Trembling-Hand Perfection and Correlation in Sequential Games

Alberto Marchesi, Nicola Gatti
Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milan, Italy
alberto.marchesi@polimi.it, nicola.gatti@polimi.it

Abstract

We initiate the study of trembling-hand perfection in sequential (*i.e.*, extensive-form) games with correlation. We introduce the *extensive-form perfect correlated equilibrium* (EFPCE) as a refinement of the classical *extensive-form correlated equilibrium* (EFCE) that amends its weaknesses off the equilibrium path. This is achieved by accounting for the possibility that players may make mistakes while following recommendations independently at each information set of the game. After providing an axiomatic definition of EFPCE, we show that one always exists since any perfect (Nash) equilibrium constitutes an EFPCE, and that it is a refinement of EFCE, as any EFPCE is also an EFCE. Then, we prove that, surprisingly, computing *an* EFPCE is *not* harder than finding an EFCE, since the problem can be solved in polynomial time for general n-player extensive-form games (also with chance). This is achieved by formulating the problem as that of finding a limit solution (as $\epsilon \rightarrow 0$) to a suitably defined *trembling* LP parametrized by ϵ, featuring exponentially many variables and polynomially many constraints. To this end, we show how a recently developed polynomial-time algorithm for trembling LPs can be adapted to deal with problems having an exponential number of variables. This calls for the solution of a sequence of (non-trembling) LPs with exponentially many variables and polynomially many constraints, which is possible in polynomial time by applying an ellipsoid against hope approach.

Introduction

Nash equilibrium (NE) (Nash 1951) computation in 2-player zero-sum games has been the flagship challenge in artificial intelligence for several years (see, *e.g.*, landmark results in poker (Brown and Sandholm 2018, 2019)). Recently, increasing attention has been devoted to multi-player games, where equilibria based on *correlation* are now mainstream.

Correlation in games is customarily modeled through a trusted external mediator that privately recommends actions to the players. The mediator acts as a *correlation device* that draws action recommendations according to a publicly known distribution. The seminal notion of *correlated equilibrium* (CE) introduced by Aumann (1974) requires that no player has an incentive to deviate from a recommendation. This is encoded by NE conditions applied to an *extended* game where the correlation device plays first by randomly selecting a profile of actions according to the public distribution; then, the original game is played with each player being informed only of the action selected for her. CEs are computationally appealing since they can be implemented in a *decentralized* way by letting players play independently according to no-regret procedures (Hart and Mas-Colell 2000).

Computing CEs in sequential (*i.e.*, extensive-form) games with imperfect information has received considerable attention in the last years (Celli et al. 2019; Farina, Bianchi, and Sandholm 2020). In this context, various CE definitions are possible, depending on the ways recommendations are revealed to the players. The one that has emerged as the most suitable for sequential games is the *extensive-form correlated equilibrium* (EFCE) of Von Stengel and Forges (2008). The key feature of EFCE is that recommendations are revealed to the players only when they reach a decision point where the action is to be played, and, if one player defects from a recommendation, then she stops receiving them in the future. Von Stengel and Forges (2008) show that EFCEs can be characterized by a polynomially-sized *linear program* (LP) in two-player games without chance. In the same restricted setting, Farina et al. (2019a) show how to find an EFCE by solving a bilinear saddle-point problem, which can be exploited to derive an efficient no-regret algorithm (Farina et al. 2019b). In general n-player games, Huang and von Stengel (2008) prove that *an* EFCE can be computed in polynomial time by means of an *ellipsoid against hope* (EAH) algorithm similar to that introduced by Papadimitriou and Roughgarden (2008) for CEs in compactly represented games (see also (Gordon, Greenwald, and Marks 2008) for another algorithm). Instead, finding a payoff-maximizing EFCE is NP-hard (Von Stengel and Forges 2008). Very recently, Celli et al. (2020) provide an efficient no-regret procedure for EFCE in n-player games.

One of the crucial weaknesses of standard equilibrium notions, such as NE, in sequential games is that they may prescribe to play sub-optimally off the equilibrium path, *i.e.*, at those information sets never reached when playing equilibrium strategies. One way to amend this issue is *trembling-hand perfection* (Selten 1975), whose rationale is to let players reasoning about the possibility that they may make mistakes in the future, playing sub-optimal actions with small, vanishing probabilities (a.k.a. trembles). This idea leads to the NE refinement known as *perfect equilibrium* (PE) (Selten 1975). Other refinements have been introduced in the literature; *e.g.*, in the *quasi-perfect equilibrium* of Van Damme (1984) players only account for opponents' future trembles (see (Van Damme 1991) for other examples). Trembles can also be introduced in normal-form games, leading to robust equilibria that rule out weakly dominated strategies (Hillas and Kohlberg 2002). Recently, equilibrium refinement has been addressed beyond the NE case, such as, *e.g.*, in Stackelberg settings (Farina et al. 2018; Marchesi et al. 2019).

Trembling-hand perfection for CEs has only been studied from a theoretical viewpoint in normal-form games, by Dhillon and Mertens (1996). The authors introduce the concept of *perfect* CE by enforcing PE conditions in the extended game, rather than NE ones. Despite equilibrium refinements in sequential games are ubiquitous, no work addressed perfection and correlation together in such setting. [1]

Original Contributions We give an axiomatic definition of *extensive-form perfect correlated equilibrium* (EFPCE), enforcing PE conditions, rather than NE ones, in the extended game introduced by Von Stengel and Forges (2008) for their original definition of EFCE. Intuitively, this accounts for the possibility that players may make mistakes while following recommendations independently at each information set of the game. Trembles are introduced on players' strategies, while the correlation device is defined as in classical CE notions. First, we show that an EFPCE always exists, since any PE constitutes an EFPCE, and that EFPCE is a refinement of EFCE, as any EFPCE is also an EFCE. Then, we show how *an* EFPCE can be computed in polynomial time in any n-player extensive-form game (also with chance). At first, we introduce a characterization of the equilibria of perturbed extended games (*i.e.*, extended games with trembles) inspired by the definition of EFCE based on *trigger agents*, introduced by Gordon, Greenwald, and Marks (2008) and Farina et al. (2019a). This result allows us to formulate the EFPCE problem as that of finding a limit solution (as $\epsilon \to 0$) to a suitably defined *trembling* LP parametrized by ϵ, featuring exponentially many variables and polynomially many constraints. To this end, we show how the polynomial-time algorithm for trembling LPs developed by Farina, Gatti, and Sandholm (2018) can be adapted to deal with problems having an exponential number of variables. This calls for the solution of a sequence of (non-trembling) LPs with exponentially many variables and polynomially many constraints, which is possible in polynomial time by applying an EAH approach. The latter is inspired by the analogous algorithm of Huang and von Stengel (2008) for EFCEs, which is adapted to deal with a different set of dual constraints, requiring a modification of the polynomial-time separation oracle of Huang and von Stengel (2008). [2]

[1] Applying the perfect CE by Dhillon and Mertens (1996) to the normal-form representation of a sequential game does *not* generally solve equilibrium weaknesses. This would lead to a correlated version of the *normal-form* PE, which is known not to guard against sub-optimality off the equilibrium path (Van Damme 1991).

[2] All the proofs are in (Marchesi and Gatti 2020).

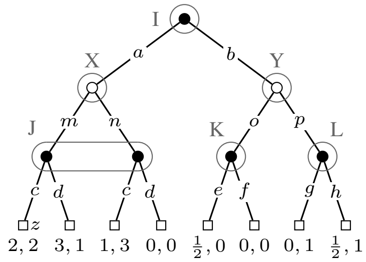
Preliminaries

Extensive-Form Games

We focus on n-player *extensive-form games* (EFGs) with imperfect information. We let $N := \{1, \dots, n\}$ be the set of players, and, additionally, we let c be the *chance* player representing exogenous stochasticity. The sequential structure is encoded by a game tree with node set H. Each node $h \in H$ is identified by the ordered sequence $\sigma(h)$ of actions encountered on the path from the root to h. We let $Z \subseteq H$ be the subset of terminal nodes, which are the leaves of the game tree. For every non-terminal node $h \in H \setminus Z$, we let $P(h) \in N \cup \{c\}$ be the player who acts at h, while $A(h)$ is the set of actions available. The function $p_c : Z \to (0, 1]$ defines the probability of reaching each terminal node given the chance moves on the path from the root to that node. For every player $i \in N$, the function $u_i : Z \to \mathbb{R}$ encodes player i's utilities over terminal nodes. Imperfect information is modeled through *information sets* (infosets). An infoset $I \subseteq H \setminus Z$ of player $i \in N$ is a group of player i's nodes indistinguishable for her, *i.e.*, for every $h \in I$, it must be the case that $P(h) = i$ and $A(h) = A(I)$, where $A(I)$ is the set of actions available at the infoset. W.l.o.g., we assume that the sets $A(I)$ are disjoint. We denote with $\mathcal{I}_i$ the collection of infosets of player $i \in N$. For every $i \in N$, we let $A_i := \bigcup_{I \in \mathcal{I}_i} A(I)$ be the set of all player i's actions. Moreover, we let $A := \bigcup_{i \in N} A_i$. We focus on EFGs with *perfect recall* in which no player forgets what she did or knew in the past. Formally, for every player $i \in N$ and infoset $I \in \mathcal{I}_i$, it must be that every node $h \in I$ is identified by the same ordered sequence $\sigma_i(I)$ of player i's actions from the root to that node. Given two infosets $I, J \in \mathcal{I}_i$ of player $i \in N$, we say that J *follows* I, written $I \prec J$, if there exist two nodes $h \in I$ and $k \in J$ such that h is on the path from the root to k. By perfect recall, $\prec$ is a partial order on $\mathcal{I}_i$. We also write $I \preceq J$ whenever either $I = J$ or $I \prec J$. For every infoset $I \in \mathcal{I}_i$, we let $\mathcal{C}(I, a) \subseteq \mathcal{I}_i$ be the set of all infosets that immediately follow I by playing action $a \in A(I)$.

Strategies A player's *pure strategy* specifies an action at every infoset of her. For every $i \in N$, the set of player i's pure strategies π_i is $\Pi_i := \bigtimes_{I \in \mathcal{I}_i} A(I)$, with $\pi_i(I) \in A(I)$ being the action at infoset $I \in \mathcal{I}_i$. Moreover, $\Pi := \bigtimes_{i \in N} \Pi_i$ denotes the set of *strategy profiles* specifying a strategy for each player, while, for $i \in N$, we let $\Pi_{-i} := \bigtimes_{j \neq i \in N} \Pi_j$ be the (partial) strategy profiles defining a strategy for each player other than i. Given $\pi_i \in \Pi_i$ and $a \in A_i$, we write $a \in \pi_i$ whenever π_i prescribes to play a. Analogously, for $\pi \in \Pi$ and $a \in A$, we write $a \in \pi$. Players are allowed to randomize over pure strategies by playing *mixed strategies*. For $i \in N$, we let $\mu_i : \Pi_i \to [0, 1]$ be a player i's mixed strategy, where $\sum_{\pi_i \in \Pi_i} \mu_i(\pi_i) = 1$. The perfect recall assumption allows to work with *behavior strategies*, which define probability distributions locally at each infoset. For $i \in N$, we let $\beta_i : A_i \to [0, 1]$ be a player i's behavior strategy, which is such that $\sum_{a \in A(I)} \beta_i(a) = 1$ for all $I \in \mathcal{I}_i$. [3]

[3] EFGs with perfect recall admit a compact strategy representation called *sequence form* (Von Stengel 1996). See the Appendix A



	I J K L		I J K L
π_1	$a\,c\,e\,g$	π_9	$b\,c\,e\,g$
π_2	$a\,c\,e\,h$	π_{10}	$b\,c\,e\,h$
π_3	$a\,c\,f\,g$	π_{11}	$b\,c\,f\,g$
π_4	$a\,c\,f\,h$	π_{12}	$b\,c\,f\,h$
π_5	$a\,d\,e\,g$	π_{13}	$b\,d\,e\,g$
π_6	$a\,d\,e\,h$	π_{14}	$b\,d\,e\,h$
π_7	$a\,d\,f\,g$	π_{15}	$b\,d\,f\,g$
π_8	$a\,d\,f\,h$	π_{16}	$b\,d\,f\,h$

$$\Pi_1(a) = \{\pi_1, \dots, \pi_8\}$$
$$\Pi_1(f) = \{\pi_3, \pi_4, \pi_7, \pi_8, \pi_{11}, \pi_{12}, \pi_{15}, \pi_{16}\}$$
$$\Pi_1(\text{I}) = \{\pi_1, \dots, \pi_{16}\}$$
$$\Pi_1(\text{J}) = \{\pi_1, \dots, \pi_8\}$$
$$\Pi_1(\text{K}) = \{\pi_9, \dots, \pi_{16}\}$$
$$\Pi_1(\text{L}) = \{\pi_9, \dots, \pi_{16}\}$$
$$\Pi_1(\text{K}, f) = \{\pi_3, \pi_4, \pi_7, \pi_8\}$$
$$\Pi_1(\text{L}, g) = \{\pi_9, \pi_{11}, \pi_{13}, \pi_{15}\}$$
$$\Pi_1(z) = \{\pi_1, \pi_2, \pi_3, \pi_4\}$$

Figure 1: (*Left*) Sample EFG. Black round nodes belong to player 1, white round nodes belong to player 2, and white square nodes are leaves (with players' payoffs specified under them). Rounded gray lines denote infosets. (*Center*) Set Π_1 of pure strategies for player 1. (*Right*) Examples of certain subsets of Π_1 used in this work.

Additional Notation We introduce some subsets of Π_i (see Figure 1 for some examples). For every action $a \in A_i$ of player $i \in N$, we define $\Pi_i(a) := \{\pi_i \in \Pi_i \mid a \in \pi_i\}$ as the set of player i's pure strategies specifying a. For every infoset $I \in \mathcal{I}_i$, we let $\Pi_i(I) \subseteq \Pi_i$ be the set of strategies that prescribe to play so as to reach I whenever possible (depending on players' moves up to that point) and *any* action whenever reaching I is *not* possible anymore. Additionally, for every action $a \in A(I)$, we let $\Pi_i(I, a) \subseteq \Pi_i(I) \subseteq \Pi_i$ be the set of player i's strategies that reach I and play a. Given a terminal node $z \in Z$, we denote with $\Pi_i(z) \subseteq \Pi_i$ the set of strategies by which player i plays so as to reach z, while $\Pi(z) := \bigtimes_{i\in N}\Pi_i(z)$ and $\Pi_{-i}(z) := \bigtimes_{j\neq i\in N}\Pi_j(z)$. We also introduce the following subsets of Z. For every $i \in N$ and $I \in \mathcal{I}_i$, we let $Z(I) \subseteq Z$ be the set of terminal nodes reachable from infoset I of player i. Moreover, $Z(I, a) \subseteq Z(I) \subseteq Z$ is the set of terminal nodes reachable by playing action $a \in A(I)$ at I, whereas $Z^{\perp}(I, a) := Z(I, a) \setminus \bigcup_{J\in\mathcal{C}(I,a)} Z(J)$ is the set of those reachable by playing a at I without traversing any other player i's infoset.

Nash Equilibrium and Its Refinements

Given an EFG, players' behavior strategies $\{\beta_i\}_{i\in N}$ constitute an NE if no player has an incentive to unilaterally deviate from the equilibrium by playing another strategy (Nash 1951). The PE defined by Selten (1975) relies on the idea of introducing *trembles* in the game, representing the possibility that players may take non-equilibrium actions with small, vanishing probability. Trembles are encoded by means of Selten's *perturbed games*, which force lower bounds on the probabilities of playing actions. Given an EFG Γ, a pair (Γ, η) defines a perturbed game, where $\eta : A \to (0, 1)$ is a function assigning a positive lower bound $\eta(a)$ on the probability of playing each action $a \in A$, with $\sum_{a\in A(I)} \eta(a) < 1$ for every $i \in N$ and $I \in \mathcal{I}_i$. Then:

Definition 1. *Given an EFG Γ, $\{\beta_i\}_{i\in N}$ is a PE of Γ if it is a limit point of NEs for at least one sequence of perturbed games $\{(\Gamma, \eta_t)\}_{t\in\mathbb{N}}$ such that, for all $a \in A$, the lower bounds $\eta_t(a)$ converge to zero as $t \to \infty$.*

There are only a few computational works on NE refinements. For instance, Miltersen and Sørensen (2010) characterize quasi-perfect equilibria of 2-player EFGs using the sequence form (see the recent work by Gatti, Gilli, and Marchesi (2020) for its extension to n-player games) and exploit this to compute an equilibrium by solving a linear complementarity problem with trembles defined as polynomials of some parameter treated symbolically. Farina and Gatti (2017) do the same for the PE. Recently, Farina, Gatti, and Sandholm (2018) provide a general framework for computing NE refinements in 2-player zero-sum EFGs in polynomial time. The authors show how to reduce the task to the more general problem of solving *trembling* LPs parametrized by some parameter ϵ, *i.e.*, finding their limit solutions as $\epsilon \to 0$. Then, they provide a general polynomial-time algorithm to find limit solutions to trembling LPs. Other works study the problem of computing (approximate) NE refinements in 2-player zero-sum EFGs by employing online convex optimization techniques (Kroer, Farina, and Sandholm 2017; Farina, Kroer, and Sandholm 2017).

in (Marchesi and Gatti 2020) for more details.

Correlation in Extensive-Form Games

We model a correlation device as a probability distribution $\mu \in \Delta_\Pi$. In the classical CE by Aumann (1974), the correlation device draws a strategy profile $\pi \in \Pi$ according to μ; then, it privately communicates π_i to each player $i \in N$. This notion of CE does *not* fit well to EFGs, as it requires the players to reason over the exponentially-sized set Π_i. Von Stengel and Forges (2008) introduced the EFCE to solve this issue. The first crucial feature of the EFCE is a different way of giving recommendations: the strategy π_i is revealed to player i as the game progresses, *i.e.*, the player is recommended to play the action $\pi_i(I)$ at infoset $I \in \mathcal{I}_i$ only when I is actually reached during play. The second key aspect characterizing EFCEs is that, whenever a player decides to defect from a recommended action at some infoset, then she may choose any move at her subsequent infosets and she stops receiving recommendations from the correlation device. The definition of EFCE introduced by Von Stengel and Forges (2008) (Definition 3) requires the introduction of the notion of *extended game* with a correlation device.

Definition 2. *Given an EFG Γ and a distribution $\mu \in \Delta_\Pi$, the extended game $\Gamma^{\text{ext}}(\mu)$ is a new EFG in which chance first selects $\pi \in \Pi$ according to μ, and, then, Γ is played with each player $i \in N$ receiving the recommendation to play $\pi_i(I)$ as a signal, whenever she reaches an infoset $I \in \mathcal{I}_i$.*

The signaling in $\Gamma^{\text{ext}}(\mu)$ induces a new infoset structure. Specifically, every infoset $I \in \mathcal{I}_i$ of the original game Γ corresponds to many, new infosets in $\Gamma^{\text{ext}}(\mu)$, one for each combination of possible action recommendations received at the infosets preceding I (this included). At each new infoset, player i can only distinguish among chance moves corresponding to strategy profiles $\pi \in \Pi$ that differ in the recommendations at infosets $J \in \mathcal{I}_i : J \preceq I$. Figure 2 shows a simple EFG with its corresponding extended game.

Definition 3. *Given an EFG Γ, $\mu \in \Delta_\Pi$ defines an EFCE of Γ if following recommendations is an NE of $\Gamma^{\text{ext}}(\mu)$.* [4]

Next, we introduce an equivalent characterization of EFCEs (Farina, Bianchi, and Sandholm 2020). It is based on the following concept of *trigger agent*, originally due to Gordon, Greenwald, and Marks (2008).

Definition 4. *Given an infoset $I \in \mathcal{I}_i$ of player $i \in N$, an action $a \in A(I)$, and a distribution $\hat{\mu}_i \in \Delta_{\Pi_i(I)}$, an* $(I, a, \hat{\mu}_i)$-trigger agent for player i *is an agent that takes on the role of player i and follows all recommendations unless she reaches I and gets recommended to play a. If this happens, she stops committing to recommendations and plays according to a strategy sampled from $\hat{\mu}_i$ until the game ends.*

Then, it follows that $\mu \in \Delta_\Pi$ is an EFCE if, for every $i \in N$, player i's expected utility when following recommendations is at least as large as the expected utility that any $(I, a, \hat{\mu}_i)$-trigger agent for player i can achieve (assuming the opponents' do not deviate). A formal statement is provided in Appendix B in (Marchesi and Gatti 2020).

Computing EFCEs in n-player EFGs The algorithm of Huang and von Stengel (2008) relies on the following LP formulation of the problem of finding an EFCE, which has exponentially many variables and polynomially many constraints (see also Appendix C in (Marchesi and Gatti 2020)).

$$\max_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{v}} \quad \sum_{\pi \in \Pi} \mu[\pi] \quad \text{s.t.} \tag{1a}$$

$$A\boldsymbol{\mu} + B\boldsymbol{v} \geq \mathbf{0}, \tag{1b}$$

where $\boldsymbol{\mu}$ is a vector of variables $\mu[\pi]$ for $\pi \in \Pi$, encoding a probability distribution $\mu \in \Delta_\Pi$. Problem 1 does not enforce any simplex constraint on variables $\mu[\pi]$, and, thus, it is either unbounded or it has an optimal solution with value zero (by setting $\boldsymbol{\mu}$ and $\boldsymbol{v}$ to zero). In the former case, any feasible $\boldsymbol{\mu}$ encodes an EFCE after normalizing it. As a result, since an EFCE always exists (Von Stengel and Forges 2008), the following dual of Problem 1 is always infeasible:

$$A^\top \boldsymbol{y} \leq -\mathbf{1} \tag{2a}$$

$$B^\top \boldsymbol{y} = \mathbf{0} \tag{2b}$$

$$\boldsymbol{y} \geq \mathbf{0}, \tag{2c}$$

[4]For EFCEs, one can restrict the attention to distributions μ over *reduced* strategy profiles, *i.e.*, those in which each player's pure strategy only specifies actions at infosets reachable given that player's moves (Vermeulen and Jansen 1998). We stick to unreduced strategy profiles since, as shown in Appendix D in (Marchesi and Gatti 2020), they are necessary for trembling-hand perfect CEs to define the players' behavior off the equilibrium path.

where $\boldsymbol{y}$ is a vector of dual variables. The EAH approach applies the ellipsoid algorithm (Grötschel, Lovász, and Schrijver 1993) to Problem 2 in order to conclude that it is infeasible. Since there are exponentially many constraints, the algorithm runs in polynomial time only if a polynomial-time separation oracle is available. This is given by the following:

Lemma 1 (Lemma 5, (Huang and von Stengel 2008))**.** *If $\boldsymbol{y} \geq \mathbf{0}$ is such that $B^\top \boldsymbol{y} = \mathbf{0}$, then there exists $\boldsymbol{\mu}$ encoding a product distribution $\mu \in \Delta_\Pi$ such that $\boldsymbol{\mu}^\top A^\top \boldsymbol{y} = 0$. Moreover, $\boldsymbol{\mu}$ can be computed in polynomial time.*

Jiang and Leyton-Brown (2015) show how, given a product distribution μ computed as in Lemma 1, it is possible to recover, in polynomial time, a violated constraint for Problem 2, corresponding to some strategy profile $\pi \in \Pi$. This, together with some additional technical tricks ensuring that $B^\top \boldsymbol{y} = \mathbf{0}$ holds (see (Huang and von Stengel 2008) for more details), allows to apply the ellipsoid algorithm to Problem 2 in polynomial time. Since the problem is infeasible, the algorithm must terminate after polynomially many iterations with a collection of violated constraints, which correspond to polynomially many strategy profiles. Then, solving (in polynomial time) Problem 1 with the variables $\boldsymbol{\mu}$ restricted to these strategy profiles gives an EFCE of the game. Let us also remark that the EFCE obtained in this way has support size polynomial in the size of the game.

Trembling-Hand Perfection and Correlation

We are now ready to show how trembling-hand perfection can be injected into the definition of EFCE so as to amend its weaknesses off the equilibrium path (see the following for an example). We generalize the approach of Dhillon and Mertens (1996) (restricted to CEs in normal-form games) to the general setting of EFCEs in EFGs. The core idea is to use the PE rather than the NE in the definition of CE. Thus:

Definition 5. *Given an EFG Γ, a distribution $\mu \in \Delta_\Pi$ is an* extensive-form perfect correlated equilibrium (EFPCE) *if following recommendations is a PE of $\Gamma^{\text{ext}}(\mu)$.*

The definition of EFPCE crucially relies on the introduction of trembles in extended games, *i.e.*, it takes into account the possibility that each player may not follow action recommendations with a small, vanishing probability. In the following, given a perturbed EFG (Γ, η) and $\mu \in \Delta_\Pi$, we denote with $(\Gamma^{\text{ext}}(\mu), \eta)$ a perturbed extended game in which the probability of playing each action is subject to a lower bound equal to the lower bound $\eta(a)$ of the corresponding action $a \in A$ in Γ. By recalling the definition of PE (Definition 1) and the structure of perturbed extended games, it is easy to infer the following characterization of EFPCEs:

Lemma 2. *Given an EFG Γ, a distribution $\mu \in \Delta_\Pi$ is an EFPCE of Γ if following recommendations constitutes NEs for at least one sequence of perturbed extended games $\{\Gamma^{\text{ext}}(\mu), \eta_t)\}_{t \in \mathbb{N}}$ such that, for all $a \in A$, the lower bounds $\eta_t(a)$ converge to zero as $t \to \infty$.*

We remark that, with an abuse of terminology, we say that players follow recommendations in a perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta)$ whenever they play strategies which place

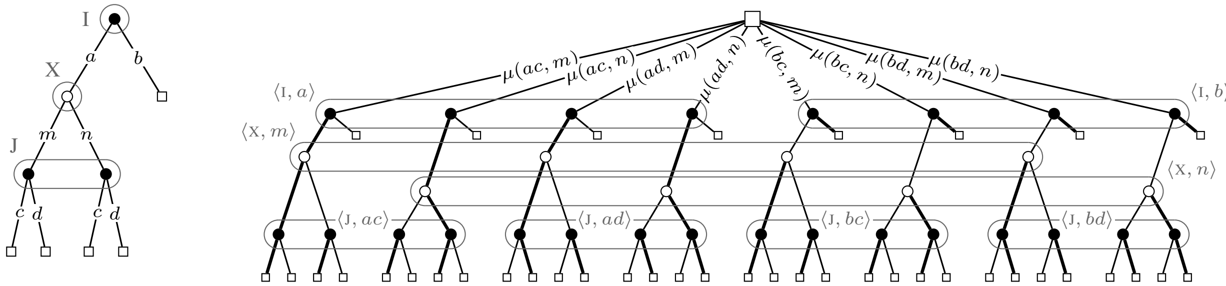


Figure 2: (*Left*) An EFG Γ. (*Right*) The extended game $\Gamma^{\text{ext}}(\mu)$. The white square node at the root is a chance node, where each action corresponds to some $\pi \in \Pi$ and is labeled with its probability $\mu(\pi)$. Infosets in $\Gamma^{\text{ext}}(\mu)$ are identified by pairs. For instance, infoset $\langle \text{J}, ac \rangle$ corresponds to J when being recommended to play a and c at I and J, respectively. Thick actions represent players' behavior when following recommendations (for the ease of reading, action names are omitted).

all the residual probability (given lower bounds) on recommended actions. In the following sections, we crucially rely on the characterization of EFPCEs given in Lemma 2 in order to derive our computational results. First, we show an example of EFPCE and prove some of its properties.

Example of EFPCE Consider the EFG in Figure 1(*Left*) and lower bounds $\eta_t : A \to (0, 1)$ for $t \in \mathbb{N}$, with $\eta_t(a) \to 0$ as $t \to \infty$ for all $a \in A$. First, notice that player 1 is always better off playing action a at the root infoset I, since she can guarantee herself a utility of 1 by selecting c at the following infoset J, while she can achieve at most $\frac{1}{2}$ by playing b. Thus, any EFPCE of the game (as well as any EFCE) must recommend a at I with probability 1. Then, in the sub-game reached when playing a at I, it is easy to check that recommending the pairs of actions (c, m), (c, n), and (d, m) each with probability $\frac{1}{3}$ is an equilibrium, as no player has an incentive to deviate from a recommendation, even with trembles (see Appendix E in (Marchesi and Gatti 2020)). The correlation device described so far is sufficient to define an EFCE, as recommendations at infosets Y, K, and L are *not* relevant given that they do not influence players' utilities at the equilibrium (b is never recommended). However, they become relevant for EFPCEs, since, in perturbed extended games, these infosets could be reached due to a tremble with probability $\eta_t(b)$. Then, player 2 must be told to play p at Y, because her utility is always 1 if she plays p, while it is always 0 for o. Moreover, with an analogous reasoning, player 1 must be recommended to play e and h at K and L, respectively. In conclusion, $\mu \in \Delta_\Pi : \mu(aceh, mp) = \mu(aceh, np) = \mu(adeh, mp) = \frac{1}{3}$ is an EFPCE.

Properties of EFPCEs We characterize the relation between EFPCEs and other equilibria, also showing that EFPCEs always exist and represent a refinement of EFCEs. [5]

Theorem 1. *This relation holds:* PE $\subseteq$ EFPCE $\subseteq$ EFCE.

Theorem 2. *The following relations hold:*

- EFPCE $\not\subset$ NE *and* NE $\not\subset$ EFPCE*;*
- EFPCE $\cap$ NE $=$ PE.

[5] In the following, we denote the sets of equilibria with their corresponding acronyms (*e.g.*, NE is the set of all NEs of a game).

NEs of Perturbed Extended Games

We provide a characterization of NEs of perturbed extended games $(\Gamma^{\text{ext}}(\mu), \eta)$, useful for our main algorithmic result on EFPCEs given in the following section. Specifically, we give a set of easily interpretable conditions which ensure that following recommendations is an NE of $(\Gamma^{\text{ext}}(\mu), \eta)$. These are crucial for the derivation of the LP exploited by our algorithm. Our characterization is inspired by that of EFCEs based on trigger agents (see Lemma 4 in Appendix B in (Marchesi and Gatti 2020)). However, the presence of trembles in extended games requires some key changes, which we highlight in the following.

First, we introduce some additional notation. Given a perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta)$, we let $\xi^\eta(z, \pi)$ be the probability of reaching a node $z \in Z$ when a strategy profile $\pi \in \Pi$ is recommended and players' obey to recommendations, in presence of trembles defined by η. Each $\xi^\eta(z, \pi)$ is obtained by multiplying probabilities of actions in $\sigma(z)$, which are those on the path from the root to z. For each $a \in \sigma(z)$, two cases are possible: either a is prescribed by the recommended π and played with its maximum probability given η, or it is *not*, which means that a tremble occurred with probability $\eta(a)$. Formally, letting $\mathbb{1}\{a \in \pi\}$ be an indicator for the event $a \in \pi$, for every $z \in Z$ and $\pi \in \Pi$:

$$\xi^\eta(z, \pi) := \prod_{a \in A : a \in \sigma(z)} \eta(a)^{\mathbb{1}\{a \in \pi\}} \tilde{\eta}(a)^{1 - \mathbb{1}\{a \in \pi\}} p_c(z),$$

where, for $a \in A(I)$, we let $\tilde{\eta}(a) := 1 - \sum_{a' \neq a \in A(I)} \eta(a')$ be the maximum probability assignable to a given η. Moreover, for every player $i \in N$, infoset $I \in \mathcal{I}_i$, terminal node $z \in Z(I)$ reachable from I, and strategy profile $\pi \in \Pi$, we let $\xi^\eta(z, I, \pi)$ be defined as $\xi^\eta(z, \pi)$ excluding player i's actions leading from I to z, *i.e.*, with the product restricted to actions $a \in \sigma_i(I) \cup (\sigma(z) \setminus A_i)$. Analogously, for every player i's strategy $\pi_i \in \Pi_i(I)$, we let $\xi^\eta(z, \pi_i)$ be defined for player i's actions $a \in A_i \cap (\sigma(z) \setminus \sigma_i(I))$ from I to z.

Following recommendations is an NE of the perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta)$ if, for every player $i \in N$, infoset $I \in \mathcal{I}_i$, and action $a \in (I)$, player i's utility when obeying to the recommendation a at I is at least as large as the utility achieved by any $(I, a, \hat{\mu}_i)$-trigger agent. The fundamental differences with respect to EFCE are: *(i)* an infoset I could be reached even when actions recommended at preceding infosets do not allow it (due to trembles); and *(ii)* trigger agents are subjects to trembles, which means that they may make mistakes while playing the strategy sampled from $\hat{\mu}_i$.

For any terminal node $z \in Z$, the probability of reaching it when following recommendations is:

$$q^\eta_\mu(z) \coloneqq \sum_{\pi \in \Pi} \xi^\eta(z, \pi)\mu(\pi),$$

where the summation accounts for the probability of reaching z for every possible π. The sum is over Π rather than $\Pi(z)$ as for EFCE, since, due to trembles, z could be reached even when $\pi \notin \Pi(z)$.

For any $(I, a, \hat{\mu}_i)$-trigger agent, the probability of reaching $z \in Z(I)$ when the agent 'gets triggered' is defined as:

$$p^{\eta,I,a}_{\mu,\hat{\mu}_i}(z) \coloneqq \left(\sum_{\substack{\pi_i \in \Pi_i(a) \\ \pi_{-i} \in \Pi_{-i}}} \xi^\eta(z, I, \pi)\mu(\pi) \right) \left(\sum_{\hat{\pi}_i \in \Pi_i(I)} \xi^\eta(z, \hat{\pi}_i)\hat{\mu}_i(\hat{\pi}_i) \right),$$

where the first summation is over $\Pi_i(a)$ instead of $\Pi_i(I, a)$ (as in the EFCE) since it might be the case that the agent is activated also when the recommended strategy π_i does *not* allow to reach infoset I. Finally, the overall probability of reaching $z \in Z(I)$ is:

$$y^{\eta,I,a}_{\mu,\hat{\mu}_i}(z) \coloneqq p^{\eta,I,a}_{\mu,\hat{\mu}_i}(z) + \sum_{\substack{\pi_i \in \Pi_i \setminus \Pi_i(a) \\ \pi_{-i} \in \Pi_{-i}}} \xi^\eta(z, \pi)\mu(\pi),$$

where the first term is for when the agent 'gets triggered', while the second term accounts for the case in which the agent is *not* activated (the two events are independent).

Theorem 3. *Given a perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta)$, following recommendations is an NE of the game if for every $i \in N$ and $(I, a, \hat{\mu}_i)$-trigger agent for player i, it holds that:*

$$\sum_{z \in Z(I)} \left[\left(\sum_{\substack{\pi_i \in \Pi_i(a) \\ \pi_{-i} \in \Pi_{-i}}} \xi^\eta(z, \pi)\mu(\pi) \right) u_i(z) \right] \geq \sum_{z \in Z(I)} p^{\eta,I,a}_{\mu,\hat{\mu}_i}(z)u_i(z).$$

Computing an EFPCE in n-player EFGs

We provide a polynomial-time algorithm to compute *an* EFPCE in n-player EFGs (also with chance). The algorithm is built on three fundamental components: *(i)* a trembling LP (with exponentially many variables and polynomially many constraints) whose limit solutions define EFPCEs; *(ii)* an adaption of the algorithm by Farina, Gatti, and Sandholm (2018) that finds such limit solutions by solving a sequence of (non-trembling) LPs; and *(iii)* a polynomial-time EAH procedure that solves these LPs.

Trembling LP for EFPCEs It resembles the EFCE LP in Problem 1. In this case, the constraints appearing in the LP ensure that following recommendations is an NE in a given sequence of perturbed extended games, by exploiting the characterization given in Theorem 3. Then, Lemma 2 allows to conclude that the limit solutions of the trembling LP define EFPCEs. In the following, we assume that a sequence of perturbed extended games $\{(\Gamma^{\text{ext}}(\mu), \eta_t)\}_{t \in \mathbb{N}}$ is given. For every player $i \in N$, infoset $I \in \mathcal{I}_p$, and action $a \in A(I)$, we introduce a variable $u[i, I, a]$ to encode player i's expected utility when following the recommendation to play a at I in the perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta_t)$. These variables are defined by the following constraints:

$$u[i, I, a] = \sum_{z \in Z(I)} \left(\sum_{\substack{\pi_i \in \Pi_i(a) \\ \pi_{-i} \in \Pi_{-i}}} \xi^\eta(z, \pi)\mu[\pi] \right) u_i(z) \qquad (3)$$
$$\forall i \in N, \forall I \in \mathcal{I}_i, \forall a \in A(I).$$

Then, we introduce constraints that recursively define variables $v[i, I, a, J]$ for every infoset $J \in \mathcal{I}_i : I \preceq J$. These encode the maximum expected utility obtained at infoset J by trigger agents associated with I and a. To this end, we also need some auxiliary non-negative variables $w[i, I, a, J, a']$, which are defined for every player $i \in N$, infoset $I \in \mathcal{I}_i$, action $a \in A(I)$, infoset $J \in \mathcal{I}_i : I \preceq J$ following I (this included), and action $a' \in A(J)$ available at J.

$$v[i, I, a, J] - w[i, I, a, J, a'] \geq \qquad (4)$$
$$\sum_{z \in Z^\perp(J,a')} \left(\sum_{\substack{\pi_i \in \Pi_i(a) \\ \pi_{-i} \in \Pi_{-i}(z)}} \xi^\eta(z, I, \pi)\mu[\pi] \right) u_i(z) +$$
$$\sum_{K \in \mathcal{C}(J,a')} \left(v[i, I, a, K] - \sum_{a'' \in A(K)} \eta_t(a'')w[i, I, a, K, a''] \right)$$
$$\forall i \in N, \forall I \in \mathcal{I}_i, \forall a \in A(I), \forall J \in \mathcal{I}_i : I \preceq J, \forall a' \in A(J).$$

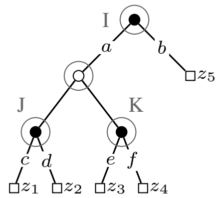
Intuitively, each auxiliary variable $w[i, I, a, J, a']$ represents a penalty on $v[i, I, a, J]$ due to the possibility of trembling by playing a (possibly) sub-optimal action $a' \in A(J)$ at J. Indeed, whenever a' is an optimal action at infoset J, then $w[i, I, a, J, a']$ is set to 0 in any solution; otherwise, $w[i, I, a, J, a']$ represents how much utility is lost by playing a' instead of an optimal action (see Figure 3(***Right***) for an example). Finally, the incentive constraints are:

$$u[i, I, a] = v[i, I, a, I] - \sum_{a' \in A(I)} \eta_t(a')w[i, I, a, I, a'] \quad (5)$$
$$\forall i \in N, \forall I \in \mathcal{I}_i, \forall a \in A(I).$$

Figure 3 provides an example of Constraints (4) and (5) to better clarify their meaning. The following theorem shows that Constraints (3), (5), and (4) correctly encode the conditions given in Theorem 3, which ensure that following recommendations is an NE in $(\Gamma^{\text{ext}}(\mu), \eta_t)$.

Theorem 4. *Given a perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta_t)$, if Constraints* (3), (4)*, and* (5) *can be satisfied for the vector $\boldsymbol{\mu}$ of variables $\mu[\pi]$ encoding the distribution μ, then following recommendations is an NE of $(\Gamma^{\text{ext}}(\mu), \eta_t)$.*

By substituting the expression of $u[i, I, a]$ (given by Constraints (3) and (5)) into Constraints (4), we can formulate



Constraints (4) and (5) for player 1, infoset I, and action a

	$u[1,\textsc{i},a] = v[1,\textsc{i},a,\textsc{i}] - \eta(a)w[1,\textsc{i},a,\textsc{i},a] - \eta(b)w[1,\textsc{i},a,\textsc{i},b]$		
$(\textsc{i},a)$:	$v[1,\textsc{i},a,\textsc{i}] - w[1,\textsc{i},a,\textsc{i},a] \geq v[1,\textsc{i},a,\textsc{j}] + v[1,\textsc{i},a,\textsc{k}] + -\eta(c)w[1,\textsc{i},a,\textsc{j},c] - \eta(d)w[1,\textsc{i},a,\textsc{j},d] - \eta(e)w[1,I\textsc{i},a,\textsc{k},e] - \eta(f)w[1,\textsc{i},a,\textsc{k},f]$		
$(\textsc{i},b)$	$v[1,\textsc{i},a,\textsc{i}] - w[1,\textsc{i},a,\textsc{i},b] \geq U_1(z_5)$		
$(\textsc{j},c)$:	$v[1,\textsc{i},a,\textsc{j}] - w[1,\textsc{i},a,\textsc{j},c] \geq U_1(z_1)$	$(\textsc{k},e)$:	$v[1,\textsc{i},a,\textsc{k}] - w[1,\textsc{i},a,\textsc{k},e] \geq U_1(z_3)$
$(\textsc{j},d)$:	$v[1,\textsc{i},a,\textsc{j}] - w[1,\textsc{i},a,\textsc{j},d] \geq U_1(z_2)$	$(\textsc{k},f)$:	$v[1,\textsc{i},a,\textsc{k}] - w[1,\textsc{i},a,\textsc{k},f] \geq U_1(z_4)$

Figure 3: (*Left*) Simple EFG. (*Right*) Example of Constraints (4)–(5); we let $U_1(z) \coloneqq \sum_{\pi_1 \in \Pi(a) \pi_2 \in \Pi_2(z)} \xi^\eta(z,\textsc{i},\pi)\mu[\pi]u_1(z)$ for every $z \in Z$. Variables $v[1,\textsc{i},a,\cdot]$ encode the optimal utility of trigger agents associated to $\textsc{i}, a$ at infosets following I (without trembles). Variables $w[1,\textsc{i},a,\cdot,\cdot]$ account for penalties due to trembles. To see this, fix $\mu[\pi]$. Assume that $U_1(z_4) > U_1(z_3)$ and consider the constraints for $(\textsc{k},e)$ and $(\textsc{k},f)$. Then, it must be $v[1,\textsc{i},a,\textsc{k}] = U_1(z_4)$ and $w[1,\textsc{i},a,\textsc{k},f] = 0$, which implies $w[1,\textsc{i},a,\textsc{k},e] = U_1(z_4) - U_1(z_3)$ (constraint for $(\textsc{i},a)$). Similarly, assuming $U_1(z_2) > U_1(z_1)$, it must be $v[1,\textsc{i},a,\textsc{j}] = U_1(z_2)$, $w[1,\textsc{i},a,\textsc{j},d] = 0$, and $w[1,\textsc{i},a,\textsc{j},c] = U_1(z_2) - U_1(z_1)$. An analogous reasoning holds at infosets upwards in the game tree.

the following trembling LP parameterized by $t \in \mathbb{N}$:

$$\max_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{v}, \boldsymbol{w} \geq \mathbf{0}} \quad \sum_{\pi \in \Pi} \mu[\pi] \quad \text{s.t.} \tag{6a}$$

$$A_t \boldsymbol{\mu} + B\boldsymbol{v} + C_t \boldsymbol{w} \geq \mathbf{0}, \tag{6b}$$

where A_t is the analogous of matrix A in Problem 1, $\boldsymbol{w}$ is a vector whose components are the variables $w[i,I,a,J,a']$, and C_t is a matrix defining the constraints coefficients for these variables. Notice that the coefficients of variables in $\boldsymbol{v}$ (as defined by B) are the same as in Problem 1.

Limit Solutions of Trembling LP Problem 6 can be cast into the framework of Farina, Gatti, and Sandholm (2018) by defining sequences of lower bounds η_t by means of vanishing polynomials in a parameter $\epsilon \to 0$. As a result, the polynomial-time algorithm by Farina, Gatti, and Sandholm (2018) can be used, with the only difference that, at each step, for a fixed value of the parameter ϵ (*i.e.*, particular lower bounds η_t), it needs to solve an instance of Problem 6 featuring exponentially many variables. Provided that the latter can be done in polynomial time, the polynomiality of the overall procedure is preserved, since the bounds on the running time provided by Farina, Gatti, and Sandholm (2018) do not depend on the number of variables in the LP.

EAH Procedure In order to solve Problem 6 for a particular lower bound function η_t in polynomial time, we can apply a procedure similar to the EAH algorithm by Huang and von Stengel (2008). Notice that Problem 6 is always unbounded, since there always exists a distribution $\mu \in \Delta_\Pi$ such that following recommendations is an NE of the perturbed extended game $(\Gamma^{\text{ext}}(\mu), \eta_t)$ (such μ is an EFCE of the corresponding perturbed, non-extended game). Thus, we only need to provide a polynomial-time separation oracle for the always-infeasible dual of Problem 6, which reads as:

$$A_t^\top \boldsymbol{y} \leq -\mathbf{1} \tag{7a}$$

$$B^\top \boldsymbol{y} = \mathbf{0} \tag{7b}$$

$$C_t^\top \boldsymbol{y} \geq \mathbf{0} \tag{7c}$$

$$\boldsymbol{y} \geq \mathbf{0}, \tag{7d}$$

where the vector of dual variables $\boldsymbol{y}$ has the same role as in Problem 2, since the constraints of the primal problems are indexed on the same sets. Notice that constraints $C_t^\top \boldsymbol{y} \geq \mathbf{0}$ are polynomially many. As a result, one can always check whether one of these constraints is violated in polynomial time and, if this is the case, output one such constraint as a violated inequality. This allows to focus on separation oracles for the other constraints. Then, the required one is given by the following lemma, an analogous of Lemma 1.

Lemma 3. *If $\boldsymbol{y} \geq \mathbf{0}$ is such that $B^\top \boldsymbol{y} = \mathbf{0}$, then there exists $\boldsymbol{\mu}$ encoding a product distribution $\mu \in \Delta_\Pi$ such that $\boldsymbol{\mu}^\top A_t^\top \boldsymbol{y} = 0$. Moreover, $\boldsymbol{\mu}$ can be computed in poly-time.*

The proof of Lemma 3 follows the same line as that of Lemma 5 by Huang and von Stengel (2008) (see (Huang 2011) for its complete version) and it is based on the CE existence proof by Hart and Schmeidler (1989).

Discussion and Future Works

We started the study of *trembling-hand perfection* in sequential games with correlation, introducing the EFPCE as a refinement of the EFCE that amends its weaknesses off the equilibrium path. This paves the way to a new research line, raising novel game-theoretic and computational challenges.

As for EFPCEs, an open question is whether compact correlated strategy representations, like the EFCE-based *correlation plan* by Von Stengel and Forges (2008), are possible in some restricted settings, such as 2-player games without chance. This would enable the optimization over the set of EFPCEs in polynomial time. The main challenge raised by EFPCEs with respect to EFCEs is that the former require to reason about general, un-reduced strategy profiles.

Another possible future work is to extend our analysis to other CE-based solution concepts, such as the *normal-form* CE and the *agent-form* CE (see (Von Stengel and Forges 2008) for their definitions). This raises the interesting question of how different trembling-hand-based CEs are able to amend weaknesses off the equilibrium path.

Finally, an interesting direction is to consider different ways of refining CE-based equilibria in sequential games, such as, *e.g.*, using *quasi-perfection* (Van Damme 1984).

Acknowledgments

This work has been partially supported by the Italian MIUR PRIN 2017 Project ALGADIMAR "Algorithms, Games, and Digital Market".

References

Aumann, R. J. 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics* 1(1): 67–96.

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

Brown, N.; and Sandholm, T. 2019. Superhuman AI for multiplayer poker. *Science* 365(6456): 885–890.

Celli, A.; Marchesi, A.; Bianchi, T.; and Gatti, N. 2019. Learning to correlate in multi-player general-sum sequential games. In *Advances in Neural Information Processing Systems*, 13076–13086.

Celli, A.; Marchesi, A.; Farina, G.; and Gatti, N. 2020. No-regret learning dynamics for extensive-form correlated equilibrium. *Advances in Neural Information Processing Systems* 33.

Dhillon, A.; and Mertens, J. F. 1996. Perfect correlated equilibria. *Journal of economic theory* 68(2): 279–302.

Farina, G.; Bianchi, T.; and Sandholm, T. 2020. Coarse Correlation in Extensive-Form Games. In *AAAI Conference on Artificial Intelligence*, 1934–1941.

Farina, G.; and Gatti, N. 2017. Extensive-Form Perfect Equilibrium Computation in Two-Player Games. In *AAAI Conference on Artificial Intelligence*, 502–508.

Farina, G.; Gatti, N.; and Sandholm, T. 2018. Practical exact algorithm for trembling-hand equilibrium refinements in games. In *Advances in Neural Information Processing Systems*, 5039–5049.

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

Farina, G.; Ling, C. K.; Fang, F.; and Sandholm, T. 2019a. Correlation in Extensive-Form Games: Saddle-Point Formulation and Benchmarks. In *Advances in Neural Information Processing Systems*, 9229–9239.

Farina, G.; Ling, C. K.; Fang, F.; and Sandholm, T. 2019b. Efficient Regret Minimization Algorithm for Extensive-Form Correlated Equilibrium. In *Advances in Neural Information Processing Systems*, 5187–5197.

Farina, G.; Marchesi, A.; Kroer, C.; Gatti, N.; and Sandholm, T. 2018. Trembling-hand perfection in extensive-form games with commitment. In *International Joint Conferences on Artificial Intelligence Organization*, 233–239.

Gatti, N.; Gilli, M.; and Marchesi, A. 2020. A characterization of quasi-perfect equilibria. *Games and Economic Behavior* 122: 240–255.

Gordon, G. J.; Greenwald, A.; and Marks, C. 2008. No-regret learning in convex games. In *International Conference on Machine Learning*, 360–367.

Grötschel, M.; Lovász, L.; and Schrijver, A. 1993. *Geometric algorithms and combinatorial optimization*. Springer.

Hart, S.; and Mas-Colell, A. 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68(5): 1127–1150.

Hart, S.; and Schmeidler, D. 1989. Existence of correlated equilibria. *Mathematics of Operations Research* 14(1): 18–25.

Hillas, J.; and Kohlberg, E. 2002. Foundations of strategic equilibrium. *Handbook of Game Theory with Economic Applications* 3: 1597–1663.

Huang, W. 2011. *Equilibrium computation for extensive games*. Ph.D. thesis, Citeseer.

Huang, W.; and von Stengel, B. 2008. Computing an extensive-form correlated equilibrium in polynomial time. In *International Workshop on Internet and Network Economics*, 506–513. Springer.

Jiang, A. X.; and Leyton-Brown, K. 2015. Polynomial-time computation of exact correlated equilibrium in compact games. *Games and Economic Behavior* 91: 347–359.

Kroer, C.; Farina, G.; and Sandholm, T. 2017. Smoothing Method for Approximate Extensive-Form Perfect Equilibrium. In *International Joint Conference on Artificial Intelligence*, 295–301.

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.

Marchesi, A.; and Gatti, N. 2020. Trembling-Hand Perfection and Correlation in Sequential Games. *arXiv preprint arXiv:2012.06528* .

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

Nash, J. 1951. Non-cooperative games. *Annals of mathematics* 286–295.

Papadimitriou, C. H.; and Roughgarden, T. 2008. Computing correlated equilibria in multi-player games. *Journal of the ACM (JACM)* 55(3): 1–29.

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

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. 1991. *Stability and perfection of Nash equilibria*, volume 339. Springer.

Vermeulen, D.; and Jansen, M. 1998. The reduced form of a game. *European journal of operational research* 106(1): 204–211.

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

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