# Computing Quantal Stackelberg Equilibrium in Extensive-Form Games

**Jakub Černý**[1], **Viliam Lisý**[2], **Branislav Bošanský**[2], **Bo An**[1]
[1] Nanyang Technological University, Singapore
[2] AI Center, FEE, Czech Technical University in Prague, Czech Republic
cerny@disroot.org, {viliam.lisy, branislav.bosansky}@fel.cvut.cz, boan@ntu.edu.sg

## Abstract

Deployments of game-theoretic solution concepts in the real world have highlighted the necessity to consider human opponents' boundedly rational behavior. If subrationality is not addressed, the system can face significant losses in terms of expected utility. While there exist algorithms for computing optimal strategies to commit to when facing subrational decision-makers in one-shot interactions, these algorithms cannot be generalized for solving sequential scenarios because of the inherent curse of strategy-space dimensionality in sequential games and because humans act subrationally in each decision point separately. We study optimal strategies to commit to against subrational opponents in sequential games for the first time and make the following key contributions: (1) we prove the problem is NP-hard in general; (2) to enable further analysis, we introduce a non-fractional reformulation of the direct non-concave representation of the equilibrium; (3) we identify conditions under which the problem can be approximated in polynomial time in the size of the representation; (4) we show how an MILP can approximate the reformulation with a guaranteed bounded error, and (5) we experimentally demonstrate that our algorithm provides higher quality results several orders of magnitude faster than a baseline method for general non-linear optimization.

## Introduction

In recent years, game theory has achieved many groundbreaking advances both in large games (e.g., super-human AI in poker (Moravčík et al. 2017; Brown and Sandholm 2018)), and practical applications (physical security (Sinha et al. 2018), wildlife protection (Fang et al. 2017), AI for social good (Yadav et al. 2016)). In deployed two-player real-world solutions, the aim is often to compute an optimal strategy to commit to while assuming that the other player plays the best response to this commitment – i.e., to find a *Stackelberg Equilibrium* (SE). While the traditional theory (e.g., the SE) assumes that all players behave entirely rationally, the real world's deployments proved that taking into account the bounded rationality of the human players is necessary to provide higher quality solutions (An et al. 2013; Fang et al. 2017). One of the most commonly used models of bounded rationality is *Quantal Response* (QR), which assumes that players choose better actions more often than worse actions instead of playing only the best action (McKelvey and Palfrey 1995). When facing a human opponent we aim to find a strategy optimal against QR, termed *Quantal Stackelberg Equilibrium* (QSE). Relying on SE is not advisable in such situations as the committing player can suffer huge losses when deploying rational strategies against boundedly rational counterparts (Yang et al. 2011).

QSE was studied extensively in Stackelberg Security Games (SSGs) (Sinha et al. 2018), and the developed algorithms were used in many real-world applications. SSG is a widely applicable game model, but it requires formulating the problem in terms of allocating limited resources to a set of targets. To solve real-world problems beyond SSGs, QSE was recently introduced also in a more general model of normal-form games (NFGs) (Černý et al. 2020). However, even NFGs model only one-shot interactions. To the best of our knowledge, QSE was never properly analyzed for extensive-form (i.e., sequential) games (EFGs), despite the fact that many real-world domains are sequential.

In this paper, we develop methods for finding QSE in EFGs. There are two main challenges that prevent us from using techniques from SSGs or NFGs directly. First, while there is only one decision point for the boundedly rational player in NFGs, any non-trivial EFG contains multiple causally linked decision points where the same player acts. The psychological studies show that humans prefer short-term, delayed heuristic decisions (Gigerenzer and Goldstein 1996) rather than long-term, premeditated decisions. This behavior arises especially in conflicts (Gray 1999) or when facing information overload caused by large decision space (Malhotra 1982). Therefore, a natural assumption is that a subrational player acts according to a QR model in each decision point separately. This behavior cannot be modeled by an equivalent NFG. Second, contrary to NFGs, the number of rational player's pure strategies in EFGs is exponential in the game size. Therefore, even if the QSE concepts would coincide in NFGs and EFGs, applying the algorithms for NFGs directly would scale much worse. We begin our analysis by showing that finding QSE is NP-hard, and the straightforward formulation of QSE in EFGs is a non-concave fractional problem that is difficult to optimize. Therefore, we derive an equivalent Dinkelbach-type formulation of QSE that does not contain any fraction, and rep-

resents the rational player's strategy linearly in the size of the game. We use the Dinkelbach formulation to identify sufficient condition for solving the problem in polynomial time. If the conditions are satisfied, the optimal solution can be found by gradient ascent. For other cases, we formulate a mixed-integer linear program (MILP) approximating the QSE through QR model's linear relaxation. We provide theoretical guarantees on the solution quality, depending on the number of segments used to linearize the QR function and the arising bilinear terms. Full proofs and additional examples can be found in the appendix (Černý et al. 2021).

In the experiments, we compare the direct formulation solved by an algorithm for general non-linear optimization to our MILP reformulation. We show that in 3.5 hours, our algorithm computes solutions that the baseline cannot reach within three days. Moreover, for solvable instances the solutions of our algorithm outperform the baseline's solutions.

**Related Solution Concepts.** Several other solution concepts study bounded rationality in EFGs. Perhaps the most well-known one is the model of Quantal Response Equilibrium (QRE) (McKelvey and Palfrey 1998). In QRE, all players act boundedly rationally, and no player has the ability to commit to a strategy. This is in contrast to QSE, where the entirely-rational leader aims to exploit the boundedly rational opponent. A similar approach for EFGs was taken in (Basak et al. 2018), where only one of the players was considered rational. However, their goal was to evaluate the performance against human participants, and the approach lacked theoretical foundations. The computed strategies did not correspond to a well-defined solution concept and were empirically suboptimal. We define and study this concept in our concurrent paper (Milec et al. 2020). Besides rationality, in QSE, the leader also benefits from the power of commitment. It is well known that the ability to commit increases the payoffs achievable by the leader (Von Stengel and Zamir 2010) and it trivially holds also against bounded rational players. Mathematically, QRE is a fixed-point, while QSE is a global optimization problem. Hence, the techniques for computing QRE cannot be applied for finding QSE.

## Background

Extensive-form games model sequential interactions between players and can be visually represented as game trees. Formally, a two-player Stackelberg EFG is defined as a tuple $G = (\mathcal{N}, \mathcal{H}, \mathcal{Z}, \mathcal{A}, u, \mathcal{C}, \mathcal{I})$: $\mathcal{N}$ is a set of 2 players: a *leader (l)* and a *follower (f)*. We use $i$ to refer to one of the players, and $-i$ to refer to his opponent. $\mathcal{H}$ denotes a finite set of *histories* in the game tree, with $h_0 \in \mathcal{H}$ being the root. $\mathcal{H}_i$ denotes the set of histories in which player $i$ acts. We use $pr(h) = (h', a)$ to identify a pair of immediately preceding $h'$ and action $a$ connecting $h'$ with $h$. For $h_0$ we set $pr(h_0) = (\emptyset, \emptyset)$. $\mathcal{A}$ denotes the set of all actions, with $\mathcal{A}_i$ denoting the actions of player $i$. $\mathcal{Z} \subseteq \mathcal{H}$ is the set of all *terminal histories* of the game. We define for each player $i$ a *utility function* $u_i : \mathcal{Z} \to \mathbb{R}$. The chance player selects actions based on a fixed probability distribution known to all players. Function $\mathcal{C} : \mathcal{A} \to [0, 1]$ denotes the probability of taking an action by chance; we write $\mathcal{C}(h)$ for the product of the probabilities of chance actions in history $h$. The set of chance histories is denoted as $\mathcal{H}_c$.

Imperfect observation of player $i$ is modeled via *information sets* $\mathcal{I}_i$ that form a partition over $h \in \mathcal{H}_i$. Player $i$ cannot distinguish between histories in any information set $I \in \mathcal{I}_i$. We overload the notation and use $\mathcal{A}(I_i)$ to denote possible actions available in each history in an information set $I_i$. Each action $a$ uniquely identifies the information set where it is available, referred as $\mathcal{I}(a)$. We assume *perfect recall*, which means that players remember the history of their own actions and all information gained during the course of the game. As a consequence, all histories in any information set $I_i$ have the same history of actions for player $i$.

*Pure strategies* $\Pi_i$ assign one action for each $I \in \mathcal{I}_i$ and we assume the actions of a strategy $\pi \in \Pi_i$ can be enumerated as $a \in \pi$. We denote the set of leafs reachable by a strategy $\pi$ as $\mathcal{Z}(\pi)$. A *behavioral strategy* $\beta_i \in B_i$ is one probability distribution over actions $\mathcal{A}(I)$ for each $I \in \mathcal{I}_i$. For any pair of strategies $\beta \in B = (B_l, B_f)$ we use $u_i(\beta) = u_i(\beta_i, \beta_{-i})$ for the expected outcome of the game for player $i$ when players follow strategies $\beta$.

Strategies in EFGs with perfect recall can be also compactly represented by using the sequence form (Koller, Megiddo, and von Stengel 1996). A *sequence* $\sigma_i \in \Sigma_i$ is an ordered list of actions taken by a single player $i$ in history $h$. $\emptyset$ stands for the empty sequence (i.e., a sequence with no actions). A sequence $\sigma_i \in \Sigma_i$ can be extended by a valid action $a$ taken by player $i$, written as $\sigma_i a = \sigma_i'$. We say that $\sigma_i$ is a *prefix* of $\sigma_i'$ ($\sigma_i \sqsubseteq \sigma_i'$) if $\sigma_i'$ is obtained by finite number (possibly zero) of extensions of $\sigma_i$. We use $seq(h)$ to denote the sequence of actions of all players leading to $h$. By using a subscript $seq_i$ we refer to the subsequence of action of player $i$. Because of perfect recall we write $seq_i(I) = seq_i(h)$ for $I \in \mathcal{I}_i, h \in I$. We use the function $inf_i(\sigma_i')$ to denote the information set in which the last action of the sequence $\sigma_i'$ is taken. For an empty sequence, function $inf_i(\emptyset)$ returns the information set of the root. Using sequences, any behavioral strategy of a player can be represented as a *realization plan* ($r_i : \Sigma_i \to \mathbb{R}$). A realization plan for a sequence $\sigma_i$ is the probability that player $i$ will play $\sigma_i$ under the assumption that the opponents play to allow the actions specified in $\sigma_i$ to be played. In other words, for a behavioral strategy $\beta_i$, $r_i(\sigma) = \prod_{a \in \sigma} \beta_i(a)$. The set of valid realization plans for player $i$ can be represented using a set of linear network-flow constraints of a size linear in the number of information sets:

$$r_i(\emptyset) = 1, \quad 0 \le r_i(\sigma) \le 1 \qquad \forall \sigma \in \Sigma_i$$
$$r_i(seq_i(I)) = \sum_{a \in \mathcal{A}(I)} r_i(seq_i(I)a) \qquad \forall I \in \mathcal{I}_i. \quad (1)$$

For a given history $h$, we shorten the notation and write $r_i(seq_i(h))$ as $r_i(h)$.

## Quantal Stackelberg Equilibrium in EFGs

We follow the earlier works on boundedly rational Stackelberg equilibria and consider two possible causes of the emergence of subrational behavior that combine into a formal definition of quantal response: (i) a subjective perception of action values and (ii) a proneness to making mistakes

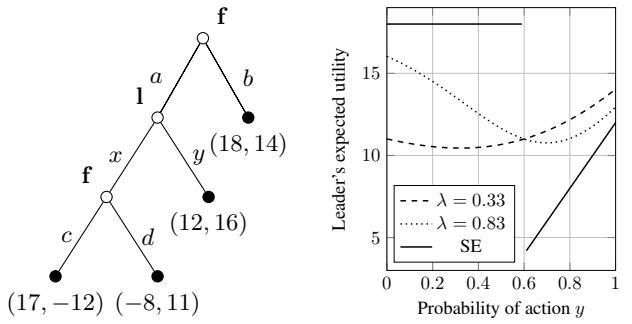


Figure 1: (Left) An example of a general-sum EFG with utilities in form $(u_l, u_f)$, and (Right) objective functions of three equilibria: QSEs with a generator $q = exp(\lambda x)$, where $\lambda \in \{0.33, 0.83\}$, and SE.

choosing an action to play. The first source of subrationality is the (possibly) imperfect evaluation of the follower's own choices. We assume an evaluation model in the form of a function $e : \mathcal{A}_f \times B_l \to \mathbb{R}$ that describes how the follower values an action given the leader's strategy. An entirely rational player uses expected utility of the best response as the evaluation function. For boundedly rational players, many examples of evaluation functions can be found in the literature: among others, the subjective utility (Nguyen et al. 2013), past-experience-based learning (Lejarraga, Dutt, and Gonzalez 2012), risk assessment (Yechiam and Hochman 2013; Kahneman and Tversky 2013), preference in simple strategies (Černý, Bošanský, and An 2020) or reasoning levels in hierarchical models (Wright and Leyton-Brown 2020). The second cause of subrationality relates to the ability of the player to pick a correct action. Entirely rational players always select the utility-maximizing option. Relaxing this assumption leads to a "statistical version" of the best response, which considers the inevitable error-proneness of humans and allows the players to make systematic errors (McFadden 1976).

**Definition 1.** *Let $e$ be a follower's evaluation function. Function $QR : B_l \to B_f$ is a canonical quantal response function if for each $I \in \mathcal{I}_f$*

$$QR(\beta_l) = \left( \frac{q(e(a, \beta_l))}{\sum_{a' \in \mathcal{A}(I)} q(e(a', \beta_l))} \right)_{a \in \mathcal{A}(I), I \in \mathcal{I}_f} \tag{2}$$

*for all $\beta_l \in B_l$ and some real-valued function $q$.*

Note that whenever $q$ is a strictly positive increasing function, the corresponding $QR$ is a valid quantal response function, which gives rise to a valid behavioral strategy of the follower. We call such functions $q$ *generators* of canonical quantal functions.

**Definition 2.** *Given an extensive-form game $G$, a behavioral strategy $\beta_l^{QS} \in B_l$ and a quantal response function $QR$ of the follower form a* Quantal Stackelberg Equilibrium *(QSE) if and only if*

$$\beta_l^{QS} = \arg\max_{\beta_l \in B_l} u_l(\beta_l, QR(\beta_l)). \tag{3}$$

QSE of an EFG is not equivalent to QSE of a normal-form representation of the EFG, because instead of picking a pure strategy in the whole game according to a given model of bounded rationality, we assume a more natural setting, when a player acts quantally in every information set they encounter separately.

**Example 1.** *Consider an EFG depicted in Figure 1 and two possible behavioral models of the follower. In both models, the follower assumes they act fully rationally and uses an expected utility of a best response as their evaluation function in both information sets. However, they are unaware that when choosing an action to play, they will act quantally according to a generator $q_U(x) = exp(\lambda x)$. In the first model, we set $\lambda = 0.33$, while in the second model $\lambda$ is equal to $0.83$. On the right of the same figure we can find the non-concave objective functions of both QSEs, and SE. The choice of the behavioral model significantly affects the solution. While with $\lambda = 0.33$ the leader commits to playing action $y$, with $\lambda = 0.83$ her strategy is completely opposite: to play the action $x$ with probability 1. Furthermore, an optimal solution in SE is any strategy with probability of action $y$ lower than 0.6. However, if the leader deploys a strategy close to this threshold against either of the two behavioral models, her utility will be, in fact, close to global minimum of the corresponding QSE. For $\lambda = 0.33$, the utility is low for all SE strategies. An example with a fixed quantal function and different evaluation functions is in the appendix.*

The example verifies the observations made in SSGs: playing SE strategies against boundedly rational opponents may inflict huge losses in utility for the leader (Yang et al. 2011). It is not difficult to design an EFG in which a unique SE is a global minimum of a QSE with arbitrarily low utility. The following straightforward mathematical program represents QSE using the direct definition of quantal response from Eq. (2):

$$\max_{\beta_l \in B_l} v(\emptyset, \emptyset) \tag{4a}$$

$$v(pr(h)) = \sum_{a \in \mathcal{A}(h)} v(h, a) C(a) \qquad \forall h \in \mathcal{H}_c \tag{4b}$$

$$v(pr(h)) = \sum_{a \in \mathcal{A}(h)} v(h, a) \beta_l(a) \qquad \forall h \in \mathcal{H}_l \tag{4c}$$

$$v(pr(h)) = \frac{\sum_{a \in \mathcal{A}(h)} v(h, a) q(e(a, \beta_l))}{\sum_{a \in \mathcal{A}(h)} q(e(a, \beta_l))} \qquad \forall h \in \mathcal{H}_f \tag{4d}$$

$$v(pr(z)) = u_l(z) \qquad \forall z \in \mathcal{Z}. \tag{4e}$$

The variable $v$ is defined for every action interconnecting two consecutive nodes in the game tree (i.e., an edge) and it serves to propagate the leader's utility from the leafs up to the root through both the chance nodes and nodes of the players. As Example 1 shows, this formulation using behavioral strategies is non-concave, might have multiple local optima, and it contains fractional terms (Eq. (4d)). Unfortunately, similarly to SE (Letchford and Conitzer 2010), also computing QSE in EFGs is an NP-hard problem.

**Theorem 1.** *Let $q$ be an exponential generator of a quantal function of the follower. Computing optimal strategy of a rational player against the quantal response opponent in two-player imperfect-information EFGs with perfect recall is an NP-hard problem.*

*Proof (Sketch).* We reduce from the 3-SAT problem using the tree structure from (Letchford and Conitzer 2010). We adapt the utilities because while the follower plays the leader's preferred action in case of indifference in SE, the quantal response chooses all actions with the same utility uniformly. We show there exists a threshold probability of reaching a target subtree separating SAT from UNSAT. Full proof is in the appendix. □

## Dinkelbach-Type Formulation of QSE

The non-concavity of formulation (4) makes it difficult to optimize and guarantee global optimality. Therefore, we search for an alternative representation of QSE that would express the problem as a single fractional criterion, instead of a set of equations. Such representation would allow us to leverage reformulation methods from fractional programming to eliminate the fraction. For this purpose we use the realization plans. Note that in (4), the utility from each leaf (Eq. (4e)) is propagated up through the variables $v$ by multiplying it by the values of a behavioral strategy (Eqs. (4c) and (4d)) and chance (Eq. (4b)) of all actions on the way to the root. Because the product of behavioral probabilities of a sequence is (by definition) equivalent to the realization of the sequence, the criterion (4a) is expressed as

$$\max_{r_l} \sum_{z \in \mathcal{Z}} u_l(z) C(z) r_l(z) QR(r_l, z), \tag{5}$$

where

$$QR(r_l, z) = \prod_{a \in seq_f(z)} \frac{q(e(a, r_l))}{\sum_{a' \in \mathcal{A}(\mathcal{I}(a)} q(e(a', r_l))}.$$

Because realization plans are equivalent to behavioral strategies, instead of $e(a, \beta_l)$ we write $e(a, r_l)$. The purpose of this reformulation is that the criterion (3) can now be expressed as a single fraction, similarly to the representation of QSE in mixed strategies in security games (Yang, Ordonez, and Tambe 2012) and normal-form games (Černý et al. 2020).We can hence use the Dinkelbach's method for solving nonlinear fractional programming problems (Dinkelbach 1967) and adapt it for finding the QSE. The key idea of the Dinkelbach's algorithm is to express a problem in a form of $\max_{x \in M} f(x)/g(x), g(x) > 0$ for some convex set $M$ and continuous, real-valued functions $f$ and $g$ as an equivalent problem of finding a unique root of function $\mathcal{D}(p) = \max_{x \in M} f(x) - pg(x)$, $p \in \mathbb{R}$. It holds that $\max_{x \in M} f(x)/g(x) = p^*$ if and only if $\mathcal{D}(p^*) = 0$. Because $\mathcal{D}$ is a maximum of functions affine in $p$, it is convex. Finding the root is hence straightforward (e.g., a binary search method can be used for this purpose) and it can be done efficiently if and only if we are able to effectively determinate the value of function $\mathcal{D}$ for any $p$. Most importantly, $D$ no longer contains any fractional term and is therefore easier to optimize.

**Theorem 2.** *Computing the maximum of the original formulation of QSE (3) is equivalent to finding a unique root of the following function $\mathcal{D}$:*

$$\mathcal{D}(p) = \max_{r_l} F(r_l, p), \tag{6}$$

*where*

$$F(r_l, p) = \sum_{\pi \in \Pi_f} \prod_{a \in \pi} q(e(a, r_l)) \Big( \sum_{z \in \mathcal{Z}(\pi)} u_l(z) C(z) r_l(z) - p \Big).$$

*Proof.* We start from the representation of QSE in form of Eq. (5). Because the follower acts quantally in every information set, the smallest common multiple over all leafs is $\prod_{I \in \mathcal{I}_f} \sum_{a \in \mathcal{A}(I)} q(e(a, r_l))$. The fractional representation of the QSE is hence

$$\max_{r_l} \frac{\sum_{z \in \mathcal{Z}} u_l(z) C(z) r_l(z) Q(z, r_l)}{\prod_{I \in \mathcal{I}_f} \sum_{a \in \mathcal{A}(I)} q(e(a, r_l))},$$

where

$$Q(z, r_l) = \prod_{a \in seq_f(z)} q(e(a, r_l)) \prod_{\substack{I \in \mathcal{I}_f \\ seq(I) \not\sqsubseteq seq(z)}} \sum_{a \in \mathcal{A}(I)} q(e(a, r_l)).$$

Because $Q(z, r_l)$ is a sum of products of functions $q$ applied to a fixed path from root to $z$ and one action in each information set outside this path, it iterates over pure strategies enabling to reach $z$. $Q(z, r_l)$ is therefore equivalent to $\sum_{\pi \in \Pi_f : z \in \mathcal{Z}(\pi)} \prod_{a \in \pi} q(e(a, r_l))$. Applying the same idea also for the denominator and swapping the sum over leafs with the sum over pure strategies in the nominator we obtain

$$\max_{r_l} \frac{\sum_{\pi \in \Pi_f} \prod_{a \in \pi} q(e(a, r_l)) \sum_{z \in \mathcal{Z}(\pi)} u_l(z) C(z) r_l(seq_l(z))}{\sum_{\pi \in \Pi_f} \prod_{a \in \pi} q(e(a, r_l))}. \tag{7}$$

By the Dinkelbach reformulation, maximizing this equation is equivalent to finding a root of

$$\max_{r_l} \sum_{\pi \in \Pi_f} \prod_{a \in \pi} q(e(a, r_l)) \sum_{z \in \mathcal{Z}(\pi)} u_l(z) C(z) r_l(z) - p \sum_{\pi \in \Pi_f} \prod_{a \in \pi} q(e(a, r_l)),$$

which is the desired equation. □

Due to the leader's strategy being represented using a realization plan, the formulation (6) has $|\Sigma_l|$ variables. The expression $e(a, r_l)$ is evaluated for every follower's action in the game tree, the number of evaluations is hence also linear in $|\mathcal{I}_f|$. The outer sum, however, enumerates the follower's pure strategies and is thus exponential in $|\mathcal{I}_f|$. In many real-world applications this fact might not be critical, as the follower's strategy space (i.e., choosing a target to attack) is often much smaller than a combinatorial strategy space of the leader (i.e., deploying/moving multiple units to different locations).

**Algorithm 1:** Dinkelbach-Type Algorithm for QSE

$UB \leftarrow \max_{z\in\mathcal{Z}} u_l(z)$, $LB \leftarrow \min_{z\in\mathcal{Z}} u_l(z)$
$r_l^* \leftarrow \arg\max_{r_l} F(r_l, LB)$
**repeat**
  $p \leftarrow (UB - LB)/2$
  $v \leftarrow \max_{r_l} F(r_l, p)$
  $r_l^p \leftarrow \arg\max_{r_l} F(r_l, p)$
  **if** $v < 0$ **then** $LB \leftarrow p, r_l^* \leftarrow r_l^p$ **else** $UB \leftarrow p$
**until** $UB - LB < \epsilon$
**return** $r_l^*$

Because function $D$ is convex, its root can be found using a binary search method, as described in Algorithm 1. We refer to formulation (6) as to the Dinkelbach subproblem of the Dinkelbach formulation of QSE. Algorithm 1 iteratively updates the upper bound ($UB$) and lower bound ($LB$) on the value of QSE according to a binary search method for finding a root of a function. For running the binary search it is essential to solve the Dinkelbach subproblems (i.e., evaluate $\mathcal{D}(p)$ for any $p$). The following proposition presents conditions under which the subproblem can be efficiently approximated.

**Proposition 3.** *Let $q$ be a twice differentiable generator of a quantal response and $e$ be twice differentiable evaluation function of the follower. The Dinkelbach subproblem for $p \in [\min_{z\in\mathcal{Z}} u_l(z), \max_{z\in\mathcal{Z}} u_l(z)]$ is concave if for any $\pi \in \Pi_f, a \in \pi$ and realization plan $r_l$*

$$\delta(\pi,a) = q'(e_a)(u_{r_l} e'^T_a + e'_a u^T_{r_l}) + (u^T_{r_l} r_l - p)\Big(e''_a q'(e_a) + \\ + e'^2_a q''(e_a) + \sum_{a'\neq a\in\pi} e'_a e'_{a'} q'(e_a) q'(e_{a'}) \prod_{a''\neq a'\neq a\in\pi} q(e_{a''})\Big)$$

*is negative semidefinite, where $e_a = e(a, r_l)$, $e'_a = e'(a, r_l)$, $e''_a = e''(a, r_l)$, and $u^T_{r_l} r_l = \sum_{z\in\mathcal{Z}(\pi)} u_l(z) C(z) r_l(z)$.*

*Proof.* The formulation of the subproblem from Eq. (6) is concave when its Hessian matrix is negative semidefinite (NSD). The Hessian matrix is of a form $\sum_{\pi\in\Pi_f} \sum_{a\in\pi} \delta(\pi, a) \prod_{a'\neq a\in\pi} q(e_{a'})$. Because a sum of NSD matrices is NSD, the generator is always positive and the definiteness is preserved under multiplication by a positive number, Eq. (6) is concave if $\delta(\pi, a)$ is NSD. □

In case the conditions are met, local-optimization algorithms (e.g., projected gradient ascent (Nesterov 2004), given Eq. (6) is L-smooth) are guaranteed to reach optimum. We discuss how useful Proposition 3 is in the appendix.

## Approximating the Dinkelbach Subproblem

In case a game does not satisfy the conditions in Proposition 3, the guarantee of convergence is lost. A solution commonly suggested in the literature is then to linearize the criterion (6) and transform the problem into an (MI)LP that can be solved using standard methods. For linearizing the criterion we need to approximate both functions of the behavioral model from Definition 1: (i) the quantal function $q$ and (ii) the utility evaluation function $e$.

We begin by approximating the quantal generator $q$. We focus on logit QR, which is the most commonly studied quantal response in the literature. In case of logit QR, function $q$ is defined as $q(x) = exp(\lambda x), \lambda \in \mathbb{R}^+$. The player becomes more rational as $\lambda$ approaches infinity. We can express the product of generator functions from Eq. (6) through a substitutional variable $x_\pi$ as

$$\prod_{a\in\pi} exp(\lambda e(a, r_l)) = exp(\lambda \sum_{a\in\pi} e(a, r_l)) \to x_\pi.$$

The $exp$ function is linearizable into K segments as

$$\overline{exp}(\lambda \sum_{a\in\pi} e(a, r_l)) = \sum_{k=0}^{K} \alpha^k t^k_\pi + exp(\underline{e}) \to \overline{x}_\pi, \qquad (8)$$

where $(\alpha^k)_{k\in[K]}, \alpha^k \in \mathbb{R}$ is a slope of the $k$-th segment, subjected to constraints

$$\begin{aligned} \sum_{k=0}^{K} t^k_\pi + \underline{e} &= \sum_{a\in\pi} e(a, r_l) \\ t^k_\pi &\le z^k_\pi(\overline{e} - \underline{e})/K \\ t^{k+1}_\pi &\ge z^k_\pi(\overline{e} - \underline{e})/K \\ 0 \le t^k_\pi &\le (\overline{e} - \underline{e})/K, \quad z^k_\pi \in \{0, 1\}, \end{aligned} \qquad (9)$$

where binary variables $z$ indicate whether the linear segment is used and real variables $t$ define what portion of the segment is active. We set $\underline{e} = \lambda|\mathcal{I}_f| \min_{a\in\mathcal{A}_f, r_l} e(a, r_l)$ and $\overline{e} = \lambda|\mathcal{I}_f| \max_{a\in\mathcal{A}_f, r_l} e(a, r_l)$. Using the bound from (Yano et al. 2013), the maximum difference in values of $exp$ and its linearization $\overline{exp}$ using K segments on interval $[\underline{e}, \overline{e}]$ can be bounded as

$$|exp(x) - \overline{exp}(x)| \le exp(\overline{e}) \frac{(\overline{e} - \underline{e})^2}{8K^2}, \quad x \in [\underline{e}, \overline{e}]. \qquad (10)$$

With linearized logit generator, the Dinkelbach subproblem (6) is expressed as

$$\mathcal{D}(p) = \max_{r_l} \left( \sum_{z\in\mathcal{Z}(\pi)} u_l(z) C(z) r_l(z) - p \right) \overline{x}_\pi.$$

Clearly, the criterion contains multiple bilinear terms $\overline{x}_\pi r_l(z)$. For linearizing the bilinear terms, we use the MDT technique (Kolodziej, Castro, and Grossmann 2013). MDT is a parametrizable method which enables controlling the error in exchange of introducing binary variables. The product $c(\pi, z) = \overline{x}_\pi r_l(z)$ is expressed using linear equations

$$\begin{aligned} c(\pi, z) = \sum_{i=0}^{b-1} \sum_{j\in\mathcal{E}} i b^j r_{i,j}, \quad \overline{x}^{\mathcal{E}}_\pi = \sum_{i=0}^{b-1} \sum_{j\in\mathcal{E}} i b^j s_{i,j} \\ 1 = \sum_{i=0}^{b-1} s_{i,j}, \quad r_l(z) = \sum_{i=0}^{b-1} r_{i,j} \qquad \forall j \in \mathcal{I} \\ s_{i,j} \in \{0, 1\}, 0 \le r_{i,j} \le s_{i,j} \forall j \in \mathcal{E}, \forall i \in [b-1], \end{aligned} \qquad (11)$$

where $\mathcal{E} \subset \mathbb{Z}$ is a finite subset controlling the error of the approximation with basis $b$. $\overline{x}^{\mathcal{E}}_\pi$ is a representation of $\overline{x}_\pi$ in MDT over $\mathcal{E}$. The following lemma identifies the approximation error introduced by the selection of $K$ and $\mathcal{E}$.

**Lemma 4.** *Let $|\mathcal{E}|$=L and $x_\pi$ be linearized with K segments.*

$$|x_\pi r_l(z) - \overline{x}^{\mathcal{E}}_\pi r_l(z)| \leq \epsilon_K + \epsilon_\mathcal{E}, \tag{12}$$

*where $\epsilon_K = exp(\overline{e})\frac{(\overline{e}-\underline{e})^2}{8K^2}$, $\epsilon_\mathcal{E} = \max(N, exp(\overline{e}) - b^{M+1} + N, N - exp(\underline{e}))$, $M = \lfloor \log_b(exp(\overline{e})) \rfloor$ and $N = b^{M-L+1}$.*

*Proof.* Let $\mathcal{E} = \{M\text{–}L\text{+}1, M\text{–}L\text{+}2, \ldots, M\}$. $\mathcal{E}$ hence defines a discretization of variable $\overline{x}_\pi$ on interval $[N, b^{M+1} - N]$ with a step of size $N$. Because $\overline{x}_\pi$ is defined on interval $[exp(\underline{e}), exp(\overline{e})]$, the maximum difference between $\overline{x}_\pi$ and $\overline{x}^{\mathcal{E}}_\pi$ is $\epsilon_\mathcal{E}$. As $r_l(z)$ is (by definition) always at most 1, we have

$$|x_\pi r_l(z) - \overline{x}^{\mathcal{E}}_\pi r_l(z)| \leq |x_\pi - \overline{x}_\pi| + |\overline{x}_\pi - \overline{x}^{\mathcal{E}}_\pi|.$$

By Eq. (10), $|x_\pi - \overline{x}_\pi| \leq \epsilon_K$, concluding the proof. □

Now we move to the second part of the QR definition: the evaluation function $e$. We present a domain-independent formulation of a common situation when the follower is not aware of their subrationality and evaluate the actions in the current information set on the basis of acting rationally in the subsequent information sets, weighted by the probability of reaching the current set. In that case, the evaluation function $e$ can be expressed as

$$\begin{aligned} e(a, r) &= v(\mathcal{I}(a)) - s(seq_f(\mathcal{I}(a))a) \\ v(inf_f(\sigma_f)) &= s(\sigma_f) + \sum_{I:seq_f(I)=\sigma_f} v(I) + \\ &\quad + \sum_{z:seq_f(z)=\sigma_f} u_l(z)C(z)r_l(z) \qquad \forall \sigma_f \in \Sigma_f \\ 0 \leq s(\sigma) &\leq M(1 - r_f(\sigma)) \qquad \forall \sigma \in \Sigma_f, \end{aligned} \tag{13}$$

where $r_f$ is the binary best-response realization plan of the follower, $v$ is the optimal expected utility contribution in an information set and $s$ is a slack variable compensating the deficiency in action's suboptimal utility. Now, we can finally state the approximation error for computing the QSE.

**Proposition 5.** *Consider a linear formulation of the Dinkelbach subproblem in form*

$$\mathcal{D}(p) = \max_{r_l} \sum_{z \in \mathcal{Z}(\pi)} u_l(z)C(z)c(\pi, z) - p\overline{x}^{\mathcal{E}}_\pi,$$

*with constraints* (1)*,* (9)*,* (11)*, and* (13) *with K segments, $|\mathcal{E}| = L$ and substitution* (8)*. Let $r^*_l$ be a realization plan computed by Algorithm 1 with precision $\epsilon_B$, and $r^{QS}_l$ be a realization plan of the leader in QSE. Then for the utility difference $d = |u_l(r^*_l, QR(r^*_l)) - u_l(r^{QS}_l, QR(r^{QS}_l))|$ it holds*

$$d \leq \epsilon_B + \frac{\overline{u}_l |\Pi_f| \epsilon_K + |\Pi_f| \max_{z \in \mathcal{Z}} |u_l(z)| (\epsilon_K + \epsilon_\mathcal{E})}{exp(\underline{e})},$$

*where $\epsilon_K$ and $\epsilon_\mathcal{E}$ are defined as in Lemma 4.*

*Proof (Sketch).* We derive specific bounds for the Dinkelbach representation of QSE based on Lemma 4 and combine them with bounds on the difference between linearized and non-linearized nominator and denominator introduced in (Černý et al. 2020), adapted for QSE in form of Eq. (7). Full proof can be found in the appendix. □

# Experimental Evaluation

We compare the Dinkelbach-type algorithm (DTA) to the standard benchmark for nonlinear optimization: the COBYLA algorithm (Powell 2007), implemented in the open-source NLOPT library. COBYLA is a gradient-free algorithm capable of handling linear equality constraints induced by realization plans. We opted for COBYLA because the follower's evaluation function is non-differentiable, possibly in infinitely many points. We apply COBYLA directly to formulation (5). For evaluating the algorithms, we used two domains: a variant of *search game* commonly used to evaluate algorithms for SE (Marchesi et al. 2019; Kroer, Farina, and Sandholm 2018; Čermák et al. 2016) and a *network game*, handcrafted to be difficult for QSE.

**Search Game.** The game is played on a directed graph, depicted in the middle of Figure 2. The attacker's goal is to reach one of the destination nodes (D0 – D7) from the starting node (S), while the defender aims to catch the attacker with one of the two units operating in the marked areas of the graph (P0 and P1). The attacker receives a different reward for reaching a different destination node (the reward is selected randomly from interval $[0, 2]$). While the defender can move freely with unit P1, unit P0 is static – placed by the defender at the beginning of the game. If the attacker evades unit P0, the defender is given $N$ steps to set unit P1. The defender receives a signal if P1 is within 1 step from the attacker. In case the defender captures the attacker, she gets a positive reward of $1 - n/(N + 1)$, where $n \leq N$ is a number of taken steps, and the attacker receives 0. We consider a version of the game in which the attacker perceives no information about the whereabouts of the defender's units, and unit P1 starts above $D_0$.

**Network Game.** The network game is played on a directed graph too. Some nodes in the graph are grouped together into mutually disjunctive areas. In the beginning, the attacker observes their areas of possible infiltration and the area the defender uses to enter the network, and selects one area for further probing. The defender then picks a node from the entering area, while the attacker chooses a node from their area to compromise. The defender is given a maximum number of steps $N$ to survey the network and discover the attacker. There is binary information: if the attacker is located within $\gamma$ steps from the defender, the defender observes it. If she captures the attacker in $n \leq N$ steps she receives a utility $1 - n/(N + 1)$, 0 otherwise. The attacker is given a reward associated with the compromised node if not found (the reward is chosen randomly from interval $[-2.5, 2.5]$–the negative utility represents nodes with compromising costs higher than the data's price or a possible honeypot), -3 otherwise. We designed three networks, shown on the left in Figure 2. Attacker's areas are depicted in thin-lined rectangles. The defender's entering area is depicted in a thick-lined rectangle in game 03 and selected randomly (4 nodes per seed) in other games. We set $\gamma = 0$ for game 01, 1 otherwise. It is a type of coordination game; hence, the vast majority of the attacker's strategies can be best responses to the defender's strategy, which makes computing Stackelberg strategies particularly difficult.

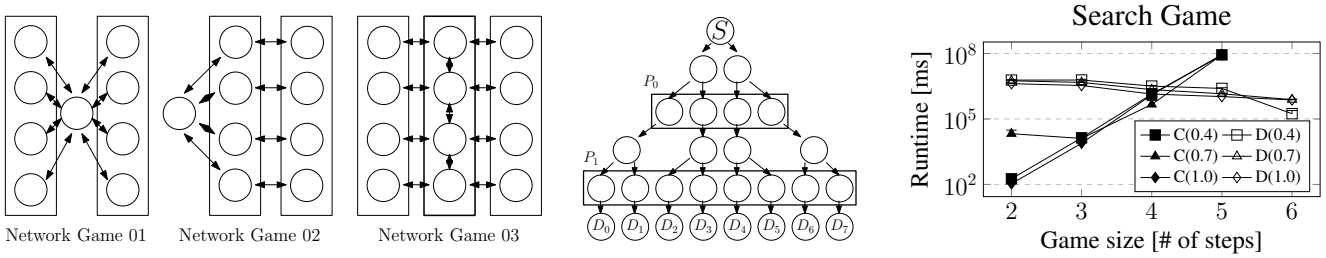


Figure 2: (Left) Graphs for network games, (Middle) Graph for search game, (Right) Mean runtimes in search game.

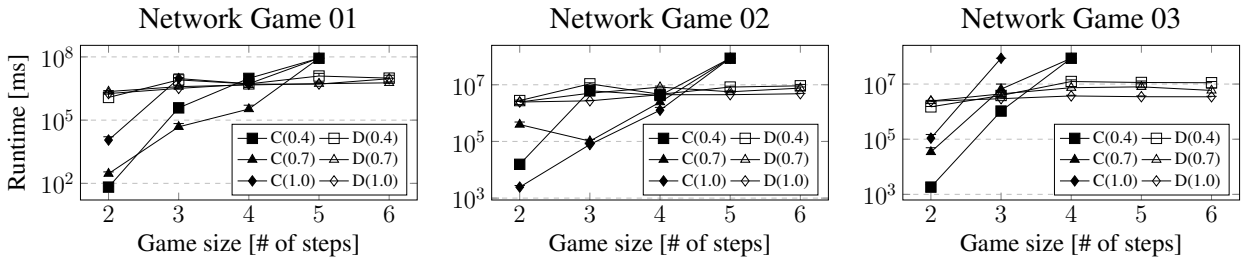


Figure 3: Mean runtimes of COBYLA and DTA in network games. Every point shows also a standard error.

**Experimental Setting.** We assume the defender acts as a leader, while the attacker assumes the follower's role. We consider three exponential generators of canonical logit functions, $\lambda \in \{0.4, 0.7, 1.0\}$. The tolerance parameter for the COBYLA algorithm in NLOPT was set to $10^{-2}$ and $\epsilon_B = 1\%$ of the leader's utility range for the DTA's binary search. The linearization uses $K = 3$, the basis of MDT is set to $b = 3$ and the size of the precision interval $\mathcal{E}$ is $L = 4$. For each combination of game size $\times$ generator function, we constructed 20 instances. All implementations were done in C++17. We used NLOPT 2.6.1, and a single-threaded IBM CPLEX 12.8 carried all MILP computations. The experiments were performed on a 3.2GHz CPU with 16GB RAM.

**Runtimes.** In Figures 2 (right) and 3, the x-axis varies the game size, while the y-axis shows the runtimes of the algorithms. Every point in the graphs corresponds to the mean over the sampled instances and shows the achieved standard error. We terminated all running seeds after 24h and depict them in the graphs with this lower bound on runtime if the computation was still ongoing. As the figures show, despite the overhead of the DTA algorithm on smaller instances, it scales significantly better than COBYLA. For 6 steps, we ran longer jobs, and COBYLA computed no game within 3 days. Interestingly, as in some other NP-hard problems (Cheeseman, Kanefsky, and Taylor 1991), increasing the search game's strategy space enables finding better strategies faster, and binary search terminates earlier.

**Solution Quality.** The relative errors of computed solutions are presented in Table 1. The values correspond to the mean ratio of the difference in the defender's expected utility computed using COBYLA and the DTA to the length of the defender's utility range in the game. Due to linear approximations used by COBYLA, it can find close-to-optimal solutions in smaller instances. In some cases, the solution is even better than the DTA because of the DTA's approximation parameters. However, as the Table reveals, the quality of COBYLA's solutions degrades with increasing game size, reaching error of $18.24\%$ for 4 steps in the search game.

| | 2 steps | 3 steps | 4 steps |
|---|---|---|---|
| **Networ Game 01** | | | |
| $\lambda = 0.4$ | 0.23% ± 0.03% | 1.27% ± 0.20% | 2.83% ± 0.46% |
| $\lambda = 0.7$ | -1.99% ± 0.38% | 0.42% ± 0.38% | 3.84% ± 0.74% |
| $\lambda = 1.0$ | -3.82% ± 0.51% | 0.50% ± 0.50% | 2.32% ± 0.92% |
| **Network Game 02** | | | |
| $\lambda = 0.4$ | 1.17% ± 0.27% | 1.16% ± 0.44% | 13.32% ± 7.35% |
| $\lambda = 0.7$ | 0.20% ± 0.21% | 1.80% ± 0.42% | 8.87% ± 3.47% |
| $\lambda = 1.0$ | -2.35% ± 0.35% | -0.22% ± 0.49% | 3.74% ± 4.05% |
| **Network Game 03** | | | |
| $\lambda = 0.4$ | 1.09% ± 0.27% | 1.30% ± 0.42% | - |
| $\lambda = 0.7$ | 1.02% ± 0.45% | 5.73% ± 1.85% | - |
| $\lambda = 1.0$ | 0.25% ± 0.68% | 8.33% ± 4.94% | - |
| **Search Game** | | | |
| $\lambda = 0.4$ | -2.13% ± 0.29% | -0.20% ± 0.42% | 16.23% ± 3.77% |
| $\lambda = 0.7$ | -7.24% ± 1.00% | -4.75% ± 0.79% | 18.24% ± 3.14% |
| $\lambda = 1.0$ | -21.24% ± 1.81% | -10.99% ± 1.64% | -1.98% ± 4.91% |

Table 1: Comparison of solution quality. A positive value indicates that DTA returned better solution than COBYLA.

## Conclusion

We study Quantal Stackelberg Equilibrium (QSE) – a strategy the rational player should commit to against a subrational player – in extensive-form games (EFGs). We show that computing QSE is NP-hard; still, QSE is useful for evaluating scalable heuristics or improving the understanding of human decision-making in experiments with human participants. We introduce the first practical algorithm for computing QSE in EFGs and show that contrary to direct formulation, our algorithm solves larger games with smaller errors.

## Acknowledgements

This research is supported by the SIMTech-NTU Joint Laboratory on Complex Systems, the Czech Science Foundation (grant no. 18-27483Y and 19-24384Y) and by the OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". Bo An is partially supported by Singtel Cognitive and Artificial Intelligence Lab for Enterprises (SCALE@NTU), which is a collaboration between Singapore Telecommunications Limited (Singtel) and Nanyang Technological University (NTU) that is funded by the Singapore Government through the Industry Alignment Fund – Industry Collaboration Projects Grant.

## References

An, B.; Shieh, E.; Yang, R.; Tambe, M.; Baldwin, C.; DiRenzo, J.; Maule, B.; and Meyer, G. 2013. A deployed quantal response based patrol planning system for the US Coast Guard. *In Interfaces* 43(5): 400–420.

Basak, A.; Černý, J.; Gutierrez, M.; Curtis, S.; Kamhoua, C.; Jones, D.; Bošanský, B.; and Kiekintveld, C. 2018. An initial study of targeted personality models in the FlipIt game. In *International Conference on Decision and Game Theory for Security*, 623–636. Springer.

Brown, N.; and Sandholm, T. 2018. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359(6374): 418–424.

Čermák, J.; Bošanský, B.; Durkota, K.; Lisý, V.; and Kiekintveld, C. 2016. Using correlated strategies for computing Stackelberg equilibria in extensive-form games. In *Proceedings of Thirtieth AAAI Conference on Artificial Intelligence*, 439–445.

Černý, J.; Bošanský, B.; and An, B. 2020. Finite state machines play extensive-form games. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, 509–533. New York, NY, USA: Association for Computing Machinery.

Černý, J.; Lisý, V.; Bošanský, B.; and An, B. 2020. Dinkelbach-type algorithm for computing Quantal Stackelberg equilibrium. In Bessiere, C., ed., *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, 246–253. IJCAI.

Černý, J.; Lisý, V.; Bošanský, B.; and An, B. 2021. Computing Quantal Stackelberg Equilibrium in Extensive-Form Games: Appendix. https://cloud.disroot.org/s/4Cin5Ny3nmZzWkR. Accessed: 2021-03-15.

Cheeseman, P. C.; Kanefsky, B.; and Taylor, W. M. 1991. Where the really hard problems are. In *IJCAI*, volume 91, 331–337.

Dinkelbach, W. 1967. On nonlinear fractional programming. *Management Science* 13(7): 492–498.

Fang, F.; Nguyen, T. H.; Pickles, R.; Lam, W. Y.; Clements, G. R.; An, B.; Singh, A.; Schwedock, B. C.; Tambe, M.; and Lemieux, A. 2017. PAWS - a deployed game-theoretic application to combat poaching. *AI Magazine* .

Gigerenzer, G.; and Goldstein, D. G. 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological review* 103(4): 650.

Gray, J. R. 1999. A bias toward short-term thinking in threat-related negative emotional states. *Personality and Social Psychology Bulletin* 25(1): 65–75.

Kahneman, D.; and Tversky, A. 2013. Prospect theory: An analysis of decision under risk. In *Handbook of the Fundamentals of Financial Decision Making: Part I*, 99–127. World Scientific.

Koller, D.; Megiddo, N.; and von Stengel, B. 1996. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior* 247–259.

Kolodziej, S.; Castro, P. M.; and Grossmann, I. E. 2013. Global optimization of bilinear programs with a multiparametric disaggregation technique. *Journal of Global Optimization* 57(4): 1039–1063.

Kroer, C.; Farina, G.; and Sandholm, T. 2018. Robust Stackelberg equilibria in extensive-form games and extension to limited lookahead. In *Proceedings of Thirty-Second AAAI Conference on Artificial Intelligence*.

Lejarraga, T.; Dutt, V.; and Gonzalez, C. 2012. Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making* 25(2): 143–153.

Letchford, J.; and Conitzer, V. 2010. Computing optimal strategies to commit to in extensive-form games. In *Proceedings of the 11th ACM conference on Electronic commerce*, 83–92.

Malhotra, N. K. 1982. Information load and consumer decision making. *Journal of consumer research* 8(4): 419–430.

Marchesi, A.; Farina, G.; Kroer, C.; Gatti, N.; and Sandholm, T. 2019. Quasi-perfect stackelberg equilibrium. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 2117–2124.

McFadden, D. L. 1976. Quantal choice analaysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4*, 363–390. NBER.

McKelvey, R. D.; and Palfrey, T. R. 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior* 10(1): 6–38.

McKelvey, R. D.; and Palfrey, T. R. 1998. Quantal response equilibria for extensive form games. *Experimental Economics* 1(1): 9–41.

Milec, D.; Černý, J.; Lisý, V.; and An, B. 2020. Complexity and Algorithms for Exploiting Quantal Opponents in Large Two-Player Games. In *Thirty-Fifth AAAI Conference on Artificial Intelligence*.

Moravčík, M.; Schmid, M.; Burch, N.; Lisý, V.; Morrill, D.; Bard, N.; Davis, T.; Waugh, K.; Johanson, M.; and Bowling, M. 2017. DeepStack: Expert-level artificial intelligence in no-limit poker. *Science* .

Nesterov, Y. 2004. *Introductory Lectures on Convex Optimization: A Basic Course*. Applied Optimization 87. Springer US, 1 edition.

Nguyen, T. H.; Yang, R.; Azaria, A.; Kraus, S.; and Tambe, M. 2013. Analyzing the effectiveness of adversary modeling in security games. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, AAAI'13, 718–724. AAAI Press.

Powell, M. J. 2007. A view of algorithms for optimization without derivatives. *Mathematics Today-Bulletin of the Institute of Mathematics and its Applications* 43(5): 170–174.

Sinha, A.; Fang, F.; An, B.; Kiekintveld, C.; and Tambe, M. 2018. Stackelberg security games: Looking beyond a decade of success. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18)*, 5494–5501. IJCAI.

Von Stengel, B.; and Zamir, S. 2010. Leadership games with convex strategy sets. *Games and Economic Behavior* 69(2): 446–457.

Wright, J. R.; and Leyton-Brown, K. 2020. A formal separation between strategic and nonstrategic behavior. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, 535–536. New York, NY, USA: Association for Computing Machinery.

Yadav, A.; Chan, H.; Xin Jiang, A.; Xu, H.; Rice, E.; and Tambe, M. 2016. Using social networks to aid homeless shelters: Dynamic influence maximization under uncertainty. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '16, 740–748. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Yang, R.; Kiekintveld, C.; Ordonez, F.; Tambe, M.; and John, R. 2011. Improving resource allocation strategy against human adversaries in security games. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, IJCAI'11, 458–464. IJCAI.

Yang, R.; Ordonez, F.; and Tambe, M. 2012. Computing optimal strategy against quantal response in security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, 847–854.

Yano, M.; Penn, J. D.; Konidaris, G.; and Patera, A. T. 2013. *Math, Numerics & Programming (for Mechanical Engineers)*. MIT.

Yechiam, E.; and Hochman, G. 2013. Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin* 139(2): 497.