 5.12–5.14 below.

Theorem 5.12. Opt-XPolicy *admits a polynomial-time* $\frac{1}{|\Theta|}$*-approximation algorithm.*

Proof sketch. The proof follows from arguments similar to those used to prove Theorem 5.7. □

Theorem 5.13. *For a fixed number of follower types,* Opt-XPolicy *can be solved in polynomial time.*

Proof sketch. We modify Program (5.3) for a formulation for Opt-XPolicy: replace every $\mathbf{x}^{\theta}$ by $\mathbf{x}^{\theta}_{j}$ wherever they are associated with p^{θ}_{j}; and relax every p^{θ}_{j} to be in $[0, 1]$. The modified program can be linearized into an MILP in the same way we linearize Program (5.3). By the same argument for Theorem 5.8, the MILP formulation yields a polynomial-time algorithm for Opt-XPolicy when the number of follower types is fixed. □

Theorem 5.14. Opt-XPolicy-IC *can be computed in polynomial time.*

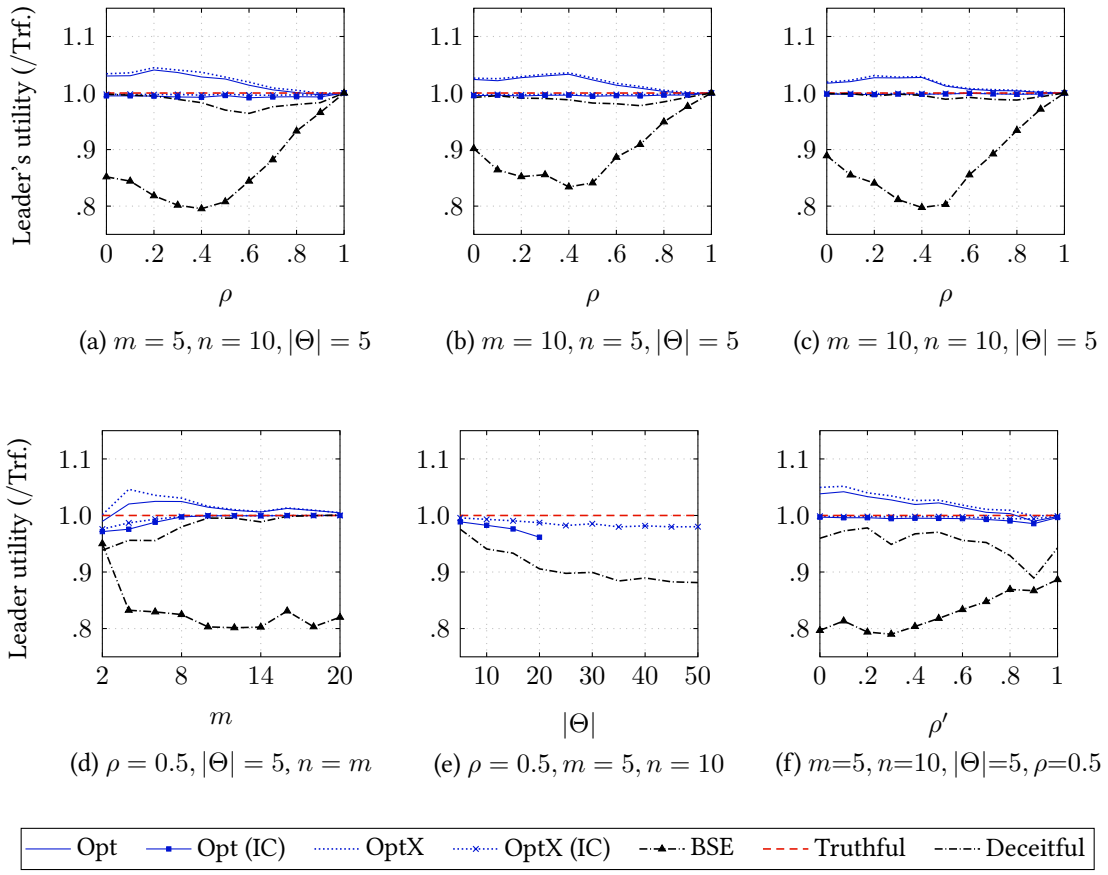
Proof sketch. The problem can be formulated as a linear program as follows, where $\tilde{\mathbf{x}}^{\theta}_{j}$ and p^{θ}_{j} are the variables ($j \in [n], \theta \in \Theta$).

$$
\begin{aligned}
\text{maximize} \quad & \textstyle\sum_{\theta} \sigma_{\theta} \sum_{j} u^{\mathsf{L}}(\tilde{\mathbf{x}}^{\theta}_{j}, j) && \\
\text{subject to} \quad & \textstyle\sum_{j} u^{\mathsf{F}}_{\theta}(\tilde{\mathbf{x}}^{\theta}_{j}, j) - \sum_{j} u^{\mathsf{F}}_{\theta}(\tilde{\mathbf{x}}^{\beta}_{j}, j) \geq 0 && \text{for all } \theta, \beta \in \Theta \\
& u^{\mathsf{F}}_{\theta}(\tilde{\mathbf{x}}^{\theta}_{j}, j) - u^{\mathsf{F}}_{\theta}(\tilde{\mathbf{x}}^{\theta}_{j}, k) \geq 0 && \text{for all } \theta \in \Theta,\ j, k \in [n] \\
& \textstyle\sum_{i} \tilde{x}^{\theta}_{ji} = p^{\theta}_{j} && \text{for all } \theta \in \Theta,\ j \in [n] \\
& \tilde{\mathbf{x}}^{\theta}_{j} \in \mathbb{R}^{m}_{\geq 0},\ (p^{\theta}_{1}, \ldots, p^{\theta}_{n}) \in \Delta^{n-1} && \text{for all } \theta, \beta \in \Theta,\ j \in [n]
\end{aligned}
$$

The strategies $\mathbf{x}^{\theta}_{j}$ in the leader policy can be extracted from the solution as $\mathbf{x}^{\theta}_{j} = \frac{1}{p^{\theta}_{j}} \cdot \tilde{\mathbf{x}}^{\theta}_{j}$. □

5.6 Empirical Evaluation

We evaluate our framework with games generated randomly using the well-known covariance game model [Nudelman et al., 2004]. We generate the players' payoffs uniformly at random from the range $[0, 1]$, and then adjust the follower's payoffs by blending them with the negative value of the leader's payoffs; the degree of blending is controlled by a parameter $\rho \in [0, 1]$, i.e., $u_\theta^{\mathsf{F}} \leftarrow (1-\rho) \cdot u_\theta^{\mathsf{F}} - \rho \cdot u^{\mathsf{L}}$. As such, the game is zero-sum when $\rho = 1$; and, when $\rho = 0$, payoffs of the leader and the follower are completely uncorrelated. We also generate the probabilities σ_θ uniformly at random.



(a) $m = 5, n = 10, |\Theta| = 5$ (b) $m = 10, n = 5, |\Theta| = 5$ (c) $m = 10, n = 10, |\Theta| = 5$

(d) $\rho = 0.5, |\Theta| = 5, n = m$ (e) $\rho = 0.5, m = 5, n = 10$ (f) m=5, n=10, $|\Theta|$=5, ρ=0.5

Figure 5.3: The leader's utility obtained with different approaches. All y-axes represent the leader's utility as a *ratio to the utility obtained in the truthful situation*. The missing data points are instances that are not solvable within one hour in our platform because of scalability issues. All results are averaged over 50 runs.

We compare the leader's utilities obtained by different policies we propose and by approaches from previous research, including: (1) Optimal pure policy (labeled *Opt*), with and without IC; (2) Optimal mixed policy (*OptX*), with and without IC; (3) Bayesian Stackelberg Equilibrium (*BSE*); (4) Optimal strategy of the leader when the follower truthfully reports his type (*Truthful*), and (5) the same strategy but when the follower strategically reports his type (*Deceitful*).

Figure 5.3 depicts our results. The first three figures, (a)–(c), show the variance of the leader's utility with ρ under different numbers of player actions. For ease of comparison, the results are

depicted as ratios to the leader's utility in the truthful situation. The figures exhibit quite similar patterns. All our proposed polices (lines in blue) improve the leader's utility significantly upon BSE. The utilities obtained are also very close to the truthful utility even when IC constraints are imposed. When IC is not required, the optimal policies even outperform the optimal truthful strategies in almost all experiments. Surprisingly, even when the leader naïvely trusts a deceitful follower, the results are very positive; the utility obtained is usually very close to the truthful utility. This is in contrast to our theoretical examples. An explanation is that in randomly generated games, the negative effect of follower deception is likely to be cancelled out by positive effect brought by benign follower types: as we have observed in Section 5.3, follower deception is not always bad for the leader.

The utility differences are most significant when ρ is around 0.5. Intuitively, this is because when ρ approaches 1 the diminishing uncertainty over follower types reduces the effect of follower deception. But when ρ is close to 0, payoffs of the leader and the follower are less correlated and their objectives less conflicted; hence, when the follower leverages deception to increase his utility, there is a lower chance that this will decrease the leader's utility.

We also vary the numbers of actions and of follower types and investigate the effects. From Figure 5.3 (d) we see that utility differences (except for BSE) quickly diminish when the number of actions increases. Similarly to the situation when ρ is close to 0, the level of randomness of the payoffs increases as the action space grows, which dissolves the negative effect of follower deception. In contrast, as shown in Figure 5.3 (e), the significance of follower deception on the leader's utility appears to increase with the number of follower types as the leader's utility drops in the deceitful setting. The utility offered by the optimal mixed IC policy, however, decreases much more slowly. Hence, the optimal mixed IC policy offers a scalable practical solution when there is a large number of follower types. At a high level, these results suggest that uncertainty over follower's payoff information is a major amplifier of the negative effect of follower deception.

To explore situations in which follower deception is likely to be more harmful, we slightly modify the covariance game model, generating the follower type set with multiple blending parameters. In Figure 5.3 (f), the results are obtained on a follower type set in which half of the types are generated with $\rho = 0.5$, and another half generated with ρ' as on the x-axis. Interestingly, the difference between the approaches suddenly becomes more distinguishable, and there is a clear gap between the optimal and the deceptive situations. An in-depth characterization of when deception tends to be more harmful to the leader and when it is not would be an interesting question for future work.

Robustness against Tie-breaking In our model, we assumed that the follower breaks ties in favor of the leader in selecting their reporting strategy. This assumption is one step further beyond the optimistic tie-breaking assumption in the SSE and poses potential inducibility issues, when a desired follower reporting strategy cannot be induced by any small perturbation. Our next set of simulation results experimentally demonstrates that, nevertheless, inducibility does *not* appear to be an issue in randomly generated instances. In the simulations, we enforce a small utility gap of at least ϵ for the follower between taking the induced action and any other action (the same with the induced reporting strategy). This removes all the ties, and makes the follower *strictly* prefer the action (by ϵ) the leader intends to induce. The results are depicted in Table 5.1, with ϵ varying from 10^{-5} to 10^{-1}. The experimental results show that our solution is robust to tie breaking and always maintains optimality for $\epsilon \leq 10^{-3}$ (recall that the follower's payoffs are in $[-1, 1]$). Moreover, desired follower actions are almost always inducible except for a small number of instances when

	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
Opt	1.00	1.00	1.00	0.97	0.63
Opt (IC)	1.00	1.00	1.00	0.97	0.61*
OptX	1.00	1.00	1.00	0.97	0.66
OptX (IC)	1.00	1.00	1.00	0.98	0.67
OptX (IC) - large	1.00	1.00	1.00	0.98	0.79

Table 5.1: The leader's utility yielded by robust solutions as a ratio to the utility under the optimistic tie-breaking assumption, with ϵ varying from 10^{-5} to 10^{-1}. Results of *OptX (IC) - large* are obtained on large instances with $\rho = 0.5$, $|\Theta| = 50$, $n = m = 20$; all other results are obtained on instances with $\rho = 0.5$, $|\Theta| = 5$, $n = 5$ and $m = 10$. Each entry is averaged over 100 runs. $(*)$ In 9 out of 100 runs in this setting there is no feasible solution.

$\epsilon = 0.1$.[1] These results empirically show that the issue of tie-breaking does not affect the validity of our approach in natural instances.

5.7 Extensions — Beyond Bayesian Analysis

Sometimes the leader faces the situation with no prior information about the distribution of follower types. In this case, worst-case (maximin) analysis is relevant; taking this perspective, the optimal policy should maximize the leader's utility against the worst possible follower type, i.e.,

$$\mathrm{WU}(\pi) = \min_{\theta \in \Theta} u^{\mathsf{L}} \left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)} \right).$$

where $\hat{\theta}(\pi) \in \arg\max_{\beta \in \Theta} u_\theta^{\mathsf{F}}(\mathbf{x}^\beta, j^\beta)$ as in (5.1). We call it Opt-Policy-WC the problem of computing a policy that maximizes $\mathrm{WU}(\pi)$.

In addition to the worst-case utility, we also propose an alternative measure of the quality of a policy, termed the *efficiency of the policy* (EoP), which takes into consideration the "hardness" of playing against each follower type in the truthful setting. As in Definition 5.15, the EoP is the worst-case *ratio* between the utility the leader obtains and what she should have obtained had the follower been truthful. A higher EoP indicates a smaller loss due to follower manipulation, and the value of the EoP always lies between 0 and 1 according to Proposition 5.16 (we shift all payoffs to be non-negative for EoP to be meaningful). We call it Opt-Policy-EoP the problem of computing a policy that maximizes the EoP.

Definition 5.15 (EoP)**.** The efficiency of a policy π on follower type θ is

$$\mathrm{EoP}_\theta(\pi) = u^{\mathsf{L}} \left(\mathbf{x}^{\hat{\theta}(\pi)}, j^{\hat{\theta}(\pi)} \right) / U(\theta),$$

where $\hat{\theta}(\pi)$ is a type-θ follower's best reporting strategy in response to π (tie-breaking in favor of maximizing the $\mathrm{EoP}_\theta(\pi)$); and $U(\theta) = \max_{\mathbf{x} \in \Delta^{m-1}, j \in \mathsf{BR}_\theta(\mathbf{x})} u^{\mathsf{L}}(\mathbf{x}, j)$ is the leader's utility in an

[1] Note that we list $\epsilon = 0.1$ only for illustrative purpose as in practice one would never want to use such a big ϵ to induce tie breaking. There are extended approaches able to generate feasible robust solutions for every $\epsilon \in [0, 1]$ (see [Pita et al., 2010]), but this is not our focus.

SSE against a type-θ follower. The (overall) efficiency of π is

$$\mathrm{EoP}(\pi) = \min_{\theta \in \Theta} \mathrm{EoP}_\theta(\pi).$$

Proposition 5.16. $\mathrm{EoP}(\pi) \in [0, 1]$ *for any policy* π.

Proof. Clearly, $\mathrm{EoP}(\pi) \geq 0$ as the payoffs are shifted to be non-negative. We show that $\mathrm{EoP}(\pi) \leq 1$.

Let $U(\theta) = \max_{\mathbf{x} \in \Delta^{m-1}, j \in \mathsf{BR}_\theta(\mathbf{x})} u^{\mathsf{L}}(\mathbf{x}, j)$ denote the leader's utility in an SSE on follower type θ, as in the statement of the definition; let β be a follower type that gives the leader the highest utility in an SSE, i.e.,

$$U(\beta) = \max_{\theta \in \Theta} U(\theta).$$

Consider the best reporting strategy γ of a type-β follower in response to π. For the outcome $(\mathbf{x}^\gamma, j^\gamma) = \pi(\gamma)$ to be feasible, we must have $j \in \mathsf{BR}_\gamma(\mathbf{x}^\gamma)$; thus,

$$u^{\mathsf{L}}(\mathbf{x}^\gamma, j^\gamma) \leq \max_{\mathbf{x} \in \Delta^{m-1}, j \in \mathsf{BR}_\gamma(\mathbf{x})} u^{\mathsf{L}}(\mathbf{x}, j) = U(\gamma) \leq U(\beta).$$

It follows that

$$\mathrm{EoP}(\pi) = \min_{\theta \in \Theta} \mathrm{EoP}_\theta(\pi) \leq \mathrm{EoP}_\beta(\pi) = u^{\mathsf{L}}(\mathbf{x}^\gamma, j^\gamma)/U(\beta) \leq 1. \qquad \square$$

5.7.1 Inapproximability with Worst-Case Analysis

It turns out that Opt-Policy-WC and Opt-Policy-EoP remain hard to approximate. We first show the hardness of Opt-Policy-EoP.

Theorem 5.17. *For any polynomial-time computable function* $\phi > 0$, Opt-Policy-EoP *does not admit a polynomial-time* $\phi(m, n, |\Theta|)$*-approximation algorithm unless P = NP, even when the follower has only three actions.*

Proof. We show that it is NP-hard to distinguish between a game instance that admits a policy π with $\mathrm{EoP}(\pi) \geq 1$ and an instance in which $\mathrm{EoP}(\pi) < \phi$ for any policy. We provide a reduction from the Vertex Cover problem, which is well-known to be NP-complete. A *vertex cover* V' of an undirected graph $G = (V, E)$ is a subset of V such that $v_1 \in V'$ or $v_2 \in V'$ for every edge $\{v_1, v_2\} \in E$. An instance of the Vertex Cover problem is given by a graph $G = (V, E)$ and an integer $k \leq |V|$. It is a yes-instance if there exists a vertex cover of G of size at most k.

For a Vertex Cover instance, we construct the following game and show that the Vertex Cover instance is a yes-instance only if there exists a leader policy π with $\mathrm{EoP}(\pi) \geq 1$, and it is a no-instance if no policy π achieves $\mathrm{EoP}(\pi) \geq \phi$. In the game, the leader has $|V| + 1$ actions $\{a_v : v \in V\} \cup \{a_0\}$. The follower has three actions $\{1^{\mathsf{F}}, 2^{\mathsf{F}}, 3^{\mathsf{F}}\}$. The set of possible follower types is $\Theta = \{\theta_1, \ldots, \theta_k\} \cup \{\theta_e : e \in E\} \cup \{\theta_0\}$. The payoffs are given in Figure 5.4.

We first make several observations about the SSEs of this game (in the truthful situation). Below we let $U(\theta)$ denote the leader's utility in an SSE when she plays against a type-θ follower, i.e., $U(\theta) = \max_{\mathbf{x} \in \Delta^{m-1}, j \in \mathsf{BR}(\mathbf{x})} u^{\mathsf{L}}(\mathbf{x}, j)$.

- The leader's utility only depends on the follower's action, with 3^{F} being the most detrimental action the leader would anyhow avoid the follower to choose, followed by 2^{F}, and 1^{F} is the most preferred follower action.

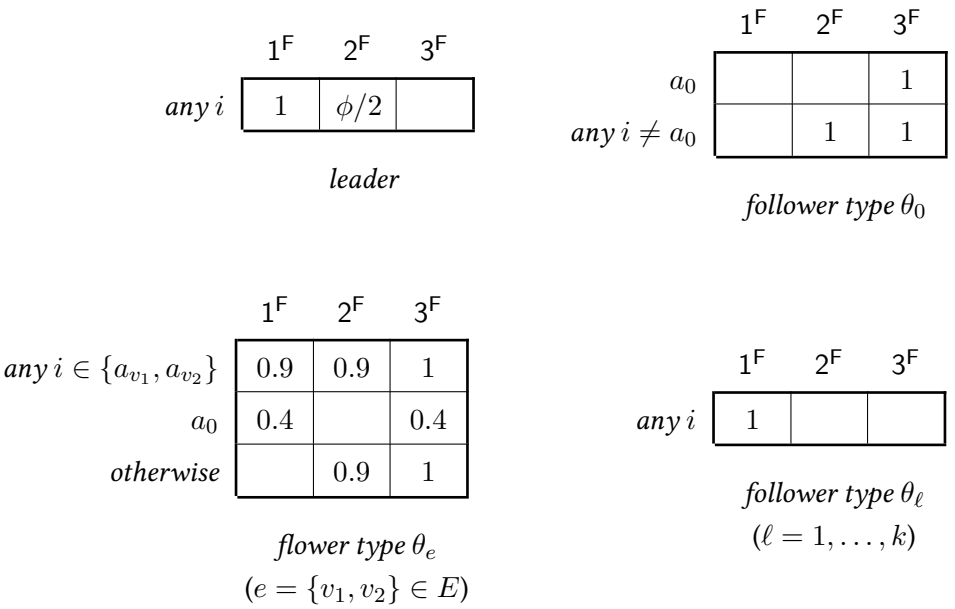


Figure 5.4: Payoff matrices (blank entries are 0).

- The *only* SSE strategy of the leader when she plays against a type-θ_e follower, $e \in E$, is the pure strategy a_0. When a_0 is played, the follower finds his best responses to be $\mathsf{BR}_{\theta_s}(a_0) = \{3^{\mathsf{F}}, 1^{\mathsf{F}}\}$ and breaks the tie in favor of the leader, playing 1^{F}; the leader obtains $u^{\mathsf{L}}(a_0, 1^{\mathsf{F}}) = 1$, which is obviously the highest possible utility she can obtain; hence, $(a_0, 1^{\mathsf{F}})$ forms an SSE and $U(\theta_e) = 1$. To see that this is the only SSE, observe that if the leader plays any other pure strategy with some probability, the follower would strictly prefer 3^{F} to 1^{F} and would not respond by playing 1^{F}, in which case the leader cannot get utility 1.
- Every strategy of the leader is an SSE strategy when the follower she plays against has type θ_ℓ, $\ell = 1, \ldots, k$, because the follower will always respond by playing 1^{F} irrespective of the strategy the leader plays, which always gives the leader utility 1. We have $U(\theta_\ell) = 1$.
- Every mixed strategy over pure strategies $i \neq a_0$ is an SSE strategy of the leader when the follower she plays against has type θ_0. When such a strategy is played, the follower finds his best response set to be $\{3^{\mathsf{F}}, 2^{\mathsf{F}}\}$ and breaks the tie in favor of the leader, playing 2^{F}; this is the best the leader can hope for because a type-θ_0 follower will never play 1^{F} as it is strictly dominated by 3^{F}. We have $U(\theta_0) = \frac{\phi}{2}$.

Suppose that $\langle G, k\rangle$ is a yes-instance; there exists a vertex cover $V' = \{v'_1, \ldots, v'_k\} \subseteq V$ of size k. The following leader policy π achieves EoP 1.

$$\pi(\theta) = \begin{cases} (a_0,\ 1^{\mathsf{F}}), & \text{for each } \theta \in \{\theta_e : e \in E\}; \\ (a_{v'_\ell},\ 1^{\mathsf{F}}), & \text{for each } \theta = \theta_\ell \in \{\theta_1, \ldots, \theta_k\}; \\ (a_{v'_1},\ 2^{\mathsf{F}}), & \text{for } \theta = \theta_0. \end{cases}$$

Clearly, all the outcomes prescribed are feasible, so π is a feasible policy. Further, it can be verified that when the leader commits to π, the optimal reporting strategy of the follower is the following.

- For every type-θ_e follower, $e \in E$, it is optimal to report a type θ_ℓ with $v'_\ell \in V'$ being an end point of e. Such a θ_ℓ always exists given that V' is a vertex cover. The leader obtains utility $u^{\mathsf{L}}(a_{v'_\ell}, 1^{\mathsf{F}}) = 1$ when θ_ℓ is reported, so $\mathrm{EoP}_{\theta_e}(\pi) \geq \frac{1}{U(\theta_e)} = 1$.
- For every type-θ_ℓ follower, $\ell \in \{1, \ldots, k\}$, it is optimal to report truthfully. The leader obtains utility $u^{\mathsf{L}}(a_{v'_\ell}, 1^{\mathsf{F}}) = 1$ when $\theta - \ell$ is reported, so $\mathrm{EoP}_{\theta_\ell}(\pi) \geq \frac{1}{U(\theta_\ell)} = 1$.
- For a type-θ_0 follower, it is optimal to report truthfully. The leader obtains utility $u^{\mathsf{L}}(a_{v'_1}, 2^{\mathsf{F}}) = \frac{\phi}{2}$ when θ_0 is reported, so $\mathrm{EoP}_{\theta_0}(\pi) \geq \frac{\phi/2}{U(\theta_0)} = 1$.

Therefore, $\mathrm{EoP}(\pi) = \min_{\theta \in \Theta} \mathrm{EoP}_\theta(\pi) \geq 1$.

Conversely, suppose that $\langle G, k \rangle$ is a no-instance. Suppose for the sake of contradiction that there exists a policy π offering $\mathrm{EoP}(\pi) \geq \phi$. Let β be the type that a type-θ_0 follower is incentivized to report in response to π (tie-breaking in favor of maximizing the EoP), and $\pi(\beta) = (\mathbf{x}^\beta, j^\beta)$. We argue that $j^\beta = 2^{\mathsf{F}}$ and $x^\beta_{a_0} = 0$. Indeed, if $j^\beta = 3^{\mathsf{F}}$ we would have $\mathrm{EoP}(\pi) \leq \mathrm{EoP}_{\theta_0}(\pi) = \frac{u^{\mathsf{L}}(\mathbf{x}^\beta, 3^{\mathsf{F}})}{U(\theta_0)} = 0 < \phi$; if $j^\beta = 1^{\mathsf{F}}$, or $j^\beta = 2^{\mathsf{F}}$ and $x^\beta_{a_0} > 0$, then by reporting β, a type-θ_0 follower obtains either $u^{\mathsf{F}}_{\theta_0}(\mathbf{x}^\beta, 1^{\mathsf{F}}) = 0$ or $u^{\mathsf{F}}_{\theta_0}(\mathbf{x}^\beta, 2^{\mathsf{F}}) < u^{\mathsf{F}}_{\theta_0}(a_0, 2^{\mathsf{F}}) = 1$, so he would be better off reporting truthfully, whereby he is guaranteed utility 1.

Therefore, a type-θ_e follower can already obtain utility 0.9 by reporting β. Let γ be the type that a type-θ_e follower is incentivized to report, $e = (v_1, v_2) \in E$, and $\pi(\gamma) = (\mathbf{x}^\gamma, j^\gamma)$; we must have $u^{\mathsf{F}}_{\theta_e}(\mathbf{x}^\gamma, j^\gamma) \geq 0.9$ for this follower to have incentive to report γ instead of β. At the same time, we need $u^{\mathsf{L}}(\mathbf{x}^\gamma, j^\gamma) \geq \phi$ in order for $\mathrm{EoP}_{\theta_0}(\pi) \geq \phi$; hence, the only possibility is $j^\gamma = 1^{\mathsf{F}}$. It follows that $x^\gamma_{a_{v_1}} + x^\gamma_{a_{v_2}} = 1$: otherwise, $u^{\mathsf{F}}_{\theta_e}(\mathbf{x}^\gamma, 1^{\mathsf{F}}) < 0.9$. Moreover, $\gamma \notin \{\theta_{e'} : e' \in E\}$: otherwise, $\mathsf{BR}_\gamma(\mathbf{x}^\gamma) = \{3^{\mathsf{F}}\} \not\ni 1^{\mathsf{F}}$ given that $x^\gamma_{a_0} = 0$, so $(\mathbf{x}^\gamma, 1^{\mathsf{F}})$ cannot be a feasible outcome. For the same reason, we also have $\gamma \neq \theta_0$, so the remaining possibility is that $\gamma = \theta_\ell$ for some $\ell \in \{1, \ldots, k\}$.

Initialize a set $V' = \varnothing$. For each $\ell = 1, \ldots, k$, suppose $\pi(\theta_\ell) = (\mathbf{x}, j)$; we let v^*_ℓ be the first vertex $v \in V$ (in lexicographical order) such that $x_{a_v} > 0$ if such a vertex indeed exists, and add v^*_ℓ into the set V'; eventually, $|V'| \leq k$. Since $x^\gamma_{a_{v_1}} + x^\gamma_{a_{v_2}} = 1$ and $\gamma = \theta_\ell$ for some $\ell \in \{1, \ldots, k\}$, we have $V' \cap e \neq \varnothing$. This holds for all $e \in E$; thus, V' forms a vertex cover of G and $|V'| \leq k$, which contradicts the assumption that $\langle G, k \rangle$ is a no-instance. □

We can adapt the above proof to show the inapproximability of Opt-Policy-WC. The problem is completely inapproximable, meaning that it is NP-hard to distinguish between an instance that admits a policy that achieves $\mathrm{WU}(\pi) > 0$ and one in which $\mathrm{WU}(\pi) = 0$ for every policy.

Theorem 5.18. Opt-Policy-WC *is completely inapproximable in polynomial time unless P = NP, even when the follower has only three actions.*

Proof sketch. We provide a reduction from the Vertex Cover problem. For a Vertex Cover instance, we construct a game similar to the one in the proof of Theorem 5.17: we remove follower type θ_0 and action 2^{F} and modify the payoffs to the ones shown in Figure 5.5.

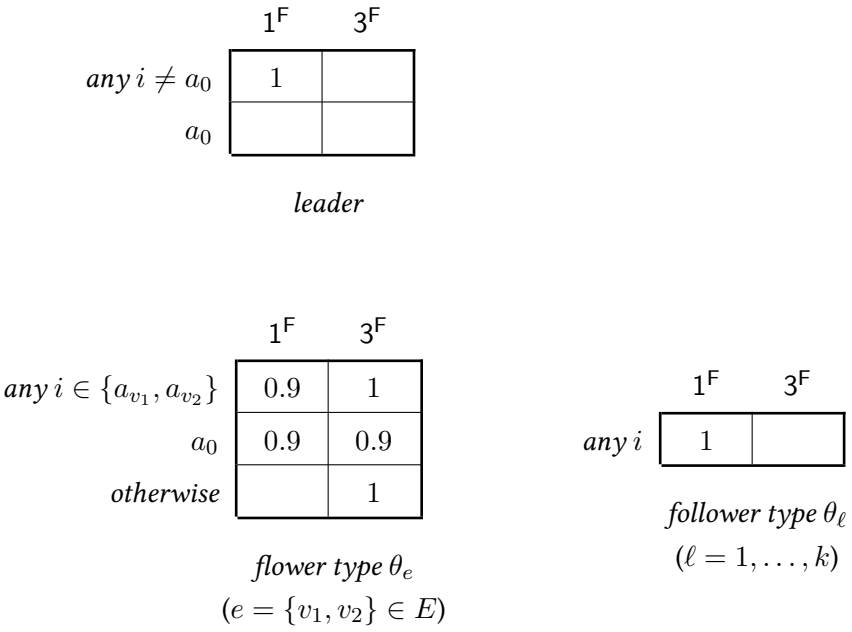


Figure 5.5: Payoff matrices (blank entries are 0).

It can be verified that, if there exists a vertex cover $V' = \{v'_1, \ldots, v'_k\} \subseteq V$ of size k, the following leader policy π achieves utility 1.

$$\pi(\theta) = \begin{cases} (a_0,\ 1^{\mathsf{F}}), & \text{for each } \theta \in \{\theta_e : e \in E\}; \\ (a_{v'_\ell},\ 1^{\mathsf{F}}), & \text{for each } \theta = \theta_\ell \in \{\theta_1, \ldots, \theta_k\}. \end{cases}$$

It is optimal for every type-θ_e follower to report θ_ℓ with $v'_\ell \in e$ and every type-θ_ℓ follower to report truthfully; the leader obtains utility 1 on every follower type.

Conversely, if there exists no vertex cover of size k, it can be shown that some type-θ_e follower will be incentivized to report truthfully no matter what policy the leader uses; the leader gets utility 0 as a result (indeed, the leader gets 0 in every SSE when the follower has type θ_e). □

5.7.2 Easiness Results and Open Questions

When IC constraints are imposed, the same LP-based approach (see Theorem 5.14) can be applied to solving Opt-Policy-WC and Opt-Policy-EoP if we allow the leader to use a mixed policy:

- For Opt-Policy-WC, we change the objective function in the LP to a variable μ and add a constraint $\mu \le \sum_{j\in[n]} u^{\mathsf{L}}(\tilde{\mathbf{x}}^\theta_j, j)$ for every $\theta \in \Theta$; the constraints will force μ to be equal to the leader's utility against the worst follower type.
- For Opt-Policy-EoP, we change the objective function in the LP to a variable μ and add a constraint $\mu \le \sum_{j\in[n]} u^{\mathsf{L}}(\tilde{\mathbf{x}}^\theta_j, j)/U(\theta)$ for every $\theta \in \Theta$, where the leader's SSE utility $U(\theta)$ can be pre-calculated using the standard technique in polynomial time. The constraints will force μ to be the EoP value of the policy.

We have not settled the complexity class for the IC variants of the Opt-Policy-WC and Opt-Policy-EoP when we restrict the leader policy to a pure one and we leave them as open questions for future work.

5.8 Summary

In this chapter, we point out potential occurrences of imitative follower deception in Stackelberg games and the risk when they are ignored. We then propose a framework to design leader policies against such deception and provide a systematic study of the computational aspects of this framework. Our results show that handling follower deception is hard in general, even when incentive compatibility constraints are imposed and when the leader is allowed to use mixed policies. While all the hardness results hold when the follower has only three actions, we do not know yet whether the hardness remains when the number of the follower's actions is reduced to two, and we leave this problem open.

Chapter 6

Deceiving the Leader Optimally

In the previous chapter, we have seen the problems follower deception can cause to a leader who learns to play a Stackelberg game. In this chapter, let us view this problem purely from the follower's perspective and examine how difficult it is for a follower to optimally deceive a leader (who is unaware of such deceptive behavior) through payoff matrix misreporting. In particular, in Chapter 5, this problem is largely simplified because of the matrices the follower can choose from are assumed to be in a finite set, the size of which is bounded by the size of the problem representation (thus, the follower can easily enumerate all the feasible choices to find out the optimal one). In a more general version of this problem, we allow the follower to choose any matrix in the space $\mathbb{R}^{m\times n}$ (for an m-by-n game). In this chapter, we solve this problem by showing that it is always possible for the follower to compute the optimal (or near-optimal) payoffs efficiently for various scenarios of learning interaction between the players.

6.1 Introduction

We explore how a follower can optimally deceive a leader in Stackelberg games by misreporting his payoff matrix, and study the tractability of the corresponding optimization problem. Our objective is to compute a fake payoff matrix according to which the follower can best respond, to make the leader learn an SSE in which the true utility of the follower is maximized. We do not impose any restrictions on the space from which the payoffs are selected or on the type of the game. By exploiting an intuitive characterization of all strategy profiles that can be induced as SSEs in Stackelberg games, we show that it is always possible for the follower to compute in polynomial time an optimal payoff matrix to misreport. Towards deriving this result, we also discovered that, interestingly, the follower is able to make a strategy profile $(\mathbf{x}, j)$ an SSE in a game parameterized u^{L} and some distorted follower payoff matrix $\tilde{u}^{\mathsf{F}} \in \mathbb{R}^{m\times n}$ as long as the leader obtains at least his maximin utility in $(\mathbf{x}, j)$.

6.1.1 Chapter Overview

Our approach involves defining several inducibility notions for an SSE. In the weakest form, the follower is allowed to use an arbitrary best-response (BR) correspondence, even the ones that cannot be realized by any payoff matrix; we present in Section 6.3 ways to construct the optimal BR correspondences for the follower in polynomial time. Then in Section 6.4, we move a step further, and require the best-response correspondence to be indeed implementable by a payoff matrix. The result we derive for arbitrary best-response correspondences helps us to construct an optimal payoff

matrix for the follower to misreport, in a more involved way, but still in polynomial time. Finally, in Section 6.5, we further strengthen our result by resolving possible equilibrium selection issues that arise when the payoff matrix (or BR correspondence) the follower misreports corresponds to multiple SSEs. We show that a payoff matrix that induces a *unique* SSE can be constructed in polynomial time; in this SSE the follower's utility is maximized up to some arbitrarily small loss.

6.2 Preliminaries

Similarly to Chapter 5, we consider an $m \times n$ Stackelberg game where the leader's and the follower's payoff matrices are given by u^{L} and u^{F}, respectively. We denote such a game by $\mathcal{G} = (u^{\mathsf{L}}, u^{\mathsf{F}})$ (we omit m and n as they are clear from context). A mixed strategy of the leader is a probability distribution over $[m]$, denoted by $\mathbf{x} \in \Delta^{m-1} = \{\mathbf{x} \geq 0 : \sum_{i\in[m]} x_i = 1\}$. We let $u^{\mathsf{L}}(\mathbf{x}, j) = \sum_{i\in[m]} x_i \cdot u^{\mathsf{L}}(i, j)$ be the *expected utility* of the leader when she plays the mixed strategy $\mathbf{x}$ and the follower plays a pure strategy j. Similarly, let $u^{\mathsf{F}}(\mathbf{x}, j) = \sum_{i\in[m]} x_i \cdot u^{\mathsf{F}}(i, j)$ be the expected utility of the follower.

For a given mixed strategy $\mathbf{x} \in \Delta^{m-1}$ of the leader, we say that $j \in [n]$ is a *follower best response* if $u^{\mathsf{F}}(\mathbf{x}, j) = \max_{\ell\in[n]} u^{\mathsf{F}}(\mathbf{x}, \ell)$. We use our previously defined notation and denote the set of all follower best responses to $\mathbf{x}$ by $\mathsf{BR}(\mathbf{x}) \subseteq [n]$. In this chapter, we will study a richer set of BR functions which the follower can use to deceive the leader; some of these functions may not actually arise from the follower's best responses as in a bi-matrix game, so we will refer to them as the *best response correspondence* of the follower. Recall that the standard solution concept for Stackelberg games, the SSE (strong Stackelberg equilibrium), is defined as follows.

Definition 6.1 (Strong Stackelberg equilibrium)**.** A strategy profile $(\mathbf{x}, j)$ is an SSE of the Stackelberg game $\mathcal{G} = (u^{\mathsf{L}}, u^{\mathsf{F}})$ if

$$(\mathbf{x}, j) \in \arg\max_{\mathbf{y}\in\Delta^{m-1},\, \ell\in\mathsf{BR}(\mathbf{y})} u^{\mathsf{L}}(\mathbf{y}, \ell).$$

6.2.1 Learning an SSE and Follower Deception

We follow the research question we put forward in Chapter 5 and consider the scenario where the leader has full knowledge of her utility matrix u^{L}, and aims to compute an SSE by interacting with the follower and gleaning information about u^{F}. For example, the leader could observe follower best responses in play (akin to having query access to BR), or observe follower payoffs at pure strategy profiles during play (akin to having query access to u^{F} as a function). Hence, this can be cast as the problem of learning an SSE with a specified notion of query access to information about the follower's incentives.

Consider a Stackelberg game $\mathcal{G} = (u^{\mathsf{L}}, u^{\mathsf{F}})$. If the follower controls the flow of information to the leader in this paradigm, he may consider perpetually interacting with the leader as if he had a different payoff matrix $\tilde{u}^{\mathsf{F}}$, which can make the leader believe that both agents are playing the game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \tilde{u}^{\mathsf{F}})$. This deceiving power provides the follower with an incentive to act according to $\widetilde{\mathcal{G}}$ for a judicious choice of $\tilde{u}^{\mathsf{F}}$, because the SSEs in $\widetilde{\mathcal{G}}$ may provide larger utility (according to u^{F}) than the SSEs in $\mathcal{G}$. More concretely, the example below shows that the follower can gain an arbitrary benefit by deceiving the leader to play a different game.

Example 6.2 (Beneficial deception). Let $\alpha \in [0, 1]$ and consider the following matrices:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad C_\alpha = \begin{pmatrix} 0 & \alpha \\ 1 & \alpha \end{pmatrix}$$

Now, suppose that $u^{\mathsf{L}} = R$ and $u^{\mathsf{F}} = C_\alpha$, and let $x \in [0, 1]$ represent the probability mass that the leader (row player) places on the first row (her first strategy); thus, $1 - x$ is the probability with which she plays her second strategy. Given this mixed strategy of the leader, the utilities that the follower expects to derive from her two strategies (columns) are $u^{\mathsf{F}}(x, 1) = 1 - x$ and $u^{\mathsf{F}}(x, 2) = \alpha$. Consequently, the first strategy is a best response of the follower when $x \in [0, 1 - \alpha]$, and the second one is a best response when $x \in (1 - \alpha, 1]$ (when $x = 1 - \alpha$, the tie is broken in favor of the leader). With this information, it is clear that the SSE of the game occurs when the leader chooses $x = 1 - \alpha$ and the follower plays his first strategy. As a result, the follower's utility is $u^{\mathsf{F}}(1 - \alpha, 1) = \alpha$.

However, for any $\alpha < 1$, the follower has an incentive to deceive the leader into playing the game $\widetilde{\mathcal{G}} = (R, C_1)$, which will improve his utility in the resulting SSE to 1. This will be an improvement by a multiplicative factor of $1/\alpha$, which can be arbitrarily large when α is arbitrarily close to 0. □

6.2.2 Inducible Strategy Profiles

The ultimate goal of the follower is to identify the SSE that maximizes his true utility, from the set of SSEs that he can deceive the leader into learning. We will refer to such SSEs as *inducible strategy profiles*. At a high level, the follower's problem can now be expressed as the following optimization problem:

$$\begin{aligned} \max_{\mathbf{x}, j} \quad & u^{\mathsf{F}}(\mathbf{x}, j), \\ \text{subject to} \quad & (\mathbf{x}, j) \text{ is inducible} \end{aligned} \tag{6.1}$$

This maximum utility for the follower is called the *optimal inducible utility*. If the maximum value is never achieved, then for every $\varepsilon > 0$, we would like to be able to find an inducible SSE that achieves a value ε-close to the supremum value.

As discussed previously, the leader can learn an SSE by gleaning information about the incentives of the follower by querying the follower's best responses to particular strategies of the leader, or more refined information about the follower's payoff matrix. Depending on the type of information queried, we can define various notions of inducible strategy profiles.

In more detail, suppose the leader can only query the best responses of the follower, who behaves according to some best response correspondence $\widetilde{\mathsf{BR}} : \Delta^{m-1} \to 2^{[n]} \setminus \{\varnothing\}$. This interaction between the leader and the follower leads to a game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \widetilde{\mathsf{BR}})$ where only information about $\widetilde{\mathsf{BR}}$ is known (instead of a payoff matrix implying $\widetilde{\mathsf{BR}}$). The definition of $\widetilde{\mathsf{BR}}$ enforces a best response answer to any possible query. Consequently, the leader learns an SSE $(\mathbf{x}, j) \in \arg\max_{\mathbf{y} \in \Delta^{m-1},\, \ell \in \widetilde{\mathsf{BR}}(\mathbf{y})} u^{\mathsf{L}}(\mathbf{y}, \ell)$, which yields the following notion of *BR-inducible* strategy profiles.

Definition 6.3 (BR-inducibility)**.** A strategy profile $(\mathbf{x}, j)$ is *BR-inducible* with respect to u^{L} if there exists a best response correspondence $\widetilde{\mathsf{BR}} : \Delta^{m-1} \to 2^{[n]} \setminus \{\varnothing\}$ such that $(\mathbf{x}, j)$ is an SSE of the game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \widetilde{\mathsf{BR}})$, in which case we say that $(\mathbf{x}, j)$ is *induced* by $\widetilde{\mathsf{BR}}$.

Next, consider the case where the leader can query information about the payoffs of the follower, who can now behave according to a fake payoff matrix $\tilde{u}^{\mathsf{F}}$. We refer to the SSEs of the resulting game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \tilde{u}^{\mathsf{F}})$ as *payoff-inducible* strategy profiles.

Definition 6.4 (Payoff-inducibility)**.** A strategy profile $(\mathbf{x}, j)$ is said to be *payoff-inducible* with respect to u^{L} if there exists $\tilde{u}^{\mathsf{F}} \in \mathbb{R}^{m \times n}$ such that $(\mathbf{x}, j)$ is an SSE in the game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \tilde{u}^{\mathsf{F}})$, in which case we say that $(\mathbf{x}, j)$ is *induced* by $\tilde{u}^{\mathsf{F}}$.

Clearly, payoff-inducibility is stricter than BR-inducibility: for every choice of $\tilde{u}^{\mathsf{F}}$, the corresponding best response correspondence $\widetilde{\mathsf{BR}}(\mathbf{y}) = \arg\max_{\ell \in [n]} \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell)$ induces the same SSEs as $\tilde{u}^{\mathsf{F}}$ does.

Note that the above definitions only require an inducible strategy profile to be a verifiable SSE, with respect to the information about the follower's incentives (either $\widetilde{\mathsf{BR}}$ or $\tilde{u}^{\mathsf{F}}$). However, it may happen that the resulting game $\widetilde{\mathcal{G}}$ has multiple SSEs, which gives rise to an equilibrium selection issue. Indeed, in practice, it is not realistic to assume that the follower has any control over which SSE is chosen by the leader (who moves first in the game). To address this, and thus completely resolve the optimal deception problem for the follower, we introduce an even stricter notion of inducibility on top of payoff-inducibility, which requires $\widetilde{\mathcal{G}}$ to have a unique SSE.

Definition 6.5 (Strong inducibility)**.** A strategy profile $(\mathbf{x}, j)$ is said to be *strongly inducible* with respect to u^{L}, if there exists a matrix $\tilde{u}^{\mathsf{F}} \in \mathbb{R}^{m \times n}$ such that $(\mathbf{x}, j)$ is the *unique* SSE of the game $\widetilde{\mathcal{G}} = (u^{\mathsf{L}}, \tilde{u}^{\mathsf{F}})$, in which case we say that $(\mathbf{x}, j)$ is *strongly induced* by $\tilde{u}^{\mathsf{F}}$.

In the next sections, we will investigate solutions to (6.1) under the inducibility notions above, from the weakest to the strongest. Our general approach is to decompose (6.1) into n sub-problems by enumerating all possible responses $j \in [n]$ of the follower. For each strategy j, we solve the corresponding optimization problem, and pick the one that yields the maximum utility for the follower.

6.3 Best Response Inducibility

Let us start our analysis by considering the case in which the leader queries the best responses of the follower. The aim of the follower is to deceive the leader towards a strategy profile that is BR-inducible; see Definition 6.3. Indeed, if the follower is allowed to use an arbitrary $\widetilde{\mathsf{BR}}$ to induce a strategy profile $(\mathbf{x}, j)$, he can simply define $\widetilde{\mathsf{BR}}$ as follows:

$$\widetilde{\mathsf{BR}}(\mathbf{y}) = \begin{cases} \{j\} & \text{if } \mathbf{y} = \mathbf{x} \\ \arg\min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell) & \text{if } \mathbf{y} \neq \mathbf{x}. \end{cases}$$

Namely, the follower threatens to choose the worst possible response against any strategy $\mathbf{y} \neq \mathbf{x}$ of the leader, so as to minimize the leader's incentive to commit to these strategies. This $\widetilde{\mathsf{BR}}$ will successfully convince the leader that $(\mathbf{x}, j)$ is an SSE of $\widetilde{\mathcal{G}}$, hence inducing $(\mathbf{x}, j)$, if the threat is powerful enough, that is, if $u^{\mathsf{L}}(\mathbf{x}, j) \geq \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell)$ for all $\mathbf{y} \in \Delta^{m-1}$. Equivalently, this means that

$$u^{\mathsf{L}}(\mathbf{x}, j) \geq M := \max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell), \tag{6.2}$$

where M is exactly the leader's *maximin utility*. Indeed, (6.2) is necessary for $(\mathbf{x}, j)$ to be BR-inducible: if on the contrary $u^{\mathsf{L}}(\mathbf{x}, j) < M$, then by committing to $\mathbf{y}^* \in$

$\arg\max_{\mathbf{y}\in\Delta^{m-1}} \min_{\ell\in[n]} u^{\mathsf{L}}(\mathbf{y},\ell)$, the leader can obtain (at least) her maximin utility, which will be strictly larger than $u^{\mathsf{L}}(\mathbf{x},j)$.

Thus, condition (6.2) gives a simple criterion for BR-inducibility. The problem is that such $\widetilde{\mathsf{BR}}$ may be far from being one that arises from a choice of $\tilde{u}^{\mathsf{F}}$. To alleviate this limitation, we impose a stricter condition on $\widetilde{\mathsf{BR}}$, which can also be viewed as an intermediate step towards our goal of studying payoff inducibility.

Polytopal BR Correspondence In a similar vein to the work of Goldberg and Marmolejo-Cossío [2018], we require that, for every $\ell \in [n]$, the set of the leader's strategies to which ℓ is a best response $\widetilde{\mathsf{BR}}^{-1}(\ell) = \{\mathbf{y} \in \Delta^{m-1} : \ell \in \widetilde{\mathsf{BR}}(\mathbf{y})\}$ is a closed convex polytope, and the union of all these sets forms a partition of Δ^{m-1} (for example, see the polytope partition of Δ^2 in Figure 6.1). Any best response correspondence $\widetilde{\mathsf{BR}}$ satisfying this assumption is called *polytopal*.

Definition 6.6 (Polytopal best response correspondence [Goldberg and Marmolejo-Cossío, 2018])**.** A best response correspondence $\widetilde{\mathsf{BR}} : \Delta^{m-1} \to 2^{[n]} \setminus \{\varnothing\}$ is *polytopal* if it also satisfies the following:

- $\widetilde{\mathsf{BR}}^{-1}(\ell)$ is a closed convex polytope for each $\ell \in [n]$, and
- For each $k \neq \ell$, either $\mathsf{relint}(\widetilde{\mathsf{BR}}^{-1}(k)) \cap \mathsf{relint}(\widetilde{\mathsf{BR}}^{-1}(\ell)) = \varnothing$ or $\widetilde{\mathsf{BR}}^{-1}(k) = \widetilde{\mathsf{BR}}^{-1}(\ell)$, where $\mathsf{relint}(H)$ denotes the relative interior of a set H.

Being polytopal is necessary for $\widetilde{\mathsf{BR}}$ to arise from some payoff matrix. Indeed, the *true* best response correspondence BR that arises from u^{F} is polytopal: clearly, each $\mathsf{BR}^{-1}(\ell)$ is a closed convex polytope defined by the hyperplanes $u^{\mathsf{F}}(\mathbf{y},\ell) \geq u^{\mathsf{F}}(\mathbf{y},k)$ for all $k \in [n]$ and the borders of Δ^{m-1}; in addition, $\cup_{\ell=1}^{n} \mathsf{BR}^{-1}(\ell) = \Delta^{m-1}$, and for any $\ell \neq k$, the polytopes $\mathsf{BR}^{-1}(\ell)$ and $\mathsf{BR}^{-1}(k)$ only intersect at their borders unless $u^{\mathsf{F}}(\cdot,\ell) = u^{\mathsf{F}}(\cdot,k)$. Thus, if the follower attempts to deceive the leader via a fake $\widetilde{\mathsf{BR}}$, the leader might spot the deception in case $\widetilde{\mathsf{BR}}$ is not polytopal.

It turns out that the following correspondence, which we denote as $\widetilde{\mathsf{BR}}_{\mathrm{P}}$, is polytopal and, as we will shortly show, it is in fact as powerful as any best response correspondence.

$$\widetilde{\mathsf{BR}}_{\mathrm{P}}(\mathbf{y}) = \begin{cases} \{j\} & \text{if } \mathbf{y} \in \Delta^{m-1} \setminus \overline{U_j(\mathbf{x})} \\ \{j\} \cup \arg\min_{\ell\in[n]\setminus\{j\}} u^{\mathsf{L}}(\mathbf{y},\ell) & \text{if } \mathbf{y} \in \overline{U_j(\mathbf{x})} \setminus U_j(\mathbf{x}) \\ \arg\min_{\ell\in[n]\setminus\{j\}} u^{\mathsf{L}}(\mathbf{y},\ell) & \text{if } \mathbf{y} \in U_j(\mathbf{x}) \end{cases}$$

where $\overline{U_j(\mathbf{x})}$ is the closure of $U_j(\mathbf{x}) = \left\{\mathbf{y} \in \Delta^{m-1} : u^{\mathsf{L}}(\mathbf{y},j) > u^{\mathsf{L}}(\mathbf{x},j)\right\}$, and $\mathbf{x}$ is the leader's strategy that we want to induce.[1] Intuitively, it is safe for the follower to respond by playing j against any strategy $\mathbf{y}$ of the leader if $u^{\mathsf{L}}(\mathbf{y},j) \leq u^{\mathsf{L}}(\mathbf{x},j)$, in which case the leader does not have a strong incentive to commit to $\mathbf{y}$ instead of $\mathbf{x}$. In response to the other strategies, however, the follower needs to play a different strategy in order to minimize the leader's incentive to commit to such a $\mathbf{y}$. Therefore, this approach will successfully induce $(\mathbf{x},j)$ if and only if the following holds:

$$u^{\mathsf{L}}(\mathbf{x},j) \geq \max_{\mathbf{y}\in\overline{U_j(\mathbf{x})}} \min_{\ell\in[n]\setminus\{j\}} u^{\mathsf{L}}(\mathbf{y},\ell), \tag{6.3}$$

[1] Note that the use of $\overline{U_j(\mathbf{x})}$, instead of the set $\left\{\mathbf{y} \in \Delta^{m-1} : u^{\mathsf{L}}(\mathbf{y},j) \geq u^{\mathsf{L}}(\mathbf{x},j)\right\}$, is important: when $u^{\mathsf{L}}(\mathbf{y},j) = u^{\mathsf{L}}(\mathbf{x},j)$ for all $\mathbf{y} \in \Delta^{m-1}$, these two sets define different behaviors.

where we use the convention that $\max \varnothing = -\infty$. It is easy to see that $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ is indeed polytopal: $\widetilde{\mathsf{BR}}_{\mathrm{P}}^{-1}(j) = \Delta^{m-1} \setminus U_j(\mathbf{x})$ is a closed convex polytope, and the same holds for the sets $\widetilde{\mathsf{BR}}_{\mathrm{P}}^{-1}(\ell)$ defined by the hyperplanes $u^{\mathsf{L}}(\mathbf{y}, \ell) \leq u^{\mathsf{L}}(\mathbf{y}, k)$, $k \in [n] \setminus \{j\}$ and the borders of $\overline{U_j(\mathbf{x})}$, which further form a partition of $\overline{U_j(\mathbf{x})}$.

In fact, (6.2) is equivalent to (6.3), meaning that the extra condition imposed on $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ does not compromise its power: if $(\mathbf{x}, j)$ can be induced by an arbitrary $\widetilde{\mathsf{BR}}$ then it can also be induced by $\widetilde{\mathsf{BR}}_{\mathrm{P}}$. We state this result in Lemma 6.7.

Lemma 6.7. *$u^{\mathsf{L}}(\mathbf{x}, j) \geq M$ if and only if $u^{\mathsf{L}}(\mathbf{x}, j) \geq \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell)$.*

Proof. Recall that we want to show that $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$ if and only if

$$u^{\mathsf{L}}(\mathbf{x}, j) \geq \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell) \tag{6.4}$$

where M is the maximin utility of the leader.

We show that (6.4) does not hold if and only if $u^{\mathsf{L}}(\mathbf{x}, j) < M$. Suppose that (6.4) does not hold. Then $u^{\mathsf{L}}(\mathbf{x}, j) < \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell)$ by definition, which implies that $U_j(\mathbf{x}) \neq \varnothing$. By the continuity of $\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\cdot, \ell)$, there exists $\mathbf{y}^* \in U_j(\mathbf{x})$ such that

$$u^{\mathsf{L}}(\mathbf{x}, j) < \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}^*, \ell).$$

By the definition of $U_j(\mathbf{x})$, we also have $u^{\mathsf{L}}(\mathbf{x}, j) < u^{\mathsf{L}}(\mathbf{y}^*, j)$. Thus,

$$u^{\mathsf{L}}(\mathbf{x}, j) < \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}^*, \ell) \leq \max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell) = M.$$

Conversely, suppose that $u^{\mathsf{L}}(\mathbf{x}, j) < M$. Let $\mathbf{y}^* \in \arg\max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell)$. Thus, $M = \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}^*, \ell)$, and we have

$$u^{\mathsf{L}}(\mathbf{x}, j) < M = \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}^*, \ell) \leq u^{\mathsf{L}}(\mathbf{y}^*, j)$$

which implies that $\mathbf{y}^* \in U_j(\mathbf{x})$. It follows that $M = \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell)$ and thus

$$u^{\mathsf{L}}(\mathbf{x}, j) < \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n]} u^{\mathsf{L}}(\mathbf{y}, \ell) \leq \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell),$$

so (6.4) does not hold. □

Using Lemma 6.7, we can efficiently compute the best strategy profile that can be induced by $\widetilde{\mathsf{BR}}_{\mathrm{P}}$, simply by solving the following LP for each $j \in [n]$:

$$\begin{aligned} \max_{\mathbf{x} \in \Delta^{m-1}} \quad & u^{\mathsf{F}}(\mathbf{x}, j) \\ \text{subject to} \quad & u^{\mathsf{L}}(\mathbf{x}, j) \geq M \end{aligned} \tag{6.5}$$

At this point, it might be tempting to think that with the polytopal constraint imposed, we would also be able to construct an explicit payoff matrix $\tilde{u}^{\mathsf{F}}$ to implement $\widetilde{\mathsf{BR}}_{\mathrm{P}}$. Unfortunately, this is not true as Example 6.8 illustrates. Surprisingly though, in the next section we will show that, even though we cannot construct a payoff matrix that implements $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ directly, every strategy profile $(\mathbf{x}, j)$ that is $\widetilde{\mathsf{BR}}_{\mathrm{P}}$-inducible is in fact payoff-inducible. We also present an efficient algorithm for computing a payoff matrix $\tilde{u}^{\mathsf{F}}$ to induce such $(\mathbf{x}, j)$.

$$u^{\mathsf{L}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1/2 & 1/2 \\ 1 & 1/2 & -1/2 \end{pmatrix}$$

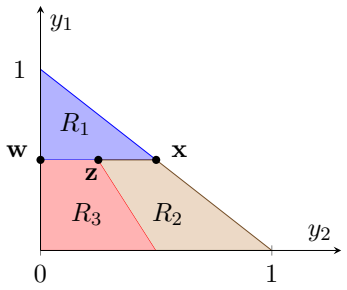


Figure 6.1: No payoff matrix $\tilde{u}^{\mathsf{F}}$ realizes the polytopal BR correspondence $\widetilde{\mathsf{BR}}_{\mathrm{P}}$, such that $\ell \in \widetilde{\mathsf{BR}}_{\mathrm{P}}$ if and only if $\mathbf{y} \in R_\ell$.

Example 6.8. Consider a 3×3 game with the leader payoff matrix given in Figure 6.1. Let $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ be a polytopal BR correspondence defined by the regions R_1, R_2, and R_3 in Figure 6.1, such that $\ell \in \widetilde{\mathsf{BR}}_{\mathrm{P}}$ if and only if $\mathbf{y} \in R_\ell$, where

- $R_1 = \{\mathbf{y} \in \Delta^2 : y_1 \geq y_2 + y_3\}$,
- $R_2 = \{\mathbf{y} \in \Delta^2 : y_1 \leq y_2 + y_3 \text{ and } y_2 \geq y_3\}$, and
- $R_3 = \{\mathbf{y} \in \Delta^2 : y_1 \leq y_2 + y_3 \text{ and } y_2 \leq y_3\}$.

This best response behavior cannot be realized by any payoff matrix.

To see this, suppose $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ is realized by some $\tilde{u}^{\mathsf{F}} \in \mathbb{R}^{3\times 3}$. Let $\mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\mathbf{w} = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$, and $\mathbf{z} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$. We have $\widetilde{\mathsf{BR}}_{\mathrm{P}}(\mathbf{z}) = \{1, 2, 3\}$ and $\widetilde{\mathsf{BR}}_{\mathrm{P}}(\mathbf{w}) = \{1, 3\}$. This means that $u^{\mathsf{L}}(\mathbf{z}, 1) = u^{\mathsf{L}}(\mathbf{z}, 3) = u^{\mathsf{L}}(\mathbf{z}, 2)$ and $u^{\mathsf{L}}(\mathbf{w}, 1) = u^{\mathsf{L}}(\mathbf{w}, 3) > u^{\mathsf{L}}(\mathbf{w}, 2)$. Since $\mathbf{x} = 2\mathbf{z} - \mathbf{w}$, by the linearity of the utility function, $u^{\mathsf{L}}(\mathbf{x}, 1) = u^{\mathsf{L}}(\mathbf{x}, 3) < u^{\mathsf{L}}(\mathbf{x}, 2)$, which contradicts the fact that $\widetilde{\mathsf{BR}}_{\mathrm{P}}(\mathbf{x}) = \{1, 2\}$. □

6.4 Payoff Inducibility

In this section, we will show that every profile strategy that can be induced by $\widetilde{\mathsf{BR}}_{\mathrm{P}}$ is also payoff-inducible, and a corresponding payoff matrix can be efficiently constructed. Recall that $M = \max_{\mathbf{y}\in\Delta^{m-1}} \min_{\ell\in[n]} u^{\mathsf{L}}(\mathbf{y}, \ell)$ is the maximin utility of the leader. We will show the following characterization as one of our key results, which enables us to use the LP in (6.5) to efficiently compute a payoff matrix that achieves the optimal inducible utility.

Theorem 6.9. *A strategy profile $(\mathbf{x}, j)$ is payoff-inducible if and only if $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$. Furthermore, a matrix $\tilde{u}^{\mathsf{F}}$ inducing $(\mathbf{x}, j)$ can be constructed in polynomial time.*

One direction of the characterization is easy to show. Indeed, if $(\mathbf{x}, j)$ is payoff-inducible, then it is also BR-inducible, and as seen in Section 6.3, it holds that $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$.

Now consider any profile $(\mathbf{x}, j)$ such that $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$. Recall that $U_j(\mathbf{x}) = \{\mathbf{y} \in \Delta^{m-1} : u^{\mathsf{L}}(\mathbf{y}, j) > u^{\mathsf{L}}(\mathbf{x}, j)\}$. Without loss of generality, in what follows, we can also assume that $U_j(\mathbf{x}) \neq \varnothing$: if $U_j(\mathbf{x}) = \varnothing$, then $(\mathbf{x}, j)$ will be an SSE if the follower always responds by playing j; this can easily be achieved by claiming that j strictly dominates all other strategies, i.e., by letting $\tilde{u}^{\mathsf{F}}(i, j) = 1$ and $\tilde{u}^{\mathsf{F}}(i, \ell) = 0$ for all $i \in [m]$ and $\ell \in [n] \setminus \{j\}$.

We begin by analyzing the following payoff function that forms the basis of our approach. Let $\widehat{S} \subseteq [n] \setminus \{j\}$ and pick $k \in \arg\min_{\ell \in \widehat{S}} u^{\mathsf{L}}(\mathbf{x}, \ell)$ arbitrarily. For all $\mathbf{y} \in \Delta^{m-1}$, let

$$\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) = \begin{cases} -u^{\mathsf{L}}(\mathbf{y}, \ell) & \text{if } \ell \in \widehat{S} \\ -u^{\mathsf{L}}(\mathbf{y}, k) - 1 & \text{if } \ell \in [n] \setminus (\widehat{S} \cup \{j\}) \\ -u^{\mathsf{L}}(\mathbf{y}, k) + \alpha \left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right) & \text{if } \ell = j \end{cases} \tag{6.6}$$

where $\alpha > 0$ is a constant. In what follows, we will let $\widetilde{\mathsf{BR}}$ denote the best response correspondence corresponding to $\tilde{u}^{\mathsf{F}}$, i.e., $\widetilde{\mathsf{BR}}(\mathbf{y}) = \arg\max_{\ell \in [n]} \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell)$. Note that we can compute the payoff matrix corresponding to $\tilde{u}^{\mathsf{F}}$ in polynomial time. Then, the hope is that with appropriately chosen $\widehat{S}$ and α, the payoff matrix will induce $(\mathbf{x}, j)$. Indeed, $\tilde{u}^{\mathsf{F}}$ has the following nice properties:

i. Strategy j is indeed a best response to $\mathbf{x}$, since, by the choice of k we have

$$\tilde{u}^{\mathsf{F}}(\mathbf{x}, j) = -u^{\mathsf{L}}(\mathbf{x}, k) \geq -\min_{\ell \in \widehat{S}} u^{\mathsf{L}}(\mathbf{x}, \ell) = \max_{\ell \in \widehat{S}} \tilde{u}^{\mathsf{F}}(\mathbf{x}, \ell).$$

ii. Any $\ell \in [n] \setminus (\widehat{S} \cup \{j\})$ cannot be a best response of the follower as it is strictly dominated by k, i.e., $\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) < \tilde{u}^{\mathsf{F}}(\mathbf{y}, k)$ for all $\mathbf{y} \in \Delta^{m-1}$. Thus, $\widetilde{\mathsf{BR}}(\mathbf{y}) \subseteq \widehat{S} \cup \{j\}$ for all $\mathbf{y} \in \Delta^{m-1}$.

iii. If j is a best response to some $\mathbf{y} \in \Delta^{m-1}$, then $u^{\mathsf{L}}(\mathbf{y}, j) \leq u^{\mathsf{L}}(\mathbf{x}, j)$. Indeed, $j \in \widetilde{\mathsf{BR}}(\mathbf{y})$ implies that

$$\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) = \max_{\ell \in [n]} \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) \geq \tilde{u}^{\mathsf{F}}(\mathbf{y}, k).$$

Substituting $\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) = -u^{\mathsf{L}}(\mathbf{y}, k) + \alpha \left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right)$ into this inequality and rearranging the terms immediately gives $u^{\mathsf{L}}(\mathbf{y}, j) \leq u^{\mathsf{L}}(\mathbf{x}, j)$.

iv. If any $\ell \in \widehat{S}$ is a best response to some $\mathbf{y} \in \Delta^{m-1}$, then it holds that $\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) = \max_{\ell' \in \widehat{S}} \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell')$, which implies that

$$u^{\mathsf{L}}(\mathbf{y}, \ell) = \min_{\ell' \in \widehat{S}} u^{\mathsf{L}}(\mathbf{y}, \ell'). \tag{6.7}$$

Therefore, if the following also holds for the $\mathbf{y}$ in (iv),

$$\min_{\ell' \in \widehat{S}} u^{\mathsf{L}}(\mathbf{y}, \ell') \leq u^{\mathsf{L}}(\mathbf{x}, j),$$

then by (6.7) we will have $u^{\mathsf{L}}(\mathbf{y}, \ell) \leq u^{\mathsf{L}}(\mathbf{x}, j)$ for every $\ell \in \widetilde{\mathsf{BR}}(\mathbf{y}) \cap \widehat{S}$. This, together with (ii) and (iii), will imply that $u^{\mathsf{L}}(\mathbf{x}, j) \geq u^{\mathsf{L}}(\mathbf{y}, \ell)$ for every $\ell \in \widetilde{\mathsf{BR}}(\mathbf{y})$. Therefore, $(\mathbf{x}, j)$ will indeed form an SSE given that $j \in \widetilde{\mathsf{BR}}(\mathbf{x})$ by (i). We state this observation as the following lemma.

Lemma 6.10. *If $\min_{\ell' \in \widehat{S}} u^{\mathsf{L}}(\mathbf{y}, \ell') \leq u^{\mathsf{L}}(\mathbf{x}, j)$ holds for all $\mathbf{y} \in \Delta^{m-1}$ such that $\widetilde{\mathsf{BR}}(\mathbf{y}) \cap \widehat{S} \neq \varnothing$, then the payoff matrix defined by (6.6) induces $(\mathbf{x}, j)$.*

The proof of Theorem 6.9 is then completed by showing the following result. We leave the complete proof to the appendix of this chapter (see Section 6.A.1).

Proposition 6.11. *If $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$ and $U_j(x) \neq \varnothing$, then we can construct $\widehat{S} \subseteq [n] \setminus \{j\}$ and $\alpha > 0$ in polynomial time, with which the condition of Lemma 6.10 holds for $\tilde{u}^{\mathsf{F}}$ as defined in (6.6).*

$$u^{\mathsf{L}} = \begin{pmatrix} 1/4 & 1 \\ 0 & 1/3 \\ 1/4 & 2/3 \end{pmatrix} \qquad u^{\mathsf{F}} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 1 & 0 \end{pmatrix}$$

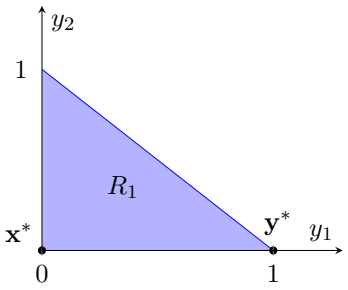


Figure 6.2: A game where the optimal inducible utility is 1, but the optimal *strongly* inducible utility is 0.

6.5 Robustness with Respect to Equilibrium Selection

As discussed in Section 6.2, a weakness of BR- and payoff-inducible strategy profiles is that the resulting games may have multiple SSEs, in which case the follower depends on the leader to choose the SSE that maximizes his utility. To avoid this, in this section, we turn our attention to strong inducibility (see Definition 6.5) and attempt to find a payoff matrix $\tilde{u}^{\mathsf{F}}$ such that $\widetilde{\mathcal{G}}$ has a unique SSE.

We begin with an example showcasing that, in general, the best strongly inducible profile can be much worse than the best payoff-inducible profile.

Example 6.12. Consider a 3×2 game $\mathcal{G} = (u^{\mathsf{L}}, u^{\mathsf{F}})$ with the payoff matrices given in Figure 6.2. Note that the follower obtains positive utility only by playing his strategy 1. Now, observe that the SSE $(\mathbf{x}^*, 1)$, $\mathbf{x}^* = (0, 0, 1) \in \Delta^2$, is payoff-inducible and yields a utility of 1 for the follower: it can be induced by any payoff matrix in which strategy 1 of the follower strictly dominates all other strategies. However, such a payoff matrix will also induce other SSEs, e.g., $(\mathbf{y}^*, 1)$ with $\mathbf{y}^* = (1, 0, 0) \in \Delta^2$. Indeed, it holds that no profile of the form $(\mathbf{y}, 1)$ can be *strongly* induced, and thus the optimal utility the follower can obtain at a strongly inducible profile is 0. To see this, first note that, as seen above, if the follower claims that strategy 1 is his unique best response for all points in Δ^2, then the SSE is not unique. On the other hand, if strategy 2 is a best response at some point $\mathbf{z} \in \Delta^2$, then $(\mathbf{y}, 1)$ will not be an SSE, since for the leader $u^{\mathsf{L}}(\mathbf{y}, 1) < u^{\mathsf{L}}(\mathbf{z}, 2)$ for any $\mathbf{y}, \mathbf{z} \in \Delta^2$. □

The problem in Example 6.12 stems from the following observation: if the follower reports a payoff matrix such that strategy 1 is the unique best response for all points in the domain, then there are multiple SSEs. This can be thought of as a "degenerate" case, since it would occur with probability 0, if the payoffs of the leader were drawn uniformly at random in $[0, 1]$. We formalize this as follows.

Definition 6.13. A leader payoff matrix u^{L} is said to be *max-degenerate*, if there exists $j \in [n]$ such that $\left|\arg\max_{i \in [m]} u^{\mathsf{L}}(i, j)\right| > 1$.

We next provide an example showing that even when u^{L} is *not* max-degenerate, we cannot hope to *exactly* achieve the optimal inducible utility via a strongly inducible profile.

$$u^{\mathsf{L}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1/2 & 1/4 \end{pmatrix} \qquad u^{\mathsf{F}} = \begin{pmatrix} 1/4 & 1/3 \\ 1/4 & 1/2 \\ 1 & 1/6 \end{pmatrix}$$

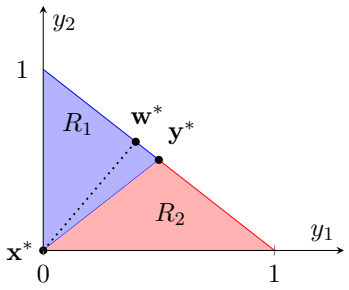


Figure 6.3: A non-max-degenerate game for which the optimal inducible utility cannot be achieved by any strongly inducible profile.

Example 6.14. Consider a 3×2 game with the leader and follower payoff matrices given in Figure 6.3. It is easy to check that u^{L} is not max-degenerate. Now, observe that the maximin utility of the leader is $M = 1/2$ and is achieved at the point $\mathbf{y}^* = \left(\frac{1}{2}, \frac{1}{2}, 0\right) \in \Delta^2$. Let $\mathbf{x}^* = (0, 0, 1) \in \Delta^2$. Since $u^{\mathsf{L}}(\mathbf{x}^*, 1) = 1/2 \geq M$, it follows that $(\mathbf{x}^*, 1)$ is payoff-inducible by Theorem 6.9. Indeed, the partition (R_1, R_2) of Δ^2 in Figure 6.3 shows how $(\mathbf{x}^*, 1)$ can be induced. Note that $u^{\mathsf{F}}(\mathbf{x}^*, 1) = 1$, while any profile different from $(\mathbf{x}^*, 1)$ yields utility strictly less than 1 for the follower. We will now show that $(\mathbf{x}^*, 1)$ cannot be strongly induced, which implies that any strongly inducible profile gives utility strictly less than 1 to the follower. Indeed, suppose that $(\mathbf{x}^*, 1)$ is induced by some $\tilde{u}^{\mathsf{F}}$. If by $\tilde{u}^{\mathsf{F}}$ strategy 1 is a best response to $\mathbf{y}^*$, then $(\mathbf{x}^*, 1)$ cannot be the unique SSE, since $u^{\mathsf{L}}(\mathbf{x}^*, 1) = u^{\mathsf{L}}(\mathbf{y}^*, 1)$. On the other hand, if strategy 2 is the only best response to $\mathbf{y}^*$, then there exists some sufficiently small $\delta > 0$ such that strategy 2 is also a best response to $\mathbf{w}^* = \left(\frac{1}{2} - \delta, \frac{1}{2} + \delta, 0\right)$ (see Figure 6.3). However, this means that $(\mathbf{x}^*, 1)$ cannot be an SSE, since $u^{\mathsf{L}}(\mathbf{x}^*, 1) = 1/2$ and $u^{\mathsf{L}}(\mathbf{w}^*, 2) = 1/2 + \delta > 1/2$. □

As a result, unlike in the previous section, here we cannot hope to solve the problem exactly. However, the next theorem shows that we can approximate the optimal utility with arbitrarily good precision.

Theorem 6.15. *If u^{L} is not max-degenerate, then for any $\varepsilon > 0$, the follower can strongly induce a profile $(\mathbf{x}, j)$ that yields the optimal inducible utility up to an additive loss of at most ε. Furthermore, a matrix $\tilde{u}^{\mathsf{F}}$ strongly inducing $(\mathbf{x}, j)$ can be constructed in time polynomial in $\log(1/\varepsilon)$ (and the size of the representation of the game).*

Proof. Let $(\mathbf{x}^*, j)$ be a payoff-inducible profile that yields the optimal inducible payoff for the follower. By Theorem 6.9, such a profile can be computed in polynomial time.

We begin by solving the following LP.

$$\begin{aligned} \max_{\delta, \mathbf{x}} \quad & \delta \\ \text{subject to} \quad & \mathbf{x} \in \Delta^{m-1} \\ & u^{\mathsf{F}}(\mathbf{x}, j) \geq u^{\mathsf{F}}(\mathbf{x}^*, j) - \varepsilon \\ & u^{\mathsf{L}}(\mathbf{x}, j) = u^{\mathsf{L}}(\mathbf{x}^*, j) + \delta \end{aligned} \tag{6.8}$$

Note that this LP can be solved in time polynomial in $\log(1/\varepsilon)$. Furthermore, note that the polytope of feasible points is not empty since $\delta = 0$ and $\mathbf{x} = \mathbf{x}^*$ satisfy all the constraints. Finally, the LP is not unbounded since δ can be at most $\max_{i \in [m]} u^{\mathsf{L}}(i, j) - u^{\mathsf{L}}(\mathbf{x}^*, j)$.

In the rest of this proof let δ and $\mathbf{x}$ denote an optimal solution to this LP. Note that we can in particular assume that $\mathbf{x}$ is a vertex of the convex polytope $P_\delta = \{\mathbf{y} \in \Delta^{m-1} : u^{\mathsf{L}}(\mathbf{y}, j) = u^{\mathsf{L}}(\mathbf{x}^*, j) + \delta\}$. Indeed, given a solution $\delta, \mathbf{x}$ to LP (6.8), if $\mathbf{x}$ is not a vertex of P_δ, then we consider the LP

$$\begin{aligned} \max_{\mathbf{y}} \quad & u^{\mathsf{F}}(\mathbf{y}, j) \\ \text{subject to} \quad & \mathbf{y} \in \Delta^{m-1} \\ & u^{\mathsf{L}}(\mathbf{y}, j) = u^{\mathsf{L}}(\mathbf{x}^*, j) + \delta \end{aligned}$$

It is known that a solution of an LP that is also a vertex of the feasible polytope can be computed in polynomial time [Grötschel et al., 1981]. Note that in this case the feasible polytope is exactly P_δ. Let $\mathbf{y}$ be an optimal solution that is a vertex of P_δ. We know that $\mathbf{x} \in P_\delta$ and $u^{\mathsf{F}}(\mathbf{x}, j) \geq u^{\mathsf{F}}(\mathbf{x}^*, j) - \varepsilon$, which implies that $u^{\mathsf{F}}(\mathbf{y}, j) \geq u^{\mathsf{F}}(\mathbf{x}^*, j) - \varepsilon$. But this means that $\delta, \mathbf{y}$ is also an optimal solution to the original LP (6.8). Thus, by letting $\mathbf{x} := \mathbf{y}$, we indeed have that $\mathbf{x}$ is a vertex of the convex polytope P_δ.

Let us first handle the case where $\delta = 0$ by showing that $(\mathbf{x}^*, j)$ itself can be strongly induced. Since $\delta = 0$, it follows that $U_j(\mathbf{x}^*) = \varnothing$. Indeed, if there exists $\hat{\mathbf{y}} \in \Delta^{m-1}$ with $u^{\mathsf{L}}(\hat{\mathbf{y}}, j) > u^{\mathsf{L}}(\mathbf{x}^*, j)$, then there exists $\mathbf{y}$ on the segment $(\mathbf{x}^*, \hat{\mathbf{y}}]$ such that $u^{\mathsf{F}}(\mathbf{y}, j) \geq u^{\mathsf{F}}(\mathbf{x}^*, j) - \varepsilon$ (when $\mathbf{y}$ is sufficiently close to $\mathbf{x}^*$) and $u^{\mathsf{L}}(\mathbf{y}, j) > u^{\mathsf{L}}(\mathbf{x}^*, j)$, a contradiction to the optimality of $\delta = 0$. Now, given that $U_j(\mathbf{x}^*) = \varnothing$, we have that $u^{\mathsf{L}}(\mathbf{y}, j) \leq u^{\mathsf{L}}(\mathbf{x}^*, j)$ for all $\mathbf{y} \in \Delta^{m-1}$. But since u^{L} is not max-degenerate (in the sense of Definition 6.13), it follows that in fact $u^{\mathsf{L}}(\mathbf{y}, j) < u^{\mathsf{L}}(\mathbf{x}^*, j)$ for all $\mathbf{y} \in \Delta^{m-1} \setminus \{\mathbf{x}^*\}$. Thus, if the follower always best responds with strategy j, then $(\mathbf{x}^*, j)$ will be the unique SSE. As seen before, it is easy to implement this behavior by reporting $\tilde{u}^{\mathsf{F}}(i, j) = 1$ and $\tilde{u}^{\mathsf{F}}(i, \ell) = 0$ for all $i \in [m]$ and $\ell \in [n] \setminus \{j\}$.

In the rest of this proof, we consider the case $\delta > 0$ and show that $(\mathbf{x}, j)$ can be strongly induced. Since $u^{\mathsf{F}}(\mathbf{x}, j) \geq u^{\mathsf{F}}(\mathbf{x}^*, j) - \varepsilon$, this means that at $(\mathbf{x}, j)$ the follower achieves the optimal inducible utility up to an additive error of ε. Using the same notation as in the proof of Proposition 6.11, we let

$$B = \{i \in [m] : \mathbf{e}_i \cdot \mathbf{x} = \beta_i\}$$

denote the set of boundary conditions of Δ^{m-1} that are tight for $\mathbf{x}$. Note that since $\mathbf{x}$ is a vertex of the polytope P_δ, it follows that $B \neq \varnothing$. We let $\mathbf{h} = \sum_{i \in B} \mathbf{e}_i$. As in the proof of Proposition 6.11, we have that for all $\mathbf{y} \in \Delta^{m-1}$ it holds that

$$\mathbf{h} \cdot (\mathbf{y} - \mathbf{x}) = \sum_{i \in B} \mathbf{e}_i \cdot (\mathbf{y} - \mathbf{x}) \geq 0. \tag{6.9}$$

Furthermore, since $\mathbf{x}$ is a vertex of P_δ, it follows that for all $\mathbf{y} \in P_\delta \setminus \{\mathbf{x}\}$ there exists $i \in B$ such that $\mathbf{e}_i \cdot (\mathbf{y} - \mathbf{x}) > 0$, and thus

$$\mathbf{h} \cdot (\mathbf{y} - \mathbf{x}) > 0. \tag{6.10}$$

Indeed, if $\mathbf{e}_i \cdot (\mathbf{y} - \mathbf{x}) = 0$ for all $i \in B$ for some $\mathbf{y} \in P_\delta \setminus \{\mathbf{x}\}$, this would contradict the fact that $\mathbf{x}$ is a vertex of P_δ (i.e. the unique point in P_δ for which the boundary conditions in B are tight).

We are now ready to construct the payoff matrix reported by the follower. Pick an arbitrary $k \in \arg\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{x}, \ell)$. For all $\mathbf{y} \in \Delta^{m-1}$ let

$$\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) = \begin{cases} -u^{\mathsf{L}}(\mathbf{y}, \ell) & \text{if } \ell \in [n] \setminus \{j\} \\ -u^{\mathsf{L}}(\mathbf{y}, k) + \alpha \left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right) - \mathbf{h} \cdot (\mathbf{y} - \mathbf{x}) & \text{if } \ell = j \end{cases} \tag{6.11}$$

where $\alpha = \left(2 \max_{i\in[m]} \max_{\ell\in[n]} \left|u^{\mathsf{L}}(i,\ell)\right| + m\right)/\delta > 0$. Note that we can compute the payoff matrix corresponding to this utility function in polynomial time. In the remainder of this proof, we show that $(\mathbf{x}, j)$ is the unique SSE of the game $(u^{\mathsf{L}}, \tilde{u}^{\mathsf{F}})$.

Clearly, j is a best response at $\mathbf{x}$, since

$$\tilde{u}^{\mathsf{F}}(\mathbf{x}, j) = -u^{\mathsf{L}}(\mathbf{x}, k) = -\min_{\ell\in[n]\setminus\{j\}} u^{\mathsf{L}}(\mathbf{x}, \ell) = \max_{\ell\in[n]\setminus\{j\}} \tilde{u}^{\mathsf{F}}(\mathbf{x}, \ell),$$

by the choice of k.

Next, let us show that if j is a best response at some $\mathbf{y} \in \Delta^{m-1} \setminus \{\mathbf{x}\}$, then $u^{\mathsf{L}}(\mathbf{y}, j) < u^{\mathsf{L}}(\mathbf{x}, j)$. Indeed, if j is a best response at $\mathbf{y}$, then in particular $\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) \geq \tilde{u}^{\mathsf{F}}(\mathbf{y}, k)$, which implies that

$$\alpha\left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right) \geq \mathbf{h}\cdot(\mathbf{y}-\mathbf{x}). \tag{6.12}$$

Since $\mathbf{h}\cdot(\mathbf{y}-\mathbf{x}) \geq 0$ by (6.9), and $\alpha > 0$, it follows that $u^{\mathsf{L}}(\mathbf{x}, j) \geq u^{\mathsf{L}}(\mathbf{y}, j)$. It remains to show that $u^{\mathsf{L}}(\mathbf{x}, j) \neq u^{\mathsf{L}}(\mathbf{y}, j)$. But if $u^{\mathsf{L}}(\mathbf{x}, j) = u^{\mathsf{L}}(\mathbf{y}, j)$, then $\mathbf{y} \in P_\delta \setminus \{\mathbf{x}\}$ and so by (6.10) we have $\mathbf{h}\cdot(\mathbf{y}-\mathbf{x}) > 0$, which contradicts (6.12).

Finally, it remains to show that if $\ell \in [n] \setminus \{j\}$ is a best response at some $\mathbf{y} \in \Delta^{m-1}$, then it must be that $u^{\mathsf{L}}(\mathbf{y}, \ell) < u^{\mathsf{L}}(\mathbf{x}, j)$: Indeed, if $\ell \in [n] \setminus \{j\}$ is a best response at $\mathbf{y}$, then in particular $\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) \leq \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell)$, which by (6.11) means that

$$\begin{aligned}
\alpha\left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right) &\leq -u^{\mathsf{L}}(\mathbf{y}, \ell) + u^{\mathsf{L}}(\mathbf{y}, k) + \mathbf{h}\cdot(\mathbf{y}-\mathbf{x}) \\
&\leq -u^{\mathsf{L}}(\mathbf{y}, \ell) + u^{\mathsf{L}}(\mathbf{y}, k) + \|\mathbf{h}\|_2\|\mathbf{y}-\mathbf{x}\|_2 \\
&\leq 2\max_{i\in[m]}\max_{\ell'\in[n]} |u^{\mathsf{L}}(i,\ell')| + \sqrt{m-1}\sqrt{m-1} \\
&\leq \alpha\delta
\end{aligned}$$

by the choice of α. Thus, we obtain that $u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j) \leq \delta$, which implies that $u^{\mathsf{L}}(\mathbf{y}, j) \geq u^{\mathsf{L}}(\mathbf{x}^*, j)$, i.e. $\mathbf{y} \in \overline{U_j(\mathbf{x}^*)}$ (since $U_j(\mathbf{x}^*) \neq \varnothing$). Since $(\mathbf{x}^*, j)$ is payoff-inducible, which means that $u^{\mathsf{L}}(\mathbf{x}^*, j) \geq M$, we can use Lemma 6.7 to obtain

$$u^{\mathsf{L}}(\mathbf{x}, j) = u^{\mathsf{L}}(\mathbf{x}^*, j) + \delta > u^{\mathsf{L}}(\mathbf{x}^*, j) \geq \min_{\ell'\in[n]\setminus\{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell') = u^{\mathsf{L}}(\mathbf{y}, \ell)$$

where the last equality comes from the fact that ℓ is a best response at $\mathbf{y}$, i.e., in particular $\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) = \max_{\ell'\in[n]\setminus\{j\}} \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell')$. □

6.6 Summary

The approach we have demonstrated for a follower to optimally deceive the leader demonstrates a connection with the leader's maximin utility. From this interesting connection it is perhaps also easy to see the main idea behind the approach — to exploit information asymmetry in the game. Indeed, the degree of exploitation can be maximum: the follower can induce any SSE as long as it provides the leader her maximin utility, which is the utility a rational player can guarantee in any case. This indicates a risk of learning to make a commitment in a Stackelberg game, especially in the face of information asymmetry; in the worst case, the deceptive behavior of the follower may essentially void the learning attempt, resulting in the leader to obtain no useful information to improve her utility. In the next chapter, we will continue our exploration but on a more specific form of Stackelberg games, the Stackelberg security games, where we discover some surprising results that do not necessarily hold in general Stackelberg games.

Appendix 6.A Omitted Proofs

6.A.1 Proof of Proposition 6.11

The proof relies on the following useful lemma.

Lemma 6.16 (Farkas' Lemma [Boyd and Vandenberghe, 2014]). *Let $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$ and $\mathbf{b} \in \mathbb{R}^{n_1}$. Then exactly one of the following statements is true:*

1. *there exists $\mathbf{z} \in \mathbb{R}^{n_2}$ such that $\mathbf{A}\mathbf{z} = \mathbf{b}$ and $\mathbf{z} \geq 0$;*
2. *there exists $\mathbf{z} \in \mathbb{R}^{n_1}$ such that $\mathbf{A}^\mathsf{T}\mathbf{z} \geq 0$ and $\mathbf{b} \cdot \mathbf{z} < 0$.*

Consider any strategy profile $(\mathbf{x}, j)$ with $u^\mathsf{L}(\mathbf{x}, j) \geq M$ and $U_j(x) \neq \varnothing$. We begin by taking care of a simple case, as an immediate corollary of Lemma 6.10.

Corollary 6.17. *A matrix $\tilde{u}^\mathsf{F}$ that induces $(\mathbf{x}, j)$ can be constructed in polynomial time if it holds that*

$$u^\mathsf{L}(\mathbf{x}, j) \geq M_{-j} := \max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in [n] \setminus \{j\}} u^\mathsf{L}(\mathbf{y}, \ell). \tag{6.13}$$

Proof. Let $\widehat{S} = [n] \setminus \{j\}$. Then, for every $\mathbf{y} \in \Delta^{m-1}$, we immediately obtain that

$$u^\mathsf{L}(\mathbf{x}, j) \geq \max_{\mathbf{y} \in \Delta^{m-1}} \min_{\ell \in [n] \setminus \{j\}} u^\mathsf{L}(\mathbf{y}, \ell) \geq \min_{\ell \in \widehat{S}} u^\mathsf{L}(\mathbf{y}, \ell)$$

By Lemma 6.10, the payoff matrix defined by (6.6) (with, say, $\alpha = 1$) then induces $(\mathbf{x}, j)$, and can clearly be computed in polynomial time. □

The more challenging case is when (6.13) does not hold (e.g., the case with the profile $(\mathbf{x}, 1)$ in Example 6.8). In what follows, we prove Proposition 6.11 by showing that there is still a choice of $\widehat{S}$ and α that leads to the condition in Lemma 6.10, even when (6.13) does not hold. Thus, from now on, we assume that

$$u^\mathsf{L}(\mathbf{x}, j) < M_{-j}. \tag{6.14}$$

We define the following useful components. By Lemma 6.7 and the assumption that $u^\mathsf{L}(\mathbf{x}, j) \geq M$, we know that

$$u^\mathsf{L}(\mathbf{x}, j) \geq V \tag{6.15}$$

where

$$V = \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^\mathsf{L}(\mathbf{y}, \ell).$$

Since $\overline{U_j(\mathbf{x})} \neq \varnothing$, there exists $\mathbf{y}^* \in \overline{U_j(\mathbf{x})}$ such that

$$\min_{\ell \in [n] \setminus \{j\}} u^\mathsf{L}(\mathbf{y}^*, \ell) = V, \tag{6.16}$$

which can be computed efficiently by solving an LP (i.e., maximize μ, subject to $\mu \leq u^\mathsf{L}(\mathbf{y}, \ell)$ for all $\ell \in [n] \setminus \{j\}$ and $\mathbf{y} \in \overline{U_j(\mathbf{x})}$). We then let

$$S = \{\ell \in [n] \setminus \{j\} \mid u^\mathsf{L}(\mathbf{y}^*, \ell) = V\}.$$

Before we proceed, we prove two useful technical results.

Lemma 6.18. $u^{\mathsf{L}}(\mathbf{y}^*, j) = u^{\mathsf{L}}(\mathbf{x}, j)$.

Proof. For the sake of contradiction, suppose that $u^{\mathsf{L}}(\mathbf{y}^*, j) \neq u^{\mathsf{L}}(\mathbf{x}, j)$. Since $\mathbf{y}^* \in \overline{U_j(\mathbf{x})}$, we have that $u^{\mathsf{L}}(\mathbf{y}^*, j) \geq u^{\mathsf{L}}(\mathbf{x}, j)$, so it must be that $u^{\mathsf{L}}(\mathbf{y}^*, j) > u^{\mathsf{L}}(\mathbf{x}, j)$.

The assumption (6.14) that $u^{\mathsf{L}}(\mathbf{x}, j) < M_{-j}$ implies that there exists $\hat{\mathbf{y}} \in \Delta^{m-1}$ such that

$$\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\hat{\mathbf{y}}, \ell) > u^{\mathsf{L}}(\mathbf{x}, j) \geq V,$$

where we also use (6.15). Now that $\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}^*, \ell) = V$ by (6.16), by the concavity of $\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\cdot, \ell)$, it follows that $\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{z}, \ell) > V$ for all $\mathbf{z}$ on the segment $[\hat{\mathbf{y}}, \mathbf{y}^*)$; $\mathbf{z} \in \Delta^{m-1}$ as Δ^{m-1} is convex. Now that we have $u^{\mathsf{L}}(\mathbf{y}^*, j) > u^{\mathsf{L}}(\mathbf{x}, j)$ under our assumption, when $\mathbf{z}$ is sufficiently close to $\mathbf{y}^*$, we can have $u^{\mathsf{L}}(\mathbf{z}, j) \geq u^{\mathsf{L}}(\mathbf{x}, j)$ and hence, $\mathbf{z} \in \overline{U_j(\mathbf{x})}$. This leads to the contradiction that

$$V = \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell) \geq \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{z}, \ell) > V. \qquad \square$$

Lemma 6.19. $\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{y}, \ell) < V$ *for all* $\mathbf{y} \in U_j(\mathbf{x})$.

Proof. For the sake of contradiction, assume that there exists $\hat{\mathbf{y}} \in U_j(\mathbf{x})$ such that

$$\min_{\ell \in S} u^{\mathsf{L}}(\hat{\mathbf{y}}, \ell) \geq V.$$

By assumption (6.14) that $u^{\mathsf{L}}(\mathbf{x}, j) < M_{-j}$, there exists $\mathbf{z} \in \Delta^{m-1}$ such that $\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{z}, \ell) > u^{\mathsf{L}}(\mathbf{x}, j) \geq V$, which immediately yields the following given that $S \subseteq [n] \setminus \{j\}$ by definition:

$$\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{z}, \ell) > V.$$

By definition, $u^{\mathsf{L}}(\mathbf{y}^*, \ell) = V$ for all $\ell \in S$, which also implies that $u^{\mathsf{L}}(\mathbf{y}^*, \ell) > V$ for all $\ell \in [n] \setminus (\{j\} \cup S)$ (otherwise, we would have $\min_{\ell \in [n] \setminus \{j\} u^{\mathsf{L}}(\mathbf{y}^*, \ell)} < V$). Thus, we have

$$\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{y}^*, \ell) = V \quad \text{and} \quad \min_{\ell \in [n] \setminus (\{j\} \cup S)} u^{\mathsf{L}}(\mathbf{y}^*, \ell) > V.$$

Now consider a point $\mathbf{w}$ on the segment $(\mathbf{y}^*, \hat{\mathbf{y}}]$. Since $\mathbf{y}^* \in \overline{U_j(\mathbf{x})}$ and $\hat{\mathbf{y}} \in U_j(\mathbf{x})$, i.e., $u^{\mathsf{L}}(\mathbf{y}^*, j) \geq u^{\mathsf{L}}(\mathbf{x}, j)$ and $u^{\mathsf{L}}(\hat{\mathbf{y}}, j) > u^{\mathsf{L}}(\mathbf{x}, j)$, we have $u^{\mathsf{L}}(\mathbf{w}, j) > u^{\mathsf{L}}(\mathbf{x}, j)$ and hence, $\mathbf{w} \in U_j(\mathbf{x})$. In addition, by continuity, when $\mathbf{w}$ is sufficiently close to $\mathbf{y}^*$, we have

$$\min_{\ell \in [n] \setminus (\{j\} \cup S)} u^{\mathsf{L}}(\mathbf{w}, \ell) > V. \tag{6.17}$$

By concavity of the function $\min_{\ell \in S} u^{\mathsf{L}}(\cdot, \ell)$, since $\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{y}, \ell) \geq V$ for both $\mathbf{y} \in \{\mathbf{y}^*, \hat{\mathbf{y}}\}$, we have

$$\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{w}, \ell) \geq V. \tag{6.18}$$

Analogously, we can find a point $\mathbf{w}' \in U_j(\mathbf{x})$ on the segment $(\mathbf{w}, \mathbf{z}]$, such that (6.17) and (6.18) hold for $\mathbf{w}'$ while (6.18) is strict, in particular. Thus, we have

$$\min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{w}', \ell) > V = \max_{\mathbf{y} \in \overline{U_j(\mathbf{x})}} \min_{\ell \in [n] \setminus \{j\}} u^{\mathsf{L}}(\mathbf{y}, \ell),$$

which is a contradiction as $\mathbf{w}' \in U_j(\mathbf{x})$. $\square$

In what follows, we use the coordinates $(y_1, \dots, y_{m-1})$ for every point $\mathbf{y} \in \Delta^{m-1}$, i.e., we have

$$\Delta^{m-1} = \left\{ (y_1, \dots, y_{m-1}) \in \mathbb{R}_{\geq 0} \,:\, \sum_{i=1}^{m-1} y_i \leq 1 \right\}.$$

Accordingly, we can write the utility function as

$$u^{\mathsf{L}}(\mathbf{y}, \ell) = \mathbf{g}_\ell \cdot \mathbf{y} + u^{\mathsf{L}}(m, \ell),$$

where $\mathbf{g}_\ell \in \mathbb{R}^{m-1}$ and its i-th component is $g_{\ell,i} = u^{\mathsf{L}}(i, \ell) - u^{\mathsf{L}}(m, \ell)$; "$\cdot$" denotes the inner product. Hence, we have

$$u^{\mathsf{L}}(\mathbf{y}, \ell) = \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) + u^{\mathsf{L}}(\mathbf{y}^*, \ell) = \begin{cases} \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) + V & \text{if } \ell \in S \\ \mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) + u^{\mathsf{L}}(\mathbf{x}, j) & \text{if } \ell = j \end{cases} \tag{6.19}$$

where $u^{\mathsf{L}}(\mathbf{y}^*, \ell) = V$ for all $\ell \in S$ by the definition of S, and $u^{\mathsf{L}}(\mathbf{y}^*, j) = u^{\mathsf{L}}(\mathbf{x}, j)$ by Lemma 6.18. Note that since $U_j(\mathbf{x}) \neq \varnothing$, it must be that $\mathbf{g}_j \neq 0$.

We also write the m boundary conditions that define Δ^{m-1} as $\mathbf{e}_i \cdot \mathbf{y} \geq \beta_i$. Namely, for each $i \in [m-1]$, let $\mathbf{e}_i \in \mathbb{R}^{m-1}$ be the i-th unit vector and $\beta_i = 0$, while $\mathbf{e}_m = (-1, \dots, -1) \in \mathbb{R}^{m-1}$ and $\beta_m = -1$. Thus, $\Delta^{m-1} = \{\mathbf{y} \in \mathbb{R}^{m-1} \,:\, \mathbf{e}_i \cdot \mathbf{y} \geq \beta_i \text{ for } i \in [m]\}$. Let

$$B = \{i \in [m] \,:\, \mathbf{e}_i \cdot \mathbf{y}^* = \beta_i\}$$

be the set of boundary conditions that are tight for $\mathbf{y}^*$. Note that for any $\mathbf{y} \in \Delta^{m-1}$ we have

$$\mathbf{e}_i \cdot (\mathbf{y} - \mathbf{y}^*) \geq 0 \quad \text{for all } i \in B. \tag{6.20}$$

We can now prove the following result using Farkas' Lemma (Lemma 6.16), which allows us to express $-\mathbf{g}_j$ as a non-negative linear combination of $\mathbf{g}_\ell$'s and $\mathbf{e}_i$'s.

Lemma 6.20. *$-\mathbf{g}_j$ can be expressed as a non-negative linear combination of $\{\mathbf{g}_\ell : \ell \in S\} \cup \{\mathbf{e}_i : i \in B\}$, i.e. $-\mathbf{g}_j = \sum_{\ell \in S} \lambda_\ell \mathbf{g}_\ell + \sum_{i \in B} \mu_i \mathbf{e}_i$, where $\lambda_\ell \geq 0$ and $\mu_i \geq 0$.*

Proof. We use Farkas' Lemma (Lemma 6.16) and let $n_1 = m - 1$ and $n_2 = |S| + |B|$. The columns of $\mathbf{A}$ are exactly the vectors $\{\mathbf{g}_\ell : \ell \in S\} \cup \{\mathbf{e}_i : i \in B\}$. We set $\mathbf{b} = -\mathbf{g}_j$. Note that the first alternative of Farkas' Lemma immediately yields the statement we want to prove. Thus, we set out to prove that the second alternative cannot hold.

Assume, for the sake of contradiction, that there exists $\mathbf{z} \in \mathbb{R}^{m-1}$ such that $\mathbf{A}^{\mathsf{T}}\mathbf{z} \geq 0$ and $\mathbf{b} \cdot \mathbf{z} < 0$, i.e., $\mathbf{g}_\ell \cdot \mathbf{z} \geq 0$ for all $\ell \in S$, $\mathbf{e}_i \cdot \mathbf{z} \geq 0$ for all $i \in B$, and $\mathbf{g}_j \cdot \mathbf{z} > 0$.

Then, by picking $\delta > 0$ sufficiently small, it holds for $\mathbf{y} = \mathbf{y}^* + \delta\mathbf{z}$ that:

- By (6.19), we have the following for all $\ell \in S$:

  $$u^{\mathsf{L}}(\mathbf{y}, \ell) = \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) + V = \delta \mathbf{g}_\ell \cdot \mathbf{z} + V \geq V.$$

  In addition,

  $$u^{\mathsf{L}}(\mathbf{y}, j) = \mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) + u^{\mathsf{L}}(\mathbf{y}^*, j) = \delta \mathbf{g}_j \cdot \mathbf{z} + u^{\mathsf{L}}(\mathbf{x}, j) > u^{\mathsf{L}}(\mathbf{x}, j).$$

- $\mathbf{y} \in \Delta^{m-1}$: For $i \in B$, we immediately obtain that $\mathbf{e}_i \cdot \mathbf{y} = \mathbf{e}_i \cdot (\mathbf{y}^* + \delta\mathbf{z}) \geq \mathbf{e}_i \cdot \mathbf{y}^* = \beta_i$, which means that these boundary conditions are satisfied. For $i \in [m] \setminus B$, we know that $\mathbf{e}_i \cdot \mathbf{y}^* > \beta_i$ and thus by picking $\delta > 0$ small enough, we can ensure that $\mathbf{e}_i \cdot \mathbf{y} = \mathbf{e}_i \cdot \mathbf{y}^* + \delta(\mathbf{e}_i \cdot \mathbf{z}) \geq \beta_i$.

Thus, it follows that $\mathbf{y} \in U_j(\mathbf{x})$ and $\min_{\ell \in S} u^{\mathsf{L}}(\mathbf{y}, \ell) \geq V$. But this cannot hold according to Lemma 6.19. □

We can now complete the proof of Proposition 6.11. We first express $-\mathbf{g}_j$ as a non-negative linear combination of the vectors $\{\mathbf{g}_\ell : \ell \in S\} \cup \{\mathbf{e}_i : i \in B\}$. By Lemma 6.20 we know that this is possible and it is easy to see that we can find the coefficients in polynomial time (e.g. by solving an LP). We thus obtain $-\mathbf{g}_j = \sum_{\ell \in S} \lambda_\ell \mathbf{g}_\ell + \sum_{i \in B} \mu_i \mathbf{e}_i$, where $\lambda_\ell \geq 0$ for every $\ell \in S$ and $\mu_i \geq 0$ for every $i \in B$. Let $\widehat{S} = \{\ell \in S : \lambda_\ell > 0\}$. We will argue that $\widehat{S} \neq \varnothing$.

Observe that since now $-\mathbf{g}_j = \sum_{\ell \in S} \lambda_\ell \mathbf{g}_\ell + \sum_{i \in B} \mu_i \mathbf{e}_i$ and, by (6.20), we have $\mathbf{e}_i \cdot (\mathbf{y} - \mathbf{y}^*) \geq 0$ for all $\mathbf{y} \in \Delta^{m-1}$ and $i \in B$, it follows that, for all $\mathbf{y} \in \Delta^{m-1}$, we have

$$\begin{aligned} -\mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) &= \sum_{\ell \in S} \lambda_\ell \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) + \sum_{i \in B} \mu_i \mathbf{e}_i \cdot (\mathbf{y} - \mathbf{y}^*) \\ &\geq \sum_{\ell \in S} \lambda_\ell \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) \\ &= \sum_{\ell \in \widehat{S}} \lambda_\ell \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*), \end{aligned} \tag{6.21}$$

where the last transition is due to the fact that $\lambda_\ell = 0$ for all $\ell \in S \setminus \widehat{S}$, as implied by the definition of $\widehat{S}$.

Since $U_j(\mathbf{x}) \neq \varnothing$, consider any $\mathbf{y} \in U_j(\mathbf{x})$. By definition, this means that $u^{\mathsf{L}}(\mathbf{y}, j) > u^{\mathsf{L}}(\mathbf{x}, j)$, which further implies that $\mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) > 0$ since $u^{\mathsf{L}}(\mathbf{y}, j) = \mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) + u^{\mathsf{L}}(\mathbf{x}, j)$ by (6.19). By (6.21), we then have

$$\sum_{\ell \in \widehat{S}} \lambda_\ell \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) < 0.$$

Hence, $\widehat{S} \neq \varnothing$.

It remains to show that with the above $\widehat{S}$ and, in particular, $\alpha = 1/\lambda_k$ (recall that $k \in \arg\min_{\ell \in \widehat{S}} u^{\mathsf{L}}(\mathbf{x}, \ell)$), the condition in Lemma 6.10 holds, i.e., we prove that $\min_{\ell \in \widehat{S}} u^{\mathsf{L}}(\mathbf{y}, \ell) \leq u^{\mathsf{L}}(\mathbf{x}, j)$ holds for all $\mathbf{y} \in \Delta^{m-1}$ such that $\widetilde{\mathsf{BR}}(\mathbf{y}) \cap \widehat{S} \neq \varnothing$.

For the sake of contradiction, suppose that there exists $\mathbf{y} \in \Delta^{m-1}$ such that $\widetilde{\mathsf{BR}}(\mathbf{y}) \cap \widehat{S} \neq \varnothing$, but $u^{\mathsf{L}}(\mathbf{y}, \ell) > u^{\mathsf{L}}(\mathbf{x}, j)$ for all $\ell \in \widehat{S}$. By (6.15), we have $u^{\mathsf{L}}(\mathbf{x}, j) \geq V$, and thus $u^{\mathsf{L}}(\mathbf{y}, \ell) > V$ for all $\ell \in \widehat{S}$. By (6.19), we have $u^{\mathsf{L}}(\mathbf{y}, \ell) = \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) + V$; thus, $\mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) > 0$ for all $\ell \in \widehat{S}$.

Using (6.21) and the fact that $k \in \widehat{S}$ by our choice, we then obtain

$$-\mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) \geq \sum_{\ell \in \widehat{S}} \lambda_\ell \mathbf{g}_\ell \cdot (\mathbf{y} - \mathbf{y}^*) \geq \lambda_k \mathbf{g}_k \cdot (\mathbf{y} - \mathbf{y}^*).$$

By (6.19), we have

$$u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j) = -\mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*).$$

Recall that it is defined that $\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) = -u^{\mathsf{L}}(\mathbf{y}, k) + \alpha \left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right)$ as in (6.6). Using the above two equations and (6.19), we then obtain the following:

$$\begin{aligned} \tilde{u}^{\mathsf{F}}(\mathbf{y}, j) &= -u^{\mathsf{L}}(\mathbf{y}, k) + \alpha \left(u^{\mathsf{L}}(\mathbf{x}, j) - u^{\mathsf{L}}(\mathbf{y}, j)\right) \\ &= -\mathbf{g}_k \cdot (\mathbf{y} - \mathbf{y}^*) - V - \alpha \mathbf{g}_j \cdot (\mathbf{y} - \mathbf{y}^*) \\ &\geq -V + (\alpha \lambda_k - 1) \mathbf{g}_k \cdot (\mathbf{y} - \mathbf{y}^*) \\ &= -V. \end{aligned}$$

However, by (6.6) we also have $\tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell) = -u^{\mathsf{L}}(\mathbf{y}, \ell)$ if $\ell \in \widehat{S}$, which implies that for all $\ell \in \widehat{S}$ it holds that

$$\tilde{u}^{\mathsf{F}}(\mathbf{y}, j) \geq -V > -u^{\mathsf{L}}(\mathbf{y}, \ell) = \tilde{u}^{\mathsf{F}}(\mathbf{y}, \ell).$$

Hence, $\widetilde{\mathsf{BR}}(\mathbf{y}) \cap \widehat{S} = \varnothing$, which contradicts our assumption.

Chapter 7

Follower Deception in SSGs and Ways to Counteract

We have investigated follower deception in general Stackelberg games in Chapters 5 and 6. In this chapter, we extend our work to SSGs. We consider the problems we have studied in the previous two chapters on this specific form of Stackelberg games: How will a strategic follower deceive a leader who is unaware of the deception in an SSG? Knowing the follower's incentives, what can the leader do to counteract the deception?

7.1 Introduction

Amongst the studies on learning to commit optimally in Stackelberg games that we reviewed in Chapter 5, the papers of Blum et al. [2014], Haghtalab et al. [2016], and Peng et al. [2019] are specialized to SSGs. A truthful attacker is assumed in these papers. Indeed, in Chapter 5 (Section 5.2), we have seen an example of an SSG in which the assumption of a truthful attacker makes learning algorithms fail in the face of a deceitful follower. We expect follower deception to be more common, yet more harmful, in security domains given the adversarial nature of the opponent and the critical applications in these domains. A key finding in this chapter is that the attacker (the follower) is often incentivized to imitate a type that makes the game zero-sum; a defender (the leader) who is unaware of such deceptive behavior of the follower would then only learn a maximin strategy (which is the optimal commitment in a zero-sum game). Effectively, the learning attempt now becomes meaningless: to compute a maximin strategy, one needs no additional information about the other player's payoffs at all!

This finding relates to the connection between follower deception and the leader's maximin utility we discovered in Chapter 6 (see Theorem 6.9). We then study what can be done to reduce the defender's loss due to the attacker's manipulation. We apply the policy-based framework we proposed in Chapter 5, whereby the defender commits to a *policy* that specifies her strategy commitment according to the learned information. A strategic attacker then takes into account the defender's policy, choosing optimally what he wants the defender to learn so that the policy outputs a strategy that benefits him the most. We make several other contributions under this framework. (i) We propose a novel quality measure of the defender's policy and argue why it is a reasonable choice in the context of SSGs. (ii) We develop a polynomial-time algorithm for computing the optimal policy with respect to this quality measure, which is in sharp contrast with the inapproximability result in generic Stackelberg games in Chapter 5. (iii) In addition, we

also propose a heuristic approach for the scenario where the attacker type space is infinite or completely unknown. This approach is inspired by the quantal response model that was initially proposed to model bounded rationality of human players. It suggests that it is rational to behave in a "boundedly rational" manner in the presence of attacker manipulation. (iv) Finally, through empirical evaluation we show that our approaches can improve the defender's utility significantly in randomly generated games, as compared to the situation when the attacker's manipulations are ignored.

Apart from the aforementioned papers on learning to commit based on the attacker's responses, several papers also take the regret-minimization perspective and design online learning algorithms for the leader to use in the adversarial setting [Balcan et al., 2015, Xu et al., 2016b]. Our work in this chapter can be seen as a middle-ground approach between the overoptimistic assumption of a truthful follower adopted by the former line of work and the pessimistic assumption of a worst-case opponent by the latter. Our approach to deal with attacker manipulation is, in a nutshell, to reduce information uncertainty by acquiring additional information while bearing in mind that information acquired may be manipulated by the attacker. In comparison to Chapter 6, the SSG model offers us an appropriate level of specificity that enables us to derive a richer set of results than for general Stackelberg games (consider that when the interests of the leader and the follower completely align, there is simply no incentive for the follower to manipulate the leader), while it also captures sufficiently many real-world applications of significant practical value.

7.1.1 Related Work on Deception in SSGs

Deception has been extensively studied in SSGs. However, previous papers mostly focused on designing deceptive strategies for the defender to use; it is the defender who has the informational advantage in these models [Brown et al., 2005, Powell, 2007, Zhuang and Bier, 2010, Yin et al., 2013, Xu et al., 2015, Guo et al., 2017, Schlenker et al., 2018]. In our model, the follower has the greater decision-pertinent knowledge and discloses it strategically. Deception from the follower's side remains largely unexplored in the literature, to the best of our knowledge. Somewhat related to our work, Xu et al. [2016a] studied signaling in Bayesian Stackelberg games and also considered the possibility that a follower may strategically misreport his type. However, the question they study is completely different from ours — they design signaling schemes for different follower types whereas we design the leader's mixed strategy for each follower type. The contrasting complexity between their model (polynomial-time solvable for normal form games) and ours (NP-hard) also highlights the intrinsic difference. Even more importantly, though, in our model type reporting takes the conceptual form of a behaviour inducing message. The common information asymmetry (see e.g. [Rabinovich et al., 2015, Xu et al., 2015, Dughmi et al., 2016]) is reversed between the leader and the follower.

More recently, Nguyen et al. [2019] studied attacker deception in finitely repeated security games. In their model, the defender also aims at learning the attacker's type from the history of their play. The attacker therefore has an incentive to play more strategically to lead the game to proceed in a favorable way. They provided equilibrium analysis for the strategic interaction between the players and proposed algorithms based on optimization programs to compute an equilibrium. Their work is conceptually very close to ours, though the finitely repeated game structure differs considerably from our one-shot game model in this chapter. In another parallel work, Nguyen and Xu [2019] studied the same one-shot game model. Upon making the same discovery about attacker deception, they did not set out to design countermeasures for the defender but studied a special

scenario where the attacker's deceptive behavior is constrained by a constraint on the distance between the misreport and ground truth.

7.1.2 Chapter Overview

We briefly describe the model in Section 7.2. Then in Section 7.3, we present our key finding, namely, the game will effectively become a zero-sum game in the presence of follower deception. We then explore countermeasures against attacker manipulation using the policy-based framework. In Section 7.4, we discuss our approach and selection of the quality measure for optimizing the defender policy. Then, in the next two sections, we present an efficient algorithm for computing the optimal policy for a finite set of attacker types (Section 7.6) and a heuristic algorithm that works with an infinite type space (Section 7.5). Finally, in Section 7.7, we evaluate our algorithms with numerical experiments.

7.2 SSG with Attacker Deception

Our model builds on the standard SSG model (see Section 2.2, Chapter 2). In this chapter, we will use the superscripts "D" and "A" to denote parameters for the defender and for the attacker, respectively, who are, respectively, the leader and the follower in the game. For simplicity, we will also write the players' payoffs as vectors: e.g., for the attacker, his payoffs are given by $(\mathbf{r}^{\mathsf{A}}, \mathbf{p}^{\mathsf{A}}) \in \mathbb{R}^{2m}$, where $\mathbf{r}^{\mathsf{A}} = (r_1^{\mathsf{A}}, \dots, r_m^{\mathsf{A}})$ and $\mathbf{p}^{\mathsf{A}} = (p_1^{\mathsf{A}}, \dots, p_m^{\mathsf{A}})$, with each r_ℓ^{A} and p_ℓ^{A} denoting the reward and penalty, respectively, that are associated to target t_ℓ. We refer to the tuple $(\mathbf{r}^{\mathsf{A}}, \mathbf{p}^{\mathsf{A}})$ as an *attacker type*.

Unless otherwise specified, when we refer to the defender's strategy in this chapter, we mean the strategy in the form of a coverage vector $\mathbf{c} = (c_1, \dots, c_m) \in \mathcal{C}_k$, where k is the number of the defender's resources. For simplicity, we omit k and write $\mathcal{C} = \mathcal{C}_k$ in this chapter since the number of resources is always k throughout. When the defender plays a strategy $\mathbf{c}$ and the attacker attacks some target $i \in T$, the defender's and the attacker's (expected) utilities $u^{\mathsf{D}}(\mathbf{c}, i)$ and $u^{\mathsf{A}}(\mathbf{c}, i)$, respectively, are given as follows.

$$u^{\mathsf{D}}(\mathbf{c}, i) = c_i \cdot r_i^{\mathsf{D}} + (1 - c_i) \cdot p_i^{\mathsf{D}} \tag{7.1}$$

$$u^{\mathsf{A}}(\mathbf{c}, i) = c_i \cdot p_i^{\mathsf{A}} + (1 - c_i) \cdot r_i^{\mathsf{A}} \tag{7.2}$$

We will also write $u^{\mathsf{D}}(c, i) = c \cdot r_i^{\mathsf{D}} + (1 - c) \cdot p_i^{\mathsf{D}}$ and $u^{\mathsf{A}}(c, i) = c \cdot p_i^{\mathsf{A}} + (1 - c) \cdot r_i^{\mathsf{A}}$ where the first variable is a scalar representing the coverage of target i. To distinguish different attacker types, we extend (7.2) and denote by

$$u_\theta^{\mathsf{A}}(\mathbf{c}, i) = c_i \cdot p_i + (1 - c_i) \cdot r_i$$

the utility function parameterized by attacker type $\theta = (\mathbf{r}, \mathbf{p}) \in \mathbb{R}^{2m}$; let $\mathsf{BR}_\theta(\mathbf{c}) = \arg\max_{i \in T} u_\theta^{\mathsf{A}}(\mathbf{c}, i)$ denote the best response set of a type-θ attacker. We will refer to an SSE in a game where the attacker's type is θ *an SSE for attacker type* θ.

The game then proceeds the same way as we described in Section 6.2.1. Here the attacker misreports a fake type θ, instead of a matrix u_θ^{A}, but effectively, this misleads the leader into playing a game $\widetilde{\mathcal{G}} = (u^{\mathsf{D}}, u_\theta^{\mathsf{A}})$ as described in Chapter 6.

7.3 Optimal Attacker Deception

In Section 5.2, we have shown that in the SSG defined by the following payoffs (see also Section 2.2.1), a type-α attacker can pretend to have type β and enjoys a utility improvement from 0 to 1 if the defender is unaware of this deception.

	t_1	t_2
$\mathbf{r}^{\mathsf{D}}$	1	0.99
$\mathbf{p}^{\mathsf{D}}$	-1	-1

defender

	t_1	t_2
$\mathbf{r}^{\mathsf{A}}_{\alpha}$	3	$1/3$
$\mathbf{p}^{\mathsf{A}}_{\alpha}$	-1	-1

attacker type α

	t_1	t_2
$\mathbf{r}^{\mathsf{A}}_{\beta}$	1	1
$\mathbf{p}^{\mathsf{A}}_{\beta}$	-1	-1

attacker type β

Observe that the above game between the defender and a type-β attacker is very close to a zero-sum game. In fact, if we allow the attacker to pretend to have any arbitrary payoffs, a type-α attacker can exploit the power of deception even further by pretending that his payoffs make the game exactly a zero-sum game. More specifically, if the defender were to play a zero-sum game, her optimal strategy commitment would be her maximin strategy, which corresponds to coverage $\overline{\mathbf{c}} = \left(\frac{1}{2} - \frac{1}{798}, \frac{1}{2} + \frac{1}{798}\right)$, and by which the defender is indifferent of which target the attacker attacks. If the type-α responds to $\overline{\mathbf{c}}$ by attacking t_1, he would obtain utility $u^{\mathsf{A}}_{\alpha}(\overline{\mathbf{c}}, t_1) = 1 + \frac{2}{399}$, which is a further improvement over the utility 1 the attacker can obtain by pretending to be of type β.

It turns out that the attacker pretending to play a zero-sum game is not a coincidence specific to this example but a general phenomenon in SSGs. We show next that it is always optimal for the attacker to mislead the defender into playing her maximin strategy. A maximin strategy is the defender's optimal commitment in a zero-sum Stackelberg game (we will formally demonstrate this in Lemma 7.5).

7.3.1 Let Us Play a Zero-sum Game

As in the previous chapters, we only consider the players' utilities in the SSE the defender learns. We view the learning process as a reporting step in which the attacker simply reports his type to the defender. The following program computes the optimal reporting strategy of a type-θ attacker.

$$\max_{\beta,\mathbf{z},t} \quad u^{\mathsf{A}}_{\theta}(\mathbf{z}, t) \tag{7.3}$$

$$\text{subject to} \quad (\mathbf{z}, t) \in \arg\max_{\mathbf{c}\in\mathcal{C},\, i\in \mathsf{BR}_{\beta}(\mathbf{c})} u^{\mathsf{D}}(\mathbf{c}, i) \tag{7.3a}$$

$$\beta \in \Theta \tag{7.3b}$$

Here $\Theta = \{(\mathbf{r}^{\mathsf{A}}, \mathbf{p}^{\mathsf{A}}) \in \mathbb{R}^{2m} : r^{\mathsf{A}}_i > p^{\mathsf{A}}_i \text{ for all } i \in T\}$ contains all types that adhere to the basic assumption that an attacker always prefers a successful attack (we will show a stronger result that allows for other specifications of Θ). In the program, the attacker reports a fake type β that results in the defender learning an SSE $(\mathbf{z}, t)$ for type β (by (7.3a)). An optimal solution thus yields a reporting strategy for a type-θ attacker that maximizes his *true* utility as the objective function specifies. Using Program (7.3), we present our key finding in Theorem 7.1. This result indicates that the defender essentially learns nothing because of the attacker's deception: she can well compute the maximin strategy without knowing any additional information about the attacker's payoffs.

Theorem 7.1. *There exists an optimal solution $(\beta, \mathbf{z}, t)$ to Program (7.3) such that $\mathbf{z}$ is a maximin strategy of the defender, i.e., $\mathbf{z} \in \arg\max_{\mathbf{c}\in\mathcal{C}} \min_{i\in T} u^{\mathsf{D}}(\mathbf{c}, i)$.*

Proof. Let $\overline{\mathbf{c}}$ be a maximin strategy of the defender and $\overline{u}$ be the corresponding maximin utility, i.e., $\overline{\mathbf{c}} \in \arg\max_{\mathbf{c}\in\mathcal{C}} \min_{i\in T} u^{\mathsf{D}}(\mathbf{c}, i)$ and $\overline{u} = \min_{i\in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, i)$. Consider the following solution $(\beta, \mathbf{z}, t)$ to Program (7.3):

- $z_i = \max\left\{0, \frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}}\right\}$ for all $i \in T$; (7.4)
- $t \in \mathsf{BR}_\beta(\mathbf{z})$ is an arbitrary best response of a type-β attacker; (7.5)
- $\beta = (\mathbf{r}, \mathbf{p})$, where $r_i = \begin{cases} -p_i^{\mathsf{D}}, & \text{if } i \neq t \\ -\min\{p_t^{\mathsf{D}}, \overline{u}\}, & \text{if } i = t \end{cases}$, and $p_i = -r_i^{\mathsf{D}}$ for all $i \in T$. (7.6)

We show that (i) $\mathbf{z}$ is a maximin strategy of the defender, (ii) $(\beta, \mathbf{z}, t)$ is a feasible solution of Program (7.3) and (iii) it is optimal.

We first focus on the case where $p_t^{\mathsf{D}} \leq \overline{u}$, and will show how the proof can be modified to show the same results when $p_t^{\mathsf{D}} > \overline{u}$. When $p_t^{\mathsf{D}} \leq \overline{u}$ we have $(r_i, p_i) = (-p_i^{\mathsf{D}}, -r_i^{\mathsf{D}})$ for all $i \in T$ by the specification in (7.6), so for any $\mathbf{c} \in \mathcal{C}$,

$$u_\beta^{\mathsf{A}}(\mathbf{c}, i) = -u^{\mathsf{D}}(\mathbf{c}, i). \tag{7.7}$$

Now we show (i)–(iii).

(i) Maximin. For all $i \in T$, since $z_i = \max\left\{0, \frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}}\right\} \geq \frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}}$, we have

$$u^{\mathsf{D}}(\mathbf{z}, i) \geq u^{\mathsf{D}}\left(\frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}}, i\right) = \frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}} \cdot r_i^{\mathsf{D}} + \left(1 - \frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}}\right) \cdot p_i^{\mathsf{D}} = \overline{u}. \tag{7.8}$$

It follows that

$$\min_{i\in T} u^{\mathsf{D}}(\mathbf{z}, i) \geq \overline{u} \geq \min_{i\in T} u^{\mathsf{D}}(\mathbf{c}, i)$$

for all $\mathbf{c} \in \mathcal{C}$, so $\mathbf{z}$ is indeed a maximin strategy. We still need to make sure that $\mathbf{z}$ is feasible, i.e., $\mathbf{z} \in \mathcal{C}$. Observe that when $z_i > 0$, (7.8) becomes an equality, so we have

$$u^{\mathsf{D}}(\mathbf{z}, i) = \overline{u} \leq u^{\mathsf{D}}(\overline{\mathbf{c}}, i)$$

which implies $z_i \leq \overline{c}_i$ by monotonicity. Thus, $\sum_{i\in T} z_i \leq \sum_{i\in T} \overline{c}_i \leq k$. To see that $0 \leq z_i \leq 1$ for all i, observe first that, by (7.4) it holds trivially that $z_i \geq 0$ for all i. As for $z_i \leq 1$, it suffices to show that $\frac{\overline{u}-p_i^{\mathsf{D}}}{r_i^{\mathsf{D}}-p_i^{\mathsf{D}}} \leq 1$; indeed, this holds as $\overline{u} = \min_{j\in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j) \leq u^{\mathsf{D}}(\overline{\mathbf{c}}, i) \leq r_i^{\mathsf{D}}$.

(ii) Feasibility. We show that $(\beta, \mathbf{z}, t)$ satisfies all the constraints of Program (7.3). Clearly, (7.3b) is satisfied because $r_i \geq -p_i^{\mathsf{D}} > -r_i^{\mathsf{D}} = p_i$ for all $i \in T$ by (7.6). To see that it also satisfies (7.3a), first observe that when $p_t^{\mathsf{D}} \leq \overline{u}$ we have

$$u_\beta^{\mathsf{A}}(\mathbf{z}, t) = u_\beta^{\mathsf{A}}\left(\frac{\overline{u}-p_t^{\mathsf{D}}}{r_t^{\mathsf{D}}-p_t^{\mathsf{D}}}, t\right) = \left(1 - \frac{\overline{u}-p_t^{\mathsf{D}}}{r_t^{\mathsf{D}}-p_t^{\mathsf{D}}}\right) \cdot (-p_t^{\mathsf{D}}) + \frac{\overline{u}-p_t^{\mathsf{D}}}{r_t^{\mathsf{D}}-p_t^{\mathsf{D}}} \cdot (-r_t^{\mathsf{D}}) = -\overline{u}. \tag{7.9}$$

Combining this with (7.7) and (7.8) gives, for all i,

$$u^{\mathsf{A}}_{\beta}(\mathbf{z}, t) = -\overline{u} \geq -u^{\mathsf{D}}(\mathbf{z}, i) = u^{\mathsf{A}}_{\beta}(\mathbf{z}, i). \tag{7.10}$$

Thus, $t \in \mathsf{BR}_{\beta}(\mathbf{z})$.

Now that $t \in \mathsf{BR}_{\beta}(\mathbf{z})$ and in (i) we have shown that $\mathbf{z} \in \mathcal{C}$, if we suppose (7.3a) is not satisfied, we would have $u^{\mathsf{D}}(\mathbf{z}', t') > u^{\mathsf{D}}(\mathbf{z}, t)$ for some $\mathbf{z}' \in \mathcal{C}$ and $t' \in \mathsf{BR}_{\beta}(\mathbf{z}')$. Applying (7.7), we find the following for all i:

$$u^{\mathsf{D}}(\mathbf{z}', i) = -u^{\mathsf{A}}_{\beta}(\mathbf{z}', i) \geq -u^{\mathsf{A}}_{\beta}(\mathbf{z}', t') = u^{\mathsf{D}}(\mathbf{z}', t'),$$

where the inequality is due to the fact that $t' \in \mathsf{BR}_{\beta}(\mathbf{z}')$. Thus,

$$u^{\mathsf{D}}(\mathbf{z}', i) \geq u^{\mathsf{D}}(\mathbf{z}', t') > u^{\mathsf{D}}(\mathbf{z}, t) \geq \min_{i \in T} u^{\mathsf{D}}(\mathbf{z}, i) = \overline{u}. \tag{7.11}$$

It follows that

$$\min_{i \in T} u^{\mathsf{D}}(\mathbf{z}', i) > \overline{u} = \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^{\mathsf{D}}(\mathbf{c}, i),$$

which is a contradiction given that $\mathbf{z}' \in \mathcal{C}$.

(iii) Optimality. Suppose that $(\beta, \mathbf{z}, t)$ is not optimal. Thus, there exists a feasible solution $(\beta', \mathbf{z}', t')$ such that $u^{\mathsf{A}}_{\theta}(\mathbf{z}', t') > u^{\mathsf{A}}_{\theta}(\mathbf{z}, t)$. By (7.5), $t \in \mathsf{BR}_{\theta}(\mathbf{z})$, so we have $u^{\mathsf{A}}_{\theta}(\mathbf{z}, t') \leq u^{\mathsf{A}}_{\theta}(\mathbf{z}, t) < u^{\mathsf{A}}_{\theta}(\mathbf{z}', t')$, which implies $z'_{t'} < z_{t'}$ by monotonicity. Since by definition $z_{t'} = \max\left\{0, \frac{\overline{u} - p^{\mathsf{D}}_{t'}}{r^{\mathsf{D}}_{t'} - p^{\mathsf{D}}_{t'}}\right\}$, now that $z_{t'} > z'_{t'} \geq 0$, it must be that $z'_{t'} < z_{t'} = \frac{\overline{u} - p^{\mathsf{D}}_{t'}}{r^{\mathsf{D}}_{t'} - p^{\mathsf{D}}_{t'}}$. Substituting this into the defender's utility function gives

$$\begin{aligned} u^{\mathsf{D}}(\mathbf{z}', t') < u^{\mathsf{D}}\left(\frac{\overline{u} - p^{\mathsf{D}}_{t'}}{r^{\mathsf{D}}_{t'} - p^{\mathsf{D}}_{t'}}, t'\right) &= \frac{\overline{u} - p^{\mathsf{D}}_{t'}}{r^{\mathsf{D}}_{t'} - p^{\mathsf{D}}_{t'}} \cdot r^{\mathsf{D}}_{t'} + \left(1 - \frac{\overline{u} - p^{\mathsf{D}}_{t'}}{r^{\mathsf{D}}_{t'} - p^{\mathsf{D}}_{t'}}\right) \cdot p^{\mathsf{D}}_{t'} \\ &= \overline{u} \\ &= \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^{\mathsf{D}}(\mathbf{c}, i) \\ &\leq \max_{\mathbf{c} \in \mathcal{C}, i \in \mathsf{BR}_{\beta}(\mathbf{c})} u^{\mathsf{D}}(\mathbf{c}, i). \end{aligned}$$

Thus, $(\beta', \mathbf{z}', t')$ violates (7.3a), contradicting the assumption that $(\beta', \mathbf{z}', t')$ is a feasible solution.

It remains to deal with the case when $p^{\mathsf{D}}_t > \overline{u}$. The only difference in this case is that now $r_t = -\overline{u} > -p^{\mathsf{D}}_t$ by (7.6), so (7.7) only holds for $i \neq t$. In our proof above, the arguments that rely on the assumption that $p^{\mathsf{D}}_t \leq \overline{u}$ and (7.7) are the equations in (7.9), (7.10), and (7.11), where the latter two now only hold for all $i \neq t$. However, observe the following:

- When $p^{\mathsf{D}}_t > \overline{u}$, we have $\frac{\overline{u} - p^{\mathsf{D}}_t}{r^{\mathsf{D}}_t - p^{\mathsf{D}}_t} < 0$ and thus $z_t = 0$ by (7.4). It follows that $u^{\mathsf{A}}_{\beta}(\mathbf{z}, t) = u^{\mathsf{A}}_{\beta}(0, t) = r_t = -\overline{u}$, so (7.9) holds as well.
- (7.10) holds trivially for $i = t$.
- When $p^{\mathsf{D}}_t > \overline{u}$, we have $u^{\mathsf{D}}(\mathbf{z}', t) \geq p^{\mathsf{D}}_t > \overline{u}$, thus establishing (7.11) for $i = t$ as well.

Therefore, (i)–(iii) hold when $p^{\mathsf{D}}_t > \overline{u}$ and the proof is complete. □

7.3.2 A Strengthened Result

Theorem 7.1 can be further strengthened under the assumption that the defender's maximin strategy is in the interior of $[0,1]^m$ (in which case the maximin strategy is also unique; see Lemma 7.6). With a slight abuse of terminology, we call such a strategy a *fully mixed* strategy (Definition 7.2). This assumption is mild as it is normally expected that no target would be so worthless that the defender would leave it wide open for the attacker to attack, while on the other hand resources are normally insufficient to allow any target to be fully protected.

The strengthening is two-fold: (i) under this additional assumption, the defender's maximin strategy is her *only* SSE strategy for any optimal attacker report, so the equilibrium selection issue that arises when a reported type induces the defender to play different strategies is avoided; (ii) one optimal attacker report, in particular, is the type that makes the game zero-sum, so our result holds even for a more stringent specification of Θ (e.g., when the defender has more precise knowledge about possible attacker types) as long as Θ contains the zero-sum attacker type. Indeed, it is very natural for an attacker to have the zero-sum type given the adversarial nature of SSGs. We present this strengthened result in Theorem 7.3.

Definition 7.2 (Fully mixed strategy)**.** A strategy $\mathbf{c} \in \mathbb{R}^m$ of the defender (as a coverage vector) is *fully mixed* if $0 < c_i < 1$ for all $i \in T$.

Theorem 7.3. *Suppose that $\overline{\mathbf{c}}$ is a maximin strategy of the defender and $(\beta, \mathbf{z}, t)$ is an arbitrary optimal solution to Program (7.3). If $\overline{\mathbf{c}}$ is fully mixed, then for every SSE $(\hat{\mathbf{c}}, \hat{i})$ for type β, it holds that $\hat{\mathbf{c}} = \overline{\mathbf{c}}$. In addition, there exists an optimal solution $(\beta', \mathbf{z}', t')$ such that $\beta' = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$.*

To prove this result we make several observations in the lemmas below. First, Lemma 7.4 provides a characterization of an SSE: in an SSE, we expect (i) the defender to fully utilize her resources to maximize the coverage, unless (ii) some target is already fully covered, in which case some fully covered target must be a best response of the attacker. Lemma 7.5 shows that when the game is zero-sum the defender's maximin strategy is also her strategy in an SSE, and Lemma 7.6 extends this result to the case of a fully mixed maximin strategy and demonstrates the uniqueness of the maximin strategy in this case. We state these lemmas below and leave the proofs to the appendix of this chapter (see Sections 7.A.1–7.A.3).

Lemma 7.4. *Suppose $(\hat{\mathbf{c}}, \hat{i})$ is an SSE for attacker type θ. The following holds:*

(i) If $\hat{c}_i < 1$ for all $i \in T$, then $\{i \in T : \hat{c}_i > 0\} \subseteq \mathsf{BR}_\theta(\hat{\mathbf{c}})$ and $\sum_{i \in T} \hat{c}_i = k$.

(ii) If $\hat{c}_i = 1$ for some $i \in T$, then there exists $j \in \mathsf{BR}_\theta(\hat{\mathbf{c}})$ such that $\hat{c}_j = 1$.

Lemma 7.5. *Suppose $\overline{\mathbf{c}}$ is a maximin strategy of the defender. Then $(\overline{\mathbf{c}}, i)$ forms an SSE in a zero-sum game for all $i \in \mathsf{BR}_\beta(\overline{\mathbf{c}})$, where $\beta = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$.*

Lemma 7.6. *Suppose $\overline{\mathbf{c}}$ is a maximin strategy of the defender and $\overline{\mathbf{c}}$ is fully mixed (i.e., $0 < \overline{c}_i < 1$ for all $i \in T$). Then for all $i \in T$, we have that $(\overline{\mathbf{c}}, i)$ forms an SSE in a zero-sum game and $u^{\mathsf{D}}(\overline{\mathbf{c}}, i) = \min_{j \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j)$. In addition, $\overline{\mathbf{c}}$ is the only maximin strategy of the defender.*

Now we are ready to prove Theorem 7.3.

Proof of Theorem 7.3. We first show that there exists an optimal solution $(\beta', \mathbf{z}', t')$ such that $\beta' = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$. In Theorem 7.1, we have shown the existence of an optimal solution containing a maximin strategy of the defender. By Lemma 7.6, $\overline{\mathbf{c}}$ is the only maximin strategy if it is fully mixed. Thus, there exists an optimal solution that contains $\overline{\mathbf{c}}$. We can then fix $\overline{\mathbf{c}}$ in Program (7.3), and show that in the resulting program, an optimal solution (β', t') is such that $\beta' = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$. Indeed, now that $\overline{\mathbf{c}}$ is fixed, we can select $t' \in \arg\max_{i \in T} u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, i)$ and it remains to show that (β', t') is feasible. Since the game with attacker type β' is a zero-sum game, by Lemma 7.6, $(\overline{\mathbf{c}}, t')$ forms an SSE for type β', so Constraint (7.3a) is satisfied. Moreover, given that $r_i^{\mathsf{D}} > p_i^{\mathsf{D}}$ for all $i \in T$, we also have $r_i^{\mathsf{A}} > p_i^{\mathsf{A}}$ if we let $r_i^{\mathsf{A}} = -p_i^{\mathsf{D}}$ and $p_i^{\mathsf{A}} = -r_i^{\mathsf{D}}$; hence, $\beta' = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}}) \in \Theta$, (7.3b) is also satisfied, and (β', t') is feasible. Therefore, $(\beta', \overline{\mathbf{c}}, t')$ is an optimal solution to Program (7.3).

Now we show the first part of the theorem, i.e., the uniqueness of the induced strategy. Suppose for a contradiction that there exists an SSE $(\hat{\mathbf{c}}, \hat{i})$ for attacker type β such that $\hat{\mathbf{c}} \neq \overline{\mathbf{c}}$. Consider the two possibilities under this condition.

Case (i). $\hat{c}_i \leq \overline{c}_i$ for all $i \in T$, and this inequality is strict for some i. It follows that $\overline{c}_i < 1$ for all $i \in T$ and $\sum_{i \in T} \hat{c}_i < \sum_{i \in T} \overline{c}_i \leq k$, which contradicts Lemma 7.4 (i).

Case (ii). $\hat{c}_j > \overline{c}_j$ for some $j \in T$. If it is also the case that $\hat{c}_i < 1$ for all $i \in T$, by Lemma 7.4 (i), we have $j \in \mathsf{BR}_\beta(\hat{\mathbf{c}})$; If otherwise $\hat{c}_i = 1$ for some $i \in T$, by Lemma 7.4 (ii), there exists $j' \in \mathsf{BR}_\beta(\hat{\mathbf{c}})$ such that $\hat{c}_{j'} = 1 > \overline{c}_{j'}$. In both cases, we find some $j \in \mathsf{BR}_\beta(\hat{\mathbf{c}})$ such that $\hat{c}_j > \overline{c}_j$; by monotonicity of $u^{\mathsf{D}}(\cdot, j)$, we have

$$u^{\mathsf{D}}(\hat{\mathbf{c}}, j) > u^{\mathsf{D}}(\overline{\mathbf{c}}, j).$$

It follows that

$$u^{\mathsf{D}}(\mathbf{z}, t) = u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i}) = \max_{i \in \mathsf{BR}_\beta(\hat{\mathbf{c}})} u^{\mathsf{D}}(\hat{\mathbf{c}}, i) \geq u^{\mathsf{D}}(\hat{\mathbf{c}}, j) > u^{\mathsf{D}}(\overline{\mathbf{c}}, j) = u^{\mathsf{D}}(\overline{\mathbf{c}}, t),$$

where the first equality $u^{\mathsf{D}}(\mathbf{z}, t) = u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i})$ is due to the fact that both $(\mathbf{z}, t)$ and $(\hat{\mathbf{c}}, \hat{i})$ are SSEs for attacker type β; and the last equality $u^{\mathsf{D}}(\overline{\mathbf{c}}, j) = u^{\mathsf{D}}(\overline{\mathbf{c}}, t)$ holds by Lemma 7.6. Thus, $u^{\mathsf{D}}(\mathbf{z}, t) > u^{\mathsf{D}}(\overline{\mathbf{c}}, t)$, so by monotonicity, we have $z_t > \overline{c}_t$ and, in turn, $u_\theta^{\mathsf{A}}(\mathbf{z}, t) < u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, t)$. This gives

$$u_\theta^{\mathsf{A}}(\mathbf{z}, t) < u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, t) \leq \max_{i \in T} u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, i) = u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, t'),$$

where $t' \in \arg\max_{i \in T} u_\theta^{\mathsf{A}}(\overline{\mathbf{c}}, i)$ as we selected above. The above inequality means that $(\beta', \overline{\mathbf{c}}, t')$ is a better solution than $(\beta, \mathbf{z}, t)$ to Program (7.3), which contradicts the assumption in the statement of the theorem.

As both cases lead to contradictions, it follows that $\hat{\mathbf{c}} = \overline{\mathbf{c}}$ and this completes the proof. □

Relation with Optimal Deception in Chapter 6

Recall that in Chapter 6, our key result is that the follower can induce a fake SSE as long as this SSE provides the leader at least his maximin utility. The above result essentially shows that in a security game, the follower has an incentive to exploit this indelibility power to the maximum degree. In other words, the optimal solution to LP (6.5) always makes the constraint $u^{\mathsf{L}}(\mathbf{x}, j) \geq M$ binding when the underlying game is an SSG (under our mild assumption), whereas this is not necessarily true for general Stackelberg games.

7.4 Commitment to a Policy

As we have analyzed in Chapter 5, the key to the success of the attacker's trick is the naive playbook the defender follows — to always play the learned optimal commitment *as is*. We adopt the same policy-based framework in Chapter 5 and study how to compute the optimal *policy* for the defender to use in an SSG.

In the context of SSGs, a policy is a function $\pi : \Theta \to \mathcal{C} \times T$ that maps a reported attacker type to an outcome $(\mathbf{c}, i)$, which consists of a strategy $\mathbf{c}$ of the defender and a best response $i \in \mathsf{BR}_\theta(\mathbf{c})$ of the reported attacker type θ. The way the defender plays when she ignores attacker manipulation can itself be viewed as a policy that maps every reported type θ to an SSE for θ; we will refer to this policy as the *SSE policy*. After knowing the defender's policy commitment, a strategic attacker of true type θ chooses to report an optimal type $\beta \in \arg\max_{\beta' \in \Theta} u^{\mathsf{A}}_\theta(\pi(\beta'))$ that will maximize his utility in the realization of the policy, where the defender plays $\mathbf{c}$ if $\pi(\beta) = (\mathbf{c}, i)$, inducing the attacker to respond by attacking i. At a higher level, this can be seen as a Stackelberg game in which the defender commits to a policy and the attacker reports optimally in response to this commitment.

The Quality Measure To find the optimal policy, we need a good measure for the quality of a policy. When there is no other prior information about the distribution of the attacker's type, a worst-case analysis seems to be appropriate and a straightforward choice of quality measure is the utility the defender obtains when playing against the worst attacker type. However, as Proposition 7.7 suggests, this measure prevents us from distinguishing the performance of many policies. Specifically, when there is a possible attacker type that makes the game zero-sum, no feasible policy can achieve anything better than the maximin utility, so if we take the worst-case utility as the measure, the quality of all policies would be hindered by this attacker type, and the SSE policy, which achieves exactly the maximin utility in the worst case, would already be the best policy we can hope for. Essentially, there will be no room for improvement other than letting the attacker lie.

Proposition 7.7. *Let $\overline{\mathbf{c}}$ be a maximin strategy of the defender, and $\overline{u} = \min_{i \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, i)$ be the maximin utility. For any policy π, let $\gamma \in \arg\max_{\theta \in \Theta} u^{\mathsf{A}}_\beta(\pi(\theta))$ be an optimal report of a type-β attacker, $\beta = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$. Then $u^{\mathsf{D}}(\pi(\gamma)) \leq \overline{u}$.*

Proof. Let $(\mathbf{z}, j) = \pi(\beta)$, which is the outcome a type-β attacker would get if he reports truthfully. By definition, $j \in \mathsf{BR}_\beta(\mathbf{z})$. By Lemma 7.5, $\overline{\mathbf{c}}$ as the defender's maximin strategy is exactly her SSE strategy in a zero-sum game and $(\overline{\mathbf{c}}, t)$ forms an SSE for any $t \in \mathsf{BR}_\beta(\overline{\mathbf{c}})$. Thus, $u^{\mathsf{D}}(\mathbf{z}, j) \leq u^{\mathsf{D}}(\overline{\mathbf{c}}, t)$ and $u^{\mathsf{A}}_\beta(\overline{\mathbf{c}}, t) = \max_{i \in T} u^{\mathsf{A}}_\beta(\overline{\mathbf{c}}, i)$. Since β makes the game zero-sum, $u^{\mathsf{A}}_\beta(\mathbf{c}, i) = -u^{\mathsf{D}}(\mathbf{c}, i)$ for any $\mathbf{c}$ and i. It follows that

$$u^{\mathsf{A}}_\beta(\mathbf{z}, j) = -u^{\mathsf{D}}(\mathbf{z}, j) \geq -u^{\mathsf{D}}(\overline{\mathbf{c}}, t) = u^{\mathsf{A}}_\beta(\overline{\mathbf{c}}, t) = \max_{i \in T} u^{\mathsf{A}}_\beta(\overline{\mathbf{c}}, i) = -\min_{i \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, i) = -\overline{u}.$$

Now suppose for the sake of contradiction that $u^{\mathsf{D}}(\pi(\gamma)) > \overline{u}$. Then we have

$$u^{\mathsf{A}}_\beta(\pi(\gamma)) = -u^{\mathsf{D}}(\pi(\gamma)) < -\overline{u} \leq u^{\mathsf{A}}_\beta(\mathbf{z}, j) = u^{\mathsf{A}}_\beta(\pi(\beta)),$$

so the attacker would be strictly better-off reporting β. This contradicts the assumption that γ is an optimal reporting strategy of a type-β attacker. □

Algorithm 7.1: Decide if there exists a policy π such that $\mathrm{EoP}(\pi) \geq \xi$.

1. For each $\theta \in \Theta$, compute an SSE $(\hat{\mathbf{c}}^\theta, \hat{i}^\theta)$ for type θ. Let $\hat{u}(\theta) = u^{\mathsf{D}}(\hat{\mathbf{c}}^\theta, \hat{i}^\theta)$.
2. Sort attacker types in Θ according to $\hat{u}(\theta)$, so that $\hat{u}(\theta_1) \geq \hat{u}(\theta_2) \geq \cdots \geq \hat{u}(\theta_\lambda)$, $\lambda = |\Theta|$.
3. For each $\ell = 1, \ldots, \lambda$:
   let $\pi(\theta_\ell) \leftarrow (\mathbf{z}, t)$, where $z_i = \min\{\hat{c}_i^{\theta_\ell},\ h_i\}$, $t = \mathsf{BR}_{\theta_\ell}(\mathbf{h})$, and
   $$h_i = \max\left\{0,\ \frac{\xi \cdot \hat{u}(\theta_\ell) - p_i^{\mathsf{D}}}{r_i^{\mathsf{D}} - p_i^{\mathsf{D}}},\ \max_{\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}} \frac{u_\theta^{\mathsf{A}}(\pi(\theta)) - r_i^\theta}{p_i^\theta - r_i^\theta}\right\}.$$
4. If $\mathrm{EoP}(\pi) \geq \xi$ and π is feasible, return π as a satisfying policy; Otherwise, claim that no such policy exists.

This is unreasonable: even in the truthful situation, it is impossible for the defender to achieve more than the maximin utility when the attacker is of the zero-sum type, so it would be unfair to underrate a policy simply because it underperforms against the zero-sum type. For this reason, we adopt the EoP (efficiency of the policy) measure defined in Chapter 5 (see Definition 5.15), to take into account the intrinsic hardness of playing against each attacker type. The EoP is the worst-case ratio between the utility the defender obtains and what she should have obtained had the attacker been truthful. For the EoP to be meaningful, we shift all payoffs to be non-negative. Without loss of generality, we assume Θ — the set of types the attacker is allowed to report — to be also the set of possible attacker types, which is common knowledge to both players. We present an efficient algorithm to compute the optimal policy with the EoP measure in the next section for a finite Θ.

7.5 Optimal Policy for a Finite Set of Attacker Types

When Θ is a finite set, a defender policy can be represented as a set of $\lambda = |\Theta|$ outcomes, i.e., $\pi = \{(\mathbf{c}^\theta, i^\theta)\}_{\theta \in \Theta}$ and $\pi(\theta) = (\mathbf{c}^\theta, i^\theta)$ for each $\theta \in \Theta$. Our analysis in Chapter 5 already demonstrated the inapproximability of the problem of computing the EoP maximizing policy in generic Stackelberg games (see Theorem 5.17), but thanks to the special utility structure of SSGs, the problem admits a polynomial-time algorithm when the underlying game is an SSG.

We consider the decision version of this optimization problem: for a given value ξ, decide whether any defender policy π achieves $\mathrm{EoP}(\pi) \geq \xi$. Trivially, once we have an efficient algorithm for this decision problem, the best EoP can be found efficiently with any degree of precision by using binary search (in particular, we already know that the value always lies in $[0, 1]$). Our algorithm for this decision problem, presented as Algorithm 7.1, is constructive; it generates a satisfying policy when there exists one. In the remainder of this section, we will let $\Theta = \{\theta_1, \ldots, \theta_\lambda\}$, where $\theta_1, \ldots, \theta_\lambda$ are ordered according to the utility they offer the defender in an SSE, as in Step 2 of Algorithm 7.1; this order can be computed efficiently given that an SSE can be computed in polynomial time. We call a policy *ℓ-compatible* if truthful reporting is incentivized for every attacker type θ_j, $j \leq \ell$ (Definition 7.8).

Definition 7.8. A policy π is *ℓ-compatible* ($0 \leq \ell \leq \lambda$), if in response to π, it is optimal for every attacker type $\theta \in \{\theta_1, \ldots, \theta_\ell\}$ to report truthfully, i.e., $u_\theta^{\mathsf{A}}(\pi(\theta)) \geq u_\theta^{\mathsf{A}}(\pi(\beta))$ for all $\beta \in \Theta$.

The correctness of Algorithm 7.1 is shown via Theorem 7.9. Intuitively, we can show that in Step 3, we have $z_t = h_t$ if there exists a satisfying policy, so the way h_t is defined ensures that:

- $z_t \geq \frac{u_\theta^{\mathsf{A}}(\pi(\theta)) - r_t^\theta}{p_t^\theta - r_t^\theta}$ for all $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$, so that $u_\theta^{\mathsf{A}}(\mathbf{z}, t) \leq u_\theta^{\mathsf{A}}\left(\frac{u_\theta^{\mathsf{A}}(\pi(\theta)) - r_t^\theta}{p_t^\theta - r_t^\theta}, t\right) = u_\theta^{\mathsf{A}}(\pi(\theta))$ and a type-θ attacker will have no incentive to misreport θ_ℓ.
- $z_t \geq \frac{\xi \cdot \hat{u}(\theta_\ell) - p_t^{\mathsf{D}}}{r_t^{\mathsf{D}} - p_t^{\mathsf{D}}}$, so that $u^{\mathsf{D}}(\mathbf{z}, t) \geq u^{\mathsf{D}}\left(\frac{\xi \cdot \hat{u}(\theta_\ell) - p_t^{\mathsf{D}}}{r_t^{\mathsf{D}} - p_t^{\mathsf{D}}}, t\right) = \xi \cdot \hat{u}(\theta_\ell)$. Given the way the types are ordered, we have $\hat{u}(\theta_\ell) \geq \hat{u}(\theta)$ for all $\theta \in \{\theta_\ell, \ldots, \theta_\lambda\}$, so if a type-$\theta$ attacker misreports type θ_ℓ, we would have

$$\mathrm{EoP}_\theta(\pi) = \frac{u^{\mathsf{D}}(\pi(\theta_\ell))}{\hat{u}(\theta)} \geq \frac{u^{\mathsf{D}}(\pi(\theta_\ell))}{\hat{u}(\theta_\ell)} = \frac{u^{\mathsf{D}}(\mathbf{z}, t)}{\hat{u}(\theta_\ell)} \geq \xi.$$

If the above two conditions hold for all $\theta \in \Theta$, the consequence is that no type-θ_j attacker would be incentivized to report any θ_ℓ with $\ell > j$; while on the contrary a type-θ_ℓ attacker misreporting θ_j would not cause the EoP value to drop below ξ. Interestingly, the policy generated by Algorithm 7.1 turns out to be *incentive compatible* (λ-compatible as in Definition 7.8); it incentivizes all attacker types to report truthfully.

Theorem 7.9. *In time* $\mathsf{poly}(k, m, |\Theta|)$*, Algorithm 7.1 either outputs a policy* π *with* $\mathrm{EoP}(\pi) \geq \xi$*, or decides correctly that no such policy exists. The policy generated is incentive compatible.*

To prove Theorem 7.9 we need the following result.

Lemma 7.10. *Let* π *be the policy generated in Step 3 of Algorithm 7.1. Suppose that there exists an* $(\ell-1)$*-compatible policy* π^**,* $\mathrm{EoP}(\pi^*) \geq \xi$*, and* $\pi^*(\theta) = \pi(\theta)$ *for all* $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$*. Then the policy* $\tilde{\pi}$*, such that* $\tilde{\pi}(\theta) = \begin{cases} \pi^*(\theta), & \text{if } \theta \in \Theta \setminus \{\theta_\ell\} \\ \pi(\theta), & \text{if } \theta = \theta_\ell \end{cases}$*, is feasible and* ℓ*-compatible, and* $\mathrm{EoP}(\tilde{\pi}) \geq \xi$*.*

We leave the proof of Lemma 7.10 to the appendix (see Section 7.A.4). Lemma 7.10 implies that Algorithm 7.1 can be viewed as a process of repeatedly replacing the ℓ-th outcome of a satisfying policy with the outcome generated in the ℓ-th iteration of Step 3. The new policy obtained after every replacement will still be a satisfying one. Hence, eventually, we will obtain a satisfying policy that consists exclusively of outcomes generated by the algorithm, which means that we do not actually need to be given a satisfying policy to begin with. Now we can prove Theorem 7.9.

Proof of Theorem 7.9. The polynomial runtime is readily seen. Trivially, if there exists no feasible policy π with $\mathrm{EoP}(\pi) \geq \xi$, then Algorithm 7.1 will decide correctly at Step 4 that no satisfying policy exists. Conversely, suppose that there exists a policy π^*, $\mathrm{EoP}(\pi^*) \geq \xi$. Let $\tilde{\pi}^0 = \pi^*$, and for each $\ell = 1, \ldots, \lambda$, we iteratively construct a policy $\tilde{\pi}^\ell$ by replacing $\tilde{\pi}^{\ell-1}(\theta_\ell)$ in policy $\tilde{\pi}^{\ell-1}$ with the outcome $\pi(\theta_\ell)$ generated in Step 3; thus, $\tilde{\pi}^\lambda = \pi$. Trivially, $\tilde{\pi}^0$ is 0-compatible as is any feasible policy. Applying Lemma 7.10 iteratively, we can also conclude that $\tilde{\pi}^\ell$ is ℓ-compatible and $\mathrm{EoP}(\tilde{\pi}^\ell) \geq \xi$ for every ℓ; in particular, $\tilde{\pi}^\lambda$ is λ-compatible and $\mathrm{EoP}(\pi^\lambda) \geq \xi$. Algorithm 7.1 outputs $\pi = \tilde{\pi}^\lambda$ as a satisfying policy. □

7.6 Beyond Finite Attacker Types

The above approach only applies to a finite type set. Therefore, we present a heuristic approach to deal with a continuous or even unknown type set. This approach is inspired by the *quantal response (QR)* model that is developed to study bounded rationality of human players [McKelvey and Palfrey, 1995]. In a QR equilibrium, players are assumed to play not only their optimal pure strategy but also every other strategy with a probability positively related to the utility the player gets from playing that strategy.

The QR policy imitates the irrational behavior in a QR equilibrium. Recall that in an SSE, a rational defender commits to an optimal strategy $\mathbf{c}$ and induces a type-θ attacker to choose a response $i^* \in \mathsf{BR}_\theta(\mathbf{c})$ that maximizes the defender's utility. The QR policy, however, induces the attacker's tie-breaking choice in an "irrational" way. It induces the attacker to choose not only i^*, but also *every* target in $\mathsf{BR}_\theta(\mathbf{c})$ with some probability; the probability of a target being chosen is positively related to the defender's utility when this target is attacked. The idea is to add some uncertainty to the outcome, so that the attacker cannot benefit from being induced to choose a particular response with certainty, which is crucial for the success of his manipulation. This also encourages truthful reporting to some extent: a truthful attacker, who reports his true type θ, is indifferent as to which response he is induced to choose in $\mathsf{BR}_\theta(\mathbf{c})$ and hence immune to such uncertainty. The QR policy is as follows.

Definition 7.11 (QR policy)**.** For each type θ, let $\hat{\mathbf{c}}^\theta$ be the defender's strategy in an SSE for attacker type θ. Given a constant $\varphi > 0$, a QR policy π^{QR} maps a report θ to a distribution σ over outcomes in $\{(\hat{\mathbf{c}}^\theta, i) : i \in \mathsf{BR}_\theta(\hat{\mathbf{c}}^\theta)\}$; the probability $\sigma(i)$ of each outcome $(\hat{\mathbf{c}}^\theta, i)$ is

$$\sigma(i) = \frac{f_i}{\sum_{j \in \mathsf{BR}_\theta(\hat{\mathbf{c}}^\theta)} f_j},$$

where $f_j = e^{\varphi \cdot u^{\mathsf{D}}(\hat{\mathbf{c}}^\theta, j)}$ for each $j \in T$.

The parameter φ in the above definition represents a player's rationality level in the QR model: when $\varphi \to 0$, a player behaves completely irrationally, playing each strategy uniformly at random; when $\varphi \to +\infty$, a player becomes perfectly rational, choosing the optimal strategy with certainty. A defender who uses π^{QR} then samples an outcome from $\pi^{\text{QR}}(\theta)$ to implement when θ is reported. The players are now concerned with their expected utility over the outcome distribution, e.g., for the defender: $u^{\mathsf{D}}(\pi^{\text{QR}}(\theta)) = \sum_{i \in \mathsf{BR}_\theta(\hat{\mathbf{c}}^\theta)} \sigma(i) \cdot u^{\mathsf{D}}(\hat{\mathbf{c}}^\theta, i)$. The EoP can also be redefined accordingly. Since $\hat{\mathbf{c}}^\theta$ and σ are independent of other types in Θ, the QR policy can be implemented on-the-fly for the type reported, and is thus able to handle infinite or unknown type sets.

Intuitively, the QR policy strikes a balance between two unaligned aspects of playing against a manipulating attacker: (1) it adds some uncertainty to discourage attacker manipulation; meanwhile, (2) the softmax function that defines σ loosely nudges the induced attacker response towards the optimal one for the defender, so that the cost of achieving (1) is prevented from being too high. We empirically evaluate the performance of the QR policy in randomly generated games in the next section.

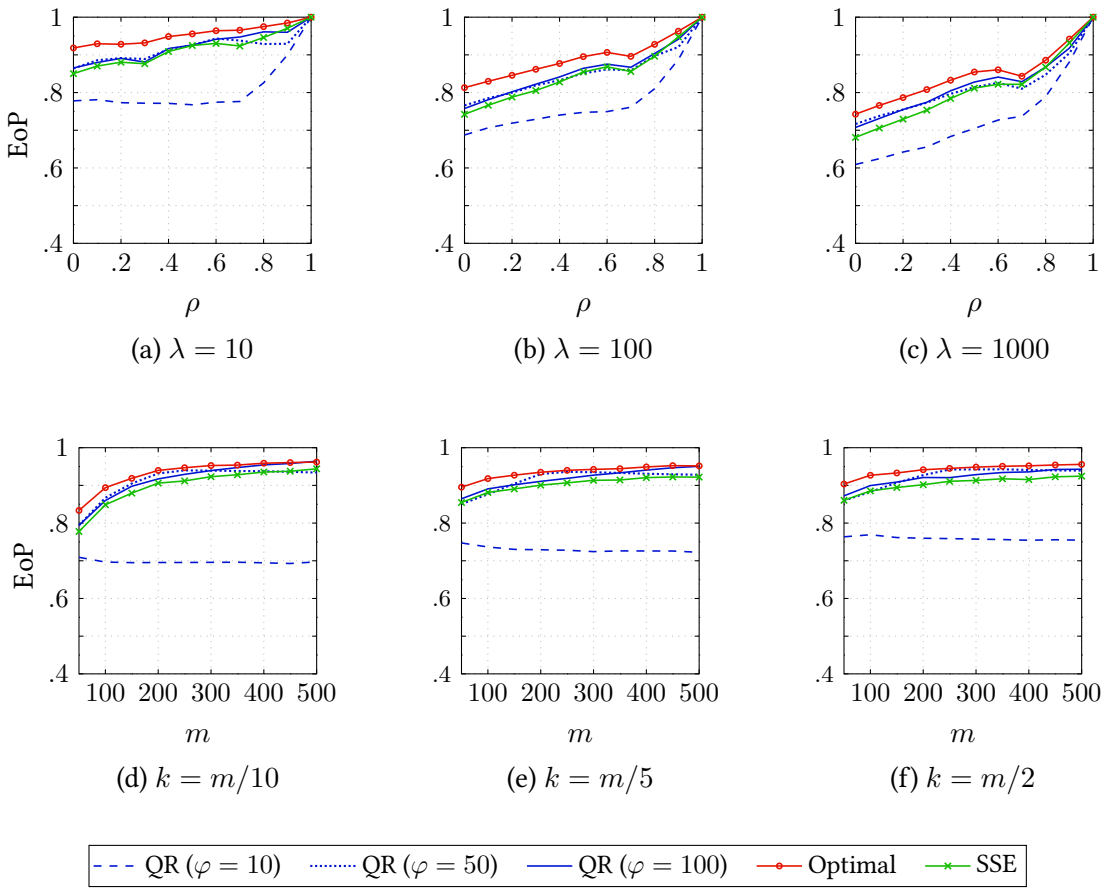
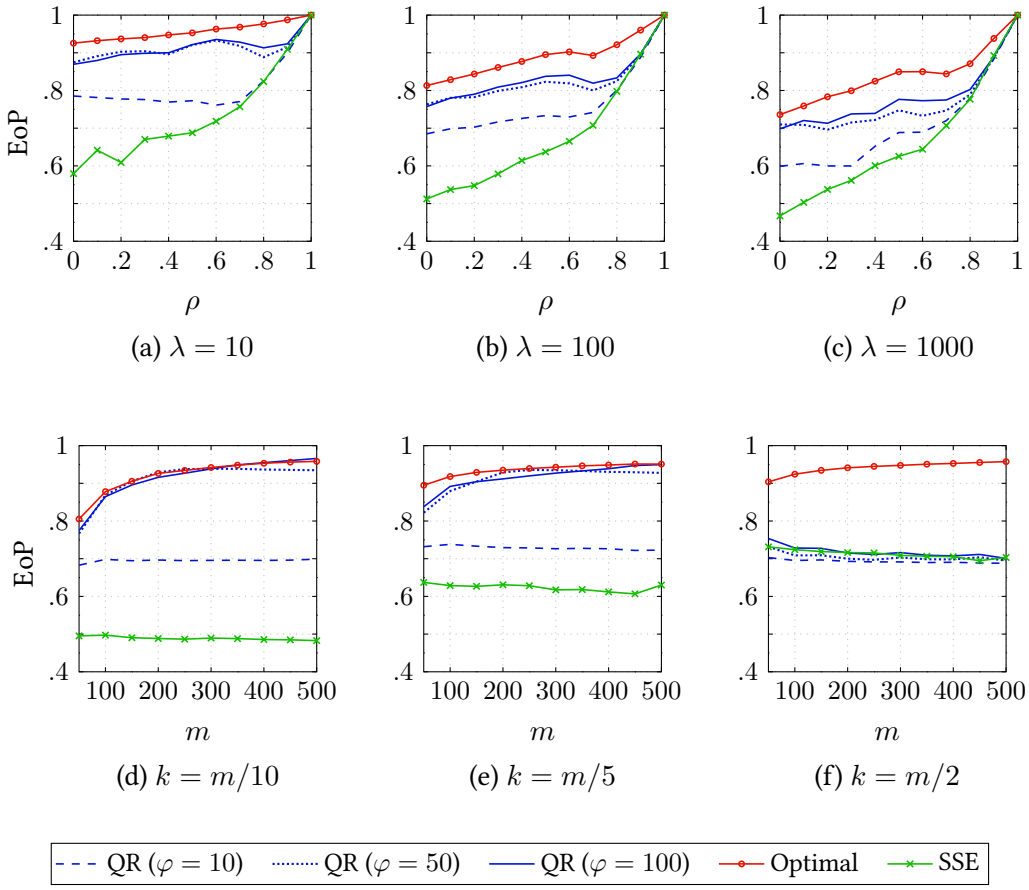


Figure 7.1: Comparison of the EoP. Other parameters are set to $m = 50$, $k = 10$ in (a)–(c); and $\rho = 0.5$, $\lambda = 100$ in (d)–(f).



(a) $\lambda = 10$ (b) $\lambda = 100$ (c) $\lambda = 1000$

(d) $k = m/10$ (e) $k = m/5$ (f) $k = m/2$

Figure 7.2: Comparison of the EoP when the zero-sum attacker type is always included in Θ in addition to types generated by the covariance model. Other parameters are set to $m = 50$, $k = 10$ in (a)–(c); and $\rho = 0.5$, $\lambda = 100$ in (d)–(f).

7.7 Empirical Evaluation

In our evaluations, attacker types are randomly generated using the covariance model [Nudelman et al., 2004], with a parameter $\rho \in [0, 1]$ to control the closeness of the generated game to a zero-sum game. That is, we shift each payoff parameter x towards the corresponding parameter y of a zero-sum attacker type, letting $x \leftarrow (1 - \rho) \cdot x + \rho \cdot y$. Thus, when $\rho = 1$ the game generated is exactly zero-sum, and when $\rho = 0$ all payoffs are generated uniformly at random. All evaluations are conducted on finite type sets. This also simulates situations with an unknown (but finite) type set, though situations with infinite type sets require more sophisticated approaches (in this case, it is unclear to us how to compute the optimal attacker report in response to the QR policy). All results shown are averaged over at least 50 runs.

We compare the EoP achieved by our optimal and heuristic policies, using the SSE policy as a benchmark (i.e., the situation where attacker manipulation is ignored). The first set of results, shown in Figure 7.1, illustrates the variance of the EoP with respect to ρ and the size of the game. Except for the QR policy with $\varphi = 10$, performance of all other policies is very close to each other, though there is a discernible gap between the optimal policy and the SSE policy. In general, in these results, the loss due to ignoring attacker manipulation appears to be very marginal.

A more interesting set of results is shown in Figure 7.2, in which we slightly tweak the randomly generated type set, by always adding a zero-sum attacker type into it. This small change leads to a very different pattern in the results. There is a wide gap between the optimal and the SSE policies, and the QR policies normally rest in between them, exhibiting good performance as well. These results corroborate our theoretical analysis, that all attacker types will be incentivized to report the zero-sum type when they are allowed to, which undermines the performance of the SSE policy significantly. The optimal policy, however, is able to achieve very high EoP, sometimes close to recovering the defender's utility in the truthful situation ($\mathrm{EoP} = 1$).

7.8 Summary

In this chapter, we investigate ways of manipulating algorithms that are designed to learn the optimal strategy to commit to in Stackelberg security games, and aim at remedying the overoptimistic assumption of a truthful attacker adopted by these algorithms. We propose exact and heuristic approaches to reduce the loss due to manipulation. The effectiveness of our approaches is evaluated both theoretically and empirically.

Appendix 7.A Omitted Proofs

7.A.1 Proof of Lemma 7.4

Since $(\hat{\mathbf{c}}, \hat{i})$ is an SSE, by definition, we have $\hat{i} \in \mathsf{BR}_\theta(\hat{\mathbf{c}})$, and $u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i}) \geq u^{\mathsf{D}}(\mathbf{c}, i)$ for all $\mathbf{c} \in \mathcal{C}, i \in \mathsf{BR}_\theta(\mathbf{c})$. We claim that $\hat{\mathbf{c}}$ is the optimal solution to the following linear program:

$$\begin{aligned}
\max_{\mathbf{c}} \quad & u^{\mathsf{D}}(\mathbf{c}, \hat{i}) && && (7.12)\\
\text{subject to} \quad & u^{\mathsf{A}}_\theta(\mathbf{c}, i) \leq u^{\mathsf{A}}_\theta(\mathbf{c}, \hat{i}) && \text{for all } i \in T \setminus \{\hat{i}\} && (7.12a)\\
& \textstyle\sum_{i \in T} c_i \leq k && && (7.12b)\\
& 0 \leq c_i \leq 1 && \text{for all } i \in T && (7.12c)
\end{aligned}$$

To see this, note that, first $\hat{\mathbf{c}}$ is a feasible solution because it satisfies all the constraints above: (7.12a) is equivalent to $\hat{i} \in \mathsf{BR}_\theta(\mathbf{c})$, and (7.12b) and (7.12c) combined are equivalent to $\mathbf{c} \in \mathcal{C}$. Second, $\hat{\mathbf{c}}$ is optimal because if it is not, there would exist another feasible solution $\mathbf{z} \neq \hat{\mathbf{c}}$, such that $u^{\mathsf{D}}(\mathbf{z}, \hat{i}) > u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i})$; this contradicts the assumption that $(\hat{\mathbf{c}}, \hat{i})$ forms an SSE.

By the Karush-Kuhn-Tucker (KKT) conditions, $\hat{\mathbf{c}}$ is an optimal solution only if there exists constants (i.e., KKT multipliers) α_i, β, γ_i and δ_i, each corresponding to an inequality constraint in (7.12a)–(7.12c), such that for all $i \in T$ (let $w^{\mathsf{A}}_i = p^\theta_i - r^\theta_i$ and $w^{\mathsf{D}}_i = r^{\mathsf{D}}_i - p^{\mathsf{D}}_i$ for each i below):

$$\begin{cases} -\ w^{\mathsf{A}}_i \cdot \alpha_i - \beta + \gamma_i - \delta_i = 0, & \text{for all } i \neq \hat{i} \\ w^{\mathsf{D}}_{\hat{i}} + w^{\mathsf{A}}_{\hat{i}} \cdot \sum_{i \neq \hat{i}} \alpha_i - \beta + \gamma_{\hat{i}} - \delta_{\hat{i}} = 0 \end{cases} \qquad \text{(by stationarity[1])} \qquad (7.13)$$

$$\alpha_i, \beta, \gamma_i, \delta_i \geq 0 \qquad \text{(by dual feasibility)} \qquad (7.14)$$

$$\begin{cases} \alpha_i \cdot \left(u^{\mathsf{A}}_\theta(\hat{\mathbf{c}}, i) - u^{\mathsf{A}}_\theta(\hat{\mathbf{c}}, \hat{i}) \right) = 0 \\ \beta \cdot \left(\sum_{j \in T} \hat{c}_j - k \right) = 0 \\ \gamma_i \cdot \hat{c}_i = 0 \\ \delta_i \cdot (\hat{c}_i - 1) = 0 \end{cases} \qquad \text{(by complementary slackness[2])} \qquad (7.15)$$

Now we prove (i) and (ii) in the statement of the lemma separately.

Part (i). Since $\hat{c}_i < 1$ for all $i \in T$, we have $\delta_i = 0$ for all $i \in T$ by the last equation in (7.15). Suppose for the sake of contradiction that $\sum_{i \in T} \hat{c}_i < k$. We would have $\beta = 0$ by the second equation in (7.15); and further, by the second equation in (7.13), $w^{\mathsf{D}}_{\hat{i}} + w^{\mathsf{A}}_{\hat{i}} \cdot \sum_{i \neq \hat{i}} \alpha_i + \gamma_{\hat{i}} = 0$, which is a contradiction as $w^{\mathsf{D}}_{\hat{i}} = r^{\mathsf{D}}_{\hat{i}} - p^{\mathsf{D}}_{\hat{i}} > 0$, $w^{\mathsf{A}}_{\hat{i}} = r^{\mathsf{A}}_{\hat{i}} - p^{\mathsf{A}}_{\hat{i}} > 0$, and $\alpha_i, \gamma_i \geq 0$ for all i by (7.13). Thus, $\beta > 0$ and $\sum_{i \in T} \hat{c}_i = k$.

Similarly, if we suppose $\hat{c}_t > 0$ for some $t \in T$, but $t \notin \mathsf{BR}_\theta(\hat{\mathbf{c}})$, we would have $u^{\mathsf{A}}_\theta(\hat{\mathbf{c}}, t) < \max_{i \in T} u^{\mathsf{A}}_\theta(\hat{\mathbf{c}}, i) = u^{\mathsf{A}}_\theta(\hat{\mathbf{c}}, \hat{i})$. Thus, $\gamma_t = 0$ and $\alpha_t = 0$ by the third and the first equations in (7.15), and then $\beta + \delta_t = 0$ by the first equation in (7.13) (note that $t \neq \hat{i}$ since $\hat{i} \in \mathsf{BR}_\theta(\hat{\mathbf{c}})$), which contradicts $\beta > 0$ and $\delta_i = 0$ for all $i \in T$ as we show above.

[1] A solution x satisfies stationarity if $\nabla f(x) = \lambda_1 \cdot \nabla g_1(x) + \cdots + \lambda_\ell \cdot \nabla g_\ell(x)$, where f is the objective function (minimization), each g_i corresponds to an inequality constraint (in the form $g_i(x) \leq 0$), and each λ_i is a KKT multiplier.

[2] A solution x satisfies complementary slackness if $\lambda_i \cdot g_i(x) = 0$ for each KKT multiplier λ_i and their corresponding inequality constraint function g_i.

Part (ii). Suppose that $\hat{c}_i = 1$ for some $i \in T$, but $\hat{c}_j < 1$ for all $j \in \mathsf{BR}_\theta(\mathbf{c})$ (in particular, $\hat{c}_{\hat{i}} < 1$). Thus, $i \notin \mathsf{BR}_\theta(\hat{\mathbf{c}})$, so $u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, i) < \max_{t \in T} u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, t) = u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, \hat{i})$. We have $\gamma_i = 0$ and $\alpha_i = 0$ by (7.15), which then implies $\beta + \delta_i = 0$ by (7.13); thus, $\beta = 0$. In addition, $\hat{c}_{\hat{i}} < 1$ implies $\delta_{\hat{i}} = 0$ by (7.15). Again, by (7.13), we end up with the contradiction that $w_{\hat{i}}^{\mathsf{D}} + w_{\hat{i}}^{\mathsf{A}} \cdot \sum_{i \neq \hat{i}} \alpha_i + \gamma_{\hat{i}} = 0$.

7.A.2 Proof of Lemma 7.5

Suppose $(\overline{\mathbf{c}}, i)$, $i \in \mathsf{BR}_\beta(\overline{\mathbf{c}})$, is not an SSE. Thus, there exists $\mathbf{z} \in \mathcal{C}$ and $t \in \mathsf{BR}_\beta(\mathbf{z})$, such that $u^{\mathsf{D}}(\mathbf{z}, t) > u^{\mathsf{D}}(\overline{\mathbf{c}}, i)$; equivalently,

$$u_\beta^{\mathsf{A}}(\mathbf{z}, t) < u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, i)$$

as β makes the game zero-sum. Since $i \in \mathsf{BR}_\beta(\overline{\mathbf{c}})$ and $t \in \mathsf{BR}_\beta(\mathbf{z})$ are the attacker's best responses, we have

$$u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, i) = \max_{j \in T} u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, j), \quad \text{and} \quad u_\beta^{\mathsf{A}}(\mathbf{z}, t) = \max_{j \in T} u_\beta^{\mathsf{A}}(\mathbf{z}, j).$$

Thus, $\max_{j \in T} u_\beta^{\mathsf{A}}(\mathbf{z}, j) < \max_{j \in T} u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, j)$. Now that $\overline{\mathbf{c}}$ is a maximin strategy, we have $\min_{j \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j) = \max_{\mathbf{c} \in \mathcal{C}} \min_{j \in T} u^{\mathsf{D}}(\mathbf{c}, j)$. This gives the following contradiction:

$$\begin{aligned} \max_{\mathbf{c} \in \mathcal{C}} \min_{j \in T} u^{\mathsf{D}}(\mathbf{c}, j) = \min_{j \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j) &= -\max_{j \in T} u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, j) \\ &< -\max_{j \in T} u_\beta^{\mathsf{A}}(\mathbf{z}, j) = \min_{j \in T} u^{\mathsf{D}}(\mathbf{z}, j). \end{aligned}$$

7.A.3 Proof of Lemma 7.6

By Lemma 7.5, $(\overline{\mathbf{c}}, i)$ forms an SSE for any $i \in \mathsf{BR}_\beta(\overline{\mathbf{c}})$ when the attacker has type $\beta = (-\mathbf{p}^{\mathsf{D}}, -\mathbf{r}^{\mathsf{D}})$. Since $\overline{\mathbf{c}}$ is fully mixed, $T = \{i \in T : \overline{c}_i > 0\}$, and by Lemma 7.4 (i), $T \subseteq \mathsf{BR}_\beta(\overline{\mathbf{c}}) \subseteq T$. Thus, $\mathsf{BR}_\beta(\overline{\mathbf{c}}) = T$, so $(\overline{\mathbf{c}}, i)$ forms an SSE for attacker type β (which makes the game zero-sum) with any $i \in T$; we have $u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, i) = \max_{j \in T} u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, j)$. It follows that

$$u^{\mathsf{D}}(\overline{\mathbf{c}}, i) = -u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, i) = -\max_{j \in T} u_\beta^{\mathsf{A}}(\overline{\mathbf{c}}, j) = \min_{j \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j).$$

To see that $\overline{\mathbf{c}}$ is the only maximin strategy. Suppose for the sake of contradiction that $\mathbf{z} = \arg\max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^{\mathsf{D}}(\mathbf{c}, i)$ is a maximin strategy and $\mathbf{z} \neq \overline{\mathbf{c}}$. Thus, either: (i) $z_i \geq \overline{c}_i$ for all $i \in T$, and this is strictly satisfied by some i; or (ii) $z_i < \overline{c}_i$ for some $i \in T$. We show either of them leads to a contradiction.

Since $\overline{\mathbf{c}}$ is a maximin strategy, it is also the defender's strategy in an SSE in a zero-sum game by Lemma 7.5; and by Lemma 7.4 (i), $\sum_{i \in T} \overline{c}_i = k$. Thus, in Case (i), it follows immediately that $\sum_{i \in T} z_i > \sum_{i \in T} \overline{c}_i = k$, which contradicts $\mathbf{z} \in \mathcal{C}$. In Case (ii), we have $u^{\mathsf{D}}(\mathbf{z}, i) < u^{\mathsf{D}}(\overline{\mathbf{c}}, i)$ by monotonicity of $u^{\mathsf{D}}(\cdot, i)$, which implies

$$\min_{j \in T} u^{\mathsf{D}}(\mathbf{z}, j) \leq u^{\mathsf{D}}(\mathbf{z}, i) < u^{\mathsf{D}}(\overline{\mathbf{c}}, i) = \min_{j \in T} u^{\mathsf{D}}(\overline{\mathbf{c}}, j),$$

where the last equality follows by Lemma 7.6. This contradicts the assumption that $\mathbf{z}$ is a maximin strategy.

7.A.4 Proof of Lemma 7.10

We first show the following useful result, which holds for any policy: the attacker will not lose any utility by reporting truthfully, no matter what outcome is prescribed for his true type.

Lemma 7.12. *Let $(\hat{\mathbf{c}}, \hat{i})$ be an SSE for attacker type $\theta \in \Theta$. For any policy π, truthful reporting in response to π guarantees a type-θ attacker his SSE utility, i.e., $u_\theta^{\mathsf{A}}(\pi(\theta)) \geq u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, \hat{i})$.*

Proof. Suppose for the sake of contradiction that $u_\theta^{\mathsf{A}}(\pi(\theta)) < u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, \hat{i})$. Let $\pi(\theta) = (\mathbf{z}, t)$. For π to be feasible, t must be a best response of a type-θ attacker against $\mathbf{z}$. Thus, $u_\theta^{\mathsf{A}}(\mathbf{z}, t) \geq u_\theta^{\mathsf{A}}(\mathbf{z}, i)$ for all $i \in T$; in particular, $u_\theta^{\mathsf{A}}(\mathbf{z}, t) \geq u_\theta^{\mathsf{A}}(\mathbf{z}, \hat{i})$. We have

$$u_\theta^{\mathsf{A}}(\hat{\mathbf{c}}, \hat{i}) > u_\theta^{\mathsf{A}}(\pi(\theta)) = u_\theta^{\mathsf{A}}(\mathbf{z}, t) \geq u_\theta^{\mathsf{A}}(\mathbf{z}, \hat{i}).$$

Since $u_\theta^{\mathsf{A}}(\mathbf{c}, \hat{i})$ changes continuously with respect to $c_{\hat{i}}$, the above inequality implies the existence of a number $\phi \in (\hat{c}_{\hat{i}}, z_{\hat{i}}]$, such that $u_\theta^{\mathsf{A}}(\phi, \hat{i}) = u_\theta^{\mathsf{A}}(\mathbf{z}, t)$.

Consider a strategy $\mathbf{z}'$ of the defender with $z'_{\hat{i}} = \phi$ and $z'_i = z_i$ for all $i \in T \setminus \{\hat{i}\}$. We have $0 \leq z_i \leq 1$ and $\sum_{i \in T} z'_i \leq \sum_{i \in T} z_i \leq k$, so $\mathbf{z}' \in \mathcal{C}$. In addition, $u_\theta^{\mathsf{A}}(\mathbf{z}, i) = u_\theta^{\mathsf{A}}(\mathbf{z}', i)$ for all $i \in T \setminus \{\hat{i}\}$. Thus,

$$u_\theta^{\mathsf{A}}(\mathbf{z}', \hat{i}) = u_\theta^{\mathsf{A}}(\phi, \hat{i}) = u_\theta^{\mathsf{A}}(\mathbf{z}, t) \geq u_\theta^{\mathsf{A}}(\mathbf{z}, i) = u_\theta^{\mathsf{A}}(\mathbf{z}', i),$$

which means $\hat{i}$ is a best response of a type-θ attacker against $\mathbf{z}'$, i.e, $\hat{i} \in \mathsf{BR}_\theta(\mathbf{z}')$. This gives rise to the following contradiction:

$$\max_{\mathbf{c} \in \mathcal{C}, i \in \mathsf{BR}_\theta(\mathbf{c})} u^{\mathsf{D}}(\mathbf{c}, i) \geq u^{\mathsf{D}}(\mathbf{z}', \hat{i}) = u^{\mathsf{D}}(\phi, \hat{i}) > u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i}) = \max_{\mathbf{c} \in \mathcal{C}, i \in \mathsf{BR}_\theta(\mathbf{c})} u^{\mathsf{D}}(\mathbf{c}, i),$$

where $u^{\mathsf{D}}(\phi, \hat{i}) > u^{\mathsf{D}}(\hat{\mathbf{c}}, \hat{i})$ since $\phi \in (\hat{c}_{\hat{i}}, z_{\hat{i}}]$. □

In what follows, we let $\pi^* = (\mathbf{c}^\theta, i^\theta)_{\theta \in \Theta}$. Let $\beta \in \Theta$ be the optimal report of a type-θ_ℓ attacker in response to π^*, i.e., $u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta)) \geq u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta'))$ for all $\beta' \in \Theta$. Let $(\hat{\mathbf{c}}^\theta, \hat{i}^\theta)$ be an SSE for each type $\theta \in \Theta$ as in Step 1 of Algorithm 7.1. Let $\mathbf{z}$, $\mathbf{h}$, and t be the variables defined in the ℓ-th iteration of Step 3. Hence, $\tilde{\pi}(\theta_\ell) = \pi(\theta_\ell) = (\mathbf{z}, t)$. We show two additional results, Lemmas 7.13 and 7.14.

Lemma 7.13. $h_{i^\beta} \leq c^\beta_{i^\beta}$.

Proof. Suppose for the sake of contradiction that $c^\beta_{i^\beta} < h_{i^\beta}$. Thus, $h_{i^\beta} > 0$ and according to the way $\mathbf{h}$ is set in Step 3, we have either $c^\beta_{i^\beta} < h_{i^\beta} = \frac{\xi \cdot \hat{u}(\theta_\ell) - p^{\mathsf{D}}_{i^\beta}}{r^{\mathsf{D}}_{i^\beta} - p^{\mathsf{D}}_{i^\beta}}$, or $c^\beta_{i^\beta} < h_{i^\beta} = \frac{u_\theta^{\mathsf{A}}(\pi(\theta)) - r^\theta_{i^\beta}}{p^\theta_{i^\beta} - r^\theta_{i^\beta}}$ for some $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$.

If it is the former case, we have

$$u^{\mathsf{D}}(\pi^*(\beta)) = u^{\mathsf{D}}(\mathbf{c}^\beta, i^\beta) < u^{\mathsf{D}}\left(\frac{\xi \cdot \hat{u}(\theta_\ell) - p^{\mathsf{D}}_{i^\beta}}{r^{\mathsf{D}}_{i^\beta} - p^{\mathsf{D}}_{i^\beta}}, i^\beta\right) = \xi \cdot \hat{u}(\theta_\ell).$$

Thus, $\mathrm{EoP}(\pi^*) \leq \mathrm{EoP}_{\theta_\ell}(\pi^*) = \frac{u^{\mathsf{D}}(\pi^*(\beta))}{\hat{u}(\theta_\ell)} < \xi$, which contradicts our assumption.

If it is the latter, we have

$$u_\theta^{\mathsf{A}}(\pi^*(\beta)) = u_\theta^{\mathsf{A}}(\mathbf{c}^\beta, i^\beta) > u_\theta^{\mathsf{A}}\left(\frac{u_\theta^{\mathsf{A}}(\pi(\theta)) - r^\theta_{i^\beta}}{p^\theta_{i^\beta} - r^\theta_{i^\beta}}, i^\beta\right) = u_\theta^{\mathsf{A}}(\pi(\theta)) = u_\theta^{\mathsf{A}}(\pi^*(\theta)),$$

so a type-θ attacker would be strictly better-off reporting type β in response to π^*, contradicting the assumption that π^* is $(\ell - 1)$-compatible. □

Lemma 7.14. *$h_t \le \hat{c}_t^{\theta_\ell}$ (hence, $z_t = \min\{\hat{c}_t^{\theta_\ell}, h_t\} = h_t$).*

Proof. Since β is the optimal report of a type-θ_ℓ attacker in response to π^*, we have $u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta)) \ge u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\theta_\ell))$. Moreover, by Lemma 7.12, $u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\theta_\ell)) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell})$. Thus,

$$u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta)) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell}) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, t), \tag{7.16}$$

where the last transition is due to the fact that $(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell})$ is an SSE, so $\hat{i}^{\theta_\ell} \in \mathsf{BR}_{\theta_\ell}(\hat{\mathbf{c}}^{\theta_\ell})$. It follows that

$$u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, t) \ge u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, i^\beta) \ge u_{\theta_\ell}^{\mathsf{A}}(\mathbf{c}^\beta, i^\beta) = u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta)) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, t), \tag{7.17}$$

where the first transition is due to $t \in \mathsf{BR}_{\theta_\ell}(\mathbf{h})$ by definition; and the second transition follows by Lemma 7.13 and monotonicity of the utility function. By monotonicity of $u_{\theta_\ell}^{\mathsf{A}}(\cdot, t)$, we have $h_t \le \hat{c}_t^{\theta_\ell}$. □

Now we prove the following results to complete the proof of Lemma 7.10.

(i) $(\mathbf{z}, t)$ is indeed feasible as an outcome prescribed for type θ_ℓ, i.e., $\mathbf{z} \in \mathcal{C}$ and $t \in \mathsf{BR}_{\theta_\ell}(\mathbf{z})$;

(ii) $\tilde{\pi}$ is ℓ-compatible;

(iii) $\mathrm{EoP}(\tilde{\pi}) \ge \xi$.

Part (i). Since $(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell})$ is an SSE for type θ_ℓ, by definition, $\hat{\mathbf{c}}^{\theta_\ell} \in \mathcal{C}$ and $\sum_{i \in T} \hat{c}_i^{\theta_\ell} \le k$. Since $z_i = \min\{\hat{c}_i^{\theta_\ell}, h_i\} \le \hat{c}_i^{\theta_\ell}$ for all $i \in T$, we have $0 \le z_i \le 1$ and $\sum_{i \in T} z_i \le \sum_{i \in T} \hat{c}_i^{\theta_\ell} \le k$. Thus, $\mathbf{z} \in \mathcal{C}$.

To see that $t \in \mathsf{BR}_{\theta_\ell}(\mathbf{z})$, suppose for the sake of contradiction that this does not hold, which means that $u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, i^*) > u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, t)$ for some $i^* \in T$. By Lemma 7.14, $z_t = h_t$, so we have

$$u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, i^*) > u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, t) = u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, t) \ge u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, i^*),$$

which implies that $z_{i^*} < h_{i^*}$ by monotonicity. Hence, $z_{i^*} = \min\{\hat{c}_{i^*}^{\theta_\ell}, h_{i^*}\} = \hat{c}_{i^*}^{\theta_\ell}$, so we have

$$u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell}) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, i^*) = u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, i^*) > u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, t) = u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, t).$$

Observe that from (7.16) and (7.17), we can also derive that $u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, t) \ge u_{\theta_\ell}^{\mathsf{A}}(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell})$, which contradicts the above inequality.

Part (ii). We show that it is optimal for every type-θ attacker, $\theta \in \{\theta_1, \dots, \theta_\ell\}$, to report truthfully in response to $\tilde{\pi}$.

First, consider a type-θ_ℓ attacker. We can establish the following inequalities for all $\beta' \in \Theta \backslash \{\theta_\ell\}$:

$$u_{\theta_\ell}^{\mathsf{A}}(\mathbf{z}, t) = u_{\theta_\ell}^{\mathsf{A}}(\mathbf{h}, t) \ge u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta)) \ge u_{\theta_\ell}^{\mathsf{A}}(\pi^*(\beta')),$$

where the first transition follows by Lemma 7.14; the second transition follows directly by (7.17); and the third transition is due to the assumption that β is the optimal reporting strategy of a type-θ_ℓ attacker.

By assumption, $\tilde{\pi}(\theta_\ell) = \pi(\theta_\ell) = (\mathbf{z}, t)$ and $\tilde{\pi}(\beta') = \pi^*(\beta') = \pi(\beta')$ for all $\beta' \in \Theta \setminus \{\theta_\ell\}$, so we have

$$u^{\mathsf{A}}_{\theta_\ell}(\tilde{\pi}(\theta_\ell)) = u^{\mathsf{A}}_{\theta_\ell}(\mathbf{z}, t) \geq u^{\mathsf{A}}_{\theta_\ell}(\pi(\beta')) = u^{\mathsf{A}}_{\theta_\ell}(\tilde{\pi}(\beta')),$$

which means that it is optimal for a type-θ_ℓ attacker to report truthfully in response to $\tilde{\pi}$.

Next, consider each type $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$. By Lemma 7.14 and the way h_t is set in Step 3 of the algorithm, we have

$$z_t = h_t \geq \frac{u^{\mathsf{A}}_\theta(\pi(\theta)) - r^\theta_t}{p^\theta_t - r^\theta_t}.$$

Thus, by monotonicity of the utility function, we have

$$u^{\mathsf{A}}_\theta(\pi(\theta)) = u^{\mathsf{A}}_\theta\left(\frac{u^{\mathsf{A}}_\theta(\pi(\theta)) - r^\theta_t}{p^\theta_t - r^\theta_t}, t\right) \geq u^{\mathsf{A}}_\theta(\mathbf{z}, t) = u^{\mathsf{A}}_\theta(\tilde{\pi}(\theta_\ell)).$$

Note that since $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$, we have $\tilde{\pi}(\theta) = \pi^*(\theta) = \pi(\theta)$. It then follows that

$$u^{\mathsf{A}}_\theta(\tilde{\pi}(\theta)) = u^{\mathsf{A}}_\theta(\pi(\theta)) \geq u^{\mathsf{A}}_\theta(\tilde{\pi}(\theta_\ell)).$$

Since π^* is $(\ell-1)$-compatible, $u^{\mathsf{A}}_\theta(\pi^*(\theta)) \geq u^{\mathsf{A}}_\theta(\pi^*(\beta'))$ for all $\beta' \in \Theta$, so for those $\beta' \neq \theta_\ell$ we have

$$u^{\mathsf{A}}_\theta(\tilde{\pi}(\theta)) = u^{\mathsf{A}}_\theta(\pi^*(\theta)) \geq u^{\mathsf{A}}_\theta(\pi^*(\beta')) = u^{\mathsf{A}}_\theta(\tilde{\pi}(\beta')).$$

Therefore, $u^{\mathsf{A}}_\theta(\tilde{\pi}(\theta)) \geq u^{\mathsf{A}}_\theta(\tilde{\pi}(\beta'))$ holds for all $\beta' \in \Theta$; it is optimal for a type-θ attacker to report truthfully in response to $\tilde{\pi}$.

Part (iii). We show that $\mathrm{EoP}_\theta(\tilde{\pi}) \geq \xi$ for every type $\theta \in \Theta$, which immediately implies $\mathrm{EoP}(\tilde{\pi}) = \min_{\theta\in\Theta} \mathrm{EoP}_\theta \geq \xi$ and complete the proof.

Since we have shown that $\tilde{\pi}$ is ℓ-compatible, truthful report is incentivized for every type in $\{\theta_1, \ldots, \theta_\ell\}$. Thus, for every $\theta \in \{\theta_1, \ldots, \theta_\ell\}$, we have $\mathrm{EoP}_\theta(\tilde{\pi}) \geq \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta))}{\hat{u}(\theta)}$. (The reason that we may have $\mathrm{EoP}_\theta(\tilde{\pi}) > \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta))}{\hat{u}(\theta)}$ is due to the optimistic tie-breaking assumption; see Definition 5.15.) For type θ_ℓ, since $z_t = h_t \geq \frac{\xi\cdot\hat{u}(\theta_\ell) - p^{\mathsf{D}}_t}{r^{\mathsf{D}}_t - p^{\mathsf{D}}_t}$ by Lemma 7.14, we have

$$\mathrm{EoP}_{\theta_\ell}(\tilde{\pi}) \geq \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta_\ell)} = \frac{u^{\mathsf{D}}(\tilde{\pi}(\mathbf{z}, t))}{\hat{u}(\theta_\ell)} \geq \frac{u^{\mathsf{D}}\left(\frac{\xi\cdot\hat{u}(\theta_\ell) - p^{\mathsf{D}}_t}{r^{\mathsf{D}}_t - p^{\mathsf{D}}_t}, t\right)}{\hat{u}(\theta_\ell)} = \frac{\xi\cdot\hat{u}(\theta_\ell)}{\hat{u}(\theta_\ell)} = \xi.$$

For each type $\theta \in \{\theta_1, \ldots, \theta_{\ell-1}\}$, we have

$$\mathrm{EoP}_\theta(\tilde{\pi}) \geq \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta))}{\hat{u}(\theta)} = \frac{u^{\mathsf{D}}(\pi^*(\theta))}{\hat{u}(\theta)} = \mathrm{EoP}_\theta(\pi^*) \geq \mathrm{EoP}(\pi^*) \geq \xi.$$

For the other types $\theta' \in \{\theta_{\ell+1}, \ldots, \theta_\lambda\}$, if their optimal reporting remains the same as that in response to π^*, then according to the same argument above, $\mathrm{EoP}_{\theta'}(\tilde{\pi}) \geq \xi$. Otherwise, since $\tilde{\pi}$ and π^* differs only in the outcomes prescribed for type θ_ℓ, if the attacker's optimal reporting strategy changes under $\tilde{\pi}$, it will only change to θ_ℓ, in which case we have

$$\mathrm{EoP}_{\theta'}(\tilde{\pi}) = \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta')} \geq \frac{u^{\mathsf{D}}(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta_\ell)} = \mathrm{EoP}_{\theta_\ell}(\tilde{\pi}) \geq \xi,$$

where we use $\hat{u}(\theta') \leq \hat{u}(\theta_\ell)$, which is due to the way the types are ordered.

Part III

Conclusion

Chapter 8

Conclusion and Future Directions

To conclude, I discuss several potential directions for future work in this chapter.

Towards Future Defense Coordination

In Chapters 3 and 4, we studied interactions in a Stackelberg game with multiple defenders and attempted to coordinate the defenders through mechanism design. While the task seems to be promising given the common enemy the defenders face, our results were mostly negative: in general, it is not possible to find a coordination mechanism that satisfies the basic properties we desire. The impossibility results indicate the necessity of other innovative ways to promote coordination, and our work can be seen as a starting point of this research. In scenarios where strategyproofness is not necessary, one open question as outlined in Section 4.5 is how to compute a correlated equilibrium with good social welfare properties. More realistically perhaps, while agents may have incentives to lie about their preferences, they may not be able to lie arbitrarily, whereby a partial notion of strategyproofness may be relevant. The scenario with only two defenders, which we have not investigated, may well be of great practical significance and worth looking into in future work.

It would be interesting to also view the problem beyond the perspective of mechanism design. Indeed, coordination is widely seen in the real world and it can often be achieved through a richer set of means such as negotiation, contracting, or even by establishing social norms. In a broader context, in the era of an increasing number of global issues in health, security, sustainability, and climate change, it is vital that cutting-edge techniques in AI and multi-agent systems are also applied to solving socio-economic problems that can benefit the most from coordination but also involve conflicts of all kinds that need to be properly accounted for.

Deception in Stackelberg Games and Beyond

In Chapters 5–7, we studied imitative follower deception in Stackelberg games. In Chapter 5 we did not settle the complexity of the extended IC variants Opt-Policy-WC and Opt-Policy-EoP (see Section 5.7). In addition, while all our hardness results hold when the follower has only three actions, we do not know whether they remain when the number of the follower's actions is reduced to two. Another open problem is how to design countermeasures when the follower is able to report a type in an infinite space (or even a finite space that is exponentially large), in which case the leader's policy cannot be represented in a concise way to allow for a computationally feasible

formulation. Indeed, in Chapter 6, we have seen that it is easy for the follower to deal with an infinite space (when no other constraints are imposed); to design a countermeasure for the leader amounts to the challenging task of finding an optimal function $f : \mathbb{R}^{m \times n} \to \Delta^{m-1}$. The easiness results for the follower make this task imperative. The QR-based heuristic approach we presented in Chapter 7 (see Section 7.6) is our first step towards addressing this task. Nevertheless, to find an exact solution and to fully understand the underlying theories remains very challenging.

While our model was built on a one-shot game, follower deception can happen in repeated games as well. In particular in finitely repeated games, the follower will likely need to take into account the cost incurred during the stage when the leader learns to play against the follower. Unlike our one-shot game model, a repeated game may not converge to any equilibrium, which also makes the structure of the problem very different. In another natural setting, the follower may not need to keep imitating the false type after the leader learns it. Indeed, the follower may not be responding optimally according to their true payoffs in the final round of interaction in our framework, so there is an incentive for them to switch back to playing their actual best response. This can be seen as another generalization of the standard Bayesian Stackelberg game model, with a signaling process where the leader attempts to elicit information from the follower and the follower sends a signal in response in a strategic way. Our preliminary thought is that the same policy-based approach is also able to guarantee the leader as least her utility in the Bayesian Stackelberg equilibrium (similar to the reasoning of Proposition 5.3). It would be interesting to understand the associated computational problems in this model.

Another extension of our work that would perhaps be of wider interest is how to leverage information asymmetry in a game with simultaneous movement. In other words, how can a player misreport their payoffs to lead the game to a different and more beneficial Nash equilibrium? Some initial thought indicates that the equilibrium selection issue plays a major role in this problem. Indeed, suppose the player who misreports payoffs is to get the *best* Nash equilibrium in the distorted game, then in the case of a two-player game, it is not hard to see that a trivial solution is to misreport that this player is indifferent among all strategies: the equilibrium selection assumption will then 'help' the player to obtain an outcome in which this player plays the optimal commitment just as in a Stackelberg equilibrium. Thus, it is more realistic, yet more challenging, to make the worst equilibrium selection assumption in future work.

In a broader sense, manipulation can happen against any learning algorithms that are designed for decision-making in strategic environments, such as no-regret algorithm, or reinforcement learning. For the former, there is some recent work on strategizing against a no-regret learner [Deng et al., 2019, Feng et al., 2020], and the latter appears to relate to inverse RL [Abbeel and Ng, 2017], where the goal is to extract a reward function from observed optimal behavior. It would be interesting to look deeper in these directions.

Bibliography

Pieter Abbeel and Andrew Y. Ng. Inverse reinforcement learning. In Claude Sammut and Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 678–682. Springer US, Boston, MA, 2017. 146

Stéphane Airiau, Haris Aziz, Ioannis Caragiannis, Justin Kruger, Jérôme Lang, and Dominik Peters. Portioning using ordinal preferences: Fairness and efficiency. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19)*, pages 11–17, 2019. 59

Bo An, Eric Shieh, Milind Tambe, Rong Yang, Craig Baldwin, Joseph DiRenzo, Ben Maule, and Garrett Meyer. PROTECT–a deployed game theoretic system for strategic security allocation for the United States Coast Guard. *AI Magazine*, 33(4):96–96, 2012. 13

Bo An, Fernando Ordóñez, Milind Tambe, Eric Shieh, Rong Yang, Craig Baldwin, Joseph DiRenzo III, Kathryn Moretti, Ben Maule, and Garrett Meyer. A deployed quantal response-based patrol planning system for the US Coast Guard. *Interfaces*, 43(5):400–420, 2013. 13

Haris Aziz, Anna Bogomolnaia, and Hervé Moulin. Fair mixing: the case of dichotomous preferences. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 753–781. ACM, 2019. 59

Yakov Babichenko. Query complexity of approximate Nash equilibria. *Journal of the ACM*, 63(4): 36:1–36:24, 2016. 83

Yakov Babichenko and Aviad Rubinstein. Communication complexity of approximate Nash equilibria. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC'17)*, pages 878–889, 2017. 83

Yang Bai, Lixing Chen, Linqi Song, and Jie Xu. Bayesian Stackelberg game for risk-aware edge computation offloading. In *Proceedings of the 6th ACM Workshop on Moving Target Defense*, pages 25–35, 2019. 13

Maria-Florina Balcan, Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Commitment without regrets: Online learning in Stackelberg security games. In *Proceedings of the 16th ACM conference on Economics and Computation (EC'15)*, pages 61–78, 2015. 124

John Bartholdi, Craig A. Tovey, and Michael A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989a. 59

John J. Bartholdi, Craig A. Tovey, and Michael A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989b. 59

Nicola Basilico, Andrea Celli, Giuseppe De Nittis, and Nicola Gatti. Computing the team-maxmin equilibrium in single-team single-adversary team games. *Intelligenza Artificiale*, 11(1):67–79, 2017. 18, 19

Omer Ben-Porat and Moshe Tennenholtz. Regression equilibrium. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, pages 173–191, 2019. 82

Dimitris Bertsimas and John N. Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997. 7

Georgios Birmpas, Jiarui Gan, Alexandros Hollender, Francisco J. Marmolejo-Cossío, Ninad Rajgopal, and Alexandros A. Voudouris. Optimally deceiving a learning leader in Stackelberg games. In *Advances in Neural Information Processing Systems*, volume 33, pages 20624–20635, 2020. 4

Avrim Blum, Jeffrey C. Jackson, Tuomas Sandholm, and Martin Zinkevich. Preference elicitation and query learning. *Journal of Machine Learning Research*, 5:649–667, 2004. 83

Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Learning optimal commitment to overcome insecurity. In *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14)*, pages 1826–1834, 2014. 82, 83, 85, 123

Liad Blumrosen and Noam Nisan. Combinatorial auctions. In *Algorithmic Game Theory*, chapter 11, pages 267–299. Cambridge University Press, 2007. 83

Anna Bogomolnaia, Hervé Moulin, and Richard Stong. Collective choice under dichotomous preferences. *Journal of Economic Theory*, 122(2):165–184, 2005. 59

Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2014. 117

Florian Brandl, Felix Brandt, Dominik Peters, Christian Stricker, and Warut Suksompong. Funding public projects: A case for the Nash product rule. Working paper, 2020. 59

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors. *Handbook of computational social choice*. Cambridge University Press, 2016. 59

George W. Brown. Some notes on computation of game solutions. *RAND corporation report*, page 78, 1949. 83

Gerald Brown, Matthew Carlyle, Douglas Diehl, Jeffrey Kline, and Kevin Wood. A two-sided optimization for theater ballistic missile defense. *Operations Research*, 53(5):745–763, September 2005. ISSN 0030-364X. 124

RenRené Caldentey and Martin Haugh. A Cournot-Stackelberg model of supply contracts with financial hedging and identical retailers. *Foundations and Trends® in Technology, Information and Operations Management*, 11(1-2):124–143, 2017. ISSN 1571-9545. 13

Colin F. Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2011. 13

Tim Carnes, Chandrashekhar Nagarajan, Stefan M. Wild, and Anke van Zuylen. Maximizing influence in a competitive social network: A follower's perspective. In *Proceedings of the 9th International Conference on Electronic Commerce (ICEC'07)*, ICEC '07, page 351–360, 2007. 1, 14

Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Be a leader or become a follower: The strategy to commit to with multiple leaders. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19)*, pages 123–129, 2019. 19

Hau Chan, Michael Ceyko, and Luis E. Ortiz. Interdependent defense games: modeling interdependent security under deliberate attacks. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI'12)*, pages 152–162, 2012. 19

Xi Chen, Yu Cheng, and Bo Tang. Well-supported versus approximate Nash equilibria: Query complexity of large games. In *Proceedings of the 8th ACM Conference on Innovations in Theoretical Computer Science (ITCS'17)*, pages 57:1–57:9, 2017. 83

Yiling Chen, Chara Podimata, Ariel D. Procaccia, and Nisarg Shah. Strategyproof linear regression in high dimensions. In *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)*, pages 9–26, 2018. 82

Yiling Chen, Yang Liu, and Chara Podimata. Grinding the space: Learning to classify against strategic agents. *CoRR*, abs/1911.04004, 2019. 83

Wolfram Conen and Tuomas Sandholm. Preference elicitation in combinatorial auctions. In *Proceedings of the 3rd ACM conference on Electronic Commerce (EC'01)*, pages 256–259, 2001. 83

Stefano Coniglio, Nicola Gatti, and Alberto Marchesi. Pessimistic leader-follower equilibria with multiple followers. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17)*, pages 171–177, 2017. 20

Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC'06)*, pages 82–90, 2006. 6, 7, 19, 82

Vincent Conitzer, Tuomas Sandholm, and Jérôme Lang. When are elections with few candidates hard to manipulate? *Journal of the ACM*, 54(3):14–es, June 2007. ISSN 0004-5411. 59

Miguel Costa-Gomes, Vincent P. Crawford, and Bruno Broseta. Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5):1193–1235, 2001. 13

Gabriele E. Danninger-Uchida. Carathéodory theorem. In Christodoulos A. Floudas and Panos M. Pardalos, editors, *Encyclopedia of Optimization*, pages 236–237. Springer US, 2001. 12

Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009. 6, 18, 38

Ofer Dekel, Felix A. Fischer, and Ariel D. Procaccia. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, 2010. 82

Francesco Maria Delle Fave, Albert Xin Jiang, Zhengyu Yin, Chao Zhang, Milind Tambe, Sarit Kraus, and John P. Sullivan. Game-theoretic patrolling with dynamic execution uncertainty and a case study on a real transit system. *Journal of Artificial Intelligence Research*, 50:321–367, 2014. 19

Victor DeMiguel and Huifu Xu. A stochastic multiple-leader Stackelberg model: analysis, computation, and application. *Operations Research*, 57(5):1220–1235, 2009. 19

Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. In *Advances in Neural Information Processing Systems (NeurIPS'19)*, pages 1577–1585, 2019. 146

Jinshuo Dong, Aaron Roth, Zachary Schutzman, Bo Waggoner, and Zhiwei Steven Wu. Strategic classification from revealed preferences. In *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)*, pages 55–70, 2018. 83

Conal Duddy. Fair sharing under dichotomous preferences. *Mathematical Social Sciences*, 73:1–5, 2015. 59

Shaddin Dughmi, David Kempe, and Ruixin Qiang. Persuasion with limited communication. In *Proceedings of the ACM Conference on Economics and Computation (EC'16)*, pages 663–680, 2016. 124

Edith Elkind, Jiarui Gan, Svetlana Obraztsova, Zinovi Rabinovich, and Alexandros A. Voudouris. Protecting elections by recounting ballots. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 259–265, 2019. 14

Edith Elkind, Jiarui Gan, Svetlana Obraztsova, Zinovi Rabinovich, and Alexandros A. Voudouris. Protecting elections by recounting ballots. *Artificial Intelligence*, 290:103401, 2021. 14

Fei Fang and Thanh H Nguyen. Green security games: Apply game theory to addressing green security challenges. *ACM SIGecom Exchanges*, 15(1):78–83, 2016. 11

John Fearnley, Martin Gairing, Paul W. Goldberg, and Rahul Savani. Learning equilibria of games via payoff queries. *Journal of Machine Learning Research*, 16:1305–1344, 2015. 83

Zhe Feng, David C. Parkes, and Haifeng Xu. The intrinsic robustness of stochastic bandits to strategic manipulation. In *Proceedings of the 37th International Conference on Machine Learning (ICML'20)*, volume 119, pages 3092–3101, 2020. 146

Jiarui Gan and Bo An. Game-theoretic considerations for optimizing taxi system efficiency. *IEEE Intelligent Systems*, 32(3):46–52, 2017. 14

Jiarui Gan, Bo An, Haizhong Wang, Xiaoming Sun, and Zhongzhi Shi. Optimal pricing for improving efficiency of taxi systems. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 2811–2818, 2013. 1, 14

Jiarui Gan, Bo An, and Chunyan Miao. Optimizing efficiency of taxi systems: Scaling-up and handling arbitrary constraints. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15)*, page 523–531, 2015a. 14

Jiarui Gan, Bo An, and Yevgeniy Vorobeychik. Security games with protection externalities. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15)*, page 914–920, 2015b. 12

Jiarui Gan, Bo An, Yevgeniy Vorobeychik, and Brian Gauch. Security games on a plane. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI'17)*, page 530–536, 2017. 12

Jiarui Gan, Edith Elkind, and Michael Wooldridge. Stackelberg security games with multiple uncoordinated defenders. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'18)*, page 703–711, 2018. 4

Jiarui Gan, Qingyu Guo, Long Tran-Thanh, Bo An, and Michael Wooldridge. Manipulating a learning defender and ways to counteract. In *Advances in Neural Information Processing Systems (NeurIPS'19)*, pages 8274–8283, 2019a. 4

Jiarui Gan, Haifeng Xu, Qingyu Guo, Long Tran-Thanh, Zinovi Rabinovich, and Michael Wooldridge. Imitative follower deception in Stackelberg games. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC'19)*, page 639–657, 2019b. 4

Jiarui Gan, Edith Elkind, Sarit Kraus, and Michael Wooldridge. Mechanism design for defense coordination in security games. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'20)*, page 402–410, 2020. 4

Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973. 59, 62

Paul W. Goldberg and Francisco J. Marmolejo-Cossío. Learning convex partitions and computing game-theoretic equilibria from best response queries. In *International Conference on Web and Internet Economics (WINE)*, pages 168–187, 2018. 83, 109

Paul W. Goldberg and Aaron Roth. Bounds for the query complexity of approximate equilibria. *ACM Transactions on Economics and Computation*, 4(4):24:1–24:25, 2016. 83

Paul W. Goldberg and Stefano Turchetta. Query complexity of approximate equilibria in anonymous games. *Journal of Computer and System Sciences*, 90:80–98, 2017. 83

Paul W. Goldberg, Francisco J. Marmolejo-Cossío, and Zhiwei Steven Wu. Logarithmic query complexity for approximate Nash computation in large games. *Theory of Computing Systems*, 63 (1):26–53, 2019. 83

Paul W. Goldberg, Edwin Lock, and Francisco Marmolejo-Cossío. Learning strong substitutes demand via queries, 2020. arXiv:2005.01496. 83

Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981. 115

Xueqiang Gu, Chengyi Zeng, and Fengtao Xiang. Applying a Bayesian Stackelberg game to secure infrastructure system: From a complex network perspective. In *Proceedings of the 2019 4th International Conference on Automation, Control and Robotics Engineering (CACRE'19)*. Association for Computing Machinery, 2019. 13

Qingyu Guo, Bo An, Yevgeniy Vorobeychik, Long Tran-Thanh, Jiarui Gan, and Chunyan Miao. Coalitional security games. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, pages 159–167, 2016a. 12, 19

Qingyu Guo, Bo An, Yair Zick, and Chunyan Miao. Optimal interdiction of illegal network flow. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, pages 2507–2513, 2016b. 12

Qingyu Guo, Bo An, Branislav Bosansky, and Christopher Kiekintveld. Comparing strategic secrecy and Stackelberg commitment in security games. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17)*, pages 3691–3699, 2017. 124

Qingyu Guo, Jiarui Gan, Fei Fang, Long Tran-Thanh, Milind Tambe, and Bo An. On the inducibility of Stackelberg equilibrium for security games. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19)*, volume 33, pages 2020–2028, 2019. 22

Nika Haghtalab, Fei Fang, Thanh Hong Nguyen, Arunesh Sinha, Ariel D. Procaccia, and Milind Tambe. Three strategies to success: Learning adversary models in security games. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, pages 362–372, 2016. 82, 83, 85, 123

John C. Harsanyi and Reinhard Selten. A generalized Nash solution for two-person bargaining games with incomplete information. *Management Science*, 18(5-Part-2):80–106, 1972. 7

Sergiu Hart and Yishay Mansour. How long to equilibrium? the communication complexity of uncoupled equilibrium procedures. *Games and Economic Behavior*, 69(1):107–126, 2010. 83

Sergiu Hart and Noam Nisan. The query complexity of correlated equilibria. *Games and Economic Behavior*, pages 401–410, 2016. 83

Dorit S. Hochbaum, Cheng Lyu, and Fernando Ordóñez. Security routing games with multivehicle Chinese postman problem. *Networks*, 64(3):181–191, 2014. 13

Safwan Hossain and Nisarg Shah. The effect of strategic noise in linear regression. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'20)*, pages 511–519, 2020. 82

Manish Jain, James Pita, Milind Tambe, Fernando Ordóñez, Praveen Paruchuri, and Sarit Kraus. Bayesian Stackelberg games and their application for security at Los Angeles International Airport. *SIGecom Exchanges*, 7(2):10:1–10:3, June 2008. 12, 82

Manish Jain, Erim Kardes, Christopher Kiekintveld, Fernando Ordónez, and Milind Tambe. Security games with arbitrary schedules: A branch and price approach. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*, pages 792–797, 2010. 12

Manish Jain, Dmytro Korzhyk, Ondřej Vaněk, Vincent Conitzer, Michal Pěchouček, and Milind Tambe. A double oracle algorithm for zero-sum security games on graphs. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'11)*, pages 327–334, 2011. 12

Manish Jain, Vincent Conitzer, and Milind Tambe. Security scheduling for real-world networks. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'13)*, pages 215–222, 2013. 12

Albert Xin Jiang, Ariel D. Procaccia, Yundi Qian, Nisarg Shah, and Milind Tambe. Defender (mis) coordination in security games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 220–226, 2013. 18, 19

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011. 83

Rohit Khandekar, Kirsten Hildrum, Deepak Rajan, and Joel Wolf. Scheduling with Setup Costs and Monotone Penalties. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'12)*, volume 18 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 185–198, 2012. 75

Christopher Kiekintveld, Towhidul Islam, and Vladik Kreinovich. Security games with interval uncertainty. In *Proceedings of The 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'13)*, pages 231–238, 2013. 81

Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. Complexity of computing optimal Stackelberg strategies in security resource allocation games. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*, page 805–810, 2010. 12

Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. *Computer Science Review*, 3 (2):65–69, May 2009. ISSN 1574-0137. 58

Ankur A. Kulkarni and Uday V. Shanbhag. A shared-constraint approach to multi-leader multi-follower games. *Set-Valued and Variational Analysis*, 22(4):691–720, 2014. 19

Sebastien M. Lahaie and David C. Parkes. Applying learning algorithms to preference elicitation. In *Proceedings of the 5th ACM conference on Electronic Commerce (EC'04)*, pages 180–188, 2004. 83

Aron Laszka, Jian Lou, and Yevgeniy Vorobeychik. Multi-defender strategic filtering against spear-phishing attacks. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*, pages 537–543, 2016. 19

Joshua Letchford, Vincent Conitzer, and Kamesh Munagala. Learning and approximating the optimal strategy to commit to. In *International Symposium on Algorithmic Game Theory (SAGT'09)*, pages 250–262, 2009. 81, 82, 83, 85

Sven Leyffer and Todd Munson. Solving multi-leader–common-follower games. *Optimisation Methods & Software*, 25(4):601–623, 2010. 19

Yuqian Li and Vincent Conitzer. Game-theoretic question selection for tests. *Journal of Artificial Intelligence Research*, 59:437–462, 2017. 1, 14

Jian Lou and Yevgeniy Vorobeychik. Equilibrium analysis of multi-defender security games. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, pages 596–602, 2015. 18, 19

Jian Lou and Yevgeniy Vorobeychik. Decentralization and security in dynamic traffic light control. In *Proceedings of the Symposium and Bootcamp on the Science of Security*, pages 90–92, 2016. 19

Jian Lou, Andrew M. Smith, and Yevgeniy Vorobeychik. Multidefender security games. *IEEE Intelligent Systems*, 32(1):50–60, 2017. 18, 19

Francisco J. Marmolejo-Cossío, Eric Brigham, Benjamin Sela, and Jonathan Katz. Competing (semi-) selfish miners in Bitcoin. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, pages 89–109, 2019. 14

Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6 – 38, 1995. 13, 134

Reshef Meir, Ariel D. Procaccia, and Jeffrey S. Rosenschein. Algorithms for strategyproof classification. *Artificial Intelligence*, 186:123–156, 2012. 83

Eiichi Miyagawa. Locating libraries on a street. *Social Choice and Welfare*, 18(3):527–541, 2001. 59

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. 59

Thanh Hong Nguyen and Haifeng Xu. Imitative attacker deception in Stackelberg security games. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19)*, pages 528–534, 2019. 124

Thanh Hong Nguyen, Amulya Yadav, Bo An, Milind Tambe, and Craig Boutilier. Regret-based optimization and preference elicitation for Stackelberg security games with uncertainty. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14)*, pages 756–762, 2014. 82

Thanh Hong Nguyen, Yongzhao Wang, Arunesh Sinha, and Michael P. Wellman. Deception in finitely repeated security games. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19)*, pages 2133–2140, 2019. 124

Noam Nisan and Ilya Segal. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory*, 129(1):192–224, 2006. 83

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007. 59

Eugene Nudelman, Jennifer Wortman, Yoav Shoham, and Kevin Leyton-Brown. Run the gamut: A comprehensive approach to evaluating game-theoretic algorithms. In *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, pages 880–887, 2004. 97, 137

Gabriele Oliva, Roberto Setola, and Marco Tesei. A Stackelberg game-theoretical approach to maritime counter-piracy. *IEEE Systems Journal*, 13(1):982–993, 2018. 13

Frank H. Page. Optimal contract mechanisms for principal-agent problems with moral hazard and adverse selection. *Economic Theory*, 1(4):323–338, 1991. 1, 13

Christos H. Papadimitriou and Tim Roughgarden. Computing correlated equilibria in multi-player games. *Journal of the ACM (JACM)*, 55(3):14, 2008. 77

Praveen Paruchuri, Jonathan P. Pearce, Milind Tambe, Fernando Ordonez, and Sarit Kraus. An efficient heuristic approach for security against multiple adversaries. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'07)*, pages 311–318, 2007. 12, 19

Praveen Paruchuri, Jonathan P Pearce, Janusz Marecki, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. Efficient algorithms to solve Bayesian Stackelberg games for security applications. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI'08)*, pages 1559–1562, 2008a. 12, 19

Praveen Paruchuri, Jonathan P. Pearce, Janusz Marecki, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. Playing games for security: An efficient exact algorithm for solving Bayesian Stackelberg games. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'08)*, pages 895–902, 2008b. 19, 82

Binghui Peng, Weiran Shen, Pingzhong Tang, and Song Zuo. Learning optimal strategies to commit to. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19)*, pages 2149–2156, 2019. 82, 83, 123

Javier Perote and Juan Perote-Peña. Strategy-proof estimators for simple regression. *Mathematical Social Sciences*, 47(2):153–176, 2004. 82

James Pita, Manish Jain, Fernando Ordónez, Christopher Portway, Milind Tambe, Craig Western, Praveen Paruchuri, and Sarit Kraus. ARMOR security for Los Angeles International Airport. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI'08)*, pages 1884–1885, 2008. 12

James Pita, Manish Jain, Fernando Ordóñez, Christopher Portway, Milind Tambe, Craig Western, Praveen Paruchuri, and Sarit Kraus. Using game theory for Los Angeles Airport security. *AI Magazine*, 30(1):43, Jan. 2009. 12

James Pita, Manish Jain, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence*, 174(15):1142–1171, 2010. 13, 19, 99

Robert Powell. Allocating defensive resources with private information about vulnerability. *American Political Science Review*, 101(04):799–809, 2007. 124

Ariel D. Procaccia. Cake cutting: not just child's play. *Communications of the ACM*, 56(7):78–87, 2013. 59

Ariel D. Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4), December 2013. 59

Zinovi Rabinovich, Albert Xin Jiang, Manish Jain, and Haifeng Xu. Information disclosure as a means to security. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15)*, pages 645–653, 2015. 124

Julia Robinson. An iterative method of solving a game. *The Annals of Mathematics*, 54(2):296–301, 1951. 83

Aaron Roth, Jonathan Ullman, and Zhiwei Steven Wu. Watch and learn: Optimizing from revealed preferences feedback. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC'16)*, pages 949–962, 2016. 82, 83, 85

Alvin E. Roth. *Who gets what—and why: The new economics of matchmaking and market design*. Houghton Mifflin Harcourt, 2015. 59

Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975. 59, 62

Aaron Schlenker, Omkar Thakoor, Haifeng Xu, Fei Fang, Milind Tambe, Long Tran-Thanh, Phebe Vayanos, and Yevgeniy Vorobeychik. Deceiving cyber adversaries: A game theoretic approach. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'18)*, pages 892–900, 2018. 124

Hanif D. Sherali. A multiple leader Stackelberg model and analysis. *Operation Research*, 32(2): 390–404, 1984. 19

Eric Shieh, Bo An, Rong Yang, Milind Tambe, Craig Baldwin, Joseph DiRenzo, Ben Maule, Garrett Meyer, and Kathryn Moretti. Protect in the ports of Boston, New York and beyond: experiences in deploying stackelberg security games with quantal response. In *Handbook of computational approaches to counterterrorism*, pages 441–463. Springer, 2013. 13

Ankur Sinha, Pekka Malo, Anton Frantsev, and Kalyanmoy Deb. Finding optimal strategies in a multi-period multi-leader-follower Stackelberg game using an evolutionary algorithm. *Computers & Operation Research*, 41:374–385, 2014. 19

Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, and Milind Tambe. Stackelberg security games: looking beyond a decade of success. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18)*, pages 5494–5501, 7 2018. 11

Andrew Smith, Yevgeniy Vorobeychik, and Joshua Letchford. Multi-defender security games on networks. *ACM SIGMETRICS Performance Evaluation Review*, 41(4):4–7, 2014. 18, 19

Jingchang Sun, Pingzhong Tang, and Yulong Zeng. Games of miners. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'20)*, pages 1323–1331, 2020. 14

Milind Tambe. *Security and Game theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, 2011. 1, 11

Kübra Tanınmış, Necati Aras, and I.K. Altınel. Influence maximization with deactivation in social networks. *European Journal of Operational Research*, 278(1):105–119, 2019. 14

Jason Tsai, Shyamsunder Rathi, Christopher Kiekintveld, Fernando Ordóñez, and Milind Tambe. IRIS-a tool for strategic security allocation in transportation networks. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'09)*, pages 37–44, 2009. 11

Jason Tsai, Zhengyu Yin, Jun-young Kwak, David Kempe, Christopher Kiekintveld, and Milind Tambe. Urban security: game-theoretic resource allocation in networked physical domains. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*, pages 881–886, 2010. 8

Jason Tsai, Yundi Qian, Yevgeniy Vorobeychik, Christopher Kiekintveld, and Milind Tambe. Bayesian security games for controlling contagion. In *Proceedings of the ASE/IEEE International Conference on Social Computing(SocialCom)*, pages 33–38. IEEE, 2013. 13, 14, 19

Ngo Van Long and Gerhard Sorger. A dynamic principal-agent problem as a feedback Stackelberg differential game. *Central European Journal of Operations Research*, 18(4):491–509, 2010. 13

Heinrich von Stackelberg. *Marktform und Gleichgewicht.* J. Springer, 1934. 1

Heinrich von Stackelberg. *Market structure and equilibrium.* Springer Science & Business Media, 2010. 1

Bernhard von Stengel and Shmuel Zamir. Leadership with commitment to mixed strategies. *Technical Report LSE-CDAM-2004-01, CDAM Research Report*, 38, 2004. 6, 22

Bo Waggoner, Rafael Frongillo, and Jacob D. Abernethy. A market framework for eliciting private data. In *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS'15)*, pages 3510–3518, 2015. 83

Xinrun Wang, Bo An, Martin Strobel, and Fookwai Kong. Catching Captain Jack: Efficient time and space dependent patrols to combat oil-siphoning in international waters. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI'18)*, pages 208–215, 2018. 18

Zhen Wang, Yue Yin, and Bo An. Computing optimal monitoring strategy for detecting terrorist plots. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*, pages 637–643, 2016. 12, 20

James R. Wright and Kevin Leyton-Brown. Beyond equilibrium: predicting human behaviour in normal form games. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*, pages 901–907, 2010. 13

Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Yeng Chai Soh. Optimal pricing for efficient electric vehicle charging station management. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, page 749–757, 2016. 1, 14

Yanhai Xiong, Jiarui Gan, Bo An, Chunyan Miao, and Ana L. C. Bazzan. Optimal electric vehicle fast charging station placement based on game theoretical framework. *IEEE Transactions on Intelligent Transportation Systems*, 19(8):2493–2504, 2017. 14

Haifeng Xu, Zinovi Rabinovich, Shaddin Dughmi, and Milind Tambe. Exploring information asymmetry in two-stage security games. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15)*, pages 1057–1063, 2015. 124

Haifeng Xu, Rupert Freeman, Vincent Conitzer, Shaddin Dughmi, and Milind Tambe. Signaling in Bayesian Stackelberg games. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, pages 150–158, 2016a. 124

Haifeng Xu, Long Tran-Thanh, and Nicholas R. Jennings. Playing repeated security games with no prior knowledge. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, pages 104–112, 2016b. 124

Rong Yang, Christopher Kiekintveld, Fernando Ordóñez, Milind Tambe, and Richard John. Improving resource allocation strategy against human adversaries in security games. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*, 2011. 13

Rong Yang, Fernando Ordonez, and Milind Tambe. Computing optimal strategy against quantal response in security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*, pages 847–854, 2012. 13

Rong Yang, Albert Xin Jiang, Milind Tambe, and Fernando Ordóñez. Scaling-up security games with boundedly rational adversaries: A cutting-plane approach. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 404–410, 2013. 13

Yue Yin and Bo An. Efficient resource allocation for protecting coral reef ecosystems. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, pages 531–537, 2016. 12

Yue Yin, Bo An, Yevgeniy Vorobeychik, and Jun Zhuang. Optimal deceptive strategies in security games: A preliminary study, 2013. AAAI 2013 Spring Symposia. 124

Hanrui Zhang, Yu Cheng, and Vincent Conitzer. When samples are strategically selected. In *Proceedings of the 36th International Conference on Machine Learning (ICML'19)*, volume 97, 2019. 83

Jun Zhuang and Vicki M. Bier. Reasons for secrecy and deception in Homeland-Security resource allocation. *Risk Analysis*, 30(12):1737–1743, 2010. 124

Martin A. Zinkevich, Avrim Blum, and Tuomas Sandholm. On polynomial-time preference elicitation with value queries. In *Proceedings of the 4th ACM Conference on Electronic Commerce (EC'03)*, pages 176–185, 2003. 83

David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'06)*, pages 681–690, 2006. 89