Quasi-Perfect Stackelberg Equilibrium

Alberto Marchesi,[1] Gabriele Farina,[2] Christian Kroer,[2] Nicola Gatti,[1] Tuomas Sandholm[2]
[1] Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milan, Italy
[2] Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
alberto.marchesi@polimi.it, gfarina@cs.cmu.edu, ckroer@cs.cmu.edu,
nicola.gatti@polimi.it, sandholm@cs.cmu.edu

Abstract

Equilibrium refinements are important in extensive-form (*i.e.*, tree-form) games, where they amend weaknesses of the Nash equilibrium concept by requiring sequential rationality and other beneficial properties. One of the most attractive refinement concepts is *quasi-perfect equilibrium*. While quasi-perfection has been studied in extensive-form games, it is poorly understood in Stackelberg settings—that is, settings where a leader can commit to a strategy—which are important for modeling, for example, security games. In this paper, we introduce the axiomatic definition of *quasi-perfect Stackelberg equilibrium*. We develop a broad class of game perturbation schemes that lead to them in the limit. Our class of perturbation schemes strictly generalizes prior perturbation schemes introduced for the computation of (non-Stackelberg) quasi-perfect equilibria. Based on our perturbation schemes, we develop a branch-and-bound algorithm for computing a quasi-perfect Stackelberg equilibrium. It leverages a perturbed variant of the linear program for computing a Stackelberg extensive-form correlated equilibrium. Experiments show that our algorithm can be used to find an approximate quasi-perfect Stackelberg equilibrium in games with thousands of nodes.

1 Introduction

The main solution concept in game theory, *Nash equilibrium (NE)*, may prescribe non-credible strategies in *extensive-form (i.e., tree-form) games (EFGs)*. To solve that problem, equilibrium refinements have been proposed for such games (Selten 1975). Among the plethora of NE refinements (see van Damme (1987) for details), the *quasi-perfect equilibrium (QPE)*, proposed by Van Damme (1984), plays a central role, and it is considered one of the most attractive NE refinement concepts, as argued, for example, by Mertens (1995). The rationale behind the QPE concept is that every player, in every information set, plays her best response to perturbed—that is, subject to trembles—strategies of the opponents. Unlike the *normal-form perfect equilibrium*, the QPE guarantees that the strategies of the players are sequentially rational, and furthermore, quasi-perfection implies normal-form perfection. Unlike the *extensive-form perfect equilibrium (EFPE)*, in a QPE every player (reasonably) assumes that she will not make mistakes in the future, and this excludes some unreasonable strategies (Mertens 1995). Computation of NE refinements has received extensive attention in the literature. In the two-player case, Miltersen and Sørensen (2010) provide algorithms for finding a QPE, while Farina and Gatti (2017) for finding an EFPE. In particular, Miltersen and Sørensen (2010) show that a strict subset of the QPEs can be found when the sequence form is subject to a specific perturbation, while Farina and Gatti (2017) do the same for the EFPE. Iterative algorithms for such perturbed games in the zero-sum EFPE setting were introduced by Kroer, Farina, and Sandholm (2017) and Farina, Kroer, and Sandholm (2017).[1]

In *Stackelberg games*, a *leader* commits to a (possibly mixed) strategy first, and then a *follower* best responds to that strategy (von Stackelberg 1934). Stackelberg games have received significant attention in recent years (Conitzer and Sandholm 2006) due to their applications, for example, in security domains (Tambe 2011).

Work on equilibrium refinements in the context of *Stackelberg extensive-form games* has only started recently. Akin to usual extensive-form game refinements, *Stackelberg equilibrium (SE)* refinements should guarantee both the optimality of the commitment off the equilibrium path and some form of robustness against small trembles of the opponent.

To our knowledge, there is only one prior study of refinements for Stackelberg extensive-form games (Farina et al. 2018). They characterize a set of SE refinements based on what solutions can be obtained by imposing a perturbation scheme on the game—where players tremble onto suboptimal strategies with some small probabilities—and tak-

[1] *Normal-form proper equilibrium* is a refinement of QPE (Van Damme 1984), but it has drawbacks: (1) it requires players to assume a very specific structure on trembles which is not necessarily well-motivated, (2) the minimum tremble magnitudes depend on the action probabilities, which begets additional computational challenges, and (3) it is unknown whether it can be represented via perturbation schemes, even in the non-Stackelberg setting. For the zero-sum case, Miltersen and Sørensen (2008) show a polynomial-time approach using the sequence form, but it is based on solving a large (possibly linear in game-size) number of LPs, and thus may not be practical. For the general-sum case, it is not even known whether the sequence form can be applied; the only known approach relies on conversion to normal form—which causes an exponential blow-up—and then applying a pivoting algorithm (Sørensen 2012).

ing the limit as the trembling probability approaches zero. They prove that, for any perturbation scheme, all the limit points of sequences of SEs in a perturbed game are SEs of the original, unperturbed game. Interestingly, they prove that when restricting attention to the common tie-breaking rules for the follower (*strong* SE assumes the follower breaks ties in the best way for the leader and *weak* SE assumes the follower breaks tie in the worst way for the leader), this is no longer the case. Their approach does not start from a game-theoretic, axiomatic definition of the refinement concept. As we show in this paper, their approach captures only a strict subset of the solutions that are consistent with our natural game-theoretically defined refinement concept. One way to view this is that their operational definition is deficient in that it does not characterize all the solutions that are consistent with the natural, axiomatic definition of the refinement concept. Another view is that they have an operational definition and we provide a generalization.

In terms of complexity, they prove that finding any SE is NP-hard. (Hardness had previously been proven for finding a strong SE (Letchford and Conitzer 2010).) Therefore, finding any SE refinement is also NP-hard.

Our contributions. In this paper, we formally define the *quasi-perfect Stackelberg equilibrium (QPSE)* refinement game theoretically in the same axiomatic fashion as QPE was defined for non-Stackelberg games (Van Damme 1984). As in the case of QPEs, our definition is based on a set of properties of the players' strategies, and it cannot be directly used to search for a QPSE. Subsequently, we define a class of perturbation schemes for the sequence form such that any limit point of a sequence of SEs in a perturbed game is a QPSE. This class of perturbation schemes strictly includes those used to find a QPE by Miltersen and Sørensen (2010). Then, we extend the algorithm by Čermák et al. (2016) to the case of QPSE computation. We derive the corresponding mathematical program for computing a *Stackelberg extensive-form correlated equilibrium (SEFCE)* when a perturbation scheme is introduced and we discuss how the individual steps of the algorithm change. In particular, the implementation of our algorithm is much more involved, requiring the combination of branch-and-bound techniques with arbitrary-precision arithmetic to deal with small perturbations. This does not allow a direct application of off-the-shelf solvers. Finally, we experimentally evaluate the scalability of our algorithm.

2 Preliminaries

Using standard notation (Shoham and Leyton-Brown 2008), a *Stackelberg extensive-form game (SEFG)* of imperfect information is a tuple $(\mathcal{N}, \mathcal{H}, \mathcal{Z}, \mathcal{A}, \rho, \chi, \mathcal{C}, u, \mathcal{I})$. $\mathcal{N} = \{\ell, f\}$ is the set of players, the leader and the follower. $\mathcal{H} = \mathcal{H}_c \cup \mathcal{H}_\ell \cup \mathcal{H}_f$ is the set of nonterminal nodes, where $\mathcal{H}_c$ is the set of chance nodes, while $\mathcal{H}_\ell$ and $\mathcal{H}_f$ are the sets of leader's and follower's decision nodes, respectively. $\mathcal{Z}$ is the set of terminal nodes. $\mathcal{A} = \mathcal{A}_c \cup \mathcal{A}_\ell \cup \mathcal{A}_f$ is the set of actions, where $\mathcal{A}_c$ contains chance moves, while $\mathcal{A}_\ell$ and $\mathcal{A}_f$ are the sets of leader's and follower's actions, respectively. $\rho : \mathcal{H} \rightarrow 2^{\mathcal{A}}$ is the action function that assigns to each nonterminal node a set of available actions. $\chi : \mathcal{H} \times \mathcal{A} \rightarrow \mathcal{H} \cup \mathcal{Z}$ is the successor function that defines the node reached when an action is performed in a nonterminal node. $\mathcal{C} : \mathcal{H} \cup \mathcal{Z} \rightarrow [0, 1]$ assigns each node with its probability of being reached given chance moves. $u = \{u_\ell, u_f\}$, where $u_\ell, u_f : \mathcal{Z} \rightarrow \mathbb{R}$ specify leader's and follower's payoffs, respectively, in each terminal node. Finally, $\mathcal{I} = \{\mathcal{I}_\ell, \mathcal{I}_f\}$, where $\mathcal{I}_\ell$ and $\mathcal{I}_f$ define partitions of $\mathcal{H}_\ell$ and $\mathcal{H}_f$, respectively, into information sets, that is, groups of nodes that are indistinguishable by the player. For every information set $I \in \mathcal{I}_i$ and nodes $h, \hat{h} \in I$, it must be the case that $\rho(h) = \rho(\hat{h}) = A(I)$, otherwise player i would be able to distinguish the two nodes. As usual, w.l.o.g., we assume that each action $a \in A$ belongs to only one set $A(I)$.

We focus on *perfect-recall* SEFGs in which no player forgets what she did or knew in the past, that is, for every $i \in \mathcal{N}$ and $I \in \mathcal{I}_i$, all nodes belonging to I share the same player i's moves on their paths from the root. Thus, we can restrict the attention to *behavioral strategies* (Kuhn 1953), which define, for every player $i \in \mathcal{N}$ and information set $I \in \mathcal{I}_i$, a probability distribution over the actions $A(I)$. For $i \in \mathcal{N}$, let $\pi_i \in \Pi_i$ be a player i's behavioral strategy, with π_{ia} denoting the probability of playing action $a \in \mathcal{A}_i$. Overloading notation, we use u_i as if it were defined over strategies instead of terminal nodes. Specifically, $u_i(\pi_\ell, \pi_f)$ is player i's expected utility when $\pi_\ell \in \Pi_\ell$ and $\pi_f \in \Pi_f$ are played.

Perfect-recall SEFGs admit an equivalent representation called the *sequence form* (Von Stengel 1996; Romanovskii 1962). Every node $h \in \mathcal{H} \cup \mathcal{Z}$ defines a *sequence* $\sigma_i(h)$ for player $i \in \mathcal{N}$, which is the ordered set of player i's actions on the path from the root to h. Let Σ_i be the set of player i's sequences. As usual, let $\sigma_\emptyset \in \Sigma_i$ be a fictitious element representing the empty sequence. In perfect-recall games, given an information set $I \in \mathcal{I}_i$, for any pair of nodes $h, \hat{h} \in I$ it holds $\sigma_i(h) = \sigma_i(\hat{h}) = \sigma_i(I)$. Given $\sigma_i \in \Sigma_i$ and $a \in A(I)$ with $I \in \mathcal{I}_i : \sigma_i = \sigma_i(I)$, we denote as $\sigma_i a$ the *extended* sequence obtained by appending a to σ_i. Moreover, for any pair $\sigma_i, \hat{\sigma}_i \in \Sigma_i$, we write $\hat{\sigma}_i \sqsubseteq \sigma_i$ whenever $\hat{\sigma}_i$ is a *prefix* of σ_i, that is, σ_i can be obtained by extending $\hat{\sigma}_i$ with a finite number of actions. Given $\sigma_i \in \Sigma_i$, we also let $I_i(\sigma_i)$ be the information set $I \in \mathcal{I}_i : \sigma_i = \sigma_i(I)a$ for some $a \in A(I)$. In the sequence form, a strategy, called a *realization plan*, assigns each sequence with its probability of being played. For $i \in \mathcal{N}$, let $r_i \in R_i$ be a player i's realization plan. In order to be well-defined, a realization plan r_i must be such that $r_i(\sigma_\emptyset) = 1$ and, for $I \in \mathcal{I}_i$, $r_i(\sigma_i(I)) = \sum_{a \in A(I)} r_i(\sigma_i(I)a)$. Finally, letting $\Sigma = \Sigma_\ell \times \Sigma_f$ be the set of sequence pairs $\sigma = (\sigma_\ell, \sigma_f)$, overloading notation, $u_i : \Sigma \rightarrow \mathbb{R}$ is player i's utility function in the sequence form, with $u_i(\sigma) = \sum_{h \in \mathcal{Z} : \sigma_\ell(h) = \sigma_\ell \wedge \sigma_f(h) = \sigma_f} u_i(h)\mathcal{C}(h)$. Moreover, we also use u_i as if it were defined over realization plans. Formally, $u_i(r_\ell, r_f) = \sum_{\sigma \in \Sigma} u_i(\sigma) r_\ell(\sigma_\ell) r_f(\sigma_f)$.

The sequence form is usually expressed with matrix notation as follows. Player i's utility function is a $|\Sigma_\ell| \times |\Sigma_f|$ matrix U_i whose entries are the utilities $u_i(\sigma)$, for $\sigma \in \Sigma$. Constraints defining $r_i \in R_i$ are expressed as $F_i r_i = f_i$, where: F_i is a $(|\mathcal{I}_i| + 1) \times |\Sigma_i|$ matrix, $f_i \in \mathbb{R}^{|\mathcal{I}_i|+1}$, and, overloading notation, $r_i \in \mathbb{R}^{|\Sigma_i|}$ is a vector representing r_i.

Specifically, introducing a fictitious information set $I_\emptyset$, the entry of F_i indexed by $(I_\emptyset, \sigma_\emptyset)$ is 1, and, for $I \in \mathcal{I}_i$ and $\sigma_i \in \Sigma_i$, the entry indexed by (I, σ_i) is -1 if $\sigma_i = \sigma_i(I)$, while it is 1 if $\sigma_i = \sigma_i(I)a$ for some $a \in A(I)$. F_i is zero everywhere else. Moreover, $f_i^T = (1\ 0 \cdots 0)$. Finally, given $r_\ell \in R_\ell$ and $r_f \in R_f$, we can write $u_i(r_\ell, r_f) = r_\ell^T U_i r_f$.

In perfect-recall games, behavioral strategies and realization plans are equally expressive. Given $r_i \in R_i$, we obtain an equivalent $\pi_i \in \Pi_i$ by setting, for all $I \in \mathcal{I}_i$ and $a \in A(I)$, $\pi_{ia} = \frac{r_i(\sigma_i(I)a)}{r_i(\sigma_i(I))}$ when $r_i(\sigma_i(I)) > 0$, while π_{ia} can be any otherwise. Similarly, $\pi_i \in \Pi_i$ has an equivalent $r_i \in R_i$ with $r_i(\sigma_i) = \prod_{a \in \sigma_i} \pi_{ia}$ for all $\sigma_i \in \Sigma_i$.[2]

The solution concept associated with SEFGs is the SE. An SEFG may have many SEs, depending on the leader's assumption on how the follower breaks ties among multiple best responses. A leader's strategy is part of an SE if it is optimal for *some* tie-breaking rule of the follower. Letting $\mathsf{BR}_\Gamma(\pi_\ell) = \arg\max_{\pi_f \in \Pi_f} u_f(\pi_\ell, \pi_f)$ be the set of follower's best responses to $\pi_\ell \in \Pi_\ell$ in an SEFG Γ, we have the following formal definition of SE.[3]

Definition 1. *Given an SEFG Γ, (π_ℓ, π_f) is an* SE *of Γ if $\pi_f \in \mathsf{BR}_\Gamma(\pi_\ell)$ and, for all $\hat\pi_\ell \in \Pi_\ell$, there exists $\hat\pi_f \in \mathsf{BR}_\Gamma(\hat\pi_\ell)$ such that $u_\ell(\pi_\ell, \pi_f) \geq u_\ell(\hat\pi_\ell, \hat\pi_f)$.*

Many papers on SEs focus on *strong SEs* (SSEs), which assume that the follower breaks ties in favor of the leader.

Definition 2. *Given an SEFG Γ, (π_ℓ, π_f) is an* SSE *of Γ if $\pi_f \in \mathsf{BR}_\Gamma(\pi_\ell)$ and, for all $\hat\pi_\ell \in \Pi_\ell$ and $\hat\pi_f \in \mathsf{BR}_\Gamma(\hat\pi_\ell)$, it holds $u_\ell(\pi_\ell, \pi_f) \geq u_\ell(\hat\pi_\ell, \hat\pi_f)$.*

Finally, SEs and SSEs can be defined analogously for SEFGs in sequence form (using the equivalence between behavioral strategies and realization plans).

3 Definition of Quasi-Perfect Stackelberg Equilibrim

In this section, we introduce QPSEs, which refine SEs in SEFGs using an approach resembling that adopted by Van Damme (1984) for defining QPEs in EFGs.

First, we provide needed additional notation. We say that $\pi_i \in \Pi_i$ is *completely mixed* if $\pi_{ia} > 0$ for all $a \in \mathcal{A}_i$. Given two information sets $I, \hat I \in \mathcal{I}_i$, we write $I \succeq \hat I$ whenever $\hat I$ follows I, *i.e.*, there exists a path from $h \in I$ to $\hat h \in \hat I$. We assume $I_\emptyset \succeq \hat I$ for all $\hat I \in \mathcal{I}_i$ such that there is no $I \neq \hat I \in \mathcal{I}_i : I \succeq \hat I$. In perfect-recall games, $\succeq$ is a partial order over $\mathcal{I}_i \cup \{I_\emptyset\}$. Given $\pi_i, \hat\pi_i \in \Pi_i$ and $I \in \mathcal{I}_i \cup \{I_\emptyset\}$, $\pi_i \big/_I \hat\pi_i$ is equal to $\hat\pi_i$ at all $\hat I \in \mathcal{I}_i : I \succeq \hat I$, while it is equal to π_i everywhere else. Moreover, for $I \in \mathcal{I}_i$, we write $\pi_i =_I \hat\pi_i$ if $\pi_{ia} = \hat\pi_{ia}$ for all $a \in A(I)$. Finally, given completely mixed strategies $\pi_\ell \in \Pi_\ell$, $\pi_f \in \Pi_f$ and $I \in \mathcal{I}_i$, $u_{i,I}(\pi_\ell, \pi_f)$ denotes player i's expected utility given that I has been reached and strategies π_ℓ and π_f are played.

Next, we introduce a fundamental building block: the idea of follower's best response at an information set $I \in \mathcal{I}_f$. Intuitively, π_f is an I-best response to π_ℓ whenever playing as prescribed by π_f at the information set I is part of some follower's best response to π_ℓ in the game following I, given that I has been reached during play. Formally:

Definition 3. *Given an SEFG Γ, a completely mixed $\pi_\ell \in \Pi_\ell$, and $I \in \mathcal{I}_f$, we say that $\pi_f \in \Pi_f$ is an I-*best response *to π_ℓ, written $\pi_f \in \mathsf{BR}_I(\pi_\ell)$, if the following holds:*

$$\max_{\substack{\hat\pi_f \in \Pi_f : \\ \pi_f =_I \hat\pi_f}} u_{f,I}\left(\pi_\ell, \pi_f \big/_I \hat\pi_f\right) = \max_{\hat\pi_f \in \Pi_f} u_{f,I}\left(\pi_\ell, \pi_f \big/_I \hat\pi_f\right).$$

For $i \in \mathcal{N}$ and $\pi_i \in \Pi_i$, let $\{\pi_{i,k}\}_{k \in \mathbb{N}}$ be a sequence of completely mixed player i's strategies with π_i as a limit point. We are now ready to define the refinement concept. In words, in a QPSE, the leader selects an optimal strategy to commit to in *all* information sets, given that the follower best responds to it at *every* information set, following *some* tie-breaking rule. Specifically, point (ii) in Definition 4 ensures that the leader's commitment is optimal also in those information sets that are unreachable in absence of players' errors. Notice that the leader only accounts for follower's future errors, while the follower assumes that only the leader can make mistakes in future. This is in line with the idea underlying QPEs in EFGs (Van Damme 1984).[4]

Definition 4. *Given an SEFG Γ, (π_ℓ, π_f) is a* quasi-perfect Stackelberg equilibrium (QPSE) *of Γ if there exist sequences $\{\pi_{i,k}\}_{k \in \mathbb{N}}$, defined for every $i \in \mathcal{N}$ and $\pi_i \in \Pi_i$, such that:*

(i) $\pi_f \in \mathsf{BR}_I(\pi_{\ell,k})$ for all $I \in \mathcal{I}_f$;

(ii) for all $I \in \mathcal{I}_\ell \cup \{I_\emptyset\}$ and $\hat\pi_\ell \in \Pi_\ell$, there exists $\hat\pi_f \in \Pi_f : \hat\pi_f \in \mathsf{BR}_{\hat I}(\pi_{\ell,k} \big/_I \hat\pi_{\ell,k})$ for all $\hat I \in \mathcal{I}_f$, with:

$$u_\ell\left(\pi_{\ell,k} \big/_I \pi_\ell, \pi_{f,k}\right) \geq u_\ell\left(\pi_{\ell,k} \big/_I \hat\pi_\ell, \hat\pi_{f,k}\right). \qquad (1)$$

As with SEs, we introduce the *strong* version of QPSEs.[5]

Definition 5. *Given an SEFG Γ, (π_ℓ, π_f) is a* quasi-perfect strong Stackelberg equilibrium (QPSSE) *of Γ if there exist $\{\pi_{i,k}\}_{k \in \mathbb{N}}$, defined for every $i \in \mathcal{N}$ and $\pi_i \in \Pi_i$, such that:*

(i) $\pi_f \in \mathsf{BR}_I(\pi_{\ell,k})$ for all $I \in \mathcal{I}_f$;

(ii) for all $I \in \mathcal{I}_\ell \cup \{I_\emptyset\}$, $\hat\pi_\ell \in \Pi_\ell$, and $\hat\pi_f \in \Pi_f : \hat\pi_f \in \mathsf{BR}_{\hat I}(\pi_{\ell,k} \big/_I \hat\pi_{\ell,k})$ for all $\hat I \in \mathcal{I}_f$, Equation (1) holds.

As we will show in Section 4, QPSEs are refinements of SEs, that is, any QPSE is also an SE.

4 Family of Perturbation Schemes for QPSE

We now introduce a family of *perturbation schemes* for SEFGs in sequence form that satisfies the following fundamental property: *limits of SEs in perturbed sequence-form SEFGs are QPSEs of the original unperturbed SEFGs as the*

[2]Here, the equivalence is in terms of probabilities that the strategies induce on terminal nodes, *i.e.*, it is *realization equivalence*.

[3]In this paper, we define SEs following a characterization introduced by Farina et al. (2018) (Lemma 2 in their paper).

[4]Van Damme (1984) defines a QPE of an n-player extensive-form game as a strategy profile $(\pi_i)_{i \in \mathcal{N}}$ obtained as a limit point of a sequence of completely mixed strategy profiles $\{(\pi_{i,k})_{i \in \mathcal{N}}\}_{k \in \mathbb{N}}$ such that $\pi_i \in \mathsf{BR}_I((\pi_{j,k})_{j \neq i \in \mathcal{N}})$ for all $i \in \mathcal{N}$ and $I \in \mathcal{I}_i$.

[5]Since Equation (1) must hold for every $\hat\pi_\ell \in \Pi_\ell$ and $\hat\pi_f \in \Pi_f : \hat\pi_f \in \mathsf{BR}_{\hat I}(\pi_{\ell,k} \big/_I \hat\pi_{\ell,k})$ for all $\hat I \in \mathcal{I}_f$, Definition 5 assumes that the follower breaks ties in favor of the leader.

magnitude of the perturbation goes to zero. In addition to being theoretically relevant, this result enables us to design an algorithm for computing QPSEs in SEFGs (Section 7).

Definition 6 (ξ-perturbation scheme)**.** *Given an SEFG Γ and $i \in \mathcal{N}$, let $\xi_i : (0,1] \times \Sigma_i \to \mathbb{R}^+$ be a function that maps a perturbation magnitude $\epsilon \in (0,1]$ and a sequence $\sigma_i \in \Sigma_i$ to a positive lower-bound $\xi_i(\epsilon, \sigma_i)$ on the probability of playing σ_i such that:*

(i) $\xi_i(\epsilon, \sigma_i)$ *is a polynomial in* ϵ*, for all* $\sigma_i \in \Sigma_i$*;*

(ii) $\lim_{\epsilon \to 0^+} \xi_i(\epsilon, \sigma_i) = 0$*, for all* $\sigma_i \in \Sigma_i \setminus \{\sigma_\emptyset\}$*;*

(iii) $\lim_{\epsilon \to 0^+} \frac{\xi_i(\epsilon, \sigma_i(I)a)}{\xi_i(\epsilon, \sigma_i(I))} = 0$*, for all* $I \in \mathcal{I}_i, a \in A(I)$*.*

Then, a ξ_i-perturbation scheme *for R_i is a function $\epsilon \mapsto R_i(\epsilon)$ defined over $\epsilon \in (0,1]$ in which $R_i(\epsilon)$ is the set of all $r_i \in R_i$ such that $r_i(\sigma_i) \geq \xi_i(\epsilon, \sigma_i)$ for all $\sigma_i \in \Sigma_i$.*

In words, the lower-bounds on sequence probabilities enjoy the following properties: (i) they are polynomials in the variable ϵ; (ii) they approach zero as ϵ goes to zero; and (iii) $\xi_i(\epsilon, \sigma_i(I)a)$ approaches zero faster than $\xi_i(\epsilon, \sigma_i(I))$.

We denote by $(\Gamma, \xi_\ell, \xi_f)$ a ξ*-perturbed* SEFG with ξ_i-perturbation schemes. We let $\Gamma(\epsilon)$ be a particular sequence-form ξ*-perturbed game instance* obtained from Γ by restricting each set of realization plans R_i to be $R_i(\epsilon)$. We denote by $r_i(\epsilon)$ any realization plan in $R_i(\epsilon)$, and we let $\xi_i(\epsilon) \in \mathbb{R}^{|Q_i|}$ be a vector whose components are the lower-bounds $\xi_i(\epsilon, \sigma_i)$. We denote by $\tilde{r}_i(\epsilon) = r_i(\epsilon) - \xi_i(\epsilon)$ the *residual* of $r_i(\epsilon)$, which represents the part of player i's strategy that is not fixed by the perturbation.[6]

Next, we state our main result about sequences of SEs in ξ-perturbed games. We postpone the proof to Section 6.

Theorem 1. *Given a ξ-perturbed SEFG $(\Gamma, \xi_\ell, \xi_f)$, let $\{\epsilon_k\}_{k \in \mathbb{N}} \to 0$ and let $\{(r_\ell(\epsilon_k), r_f(\epsilon_k))\}_{k \in \mathbb{N}}$ be a sequence of SEs in $\Gamma(\epsilon_k)$. Then, any limit point (π_ℓ, π_f) of the sequence $\{(\pi_{\ell,k}, \pi_{f,k})\}_{k \in \mathbb{N}}$ is a QPSE of Γ, where $(\pi_{\ell,k}, \pi_{f,k})$ are equivalent to $(r_\ell(\epsilon_k), r_f(\epsilon_k))$ for all $k \in \mathbb{N}$.*

Theorem 1 also allows us to conclude the following, as a consequence of Theorem 1 of Farina et al. (2018).

Corollary 1. *Any QPSE of an SEFG Γ is an SE of Γ.*

Retirements (ii)-(iii) in Definition 6 cannot be removed:

Observation 1. *There are ξ-perturbed SEFGs $(\Gamma, \xi_\ell, \xi_f)$ with ξ_i-perturbation schemes that violate point (ii) or (iii) in Definition 6 for which Theorem 1 does not hold.*

Proof. Consider the SEFG in Figure 1b with $\xi_\ell(\epsilon, a_\ell^1) = \xi_\ell(\epsilon, a_\ell^2) = \epsilon$ and $\xi_\ell(\epsilon, a_\ell^2 a_\ell^3) = \xi_\ell(\epsilon, a_\ell^2 a_\ell^4) = \frac{\epsilon}{3}$, which violates requirement (iii) in Definition 6. Clearly, any SE of $\Gamma(\epsilon)$ requires $r_\ell(\epsilon, a_\ell^1) = 1 - \epsilon$, $r_\ell(\epsilon, a_\ell^2) = \epsilon$, $r_\ell(\epsilon, a_\ell^2 a_\ell^3) = \frac{\epsilon}{3}$, and $r_\ell(\epsilon, a_\ell^2 a_\ell^4) = \frac{2\epsilon}{3}$. Thus, any limit point of a sequence of SEs has $\pi_\ell(a_\ell^3) = \frac{1}{3}$ and $\pi_\ell(a_\ell^4) = \frac{2}{3}$, which cannot be the case in a QPSE of Γ, as the leader's optimal strategy at $\ell.2$ is to play a_ℓ^4. As for requirement (ii), we can build a similar example by setting $\xi_\ell(\epsilon, a_\ell^2) = \frac{1}{3}$. □

[6] We assume without loss of generality that $\Gamma(\epsilon)$ is well-defined, that is, each set $R_i(\epsilon)$ is non-empty for every $\epsilon \in (0,1]$.

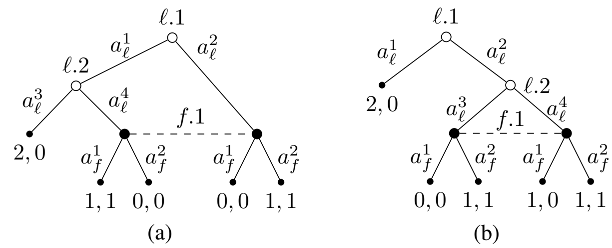

Figure 1: Examples SEFGs.

Miltersen and Sørensen (2010) introduced the idea of perturbing sequence-form EFGs in order to find a QPE. Our perturbation scheme generalizes theirs, where $\xi_i(\epsilon, \sigma_i) = \epsilon^{|\sigma_i|}$ for all $\sigma_i \in \Sigma_i \setminus \{\sigma_\emptyset\}$, with $|\sigma_i|$ being the number of actions in σ_i. There are games where our perturbation captures QPSEs that are *not* obtainable with theirs. For instance, in the SEFG in Figure 1a, (π_ℓ, π_f), with $\pi_\ell(a_\ell^1) = \pi_\ell(a_\ell^3) = 1$, $\pi_\ell(a_\ell^2) = \pi_\ell(a_\ell^4) = 0$, and $\pi_f(a_f^1) = \pi_f(a_f^2) = \frac{1}{2}$, is a QPSE that cannot be obtained with their perturbation scheme while it is reachable by setting $\xi_\ell(\epsilon, a_\ell^2) = \epsilon^2$. We observe that (π_ℓ, π_f) is also a QPE when we look at the game as an EFG without commitment; this shows that our perturbation scheme generalizes theirs also for QPEs.

Finally, when restricting the attention to SSEs, one can state the following: limits of SSEs in ξ-perturbed games are QPSSEs of the unperturbed game. This is made formal by Theorem 4 in (Marchesi et al. 2018).

5 Best Responses in ξ-Perturbed Games

We now study properties of the follower's best responses to the leader's strategy in ξ-perturbed games. These properties will be useful for proving our results later in the paper.

In the following, letting $\Sigma_i(a) = \{\sigma_i \in \Sigma_i \mid a \in \sigma_i\}$ for all $a \in \mathcal{A}_i$, $\Sigma_i(I) = \bigcup_{a \in A(I)} \Sigma_i(a)$ denotes player i's sequences that pass through information set $I \in \mathcal{I}_i$. For ease of presentation, given $I \in \mathcal{I}_i$, $g_{i,I}(r_\ell, r_f) = \sum_{\sigma \in \Sigma : \sigma_i \in \Sigma_i(I)} u_i(\sigma) r_\ell(\sigma_\ell) r_f(\sigma_f)$ denotes player i's expected utility contribution from terminal nodes reachable from I. Finally, for $I \in \mathcal{I}_i$, let $R_i(I) \subseteq R_i$ be the set of $r_i \in R_i : r_i(\sigma_i(I)) = 1$, while, for $a \in A(I)$, $R_i(a) \subseteq R_i(I)$ is the set of $r_i \in R_i : r_i(\sigma_i(I)a) = 1$.

Let $\mathsf{BR}_{\Gamma(\epsilon)}(r_\ell(\epsilon)) = \arg\max_{r_f(\epsilon) \in R_f(\epsilon)} u_f(r_\ell(\epsilon), r_f(\epsilon))$ be the set of follower's best responses to $r_\ell(\epsilon) \in R_\ell(\epsilon)$ in $\Gamma(\epsilon)$. The next lemma gives a mathematical programming formulation of the follower's best-response problem in $\Gamma(\epsilon)$.

Lemma 1. *For every $r_\ell(\epsilon) \in R_\ell(\epsilon)$, $r_f(\epsilon) \in \mathsf{BR}_{\Gamma(\epsilon)}(r_\ell(\epsilon))$ if and only if $\tilde{r}_f(\epsilon)$ is optimal for Problem $\mathcal{P}(\epsilon)$ below.*

$$\mathcal{P}(\epsilon) \ : \ \begin{cases} \max\limits_{\tilde{r}_f} & r_\ell(\epsilon)^T U_f \tilde{r}_f \\ \text{s.t.} & F_f \tilde{r}_f = f_f - F_f \xi_f(\epsilon), \quad \tilde{r}_f \geq 0. \end{cases}$$

All omitted proofs are in (Marchesi et al. 2018).

The dual of Problem $\mathcal{P}(\epsilon)$ above is as follows.

Proposition 1. *For $r_\ell(\epsilon) \in R_\ell(\epsilon)$, Problem $\mathcal{D}(\epsilon)$ below is the dual of Problem $\mathcal{P}(\epsilon)$, where $v_f \in \mathbb{R}^{|\mathcal{I}_f|+1}$ is the vector* *of dual variables.*

$$\mathcal{D}(\epsilon)\ :\ \begin{cases}\min\limits_{v_f} & \left(f_f - F_f\xi_f(\epsilon)\right)^T v_f \quad\quad (2a)\\ \text{s.t.} & F_f^T v_f \geq U_f^T r_\ell(\epsilon). \quad\quad (2b)\end{cases}$$

Remark 1. *Constraints* (2b) *in Problem $\mathcal{D}(\epsilon)$ defined above ensure that, for every $I \in \mathcal{I}_f$ and $a \in A(I)$, we have*

$$v_{f,I} \geq \sum_{\sigma\in\Sigma:\sigma_f=\sigma_f(I)a} u_f(\sigma) r_\ell(\epsilon,\sigma_\ell) + \sum_{\hat{I}\in\mathcal{I}_f:\sigma_f(\hat{I})=\sigma_f(I)a} v_{f,\hat{I}}. \quad (3)$$

The optimal solutions to Problem $\mathcal{D}(\epsilon)$ enjoy important properties that are stated in the following lemmas. The first one says that, in an optimal solution, each variable $v_{f,I}$ must equal the maximum possible expected utility the follower can achieve following information set $I \in \mathcal{I}_f$. The second lemma says that if an optimal solution to Problem $\mathcal{D}(\epsilon)$ satisfies Constraint (3) with equality for an information set $I \in \mathcal{I}_f$ and an action $a \in A(I)$, then playing a at I is optimal in the game following I.

Lemma 2. *For every $r_\ell(\epsilon) \in R_\ell(\epsilon)$, if $v_f^* \in \mathbb{R}^{|\mathcal{I}_f|+1}$ is optimal for Problem $\mathcal{D}(\epsilon)$, then for every $I \in \mathcal{I}_f$:*

$$v^*_{f,I} = \max_{\hat{r}_f\in R_f(I)} g_{f,I}(r_\ell(\epsilon),\hat{r}_f). \quad (4)$$

Lemma 3. *For every $r_\ell(\epsilon) \in R_\ell(\epsilon)$, $I \in \mathcal{I}_f$, and $a \in A(I)$, if Constraint* (3) *holds with equality in an optimal solution to Problem $\mathcal{D}(\epsilon)$, then*

$$\max_{\hat{r}_f\in R_f(a)} g_{f,I}(r_\ell(\epsilon),\hat{r}_f) = \max_{\hat{r}_f\in R_f(I)} g_{f,I}(r_\ell(\epsilon),\hat{r}_f). \quad (5)$$

Now we are ready to prove a fundamental property of the follower's best responses in ξ-perturbed game instances $\Gamma(\epsilon)$. Intuitively, in a perturbed game instance, the follower best responds playing sequence $\sigma(I_f)a$ with probability strictly greater than its lower-bound $\xi_f(\epsilon,\sigma_f(I)a)$ only if playing a is optimal in the game following I. Theorem 2 formally expresses the idea that, in a perturbed game instance $\Gamma(\epsilon)$, when the follower decides how to best respond to a leader's commitment in a given information set, she does not take into account her future trembles, but only opponents' ones.

Theorem 2. *Given $r_\ell(\epsilon) \in R_\ell(\epsilon)$, $r_f(\epsilon) \in \mathsf{BR}_{\Gamma(\epsilon)}(r_\ell(\epsilon))$, $I \in \mathcal{I}_f$, and $a \in A(I)$, if $r_f(\epsilon,\sigma_f(I)a) > \xi_f(\epsilon,\sigma_f(I)a)$, then*

$$\max_{\hat{r}_f\in R_f(a)} g_{f,I}(r_\ell(\epsilon),\hat{r}_f) = \max_{\hat{r}_f\in R_f(I)} g_{f,I}(r_\ell(\epsilon),\hat{r}_f).$$

Proof. By Lemma 1, $r_f(\epsilon) \in \mathsf{BR}_{\Gamma(\epsilon)}(r_\ell(\epsilon))$ if and only if $\tilde{r}_f(\epsilon) = r_f(\epsilon) - \xi_f(\epsilon)$ is optimal for Problem $\mathcal{P}(\epsilon)$. By applying the complementarity slackness theorem to Problems $\mathcal{P}(\epsilon)$ and $\mathcal{D}(\epsilon)$ we have that, if $\tilde{r}_f(\epsilon)$ and $v_f^* \in \mathbb{R}^{|\mathcal{I}_f|+1}$ are optimal, then, whenever $\tilde{r}_f(\epsilon,\sigma_f(I)a) > 0$, that is, $r_f(\epsilon,\sigma_f(I)a) > \xi_f(\epsilon,\sigma_f(I)a)$, Constraint (3) for information set I and action a must hold with equality, which, by Lemma 3, yields Equation (5). □

6 Limits of SEs in ξ-Perturbed Games are QPSEs of the Unperturbed Games

Here, we prove Theorem 1. First, we introduce two lemmas.

The first provides a characterization of I-best responses in terms of sequence form. Intuitively, a follower's strategy π_f is an I-best response to π_ℓ if and only if it places positive probability only on actions $a \in A(I)$ that are part of some best response of the follower below information set I.

Lemma 4. *Given an SEFG Γ, a completely mixed $\pi_\ell \in \Pi_\ell$ and $I \in \mathcal{I}_f$, $\pi_f \in \mathsf{BR}_I(\pi_\ell)$ if for every $a \in A(I)$:*

$$\pi_{ia} > 0 \implies \max_{\hat{r}_f\in R_f(a)} g_{f,I}(r_\ell,\hat{r}_f) = \max_{\hat{r}_f\in R_f(I)} g_{f,I}(r_\ell,\hat{r}_f),$$

where $r_\ell \in R_\ell$ is equivalent to π_ℓ.

The next lemma shows that any limit point of a sequence of follower's best responses in ξ-perturbed games is a follower's best response at every information set in Γ.

Lemma 5. *Given a ξ-perturbed SEFG (Γ,ξ_ℓ,ξ_f), let $\{\epsilon_k\}_{k\in\mathbb{N}} \to 0$ and let $\{(r_\ell(\epsilon_k), r_f(\epsilon_k))\}_{k\in\mathbb{N}}$ be a sequence of realization plans in $\Gamma(\epsilon_k)$ with $r_f(\epsilon_k) \in \mathsf{BR}_{\Gamma(\epsilon_k)}(r_\ell(\epsilon_k))$. Then, any limit point (π_ℓ,π_f) of $\{(\pi_{\ell,k},\pi_{f,k})\}_{k\in\mathbb{N}}$ is such that, eventually, $\pi_f \in \mathsf{BR}_{I_f}(\pi_{\ell,k})$ for all $I \in \mathcal{I}_f$, where $(\pi_{\ell,k},\pi_{f,k})$ are equivalent to $(r_\ell(\epsilon_k), r_f(\epsilon_k))$ for all $k \in \mathbb{N}$.*

Finally, we can prove Theorem 1.

Proof of Theorem 1. First, since $r_f(\epsilon_k) \in \mathsf{BR}_{\Gamma(\epsilon_k)}(r_\ell(\epsilon_k))$ for all $k \in \mathbb{N}$, Lemma 5 allows us to conclude that requirement (i) in Definition 4 holds. Therefore, in order to prove Theorem 1, we need to show that requirement (ii) holds as well. For contradiction, suppose that point (ii) does not hold, that is, no matter how we choose sequences $\{\pi_{i,k}\}_{k\in\mathbb{N}}$, for $i \in \mathcal{N}$ and $\pi_i \in \Pi_i$, there is an information set $I \in \mathcal{I}_\ell \cup \{I_\emptyset\}$ and a leader's strategy $\hat{\pi}_\ell \in \Pi_\ell$ such that, for every $\hat{\pi}_f \in \Pi_f : \hat{\pi}_f \in \mathsf{BR}_{\hat{I}}(\pi_{\ell,k}/_I\hat{\pi}_{\ell,k})$ for all $\hat{I} \in \mathcal{I}_f$, we have:

$$u_\ell(\pi_{\ell,k}/_I\pi_\ell,\pi_{f,k}) < u_\ell(\pi_{\ell,k}/_I\hat{\pi}_\ell,\hat{\pi}_{f,k}).$$

By continuity, there must exist an index $\bar{k} \in \mathbb{N}$ such that, for all $k \in \mathbb{N} : k \geq \bar{k}$, the following holds:

$$u_\ell(\pi_{\ell,k}/_I\pi_{\ell,k},\pi_{f,k}) < u_\ell(\pi_{\ell,k}/_I\hat{\pi}_{\ell,k},\hat{\pi}_{f,k}).$$

Moreover, $u_\ell(\pi_{\ell,k}/_I\pi_{\ell,k},\pi_{f,k}) = u_\ell(\pi_{\ell,k},\pi_{f,k})$. Let sequence $\{\hat{\pi}_{\ell,k}\}_{k\in\mathbb{N}}$ be such that $\hat{r}_\ell(\epsilon_k) \in R_\ell(\epsilon_k)$ for all $k \in \mathbb{N}$, where each realization plan $\hat{r}_\ell(\epsilon_k)$ is equivalent to the strategy $\pi_{\ell,k}/_I\hat{\pi}_{\ell,k}$. This is always possible since requirement (iii) in Definition 6 is satisfied. Consider a sequence $\{(\hat{r}_\ell(\epsilon_k),\hat{r}_f(\epsilon_k)\}_{k\in\mathbb{N}}$ with $\hat{r}_f(\epsilon_k) \in \mathsf{BR}_{\Gamma(\epsilon_k)}(\hat{r}_\ell(\epsilon_k))$, and let $\{(\pi_{\ell,k}/_I\hat{\pi}_{\ell,k},\hat{\pi}_{f,k})\}_{k\in\mathbb{N}}$ be a sequence such that each strategy $\hat{\pi}_{f,k}$ is equivalent to $\hat{r}_f(\epsilon_k)$. By Lemma 5, any limit point $(\pi_\ell/_I\hat{\pi}_\ell,\hat{\pi}_f)$ of $\{(\pi_{\ell,k}/_I\hat{\pi}_{\ell,k},\hat{\pi}_{f,k})\}_{k\in\mathbb{N}}$ satisfies $\hat{\pi}_f \in \mathsf{BR}_{\hat{I}}(\pi_{\ell,k}/_I\hat{\pi}_{\ell,k})$ for all $\hat{I} \in \mathcal{I}_f$. Thus, using the equivalence between strategies and realization plans, for all $k \in \mathbb{N} : k \geq \bar{k}$ we have the following:

$$u_\ell(r_\ell(\epsilon_k), r_f(\epsilon_k)) < u_\ell(\hat{r}_\ell(\epsilon_k),\hat{r}_f(\epsilon_k)).$$

Notice that this holds no matter how we choose $\hat{r}_f(\epsilon_k) \in \mathsf{BR}_{\Gamma(\epsilon_k)}(\hat{r}_\ell(\epsilon_k))$, which contradicts the fact that $(r_\ell(\epsilon_k), r_f(\epsilon_k))$ is an SE of $\Gamma(\epsilon_k)$. □

7 Algorithm

One can use our perturbation scheme to compute an (approximate) QPSE. We do this by developing an LP for computing an SEFCE in a given ξ-perturbed game instance, where we maximize the leader's value. We then conduct a *branch-and-bound* search on this SEFCE LP. It branches on which actions to *force* be recommended to the follower (by the correlation device of the SEFCE). The idea is that, as long as we only recommend a single action to the follower at any given information set, we get an SE of the perturbed game (specifically an SSE, and an SSE has maximum value among all SEs), and, thus, according to Theorem 1, a QPSE (specifically QPSSE) if we take the limit point of the perturbations. As in prior papers on EFCE computation in general-sum games, we focus on games without chance nodes (von Stengel and Forges 2008; Čermák et al. 2016).

For computing an SEFCE we need to specify joint probabilities over sequence pairs $(\sigma_\ell, \sigma_f) \in \Sigma$. However, not all pairs need to specify probabilities, only pairs such that choosing σ_f is affected by the probability put on σ_ℓ (we do not need to care about the converse of this, as only the follower needs to be induced to follow the recommended strategy). Intuitively, the set of the leader's sequences relevant to a given $\sigma_f \in \Sigma_f$ is made of those sequences that affect the expected value of σ_f or any alternative sequence $\hat{\sigma}_f \in \Sigma_f$ whose last action is available at $I_f(\sigma_f)$.

Definition 7 (Relevant sequences)**.** *A pair $(\sigma_\ell, \sigma_f) \in \Sigma$ is* relevant *if either $\sigma_\ell = \sigma_\emptyset$ or there exists $h, \hat{h} \in \mathcal{H}$ s.t. $\hat{h}$ precedes h, $h \in I_f(\sigma_f)$, and $\hat{h} \in I_\ell(\sigma_\ell)$, or if the condition holds with the roles of σ_ℓ and σ_f reversed.*

For every information set $I \in \mathcal{I}_i$, we let $rel(I)$ be the set of sequences relevant to each child sequence $\sigma_i(I)a$ for $a \in A(I)$. We let $p(\sigma_\ell, \sigma_f)$ be the probability that we recommend that the leader plays sequence σ_ℓ, and that the follower sends her *residual* (*i.e.*, the probability that is not fixed by the perturbation) to σ_f. Moreover, we let $\eta(\sigma_f)$ be the maximum probability that the follower can put on a sequence $\sigma_f \in \Sigma_f$ given the ξ_f-perturbation scheme.

First, we introduce a new value function representing the value to the leader of the sequence pair $(\sigma_\ell, \sigma_f) \in \Sigma$ given that σ_f represents an assignment of residual probability:

$$u_\ell^\epsilon(\sigma_\ell, \sigma_f) = \sum_{h \in \mathcal{Z}: \sigma_\ell(h) = \sigma_\ell \wedge \sigma_f(h) = \sigma_f} \eta(\sigma_f) u_\ell(h) + \sum_{\hat{\sigma}_f \in \Sigma_f} \xi_f(\epsilon, \hat{\sigma}_f) u_\ell(\sigma_\ell, \hat{\sigma}_f).$$

The following LP finds an SEFCE in a ξ-perturbed SEFG.

$$\max_{p,v} \sum_{(\sigma_\ell, \sigma_f) \in \Sigma} p(\sigma_\ell, \sigma_f) u_\ell^\epsilon(\sigma_\ell, \sigma_f) \quad \text{s.t.} \tag{6a}$$

$$p(\sigma_\emptyset, \sigma_\emptyset) = 1, \quad p(\sigma_\ell, \sigma_f) \geq 0 \qquad \forall (\sigma_\ell, \sigma_f) \in \Sigma \tag{6b}$$

$$\sum_{\sigma_f \in rel(\sigma_\ell)} p(\sigma_\ell, \sigma_f) \geq \xi_\ell(\epsilon, \sigma_\ell) \qquad \forall \sigma_\ell \in \Sigma_\ell \tag{6c}$$

$$p(\sigma_\ell(I), \sigma_f) = \sum_{a \in A(I)} p(\sigma_\ell(I)a, \sigma_f) \quad \forall I \in \mathcal{I}_\ell, \sigma_f \in rel(I) \tag{6d}$$

$$p(\sigma_\ell, \sigma_f(I)) = \sum_{a \in A(I)} p(\sigma_\ell, \sigma_f(I)a) \quad \forall I \in \mathcal{I}_f, \sigma_\ell \in rel(I) \tag{6e}$$

$$v(\sigma_f) = \eta(\sigma_f) \sum_{\sigma_\ell \in rel(\sigma_f)} p(\sigma_\ell, \sigma_f) u_f(\sigma_\ell, \sigma_f) \; + \; \sum_{I \in \mathcal{I}_f : \sigma_f(I) = \sigma_f} \sum_{a \in A(I)} v(\sigma_f a) \qquad \forall \sigma_f \in \Sigma_f \tag{6f}$$

$$v(I, \sigma_f) \geq \eta(\sigma_f(I)a) \sum_{\sigma_\ell \in rel(\sigma_f)} p(\sigma_\ell, \sigma_f) u_f(\sigma_\ell, \sigma_f(I)a) + \sum_{\hat{I} \in \mathcal{I}_f ; \sigma_f(\hat{I}) = \sigma_f(I)a} v(\hat{I}, \sigma_f) \qquad \forall I \in \mathcal{I}_f, a \in A(I), \sigma_f \in prec(I) \tag{6g}$$

$$v(\sigma_f(I)a) = v(I, \sigma_f(I)a) \qquad \forall I \in \mathcal{I}_f, a \in A(I). \tag{6h}$$

In (6g) of this LP, $prec(I)$, where $I \in \mathcal{I}_f$, is the set of follower's sequences σ_f that precede I in the sense that there is $\hat{I} \in \mathcal{I}_f$ with $\sigma_f(\hat{I}) \sqsubseteq \sigma_f(I)$ and $\sigma_f = \sigma_f(\hat{I})a$ for some $a \in A(\hat{I})$. This LP is a modification of the SEFCE LP given by Čermák et al. (2016). The new LP has two modifications to allow perturbation. First, it has constraints (6c) to ensure that the sum of recommendation probabilities on any leader's sequence is at least $\xi_\ell(\epsilon, \sigma_\ell)$. Second, because we are now recommending where to send residual probability for the follower, we must modify the objective in order to give the correct expected value for the leader.[7]

We can branch-and-bound on recommendations to the follower in a way that ensures that the final outcome is an SSE. That is guaranteed by the following theorem, which shows that we can add and remove constraints on which follower actions to recommend in a way that guarantees an SSE of the perturbed game as long as the follower is recommended a "pure" strategy with respect to the residual probabilities.

Theorem 3. *If a solution to LP (6) is such that for all $I \in \mathcal{I}_f$ there exists $a \in A(I)$ such that $p(\sigma_\ell, \sigma_f(I)\hat{a}) = 0$ for all $\hat{a} \in A(I), \sigma_\ell \in rel(\sigma_f(I)a)$ with $\hat{a} \neq a$, then a strategy profile can be extracted in polynomial time such that it is an SSE of the perturbed game instance.*

Now it is obvious that the LP (6) upper bounds the value of any SSE since the SSE is a feasible solution to the LP.

Theorem 3 shows that one way to find an SSE is to find a solution to LP (6) where the follower is recommended a pure strategy with respect to the residual probabilities. Since any SSE represents such a solution, we can branch on which actions we make pure at each information set, and use branch-and-bound to prune the space of possible solutions. This approach was proposed by Čermák et al. (2016) for computing SSEs in unperturbed games, where they showed that it performs better than a single MIP. Because our LP for perturbed games uses residual probabilities for the follower, we can apply the branching methodology of Čermák et al. (2016). At each node in the search we choose some information set I where more than one action is recommended. We then

[7] We use the definition of relevant sequences and the LP from von Stengel and Forges (2008) rather than those of Čermák et al. (2016). The latter are not well defined for (6d) and (6e).

branch on which action in $A(I)$ to recommend. Forcing a given action is accomplished by requiring all other action probabilities be zero. Our branch-and-bound chooses information sets according to depth, always branching on the shallowest one with at least two recommended action. We explore actions in descending order of mass, where the mass on $a \in A(I)$ (with sequence σ_f) is $\sum_{\sigma_\ell \in rel(\sigma_f)} p(\sigma_\ell, \sigma_f)$.

The algorithm finds an SSE of the perturbed game. In the limit as the perturbation approaches zero, this yields a QPSE. No algorithm is currently known for computing such an exact limit. In practice, we pick a small perturbation and solve the branch-and-bound using that value. This immediately leads to an approximate notion of QPSE (akin to approximate refinement notions in non-Stackelberg EFGs (Farina, Kroer, and Sandholm 2017; Kroer, Farina, and Sandholm 2017)). Another approach is to use our algorithm as an anytime algorithm where one runs it repeatedly with smaller and smaller perturbation values.

8 Experiments

We conducted experiments with our algorithm on two common benchmark EFGs. The first is a search game played on the graph shown in Figure 2. It is a simultaneous-move game (which can be modeled as a turn-taking EFG with appropriately chosen information sets). The leader controls two patrols that can each move within their respective shaded areas (labeled P1 and P2), and at each time step the controller chooses a move for both patrols. The follower is always at a single node on the graph, initially the leftmost node labeled S and can move freely to any adjacent node (except at patrolled nodes, the follower cannot move from a patrolled node to another patrolled node). The follower can also choose to wait in place for a time step in order to clean up their traces. If a patrol visits a node that was previously visited by the follower, and the follower did not wait to clean up their traces, they can see that the follower was there. If the follower reaches any of the rightmost nodes they received the respective payoff at the node (5 and 10, respectively). If the follower and any patrol are on the same node at any time step, the follower is captured, which leads to a payoff of 0 for the follower and a payoff of 1 for the leader. Finally, the game times out after k simultaneous moves, in which case the leader receives payoff 0 and the follower receives $-\infty$ (because we are interested in games where the follower attempts to reach an end node). This is the game considered by Kroer, Farina, and Sandholm (2018) except with the bottom layer removed, and is similar to games considered by Bošanskỳ et al. (2014) and Bošanskỳ and Čermák (2015).

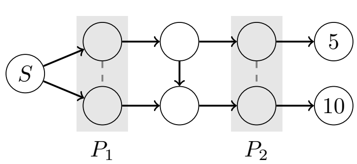

Figure 2: The graph on which the search game is played.

The second game is a variant on Goofspiel (Ross 1971), a bidding game where each player has a hand of cards numbered 1 to 3. There are 3 prizes worth $1, \ldots, 3$. In each turn, the prize is the smallest among the remaining prizes. Within the turn, the each of two players simultaneously chooses some private card to play. The player with the larger card wins the prize. In case of a tie, the prize is discarded, so this is not a constant-sum game. The cards that were played get discarded. Once all cards have been played, a player's score is the sum of the prizes that she has won.

The LP solver we used is GLPK 4.63 (GLPK 2017). We had to make the following changes to GLPK. First, we had to expose some internal routines so that we could input to the solver rational numbers rather than double-precision numbers. Second, we fixed a glitch in GLPK's rational LP solver in its pivoting step (it was not correct when the rational numbers were too small). Our code and GLPK use the GNU GMP library to provide arbitrary-precision arithmetic. The code, written in the C++14 language, was compiled with the g++ 7.2.0 compiler. It was run on a single thread on a 2.3 GHz Intel Xeon processor. The results are shown in Figure 3.

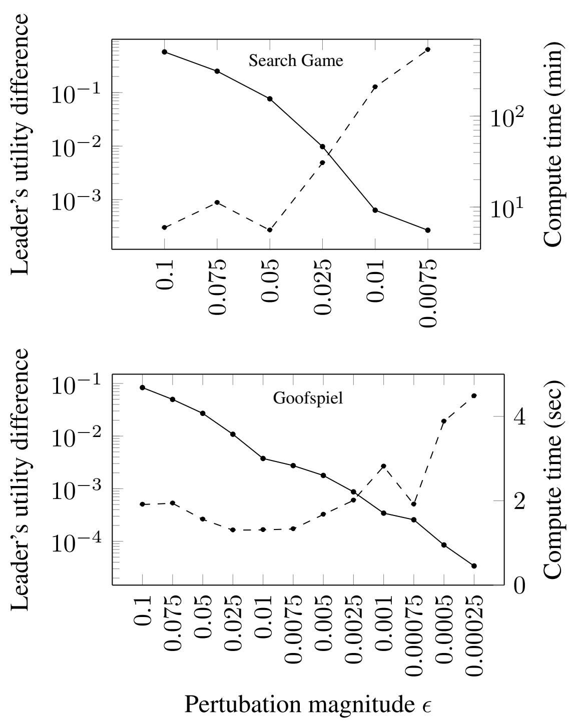

Figure 3: Experiments. Dashed lines show compute time. Solid lines show the loss in the leader's utility compared to the SSE value in the unperturbed game.

9 Conclusions and Future Research

Quasi-perfect equilibrium has been studied in extensive-form games, but was poorly understood in Stackelberg settings. We provided a game-theoretic, axiomatic definition of *quasi-perfect Stackelberg equilibrium (QPSE)*. We developed a family of game perturbation schemes that lead to a

QPSE in the limit. Our family generalizes prior perturbation schemes introduced for finding (even non-Stackelberg) quasi-perfect equilibria. Using our perturbation schemes, we developed a branch-and-bound algorithm for QPSE. It leverages a perturbed variant of the linear program for computing a Stackelberg extensive-form correlated equilibrium. Experiments show that our algorithm can be used to find an approximate QPSE in games with thousands of nodes.

We showed that some perturbation schemes outside our family do not lead to QPSEs in some games. It remains an open question whether our perturbation family fully characterizes the whole set of QPSEs. As to requirement (i) in Definition 6, can all QPSEs be captured by perturbation schemes that only use polynomial lower bounds on trembles?

It was recently shown that in non-Stackelberg extensive-form games, there exists a perturbation size that is small enough (while still strictly positive) that an exact refined (e.g., quasi-perfect) equilibrium can be found by solving a mathematical program with that perturbation size (Miltersen and Sørensen 2010; Farina and Gatti 2017; Farina, Gatti, and Sandholm 2018), and Farina, Gatti, and Sandholm (2018) provide an algorithm for checking whether a given guess of perturbation size is small enough. That obviates the need to try to explicitly compute a limit of a sequence. It would be interesting to see whether such theory can also be developed for Stackelberg extensive-form games—and for our perturbation family in particular.

Acknowledgements

This material is based on work supported by the National Science Foundation under grants IIS-1718457, IIS-1617590, and CCF-1733556, and the ARO under award W911NF-17-1-0082. Christian Kroer was also sponsored by a Facebook Fellowship.

References

Bošanskỳ, B., and Čermák, J. 2015. Sequence-form algorithm for computing Stackelberg equilibria in extensive-form games. In *AAAI*, 805–811.

Bošanskỳ, B.; Kiekintveld, C.; Lisý, V.; and Pěchouček, M. 2014. An exact double-oracle algorithm for zero-sum extensive-form games with imperfect information. *Journal of Artificial Intelligence Research* 829–866.

Čermák, J.; Bošanskỳ, B.; Durkota, K.; Lisý, V.; and Kiekintveld, C. 2016. Using correlated strategies for computing Stackelberg equilibria in extensive-form games. In *AAAI*, 439–445.

Conitzer, V., and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *EC*, 82–90.

Farina, G., and Gatti, N. 2017. Extensive-form perfect equilibrium computation in two-player games. In *AAAI*, 502–508.

Farina, G.; Marchesi, A.; Kroer, C.; Gatti, N.; and Sandholm, T. 2018. Trembling-hand perfection in extensive-form games with commitment. In *IJCAI*, 233–239.

Farina, G.; Gatti, N.; and Sandholm, T. 2018. Practical exact algorithm for trembling-hand equilibrium refinements in games. In *AAMAS-IJCAI workshop on Agents and Incentives in Artificial Intelligence (AI3)*.

Farina, G.; Kroer, C.; and Sandholm, T. 2017. Regret minimization in behaviorally-constrained zero-sum games. In *ICML*, 1107–1116.

GLPK. 2017. Gnu linear programming kit, version 4.63.

Kroer, C.; Farina, G.; and Sandholm, T. 2017. Smoothing method for approximate extensive-form perfect equilibrium. In *IJCAI*, 295–301.

Kroer, C.; Farina, G.; and Sandholm, T. 2018. Robust Stackelberg equilibria in extensive-form games and extension to limited lookahead. In *AAAI*, 1130–1137.

Kuhn, H. W. 1953. Extensive games and the problem of information. *Annals of Mathematics Study 28* 193–216.

Letchford, J., and Conitzer, V. 2010. Computing optimal strategies to commit to in extensive-form games. In *EC*, 83–92.

Marchesi, A.; Farina, G.; Kroer, C.; Gatti, N.; and Sandholm, T. 2018. Quasi-perfect stackelberg equilibrium. *CoRR* abs/1811.03871.

Mertens, J.-F. 1995. Two examples of strategic equilibrium. *Games and Economic Behavior* 8(2):378–388.

Miltersen, P. B., and Sørensen, T. B. 2008. Fast algorithms for finding proper strategies in game trees. In *SODA*, 874–883.

Miltersen, P. B., and Sørensen, T. B. 2010. Computing a quasi-perfect equilibrium of a two-player game. *Economic Theory* 42(1):175–192.

Romanovskii, I. 1962. Reduction of a game with complete memory to a matrix game. *Soviet Mathematics* 3.

Ross, S. M. 1971. Goofspielthe game of pure strategy. *Journal of Applied Probability* 8(3):621–625.

Selten, R. 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory* 4(1):25–55.

Shoham, Y., and Leyton-Brown, K. 2008. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

Sørensen, T. B. 2012. Computing a proper equilibrium of a bimatrix game. In *EC*, 916–928.

Tambe, M. 2011. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press.

Van Damme, E. 1984. A relation between perfect equilibria in extensive form games and proper equilibria in normal form games. *International Journal of Game Theory* 13(1):1–13.

van Damme, E. 1987. *Stability and Perfection of Nash Equilibria*. Berlin, Heidelberg: Springer-Verlag.

von Stackelberg, H. 1934. *Marktform und Gleichgewicht*. Springer, Vienna.

von Stengel, B., and Forges, F. 2008. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research* 33(4):1002–1022.

Von Stengel, B. 1996. Efficient computation of behavior strategies. *Games and Economic Behavior* 14(2):220–246.