# Monotonic Maximin: A Robust Stackelberg Solution Against Boundedly Rational Followers

Albert Xin Jiang[1], Thanh H. Nguyen[1], Milind Tambe[1], and Ariel D. Procaccia[2]

[1] University of Southern California, Los Angeles, USA
[jiangx,thanhhng,tambe]@usc.edu
[2] Carnegie Mellon University, Pittsburgh, USA
arielpro@cs.cmu.edu

**Abstract.** There has been recent interest in applying Stackelberg games to infrastructure security, in which a defender must protect targets from attack by an adaptive adversary. In real-world security settings the adversaries are humans and are thus boundedly rational. Most existing approaches for computing defender strategies against boundedly rational adversaries try to optimize against specific behavioral models of adversaries, and provide no quality guarantee when the estimated model is inaccurate. We propose a new solution concept, *monotonic maximin*, which provides guarantees against *all* adversary behavior models satisfying *monotonicity*, including all in the family of Regular Quantal Response functions. We propose a mixed-integer linear program formulation for computing monotonic maximin. We also consider top-monotonic maximin, a related solution concept that is more conservative, and propose a polynomial-time algorithm for top-monotonic maximin.

## 1 Introduction

Stackelberg games have been used to model resource allocation problems in infrastructure security, in which a defender must allocate limited security resources to protect targets from attack by an adversary [16, 10, 1, 2]. Due to surveillance by the adversary, any pure strategy by the defender can be exploited. The defender thus should commit to a mixed strategy as the leader in this Stackelberg game, taking into account the response by the adversary who is the follower. Classical solution concepts such as Strong Stackelberg Equilibrium assume that the follower is perfectly rational. However, in real-world security settings the adversaries are humans and thus this perfect rationality assumption is problematic. There has been much recent progress on optimal defender strategies for Stackelberg security games against boundedly rational adversaries, for various behavior models including epsilon-best response, anchoring bias, prospect theory and logit quantal response models [14, 18].

The quantal response (QR) model is well-supported by the social and behavioral science literature [11–13] and has performed well in laboratory experiments for the Stackelberg game setting [18]. Within the QR framework, there is some freedom in the choice of functional families (logit, probit, etc.) and parameter values, e.g., the parameter $\lambda$ in the logit QR model which measures the adversary's level of rationality. Once the function form is selected and parameter estimated (e.g., from real-world data or lab experiments), optimal defender strategies can be computed using optimization algorithms such as BRQR and PASAQ [18, 19].

However, there is some uncertainty about the best modeling parameters to use in real-world settings. In particular, real-world data on terrorist attacks are difficult to obtain. One can try to overcome this by running laboratory experiments, but models and parameters that give good fits in laboratory settings might not perform as well in actual security settings, due to factors such as different populations and different environments. And when the parameter estimate is inaccurate, *current algorithms provide no worst-case guarantee with respect to the solution quality*.

At the other extreme, there is the maximin solution: a leader strategy that maximizes leader expected utility when the follower is playing the worst-case strategy, i.e., play as if the follower is trying to minimize the leader's utility, even though the game is generally not zero-sum. The maximin solution provides utility guarantee without making any assumption on the attackers' behavior model. The maximin solution is computationally tractable: it can be solved by linear programming. However, the solution concept may be too conservative; in particular, the leader is disregarding any knowledge she may have about the follower's utilities in the game.

Are there robust solutions that do make use of recent advances in behavioral sciences, but are less sensitive to the choice of modeling parameters? In this paper we propose an approach that, instead of optimizing against a particular QR model, aims to guarantee good defender utility against all "reasonable" QR attackers. We note that QR in its most general form [13] covers all possible player behavior [6], so restriction to some notion of "reasonableness" is necessary. Goeree, Holt and Palfrey [5] proposed four properties that all reasonable QR models should satisfy, and called models satisfying all four properties Regular Quantal Response. In this paper, we impose constraints on attacker strategies that correspond to a relaxed version of Regular QR. Specifically, we assume that the attacker's strategies satisfy one of the four Regular QR properties, namely *monotonicity*, which is the property that actions with lower expected utility are played with smaller probability. (We further discuss the choice of monotonicity in Section 3.2.) We propose the following "monotonic maximin" solution concept to Stackelberg games: a defender plays a mixed strategy that maximizes defender expected utility, against the worst-case monotonic attacker mixed strategy. Since all Regular QR attackers satisfy monotonicity, monotonic maximin provides utility guarantees against all Regular QR attackers. Monotonic maximin is a robust alternative to the optimal Stackelberg strategy against specific QR models: it provides utility guarantees against all "reasonably rational" attackers (as defined by Regular QR) without making assumptions about parameters. This can be thought of as a "model-free" or "non-parametric" approach to Stackelberg games with boundedly rational followers.

The resulting computational problem might appear similar to a standard maximin problem, but is more challenging because the constraints for attacker's monotonicity now depend on the defender strategy. In this paper we propose an algorithm for this problem, based on LP duality and mixed-integer programming.

It is also interesting to consider attackers satisfying relaxations of the monotonicity constraint: the resulting defender strategies are *more* robust, as we are considering a larger set of possible attacker strategies. Specifically, we consider top monotonicity, the property that the follower's probability of playing each best response action is no less that that of any other action. We propose a polynomial-time algorithm for computing the resulting top-monotonic maximin solution concept.

We ran computational experiments to compare monotonic maximin and top-monotonic maximin against previously-proposed solution concepts including strong Stackelberg equilibrium [4], maximin, MATCH [15], as well as logit QR models with various parameter settings. Overall monotonic maximin is significantly more robust against monotonic adversaries compared to the previously-proposed solution concepts.

### 1.1 Related Work

There has been some recent work on designing defender strategies in security games that are robust against uncertainties, including uncertainties about the opponent as well as about the environment. One line of work is based on probabilistic models of uncertainties, and aims for security strategies that maximize the defender's expected utility under such probabilistic models. These include approaches based on specific models of bounded rationality, such as logit quantal response, prospect theory, and anchoring bias [18, 14]. A drawback of such approaches is the requirement on the availability and accuracy of probabilistic models; if an inaccurate probabilistic model is chosen, there is no quality guarantee with respect to the resulting security solution.

Another line of work, which includes our approach in this paper, adopts the robust optimization framework [17, 3] from Operations Research: define an *uncertainty set* that represents the space of likely models, and compute a security strategy that maximizes defender's utility under the worst case choice of models from that uncertainty set. For example, the BRASS algorithm [14] was designed to be robust against all adversaries playing epsilon-best response. An algorithm that is related to our approach is MATCH [15], which aims to provide a robust approach to Stackelberg security games against human attackers. MATCH is based on a similar intuition as our approach, that places less importance on attacker's actions with worse expected utilities. Specifically, MATCH bounds the defender's potential loss due to attacker's irrational behavior by a $\beta$-multiple of the attacker's loss due to his irrational behavior. Thus the robustness guarantee provided by MATCH is relative to the amount of the attacker's loss due to irrational behavior, and gets worse against less rational attackers. In contrast, our approach provides guarantees on defender utility against all Regular QR attackers.

At a high level, one drawback of these previous robust approaches is that they are still dependent on their parameter settings to define the sizes of their uncertainty sets. If the parameters are set so that the uncertainty sets are too small, the resulting solutions will be insufficiently robust. If the parameters are set so that the uncertainty sets are too large, the resulting solutions approach the maximin solution and are thus too conservative. While for certain cases it may be possible to come up with suitable parameters, our monotonic maximin approach avoids the requirement for parameters altogether. On the other hand, one could ask the same question about the uncertainty set defined by monotonic maximin: does the uncertainty set have the "right" shape and size? In particular, one potential criticism against monotonic maximin would be that it may be too conservative, because it uses only one of the four Regular QR conditions. In Section 3.2 we show that the uncertainty set for monotonic maximin is tight for Regular QR attackers, that is, any point in the uncertainty set could be arbitrarily approached by the behavior of a Regular QR attacker.

Finally, we mention work on modeling the game's uncertainties in aspects other than the adversary's behavior. Bayesian games were proposed to model players' probabilistic uncertainty about payoffs of the game [7]. There is also work that uses Bayesian games to model probabilistic uncertainties about defender's ability to execute the strategies as well as attacker's observation of defender strategies [22]. Within the robust optimization framework, The RECON algorithm [20] was designed to be robust against observation and execution uncertainties within a certain (hyperrectangular) error bound. Kiekintveld et al. [8] proposed robust solutions for security games against interval payoff uncertainties. While our paper's focus is on the behavior of the adversary, in Section 3.3 we briefly mention how our approach can be applied to achieve robustness against certain types of payoff uncertainty.

## 2 Preliminaries

Let $\mathbf{1}$ be a vector of 1's, the dimension of which will be clear from context. Let $\boldsymbol{e_i}$ be the $i$-th basis vector. Denote by $[n]$ the set $\{1, \ldots, n\}$.

We consider a two-player Stackelberg game between a leader and a follower. Leader's mixed strategy is denoted by $\boldsymbol{x} \in X \subset \mathbb{R}^m$ where $X = \{\boldsymbol{x} \in \mathbb{R}^m | C\boldsymbol{x} \leq d\}$ is a polytope. This includes the standard case where $\boldsymbol{x}$ is the distribution over $m$ leader actions, when $X$ is the simplex $\{\boldsymbol{x} \in \mathbb{R}^m | \boldsymbol{x} \geq 0, \mathbf{1}^T\boldsymbol{x} = 1\}$; it also includes cases where $\boldsymbol{x}$ is a compact representation of mixed strategy as marginal probabilities (e.g., marginal coverage on targets [9], or marginal flow on a network [21]). Follower has $n$ actions, labeled from 1 to $n$; i.e., his set of actions is $[n]$. Follower's mixed strategy is denoted by $\boldsymbol{y} \in Y$, where $Y = \{\boldsymbol{y} \in R^n | \boldsymbol{y} \geq 0, \mathbf{1}^T\boldsymbol{y} = 1\}$ is the standard simplex. The game's payoff matrices are $A, B \in \mathbb{R}^{m \times n}$. Expected utilities for the leader and the follower are $\boldsymbol{x}^T A\boldsymbol{y}$ and $\boldsymbol{x}^T B\boldsymbol{y}$ respectively. The game is general-sum: the sum of the players' utilities is not necessarily a constant.

**Stackelberg Security Games.** Although the solution concepts proposed in this paper apply to two-player Stackelberg games in general, we will frequently consider Stackelberg Security Games (SSGs) [9], a class of games with utility structure corresponding to the real-world problem of infrastructure security. Specifically, an SSG is a two-player Stackberg game between a defender (the leader) and an an adversary/attacker (the follower). There is a set of $n$ targets $T = [n]$. The defender can deploy resources to cover some of the targets. Let $Z \subset \{0,1\}^n$ be the set of feasible allocations of defender resources to targets, where for each allocation $z \in Z$ and target $j \in T = [n]$, $z_j = 1$ means the target is covered by the defender, and $z_j = 0$ means the target is not covered. Defender's set of mixed strategies $X$ can then be represented by the convex hull of feasible allocations $Z$: $X = \text{conv}(Z) \subset \mathbb{R}^n$. The attacker chooses one target to attack, i.e., his set of mixed strategies $Y$ is the standard simplex $\{\boldsymbol{y} \in R^n | \boldsymbol{y} \geq 0, \mathbf{1}^T\boldsymbol{y} = 1\}$.

The payoffs to the players depend only on which target is attacked, and whether that target is covered by the defender. In other words, whether the defender covers an un-attacked target does not affect the payoffs. Specifically, for each target $t \in T$, we denote by $U_d^u(t)$ the defender's utility for an uncovered attack on $t$, and $U_d^c(t)$ for a covered attack. Similarly, $U_a^u(t)$ and $U_a^c(t)$ are the attacker's payoffs for uncovered and covered attacks on $t$, respectively. In terms of the payoff matrices $A, B \in \mathbb{R}^{n \times n}$, this means that $A_{ij}$ is equal to $U_d^c(j)$ if $i = j$, and $U_d^u(j)$ otherwise; while $B_{ij}$ is equal to $U_a^c(j)$ if $i = j$ and $U_a^u(j)$ otherwise. We further assume that $U_d^c(t) > U_d^u(t)$ and $U_a^c(t) < U_a^u(t)$ for all $t \in T$.

In this paper we will focus on SSGs in which the set of feasible defender allocations $Z$ has a simple structure: the defender has $r$ resources, and each resource can protect any single target. Thus any allocation that uses $r$ resources is feasible. The corresponding convex hull $X$ can be described using a small number of constraints: $X = \{x \in \mathbb{R}^n | \mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}, \mathbf{1}^T \boldsymbol{x} = r\}$. We call such a game an *SSG with $r$ resources*.

**Table 1.** An example 3-target Stackelberg security game

| | Target 1 | Target 2 | Target 3 |
|---|---|---|---|
| $U_d^c$ | 7 | 10 | 2 |
| $U_d^u$ | -10 | -8 | -10 |
| $U_a^u$ | 3 | 10 | 4 |
| $U_a^c$ | -10 | -4 | -10 |

*Example 1.* Table 1 shows the payoffs of an example Stackelberg security game with 3 targets. Specifically, the columns represent the targets and for each column the defender's utilities for covered attack ($U_d^c$) and uncovered attack ($U_d^u$), and the attacker's utilities for uncovered attack ($U_a^u$) and covered attack ($U_a^c$) are given.

**Strong Stackelberg Equilibrium (SSE)** is one of the standard solution concepts of Stackelberg games. In an SSE, the leader is maximizing her expected utility, assuming that the follower plays a best response. When the follower has multiple best responses, he is assumed to break ties in favor the leader. Formally, the SSE strategy for the leader is $\arg\max_{x \in X, y \in BR(x)} x^T Ay$, where $BR(x) = \arg\max_{y \in Y} x^T By$ is the set of best responses of the follower given leader strategy $x$.

**Quantal Response** is in general defined by a function $P : \mathbb{R}^n \to Y$ from the vector of expected payoffs of an agent's actions to a probability distribution over the actions. Denote by $P_j(\boldsymbol{u})$ the probability of playing action $j$ given the vector $\boldsymbol{u} \in \mathbb{R}^n$ of expected payoffs. For example, the logit quantal response function has the form $P_j(\boldsymbol{u}) = \frac{e^{\lambda u_j}}{\sum_{j'} e^{\lambda u_{j'}}}$ where $\lambda \geq 0$ is a parameter. Other examples of $P$ include probit, and the constant mapping to the uniform distribution.

Quantal Response Equilibrium (QRE) [13] is a solution concept for simultaneous games in which all players are playing quantal response strategies. In security domains, the adversary is human (and therefore not perfectly rational) while the defender can be assumed to be a rational decision maker aided by computers. This "Stackelberg against Quantal Response" model has been studied by Yang *et al* [18], who assumed that the adversary's quantal response function is known to the defender. In this paper we consider the case where the defender knows that the adversary behaves according to some quantal response model but does not know the specific quantal response function $P$.

**Regular QRE**. Goeree, Holt and Palfrey [5] proposed constraints that all reasonable QRE models should satisfy. Formally, $P$ is a *regular quantal response function* if it satisfies the following:

1. Interiority: $P_j(\boldsymbol{u}) > 0$ for all $j$.
2. Continuity: $P_j(\boldsymbol{u})$ is continuously differentiable.
3. Responsiveness: $\frac{\partial P_j(\boldsymbol{u})}{\partial u_j} > 0$ for all $j$.
4. Monotonicity: $u_j > u_k \Rightarrow P_j(\boldsymbol{u}) > P_k(\boldsymbol{u})$ for all $j, k$.

They also point out that Continuity and Monotonicity imply $u_j = u_k \Rightarrow P_j(\boldsymbol{u}) = P_k(\boldsymbol{u})$. The logit and the probit distributions are examples of regular quantal response functions. On the other hand, choosing a best response is not a regular quantal response function because it does not satisfy Interiority and Continuity.

**The maximin solution** is the optimal defender strategy assuming that the attacker is choosing the strategy that is worst for the defender: $\arg\max_{\boldsymbol{x}\in X} \min_{\boldsymbol{y}\in Y} \boldsymbol{x}^T A\boldsymbol{y}$. This solution concept is extremely conservative: the defender has to take into account an attacker that does completely arbitrary things, and as a result is disregarding his knowledge about the attacker payoff matrix $B$ and treating the game as a zero-sum game.

## 3 Monotonic Maximin

Our overall approach is to modify maximin by imposing constraints on the attacker strategy. Specifically, we assume that the attacker strategy satisfies monotonicity. Since all Regular QR attackers satisfy monotonicity, our approach is able to provide guarantee against all Regular QR attackers. For computational convenience we will use the following form of monotonicity.

**Definition 1.** *Given $\boldsymbol{x} \in X, \boldsymbol{y} \in Y$, we say $\boldsymbol{y}$ satisfies* closed monotonicity *if for all $i, j \in [n]$, $\boldsymbol{x}^T B\boldsymbol{e_i} \geq \boldsymbol{x}^T B\boldsymbol{e_j} \Rightarrow y_i \geq y_j$.*

Recall that $\boldsymbol{x}^T B\boldsymbol{e_i}$ is the follower's expected utility of choosing action $i$, given that the leader plays $\boldsymbol{x}$. There are strategies that are closedly monotonic but not monotonic, for example the uniformly random strategy. It is straightforward to show the following:

**Proposition 1.** *If attacker is acting according to a regular quantal response function, then his mixed strategy $\boldsymbol{y}$ satisfies closed monotonicity.*

*Proof.* We need to show that for all $i, j$, $\boldsymbol{x}^T B\boldsymbol{e_i} \geq \boldsymbol{x}^T B\boldsymbol{e_j} \Rightarrow y_i \geq y_j$. Given $i, j$, suppose $\boldsymbol{x}^T B\boldsymbol{e_i} > \boldsymbol{x}^T B\boldsymbol{e_j}$. Then by the assumption of Monotonicity we have $y_i > y_j$ which implies $y_i \geq y_j$. Now suppose $\boldsymbol{x}^T B\boldsymbol{e_i} = \boldsymbol{x}^T B\boldsymbol{e_j}$. Then by Continuity and Monotonicity, $y_i = y_j$ which implies $y_i \geq y_j$.

Observe that closed monotonicity is not necessarily a weaker version of Monotonicity; nevertheless it is a consequence of Continuity and Monotonicity.

Let $Q(\boldsymbol{x}) \subseteq Y$ be the set of follower mixed strategies that satisfy closed monotonicity given $\boldsymbol{x}$. Then $Q(\boldsymbol{x}) = \{\boldsymbol{y} \in Y | \forall (i,j) \in E(\boldsymbol{x}), y_i \geq y_j\}$, where $E(\boldsymbol{x}) = \{(i,j) \in [n] | \boldsymbol{x}^T B\boldsymbol{e_i} \geq \boldsymbol{x}^T B\boldsymbol{e_j}\}$.

**Definition 2.** *The* monotonic maximin *solution is*

$$\arg\max_{\boldsymbol{x}\in X} \min_{\boldsymbol{y}\in Q(\boldsymbol{x})} \boldsymbol{x}^T A\boldsymbol{y}. \tag{1}$$

Let the monotonic maximin value be the corresponding objective value: $\max_{\boldsymbol{x} \in X} \min_{\boldsymbol{y} \in Q(\boldsymbol{x})} \boldsymbol{x}^T A \boldsymbol{y}$. By definition, the monotonic maximin solution provides guaranteed expected utility of at least the monotonic maximin value against all attacker strategies satisfying the monotonicity property.

*Example 2.* Consider the 3-target Stackelberg security game from Example 1. Suppose the defender has one resource. The defender's strategies generated by monotonic maximin, maximin, and SSE are shown in Table 2. For example, the second row indicates the defender's strategy generated by monotonic maximin, i.e., target 1 will be covered by the defender 37% of the time while there are 53% and 10% that target 2 and 3 will be covered by the defender, respectively.

When the strategy of the defender is generated by monotonic maximin, the defender's expected utility is -3.65 given a worst-case monotonic attacker strategy. Multiple monotonic attacker strategies tied for the worst case, including $(\frac{1}{2}, 0, \frac{1}{2})$ and $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. On the other hand, when maximin is used, the defender's expected utility is -4.38 for any actions of the monotonic attacker. Finally, when SSE is used, the attacker's expected utilities for all targets are the same and equal to 1.05. Thus the only feasible action for the monotonic attacker is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The defender's expected utility in this case is -3.8.

**Table 2.** Defender's strategy

| | Target 1 | Target 2 | Target 3 |
|---|---|---|---|
| Monotonic maximin | 0.3732 | 0.5277 | 0.0991 |
| Maximin | 0.3306 | 0.2011 | 0.4683 |
| SSE | 0.15 | 0.6393 | 0.2107 |

The following proposition shows that the monotonic maximin concept is of most interest when the game is not zero sum.

**Proposition 2.** *For zero-sum games, the monotonic maximin solution coincides with maximin solution.*

Intuitively, if we consider e.g., a logit QR follower with $\lambda \to \infty$, then his behavior approaches that of a perfectly rational player and the leader can do no better than the maximin solution in a zero-sum game.

### 3.1 Existence of Monotonic Maximin Solutions

The standard Extreme Value Theorem states that a continuous function on a compact domain has a maximum. Since the set of monotonic follower strategies $Q(\boldsymbol{x})$ is not continuous in $\boldsymbol{x}$, the value of the inner minimization $\min_{\boldsymbol{y} \in Q(\boldsymbol{x})} \boldsymbol{x}^T A \boldsymbol{y}$ is not necessarily continuous in $\boldsymbol{x}$. A natural question arises: does the monotonic maximin solution always exist? Of course if the maximum does not exist we could take the supremum instead, but the corresponding defender strategy would no longer be guaranteed to be robust.

**Proposition 3.** *The monotonic maximin solution exists in all Stackelberg games.*

This will be a direct consequence of Proposition 6 in Section 4, which provides an algorithm for monotonic maximin.

### 3.2 Optimality against Interiority, Continuity and Responsiveness

One potential criticism is that by focusing on monotonicity (and not the other conditions of Regular QR), monotonic maximin may be too conservative as it does not take advantage of all information about the follower behavior provided by the Regular QR model.

**Proposition 4.** *The monotonic maximin solution is arbitrarily close to optimal against an attacker who chooses the worst (for the defender) strategy satisfying both closed monotonicity and interiority.*

*Proof (sketch).* It is sufficient to show that given any $x$, $\min_{y \in Q(x)} x^T Ay = \inf_{y \in Q(x) \cap \text{Int}} x^T Ay$ where Int is the set of strategies satisfying interiority. Given an attacker strategy $y$ that does not satisfy interiority (say a solution of the LHS), we can construct another strategy $y'$ that satisfies interiority by re-assigning a small amount of probability mass to actions with zero probability in $y$. It is also straightforward to show that this can be done in a way that preserves closed monotonicity. $y$ and $y'$ achieve almost the same expected payoffs for both players.

Let us now consider continuity and responsiveness. Unlike monotonicity and interiority, which can be expressed as "local constraints" on $y$, continuity and responsiveness are properties of the response function $P(\boldsymbol{u})$ and correspond to constraints on the values of $P$ given multiple inputs.

Consider the inner minimization problem of monotonic maximin: $q(x) = \arg\min_{y \in Q(x)} x^T Ay$. This defines a response function $P(\boldsymbol{u})$ for the attacker,[3] which likely violates continuity and responsiveness. But is that necessarily the response function of the attacker we face? In particular does the attacker's response function have to be the same regardless of the defender's mixed strategy? Instead, we allow the attacker to "pick a response function" after observing defender's mixed strategy $x$, which is consistent with our overall robust optimization approach. It turns out that it is possible to pick the response function in a way that satisfies all conditions of regular quantal response, at the same time outputting the worst-case closedly monotonic strategy given $x$. This shows that monotonic maximin remains the optimal solution concept even when we consider attackers that satisfy all conditions of regular quantal response.

**Proposition 5.** *Given $x \in X$, there exists a regular quantal response function $P : \mathbb{R}^n \to Y$ such that $P(x^T B)$ is arbitrarily close to* $\arg\min_{y \in Q(x)} x^T Ay$.

We give a proof in Section 5.2.

### 3.3 Capturing other Behavioral and Uncertainty Models

In this section we show that monotonic maximin provides guarantees not only against regular quantal response attackers, but also other models of attacker behavior. Furthermore, if uncertainties in the game model (e.g., in the game's payoffs, attacker's capabilities, defender's execution, etc.) result in attacker behavior that is monotonic, then we can use monotonic maximin as a robust solution concept against such uncertainties.

[3] Actually for $P$ to be well-defined, it requires that $q(x) = q(x')$ whenever $x^T B = x'^T B$. This holds for Stackelberg security games, in which the players' utilities depend on the coverage on targets.

**Behavioral Models.** A mixture (i.e., convex combination) of regular quantal response models is also a regular quantal response model, and therefore satisfies closed monotonicity. For example, one can have some probabilistic prior belief over the values of parameter $\lambda$ in logit QR models, resulting in a mixture of logit QR models. As another example, consider a mixture of a regular QR model with the model that attacker plays a uniformly random mixed strategy. This is also a mixture of Regular QR models because the uniformly random strategy is a special case of Regular QR. Monotonic maximin provides utility guarantees against all such models.

Such guarantees are also applicable to behavior models that are not Regular QR but satisfy closed monotonicity. For example, consider the following "uniform best response" model: the attacker chooses a best response; if there are multiple pure-strategy best responses the attacker uniformly randomizes among those best responses. This is not a Regular QR function since it is not continuously differentiable; but it satisfies closed monotonicity. More generally, consider the *uniform top-$K$* strategy, in which the top $K$ actions in terms of expected utilities are played, each with equal probability of $1/K$. These are not Regular QR but are nevertheless closedly monotonic. We will see later that these strategies have importance in monotonic maximin solutions.

**Payoff Uncertainty.** A simple consequence of [13] is that if we add i.i.d. noise with a smooth distribution of zero mean to the entries of the follower's payoff matrix $B$, and assuming that the follower plays a best response, then the resulting average follower strategy is monotonic. However this kind of noise does not preserve the structure of Stackelberg security games. For Stackelberg security games, consider the following type of payoff noise: i.i.d. with a smooth distribution of zero mean, added for each target to the payoffs of covered and uncovered attacks. For each instantiation of this noise the resulting game is still a Stackelberg security game. Given a defender mixed strategy, this would result in zero-mean i.i.d. noise over the expected attacker utilities of attacking each target. By the same argument as in [13], if the follower plays a best response, then the resulting average follower strategy is monotonic.

However, if the follower has a monotonic response function (not just a best response), the resulting average strategy under i.i.d. payoff noise is not guaranteed to be monotonic. Indeed, our numerical experiments show that when the follower plays the worst-case monotonic strategy with respect to perturbed utilities, the resulting average strategy is not always monotonic. It is possible to show that monotonicity is preserved under such noise if we assume the follower's response function is symmetric with respect to actions; we leave the detailed discussion to a future extended version of the paper.

### 3.4 Top-Monotonic Maximin

We define *top monotonicity* to be the property that for each best response action of the attacker, the probability of that action is no less than that of any other action. Formally, $\boldsymbol{y}$ satisfies top monotonicity given $\boldsymbol{x}$ if for all $i \in [n]$,

$$\boldsymbol{x}^T B \boldsymbol{e_i} \geq \boldsymbol{x}^T B \boldsymbol{e_j} \; \forall j \Rightarrow y_i \geq y_j \; \forall j.$$

We denote by $\widehat{Q}(x) \subset Y$ the set of top-monotonic follower strategies given $\boldsymbol{x}$.

Top-monotonicity is a relaxation of closed monotonicity: the inequality $y_i \geq y_j$ only needs to hold between the best response action $i$ and each of the other actions $j$. In other words, the corresponding *top-monotonic maximin* solution

$$\arg\max_{x \in X} \min_{y \in \widehat{Q}(x)} x^T A y$$

is more conservative than monotonic maximin.

Top-monotonic maximin is interesting partially because there have been extensive studies on various solution concepts that focus on pairwise comparisons between the best-response action against possible deviations.[4] Furthermore, we will show later that top-monotonic maximin can be computed in polynomial time.

# 4 Computation of Monotonic Maximin

## 4.1 Multiple-LP Formulation

Unlike the Maximin problem, we cannot directly use linear programming to solve (1). This is because the feasible set $Q(\boldsymbol{x})$ for $\boldsymbol{y}$ now depends on $\boldsymbol{x}$. Fortunately, $Q(\boldsymbol{x})$ depends only on $E(\boldsymbol{x})$, which is essentially the ordering of attacker actions in terms of attacker utilities. So in theory we could solve an LP for each possible ordering, and return the one with best defender utility.

Since the attacker utilities are real numbers, the binary relation $E(\boldsymbol{x}) \subset [n] \times [n]$ satisfies the constraints of a *total order*, i.e., transitivity: $(i,j),(j,k) \in E(\boldsymbol{x}) \Rightarrow (i,k) \in E(\boldsymbol{x})$ and totality: $(i,j) \in E(\boldsymbol{x}) \vee (j,i) \in E(\boldsymbol{x})$. Given a total order $\mathcal{E} \subset [n] \times [n]$, let $E^{-1}(\mathcal{E}) = \{\boldsymbol{x} \in X : \mathcal{E} = E(\boldsymbol{x})\}$, i.e., the set of leader strategies inducing the order $\mathcal{E}$ on follower expected utilities.

Thus, for each $\mathcal{E}$ that corresponds to a total order, we solve

$$\max_{\boldsymbol{x} \in E^{-1}(\mathcal{E})} \min_{\boldsymbol{y} \in Q(x)} \boldsymbol{x}^T A \boldsymbol{y}, \tag{2}$$

and output the solution that achieves the best objective value. However, the set

$$\begin{aligned} E^{-1}(\mathcal{E}) &= \{\boldsymbol{x} : \forall (i,j) \in \mathcal{E}, (i,j) \in E(\boldsymbol{x});\ \forall (i,j) \notin \mathcal{E}, (i,j) \notin E(\boldsymbol{x})\} \\ &= \{\boldsymbol{x} : \forall (i,j) \in \mathcal{E}, \boldsymbol{x}^T B \boldsymbol{e_i} \geq \boldsymbol{x}^T B \boldsymbol{e_j};\ \forall (i,j) \notin \mathcal{E}, \boldsymbol{x}^T B \boldsymbol{e_i} < \boldsymbol{x}^T B \boldsymbol{e_j}\} \end{aligned}$$

is not closed in general, since it involves strict inequalities for pairs not in $\mathcal{E}$. This presents problems such as potential nonexistence of solutions of (2). We instead use the closure of $E^{-1}(\mathcal{E})$, which is

$$\mathrm{cl}E^{-1}(\mathcal{E}) = \{\boldsymbol{x} : \forall (i,j) \in \mathcal{E}, \boldsymbol{x}^T B \boldsymbol{e_i} \geq \boldsymbol{x}^T B \boldsymbol{e_j};\ \forall (i,j) \notin \mathcal{E}, \boldsymbol{x}^T B \boldsymbol{e_i} \leq \boldsymbol{x}^T B \boldsymbol{e_j}\}.$$

[4] For example, in a Strong Stackelberg Equilibrium, any non-best response of the follower receives zero probability; for the epsilon-best responses considered in the BRASS algorithm [14], any follower action that is more than $\epsilon$ worse than the best response receives zero probability; In a MATCH solution [15], if an adversary action $j$ is $\epsilon$ worse than the best response, the defender's potential loss if the adversary chooses $j$ instead of the best response is bounded by $\beta\epsilon$, where $\beta > 0$ is a parameter of the solution concept.

Since $\mathcal{E}$ is a total order, $(i,j) \notin \mathcal{E}$ implies that $(j,i) \in \mathcal{E}$, so the above can be simplified to $\{\boldsymbol{x} : \forall (i,j) \in \mathcal{E}, \boldsymbol{x}^T B \boldsymbol{e_i} \geq \boldsymbol{x}^T B \boldsymbol{e_j}\}$. Given $\mathcal{E}$, define the matrix $F \in \mathbb{R}^{n \times n(n-1)}$ such that its $(i,j)$-th column $F_{(i,j)}$ is $\boldsymbol{e_i} - \boldsymbol{e_j}$ if $(i,j) \in \mathcal{E}$ and the 0 vector otherwise. Then cl$E^{-1}(\mathcal{E})$ can be written as $\{\boldsymbol{x} : \boldsymbol{x}^T BF \geq 0\}$. We will show below that replacing $E^{-1}(\mathcal{E})$ with cl$E^{-1}(\mathcal{E})$ will not introduce incorrect solutions.

The inner minimization problem of (2) can then be written as

$$\min \boldsymbol{x}^T A \boldsymbol{y} \tag{3}$$

$$F^T \boldsymbol{y} \geq 0 \tag{4}$$

$$\mathbf{1}^T \boldsymbol{y} = 1 \tag{5}$$

$$\boldsymbol{y} \geq 0 \tag{6}$$

where $F^T \boldsymbol{y} \geq 0$ is the matrix form for constraints $y_i \geq y_j \; \forall (i,j) \in \mathcal{E}$.

Given $\boldsymbol{x}$, the above is an LP. By LP duality, its optimal solution is equal to that of its dual LP

$$\max t \tag{7}$$

$$F\boldsymbol{\lambda} + t\mathbf{1} \leq A^T \boldsymbol{x} \tag{8}$$

$$\boldsymbol{\lambda} \geq 0. \tag{9}$$

Now that the inner min becomes an max, max-min becomes a max-max problem. Recall that $\boldsymbol{x} \in X$ can be expressed as the linear constraints $C\boldsymbol{x} \leq d$, and $\boldsymbol{x} \in$ cl$E^{-1}(\mathcal{E})$ can be expressed as the linear constraints $\boldsymbol{x}^T BF \geq 0$. Then (2) can be written as

$$V_F = \max_{\boldsymbol{x}, \boldsymbol{\lambda}, t} \; t \tag{10}$$

$$C\boldsymbol{x} \leq d \tag{11}$$

$$\boldsymbol{x}^T BF \geq 0 \tag{12}$$

$$F\boldsymbol{\lambda} + t\mathbf{1} \leq A^T \boldsymbol{x} \tag{13}$$

$$\boldsymbol{\lambda} \geq 0 \tag{14}$$

which is an LP.

**Proposition 6.** *Consider the following Multiple-LP algorithm: given a two-player general sum game, solve the LP* (10) *for each $F$ corresponding to a total order over the set of attacker actions. For the LP achieving the highest objective value $V_F$, output its solution $\boldsymbol{x}$. Then $\boldsymbol{x}$ is a monotonic maximin solution of the game.*

*Proof (sketch).* By construction, $\{E^{-1}(\mathcal{E}) : \mathcal{E}$ is a total order$\}$ partitions $X$; however each of the $E^{-1}(\mathcal{E})$ is not necessarily closed. Instead, the feasible sets for $\boldsymbol{x}$ in the LP instances (10) are closures of $E^{-1}(\mathcal{E})$, and thus cover $X$. These feasible sets have overlaps: consider total orders $\mathcal{E}$ and $\mathcal{E}'$, corresponding to matrices $F$ and $F'$ respectively, such that $\mathcal{E}' \subset \mathcal{E}$. Then cl$E^{-1}(\mathcal{E}) \subseteq$ cl$E^{-1}(\mathcal{E}')$. Such overlap presents potential problems if the LP for $F'$ has a solution $\boldsymbol{x'} \in$ cl$E^{-1}(\mathcal{E})$, with an objective value greater than $V_F$, the optimal objective of the LP for $F$; this is because such a solution would mask the correct solution for the region cl$E^{-1}(\mathcal{E})$. We claim that this masking will never happen. Take this $\boldsymbol{x'}$, which is feasible for the LPs for $F$ and $F'$, and compare the objective values achieved by $\boldsymbol{x'}$ in the two LPs. Given $\boldsymbol{x'}$, the objective for the LP for $F$ will be higher, intuitively because $\mathcal{E}' \subset \mathcal{E}$ and closed monotonicity implies that the follower is subject to more constraints in the case of $\mathcal{E}$, which makes the leader better off. Therefore the objective value of $\boldsymbol{x'}$ can never be higher than $V_F$, and the output of the Multiple-LP algorithm will be the monotonic maximin solution.

A direct consequence of this result is the existence of monotonic maximin solutions in all Stackelberg games (Proposition 3).

However, we would need to solve one LP for each total order on $[n]$. The following proposition shows that we only need to consider the *strict* orderings on $[n]$, i.e., those $\mathcal{E}$ in which for each pair of actions $i, j$, exactly one of $(i, j)$ and $(j, i)$ is in $\mathcal{E}$.

**Proposition 7.** *Consider a "non-strict" total order $\mathcal{E}^t$, with corresponding matrix $F^t$, i.e., there exists $(i, j)$ such that $(i, j), (j, i) \in \mathcal{E}^t$. We say total order $\mathcal{E}^c$ (with corresponding matrix $F^c$) is a* sharpening *of $\mathcal{E}^t$ if $\mathcal{E}^c \subset \mathcal{E}^t$ and $\mathcal{E}^c$ is a strict order; i.e., for every pair $(i, j), (j, i) \in \mathcal{E}^t$, either $(i, j)$ or $(j, i)$ belongs to $\mathcal{E}^c$, not both. Let $\mathcal{F}(\mathcal{E}^t)$ be the set of matrices corresponding to sharpenings of $\mathcal{E}^t$. Then $V_{F^t} \leq \max_{F^c \in \mathcal{F}(\mathcal{E}^t)}\{V_{F^c}\}$.*

The proof is given in an online appendix available at https://www.dropbox.com/s/f7wjpwmqa68u3x1/appendix.pdf.

This reduces the number of orderings we need to consider, but there are still $n!$ strict orderings to consider, corresponding to permutations of $[n]$. One approach to overcome this is to formulate the problem as a mixed-integer linear program (MILP).

### 4.2 MILP Formulation

The main idea is to have a binary integer variable $z_{ij}$ that indicates whether $(i, j) \in E$. Then $F_{(i,j)} = z_{ij}(\boldsymbol{e_i} - \boldsymbol{e_j})$. To ensure that $E$ corresponds to a total order, we can have constraints $z_{ij} + z_{ji} \geq 1$ and $(1 - z_{ij}) + (1 - z_{jk}) + z_{ik} \geq 1$. Then $\boldsymbol{x}^T BF \geq 0$ can be expressed as

$$\boldsymbol{x}^T B\boldsymbol{e_i} + M(1 - z_{ij}) \geq \boldsymbol{x}^T B\boldsymbol{e_j},\ \forall i, j \tag{15}$$

where $M$ is a sufficiently large positive constant that upper bounds $|\boldsymbol{x}^T B(\boldsymbol{e_i} - \boldsymbol{e_j})|$, e.g. $M = (\max_{i\in[m],j\in[n]} B_{ij} - \min_{i\in[m],j\in[n]} B_{ij}) \max_{x\in X} \|x\|_1$.

One issue is that $F\boldsymbol{\lambda} = \sum_{i,j} \lambda_{ij} F_{(i,j)} = \sum_{i,j} \lambda_{ij} z_{ij}(\boldsymbol{e_i} - \boldsymbol{e_j})$, which now involves quadratic terms. We can transform this quadratic expression into MILP constraints using standard techniques, by replacing $\lambda_{ij} z_{ij}$ with a new variable $w_{ij}$ satisfying the following constraints:

$$w_{ij} \geq 0 \tag{16}$$
$$w_{ij} \geq \lambda_{ij} - (1 - z_{ij})N \tag{17}$$
$$w_{ij} \leq \lambda_{ij} \tag{18}$$
$$w_{ij} \leq z_{ij} N \tag{19}$$

where $N$ is a large positive constant. In fact we can eliminate $\boldsymbol{\lambda}$ since it is not used elsewhere, i.e., we do not need to include the constraints (17) and (18). Taking these together, we have a polynomial-sized MILP

$$\max_{\boldsymbol{x},\boldsymbol{w},t,\boldsymbol{z}} \; t \tag{20}$$

$$C\boldsymbol{x} \le d \tag{21}$$

$$\boldsymbol{x}^T B\boldsymbol{e_i} + M(1 - z_{ij}) \ge \boldsymbol{x}^T B\boldsymbol{e_j},\; \forall i, j \tag{22}$$

$$\sum_{i,j} w_{ij}(\boldsymbol{e_i} - \boldsymbol{e_j}) + t\mathbf{1} \le A^T\boldsymbol{x} \tag{23}$$

$$0 \le w_{ij} \le z_{ij}N \tag{24}$$

$$z_{ij} \in \{0, 1\} \tag{25}$$

$$z_{ij} + z_{ji} \ge 1 \tag{26}$$

$$(1 - z_{ij}) + (1 - z_{jk}) + z_{ik} \ge 1. \tag{27}$$

### 4.3 Computing Top-Monotonic Maximin

Top-monotonic maximin can be efficiently computed by solving a small number of LPs. There are $n$ possible best response actions of the attacker corresponding to $n$ LPs: for each $i \in [n]$, let $E = \{(i, j)|j \in [n], j \neq i\}$ and solve (10). These correspond to partial orders as opposed to the total orders used previously. What about the cases with multiple best responses? The same argument as in Proposition 7 shows that we only need to consider the case with a single best response.

**Proposition 8.** *Top-monotonic maximin can be computed in time polynomial in $n$, $m$, and the number of constraints that define $X$.*

We can define similar relaxations of closed monotonicity, that focus on the best $L$ actions: $\boldsymbol{y}$ is top-$L$ monotonic for positive integer $L \ge 2$ if the closed monotonicity condition holds for any pair of actions $(i, j)$ in which at least one of $i$ and $j$ is among the top $L$ actions.[5] Top-$L$ monotonic maximin is defined analogously. For the corresponding computational problem, we only need to solve $\frac{n!}{(n-L)!}$ LPs, one for each way of selecting an ordered $L$-tuple from $n$ actions as the top $L$ actions. This number of LPs is polynomial when $L$ is fixed to be a constant.

**Proposition 9.** *For constant $L$, top-$L$ monotonic maximin can be computed in time polynomial in $n$, $m$, and the number of constraints that define $X$.*

## 5 Structure of Monotonic Maximin Solutions

### 5.1 Extreme Points of the Set of Monotonic Follower Strategies

Given $\boldsymbol{x}$, the attacker's feasible region $Q(\boldsymbol{x})$ is a polytope. Consider the inner minimization problem $min_{y \in Q(x)} x^T Ay$. Since the objective is linear, it is sufficient to consider only vertices of the polytope $Q(\boldsymbol{x})$. These vertices correspond to points $\boldsymbol{y}$ for which a sufficient number of inequalities of $Q(\boldsymbol{x})$ become tight. The inequalities of $Q(\boldsymbol{x})$ are of the form $y_i \ge 0$ and $y_i \ge y_j$. Therefore, we have the following:

[5] The notion of top $L$ actions is well-defined: we do not need to consider the case of ties, by the same argument as in Proposition 7.

**Lemma 1.** *Let $\boldsymbol{y}$ be a vertex of the polytope $Q(\boldsymbol{x})$. Then each action that is played with positive probability in $\boldsymbol{y}$ is played with equal probability.*

Thus a vertex is specified by a support set $R \subseteq [n]$, which is the set of attacker actions played with positive probability. Given $R$, the vertex $\boldsymbol{y}^R$ has $y_i^R = 0$ for all $i \notin R$, and $y_i^R = 1/|R|$ for $i \in R$.

The support set $R$ of a vertex of $Q(\boldsymbol{x})$ has the following properties:

1. All best responses of the follower are required to be in $R$.
2. If $i \in R$ then all pure strategies that are better for the follower than $i$ are also in $R$.
3. Since $\boldsymbol{x}$ induces a total ordering on $[n]$ of the follower expected utilities, there exists an action $j \in [n]$ such that $i \in R$ iff $x^T B_i \geq x^T B_j$. We call such an action $j$ a *threshold action*.

In other words, given $\boldsymbol{x}$, threshold action determines the support set. Since there are at most $n$ possible threshold actions, we have the following characterization of the polytope $Q(\boldsymbol{x})$:

**Proposition 10.** *$Q(\boldsymbol{x})$ has at most $n$ vertices, each of which is specified by a threshold action.*

These vertices correspond to *uniform top $K$* strategies for various $K$ (recall that these are strategies in which the top $K$ actions in terms of expected utilities are played, each with equal probability of $1/K$). This ranges from the “rational” uniform best response strategy, to the completely mixed uniform random strategy.

### 5.2 Proof of Proposition 5

In this section, we give a proof of Proposition 5, making use of the structure of $Q(\boldsymbol{x})$ that we showed in Section 5.1.

*Proof (of Proposition 5).* Recall that given $\boldsymbol{x} \in X$, Proposition 5 asks for a regular quantal response function $P$ such that $P(\boldsymbol{x}^T B)$ is arbitrarily close to $\arg\min_{y \in Q(x)} x^T Ay$. Given $\boldsymbol{x}$, we know that at least one attacker strategy in $\arg\min_{y \in Q(x)} x^T Ay$ is a vertex of $Q(\boldsymbol{x})$, and is therefore a uniform top-$K$ strategy for some $K$. Now fix this $K$, and consider an attacker who plays the uniform top-$K$ strategy against any $\boldsymbol{x}$. Since the top $K$ actions depends only on the attacker utilities $\boldsymbol{x}^T B$, this defines a response function $P'$. However, $P'$ does not satisfy interiority and continuity. We instead consider $P^\lambda$, a smooth version of $P'$, defined as follows: given a vector of expected utilities $\boldsymbol{u} \in \mathbb{R}^n$, $P^\lambda$ first selects a subset $S$ of $K$ actions according to the probability distribution

$$\Pr(S) = \frac{e^{\lambda \sum_{j \in S} u_j}}{\sum_{S' \subset [n]: |S'| = K} e^{\lambda \sum_{j' \in S'} u_{j'}}},$$

then randomize uniformly over the $K$ selected actions. For any $\lambda > 0$, $P^\lambda$ is a response function satisfying monotonicity, interiority, continuity and responsiveness. As we take $\lambda \to \infty$, $P^\lambda$ becomes arbitrarily close to $P$.

# 6 Evaluation

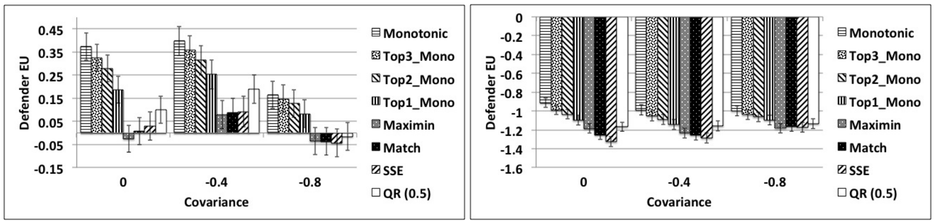
We ran computational experiments to compare the performance of monotonic maximin and its variants (top-$L$ monotonic maximin) against other previously-proposed solution concepts, including Strong Stackelberg Equilibrium (SSE), MATCH [15], maximin, and logit Quantal Response. Both the solution quality and the runtime performance are examined, on instances of Stackelberg security games across a wide range of number of targets and number of defender resources.

## 6.1 Payoff structures

The performances of solution concepts are affected by payoffs of the game. In particular, it is known that for zero-sum games, SSE, MATCH and maximin solutions coincide [15]; indeed monotonic maximin also coincides with maximin for zero-sum games (Proposition 2). We generated payoff structures for Stackelberg security games with different covariance values, by adapting the covariance game generator of the GAMUT package. The covariance value $r$, which is chosen within the range $[-1.0, 0.0]$, measures the correlation between the defender's payoff and the adversary's payoff. For example, when $r = -1.0$, the game becomes zero-sum whereas there will be no correlation between the defender and the adversary's payoffs when $r = 0.0$. The rewards for success of both the defender and the adversary are positive integers which lie within the range $[1, 10]$. On the other hand, the penalties for failure are negative integers within the range $[-10, -1]$.

## 6.2 Solution quality against worst-case monotonic attackers

In the first set of experiments, we compared the solution quality (i.e., defender expected utility) of the different solution concepts against the worst-case closedly monotonic attacker. That is, given a defender strategy $x$ provided by one of the solution concepts, the attacker chooses $\boldsymbol{y} \in \arg\min_{\boldsymbol{y} \in Q(\boldsymbol{x})} \boldsymbol{x}^T A \boldsymbol{y}$. We would expect monotonic maximin to achieve the best performance, since it is by definition the optimal solution in this measure. The purpose of these experiments is to observe the magnitudes of differences in performance between monotonic maximin and others, and to check whether our top-$L$ monotonic maximin algorithms provide good approximations to monotonic maximin. Specifically, we compared the performances among monotonic maximin, top-3, top-2, top-1 monotonic maximin, Maximin, MATCH, SSE, and logit Quantal Response. For logit Quantal Response, we tried a number of different values for $\lambda$, ranging from 1/32 to 8, but will only present the results for the best-performing value, which is $\lambda = 0.5$. The results for 6-target and 8-target games with 3 defender resources are shown in Figure 1. The x-axis represents the covariance value ranging from 0 to -0.8 with the step size of 0.4 while the y-axis shows the average of the defender expected utility when the adversary chooses the worst monotonic strategy. These computed values are averaged across 200 generated payoff structures for each covariance value, and error bars indicate standard deviations. As shown in Figure 1, monotonic maximin obtains a much higher defender's expected utility than Maximin, MATCH, SSE, and logit Quantal Response. In particular, for logit Quantal Response, even though the key parameter $\lambda$ is carefully



(a) 6 Targets, 3 Defender Resources

(b) 8 Targets, 3 Defender Resources

**Fig. 1.** The defender expected utility against the monotonic adversary, exact payoff structures

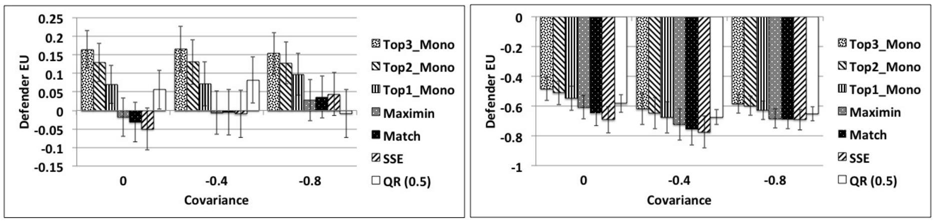
selected, it still performs poorly in comparison with monotonic maximin, implying its non-robustness against a monotonic adversary. For example, in the case of 6 targets and 3 resources (Figure 1a), when the covariance $r = 0$, while the defender achieves an average of expected utility of 0.37 using monotonic maximin, her expected utility is only -0.026, 0.007, 0.029, and 0.099 when using Maximin, MATCH, SSE, and logit Quantal Response, respectively. Furthermore, Figure 1 shows that top-1, top-2, and top-3 also significantly outperform Maximin, MATCH, SSE, and Quantal Response while their performance is in turn closer to monotonic maximin when the number of targets in the top set increases.

As predicted, when the games are zero-sum games (i.e., $r = -1.0$), the defender strategies generated by all algorithms excepts for Quantal Response turn out to be the same. In addition, as shown in Figure 1, when the covariance value $r$ is closer to -1, the differences in defender expected utilities obtained by the compared algorithms tend to be smaller. Indeed, we observed that when the games are close to zero-sum games, the defender strategies generated by the algorithms are similar, thus the difference in their performances becomes small.

The promising performance of our algorithms in the case of small games motivated us to investigate their performance in larger games. In the next experiment, we evaluated the performance of top-1, top-2, and top-3 as well as Maximin, MATCH, SSE, and Quantal Response in 12-target and 14-target games with 6 defender resources. In this experiment, we did not examine monotonic maximin due to its runtime limitation which we will describe in detail later. The results are shown in Figure 2. For each covariance value, 50 payoff structures are generated.

Figure 2 clearly shows that our top-$L$ monotonic maximin algorithms with L = 1, 2, 3 outperform Maximin, MATCH, SSE, and logit Quantal Response in terms of the obtained defender expected utility. For example, in Figure 2b, when the covariance value r = 0, the defender expected utility obtained by top-3, top-2, and top-1 are in turn -0.48, -0.51, and -0.54 while the defender expected utility obtained by Maximin, MATCH, SSE, and logit Quantal Response are -0.61, -0.65, -0.69, and -0.58, respectively. This result demonstrates that our algorithms still perform much better than the other compared algorithms in large game scenarios. It also suggests that our top-$L$ algorithm is a promising approach for handling monotonic adversaries in large games.

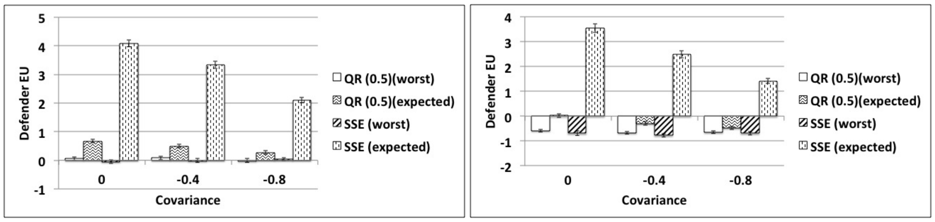
In this paper we have argued that previous algorithms such as SSE and logit Quantal Response are not robust because such algorithms only attempt to address a specific type



(a) 12 Targets, 6 Defender Resources

(b) 14 Targets, 6 Defender Resources

**Fig. 2.** The defender expected utility against the monotonic adversary, large games



(a) 12 Targets, 6 Defender Resources

(b) 14 Targets, 6 Defender Resources

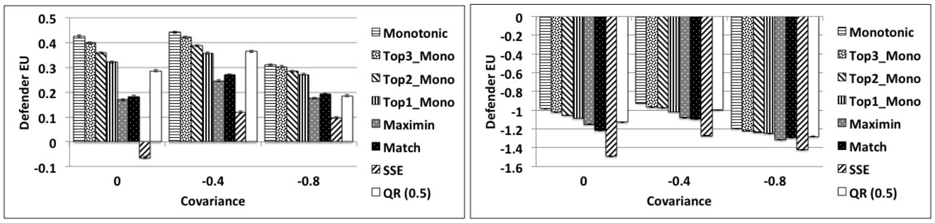
**Fig. 3.** Comparison results between the monotonic adversary and the expected adversary

of adversary which could lead to deterioration in their performance when their assumptions are inaccurate. To check whether this is confirmed by our experiments, for both SSE and logit Quantal Response we compared the defender's expected utility of the expected objective (assuming correct model) and the expected utility against the worst case monotonic adversary. As shown in Figure 3, these two algorithms' performance against the monotonic adversary is significantly worse than the performance that they expected. For example, in Figure 3a, when the covariance $r = 0$, Quantal Response obtains the defender expected utility of only 0.05 against the worst case monotonic adversary while its expected objective value is 0.67. Also, SSE obtains only -0.05 against the worst case monotonic adversary while its expected value is 4.09.

### 6.3 Solution quality against non-monotonic attackers

In the second set of experiments, we compared the solution concepts when the attacker is playing a non-monotonic strategy. The motivation for such experiments is that unlike the setting of our previous experiment, in practice our estimates about the payoffs of the adversary may be inaccurate. Recall from Section 3.3 that monotonicity is generally not preserved under payoff uncertainty, even if the payoff noise is zero-mean and i.i.d.

We added i.i.d. zero-mean noise to the reward and penalty of the adversary at every target, and calculated defender expected utilities given that the adversary responds with the worst-case monotonic strategy in each of the perturbed games. The defender computes her strategy with respect to the non-perturbed game. Specifically, the noise distribution we used is a uniform mixture of 10 Gaussians with zero mean and standard deviation values from 0.01 to 0.10 with a step size of 0.01. We used 6-target and 8-target games with 3 defender resources for evaluating the performance of the compared algorithms, and showed the results according to different covariance values $r$. For each covariance value, 50 payoff structures are generated; for each payoff structure, for each of the 10 standard deviation values, 100 samples of the adversary's payoff noise are drawn from a zero-mean Gaussian distribution with the corresponding standard deviation. That is, 50 x 10 x 100 = 50000 samples are generated for each covariance value. The result is shown in Figure 4, in which we plot the average defender expected utility across all noisy samples. As shown in Figure 4, monotonic maximin still outperforms the other compared algorithms although the differences are smaller than in the previous experiment with exact payoffs for the adversary. In addition, the top-$L$ monotonic maximin algorithms with L = 1, 2, and 3 also obtained higher defender expected utilities than Maximin, MATCH, and SSE. This result indicates that our algorithms are robust to some amount of noise in the adversary's payoffs, even when such noise induces non-monotonic behavior from the adversary.



(a) 6 Targets, 3 Defender Resources

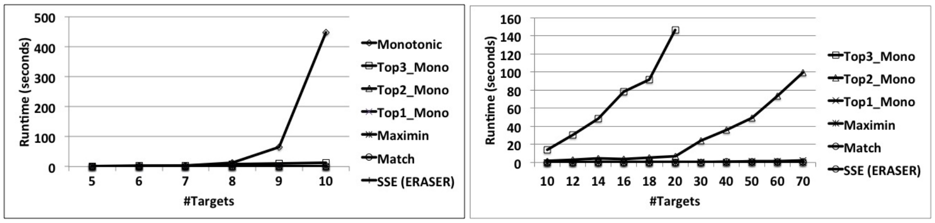
(b) 8 Targets, 3 Defender Resources

**Fig. 4.** The defender expected utility against the monotonic adversary, uncertainty in the adversary's payoffs

On the other hand, as we increase the magnitude (i.e., standard deviation) of noise, monotonic maximin no longer has a significant advantage over the previous solution concepts. Intuitively, as the noise becomes larger, the resulting average strategy of the attacker becomes farther away from the set of monotonic strategies, which are the strategies that monotonic maximin is designed to be robust against.

### 6.4 Runtime performance

Finally, we tested the runtime scaling behavior of the algorithms. The results are shown in Figure 5. Figure 5a shows the runtime comparison between monotonic maximin, top-3, top-2, top-1, Maximin, MATCH, and SSE (as implemented by the ERASER algorithm [9]) in small games, i.e., 5-10 target games with 3 defender resources. The x-axis indicates the number of targets and the y-axis shows the average runtime in seconds for each algorithm to compute the defender's strategy given a payoff structure. The runtime is averaged over 300 different payoff structures. As shown in this figure, monotonic maximin's runtime grows very quickly with regard to the number of targets compared to other algorithms. When the number of targets increases to 10, its runtime



(a) 5-10 Targets, 3 Defender Resources

(b) 10-70 Targets, 6 Defender Resources

**Fig. 5.** Runtime comparison

reaches 446 seconds while top-3, top-2, and top-1 require only 11.91 seconds, 1.53 seconds, and 0.19 seconds, respectively. In this case, it takes Maximin, MATCH, and SSE only 0.02 seconds, 0.1 seconds, and 0.1 seconds, respectively.

In figure 5b, the runtime of top-3, top-2, top-1, Maximin, MATCH, and SSE in large game scenarios (i.e., 10-70 targets and 6 defender resources) are illustrated. This figure shows that when the number of targets is up to 20 targets, the runtime of top-3 increases to 146.49 seconds. In the case of 70 targets, top-2's runtime reaches 99.68 seconds while the runtime of top-1 is about 1.83 seconds, and the runtime of Maximin, MATCH, and SSE are all less than 1 second.

Overall, monotonic maximin and its variants have been shown to outperform Maximin, MATCH, logit QR and SSE in various game settings, i.e., different number of targets and different number of defender resources, and different groups of payoff structures with corresponding covariance values. In terms of scalability, even though our algorithms are not as fast as algorithms for these existing solution concepts, we have shown that our approach (especially top-$L$ monotonic maximin) is feasible for large game scenarios. Among different variants of monotonic maximin, there is a trade-off between solution quality and runtime performance.

## 7 Conclusion and Future Work

We proposed monotonic maximin, a novel robust solution concept for Stackelberg games with boundedly rational followers. We showed both theoretically and through numerical experiments on security games that monotonic maximin provides defender strategies that are robust against all regular quantal response attackers.

Our work points the way to a variety of new research challenges and potential future directions, including extending our robust optimization approach to other behavior models such as risk averseness, as well as applying the solution concept to games with multiple followers.

**Acknowledgments.** This research was supported by MURI Grant W911NF-11-1-0332.

## References

1. Agmon, N., Kraus, S., Kaminka, G.A.: Multi-robot perimeter patrol in adversarial settings. In: ICRA (2008)

2. Basilico, N., Gatti, N., Amigoni, F.: Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In: AAMAS (2009)
3. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust optimization. Princeton University Press (2009)
4. Conitzer, V., Sandholm, T.: Computing the optimal strategy to commit to. In: EC: Proceedings of the ACM Conference on Electronic Commerce (2006)
5. Goeree, J.K., Holt, C.A., Palfrey, T.R.: Regular quantal response equilibrium. Experimental Economics 8(4), 347–367 (2005)
6. Haile, P.A., Hortaçsu, A., Kosenok, G.: On the empirical content of quantal response equilibrium. The American Economic Review pp. 180–200 (2008)
7. Harsanyi, J.: Games with incomplete information played by "Bayesian" players, i-iii. part i. the basic model. Management science 14(3), 159–182 (1967)
8. Kiekintveld, C., Islam, T., Kreinovich, V.: Security games with interval uncertainty. In: AAMAS (2013)
9. Kiekintveld, C., Jain, M., Tsai, J., Pita, J., Ordóñez, F., Tambe, M.: Computing optimal randomized resource allocations for massive security games. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems. pp. 689–696. International Foundation for Autonomous Agents and Multiagent Systems (2009)
10. Korzhyk, D., Conitzer, V., Parr, R.: Complexity of computing optimal stackelberg strategies in security resource allocation games. In: Proc. of The 24th AAAI Conference on Artificial Intelligence. pp. 805–810 (2010)
11. Luce, R.D.: Individual Choice Behavior: A Theoretical Analysis. Wiley (1959)
12. McFadden, D.: Conditional logit analysis of qualitative choice behavior. Frontiers of Econometrics pp. 105–142 (1974)
13. McKelvey, R.D., Palfrey, T.R.: Quantal Response Equilibria for Normal Form Games. Games and Economic Behavior 10(1), 6–38 (1995)
14. Pita, J., Jain, M., Ordóñez, F., Tambe, M., Kraus, S., Magori-Cohen, R.: Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In: Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (2009)
15. Pita, J., John, R., Maheswaran, R., Tambe, M., Kraus, S.: A robust approach to addressing human adversaries in security games. In: ECAI (2012)
16. Tambe, M.: Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned. Cambridge University Press (2011)
17. Wald, A.: Statistical decision functions which minimize the maximum risk. The Annals of Mathematics 46(2), 265–280 (1945)
18. Yang, R., Kiekintveld, C., Ordonez, F., Tambe, M., John, R.: Improving Resource Allocation Strategy Against Human Adversaries in Security Games. In: IJCAI (2011)
19. Yang, R., Ordonez, F., Tambe, M.: Computing optimal strategy against quantal response in security games. In: AAMAS (2012)
20. Yin, Z., Jain, M., Tambe, M., Ordonez, F.: Risk-Averse Strategies for Security Games with Execution and Observational Uncertainty. In: Proc. of The 25th AAAI Conference on Artificial Intelligence (AAAI). pp. 758–763 (2011)
21. Yin, Z., Jiang, A.X., Johnson, M.P., Tambe, M., Kiekintveld, C., Leyton-Brown, K., Sandholm, T., Sullivan, J.: TRUSTS: Scheduling randomized patrols for fare inspection in transit systems. In: IAAI (2012)
22. Yin, Z., Tambe, M.: A unified method for handling discrete and continuous uncertainty in Bayesian Stackelberg games. In: AAMAS (2012)