

---

# Connecting Optimal Ex-Ante Collusion in Teams to Extensive-Form Correlation: Faster Algorithms and Positive Complexity Results

---

Gabriele Farina<sup>1</sup> Andrea Celli<sup>2</sup> Nicola Gatti<sup>2</sup> Tuomas Sandholm<sup>1 3 4 5</sup>

## Abstract

We focus on the problem of finding an optimal strategy for a team of players that faces an opponent in an imperfect-information zero-sum extensive-form game. Team members are not allowed to communicate during play but can coordinate before the game. In this setting, it is known that the best the team can do is sample a profile of potentially randomized strategies (one per player) from a joint (a.k.a. correlated) probability distribution at the beginning of the game. In this paper, we first provide new modeling results about computing such an optimal distribution by drawing a connection to a different literature on extensive-form correlation. Second, we provide an algorithm that allows one for capping the number of profiles employed in the solution. This begets an anytime algorithm by increasing the cap. We find that often a handful of well-chosen such profiles suffices to reach optimal utility for the team. This enables team members to reach coordination through a simple and understandable plan. Finally, inspired by this observation and leveraging theoretical concepts that we introduce, we develop an efficient column-generation algorithm for finding an optimal distribution for the team. We evaluate it on a suite of common benchmark games. It is three orders of magnitude faster than the prior state of the art on games that the latter can solve and it can also solve several games that were previously unsolvable.

## 1. Introduction

Much of the computational game theory literature has focused on finding strong strategies for large two-player zero-sum extensive-form games. In that setting, perfect game playing corresponds to playing strategies that belong to a Nash equilibrium, and such strategies can be found in polynomial time in the size of the game. Recent landmark results, such as superhuman agents for heads-up limit and no-limit Texas hold'em poker (Bowling et al., 2015; Brown & Sandholm, 2019; Moravčík et al., 2017) show that the problem of computing strong strategies in two-player zero-sum games is well understood both in theory and in practice. The same cannot be said for almost any type of strategic multi-player interaction, where computing strong strategies is generally hard in the worst case.

In this paper, we study *adversarial team games*, that is, games in which a team of coordinating (colluding) players faces an opponent. We will focus on a two-player team coordinating against a third player. Team members can plan jointly at will before the game, but are not allowed to communicate during the game (other than through their actions in the game). These games are a popular middle ground between two-player zero-sum games and multiplayer games (von Stengel & Koller, 1997; Celli & Gatti, 2018). They can be used to model many strategic interactions of practical relevance. For example, how should two players colluding against a third at a poker table play? Or, how would the two defenders in Bridge (who are prohibited from communicating privately during the game) play optimally against the declarer? Even though adversarial team games are conceptually zero-sum interactions between two entities—the team and the opponent—computing optimal strategies is hard in this setting. Even finding a best-response strategy for the team given a fixed strategy for the opponent is hard (Celli & Gatti, 2018).

One might think that finding the optimal strategy for the team simply amounts to finding an optimal profile of potentially mixed (a.k.a. randomized) strategies, one strategy per team members. A solution of this type that yields maximum expected sum of utilities for the team players against a rational (that is, best-responding) opponent is known as a *team-maxmin equilibrium* (TME) strategy (Basilico et al.,

---

<sup>\*</sup>Equal contribution <sup>1</sup>Computer Science Department, Carnegie Mellon University, Pittsburgh PA 15213 <sup>2</sup>DEIB, Politecnico di Milano, Milano, Italy <sup>3</sup>Strategic Machine, Inc. <sup>4</sup>Strategy Robot, Inc. <sup>5</sup>Optimized Markets, Inc.. Correspondence to: Gabriele Farina <gfarina@cs.cmu.edu>, Andrea Celli <andrea.celli@polimi.it>, Nicola Gatti <nicola.gatti@polimi.it>, Tuomas Sandholm <sandholm@cs.cmu.edu>.

2017; Zhang & An, 2020a;b).

In this paper, we are interested in a more powerful model. Before the game starts, the team members are able to sample a profile from a joint (a.k.a. correlated) *distribution*. This form of *ex-ante coordination* is known to be the best a team can do and comes with two major advantages. First, it offers the team larger (or equal) expected utility than TME—sometimes with dramatic gains (Celli & Gatti, 2018). Second, it makes the problem of computing the optimal team strategy convex—and thus more amenable to the plethora of convex optimization algorithms that have been developed over the past 80 years—whereas the problem of computing a TME strategy is not convex. In our model, an optimal distribution for the team is known as a *team-maxmin equilibrium with coordination device* (TMECor) strategy (Celli & Gatti, 2018; Farina et al., 2018).

**Our contributions.** We introduce the notion of *semi-randomized correlation plan*, and propose a natural formulation for the problem of finding a TMECor strategy by drawing connections with the extensive-form strategy polytope defined by von Stengel & Forges (2008). Second, we propose an algorithm for computing a TMECor strategy when only a fixed number of pairs of semi-randomized correlation plans is allowed. This begets an anytime algorithm by increasing that fixed number. Surprisingly, we find that often a handful of well-chosen semi-randomized correlation plans is enough to reach optimal utility. This enables team members to reach coordination through simple and understandable strategies. Finally, by leveraging our new representation, we develop a column-generation algorithm for finding a TMECor strategy. The core of our algorithm is a new best-response (BR) oracle for computing joint team best-response strategies. We show that, in contrast with the previous state-of-the-art BR oracles for that problem, our oracle enables provably polynomial-time computation of a TMECor in some notable classes of games including, for example, Goofspiel. This result constitutes the first example of efficient computation of optimal ex-ante coordinated strategies in adversarial team games, and cannot be achieved by employing previous BR oracles. We evaluate our column-generation algorithm on a suite of common benchmark games. It is three orders of magnitude faster than the prior state of the art on games that the latter can solve. It can also solve many games that were previously unsolvable.

## 2. Preliminaries

Extensive-form games (EFGs) model games that are played on a game tree, and can capture both sequential and simultaneous moves, as well as private information. In this paper, we focus on three-player zero-sum games where two players—**T1** and **T2**—play as a team against the opponent

player, denoted by **O**.

Each node  $v$  in the game tree belongs to exactly one player  $i \in \{\mathbf{T1}, \mathbf{T2}, \mathbf{O}\} \cup \{\mathbf{C}\}$  whose turn is to move. Player **C** is a special player, called the *chance player*. It models exogenous stochasticity in the environment, such as drawing a card from a deck or tossing a coin. The edges leaving  $v$  represent the actions available at that node. Any node without outgoing edges is called a *leaf* and represents an end state of the game. We denote the set of such nodes by  $Z$ . Each  $z \in Z$  is associated with a tuple of payoffs specifying the payoff  $u_i(z)$  of each player  $i \in \{\mathbf{T1}, \mathbf{T2}, \mathbf{O}\}$  at  $z$ . The product of the probabilities of all actions of **C** on the path from the root of the game to leaf  $z$  is denoted by  $p_{\mathbf{C}}(z)$ .

Private information is represented via *information set* (infoset). In particular, the set of nodes belonging to  $i \in \{\mathbf{T1}, \mathbf{T2}, \mathbf{O}\}$  is partitioned into a collection  $\mathcal{I}_i$  of non-empty sets: each  $I \in \mathcal{I}_i$  groups together nodes that Player  $i$  cannot distinguish among, given what they have observed. Necessarily, for any  $I \in \mathcal{I}_i$  and  $v, w \in I$ , nodes  $v$  and  $w$  must have the same set of available actions. Consequently, we denote the set of actions available at all nodes of  $I$  by  $A_I$ . As it is customary in the related literature, we assume *perfect recall*, that is, no player forgets what he/she knew earlier in the game. Finally, given players  $i$  and  $j$ , two infosets  $I_i \in \mathcal{I}_i, I_j \in \mathcal{I}_j$  are *connected*, denoted by  $I_i \rightleftharpoons I_j$ , if there exist  $v \in I_i$  and  $w \in I_j$  such that the path from the root to  $v$  passes through  $w$  or vice versa.

**Sequences.** The set of *sequences* of Player  $i$ , denoted by  $\Sigma_i$ , is defined as  $\Sigma_i := \{(I, a) : I \in \mathcal{I}_i, a \in A_I\} \cup \{\emptyset\}$ , where the special element  $\emptyset$  is called the *empty sequence* of Player  $i$ . The *parent sequence* of a node  $v$  of Player  $i$ , denoted  $\sigma(v)$ , is the last sequence (information set-action pair) for Player  $i$  encountered on the path from the root of the game to that node. Since the game has perfect recall, for each  $I \in \mathcal{I}_i$ , nodes belonging to  $I$  share the same *parent sequence*. So, given  $I \in \mathcal{I}_i$ , we denote by  $\sigma(I) \in \Sigma_i$  the unique parent sequence of nodes in  $I$ . Additionally, we let  $\sigma(I) = \emptyset$  if Player  $i$  never acts before infoset  $I$ .

**Relevant sequences.** A pair of sequences  $\sigma_i \in \Sigma_i, \sigma_j \in \Sigma_j$  is *relevant* if either one is the empty sequence, or if they can be written as  $\sigma_i = (I_i, a_i)$  and  $\sigma_j = (I_j, a_j)$  with  $I_i \rightleftharpoons I_j$ . We write  $\sigma_i \bowtie \sigma_j$  to denote that they form a pair of relevant sequences. Given two players  $i$  and  $j$ , we let  $\Sigma_i \bowtie \Sigma_j := \{(\sigma_i, \sigma_j) : \sigma_i \in \Sigma_i, \sigma_j \in \Sigma_j, \sigma_i \bowtie \sigma_j\}$ . Similarly, given  $\sigma_i$  and  $I_j \in \mathcal{I}_j$ , we say that  $(\sigma_i, I_j)$  forms a relevant sequence-information set pair  $(\sigma_i \bowtie I_j)$ , if  $\sigma_i = \emptyset$  or if  $\sigma_i = (I_i, a_i)$  and  $I_i \rightleftharpoons I_j$ .

**Reduced-normal-form plans.** A *reduced-normal-form* plan  $\pi_i$  for Player  $i$  defines a choice of action for every information set  $I \in \mathcal{I}_i$  that is still reachable as a result of the other choices in  $\pi$  itself. The set of reduced-normal-

form plans of Player  $i$  is denoted  $\Pi_i$ . We denote by  $\Pi_i(I)$  the subset of reduced-normal-form plans that prescribe all actions for Player  $i$  on the path from the root to information set  $I \in \mathcal{I}_i$ . Similarly, given  $\sigma = (I, a) \in \Sigma_i$ , let  $\Pi_i(\sigma) \subseteq \Pi_i(I)$  be the set of reduced-normal-form plans belonging to  $\Pi_i(I)$  where Player  $i$  plays action  $a$  at  $I$ , and let  $\Pi_i(\emptyset) := \Pi_i$ . Finally, given a leaf  $z \in Z$ , we denote with  $\Pi_i(z) \subseteq \Pi_i$  the set of reduced-normal-form plans where Player  $i$  plays so as to reach  $z$ .

**Sequence-form strategies.** A *sequence-form strategy* is a compact strategy representation for perfect-recall players in EFGs (Romanovskii, 1962; Koller et al., 1996). Given a player  $i \in \{\mathsf{T1}, \mathsf{T2}, \mathsf{O}\}$  and a normal-form strategy  $\mu \in \Delta(\Pi_i)$ ,<sup>1</sup> the sequence-form strategy induced by  $\mu$  is the real vector  $\mathbf{y}$ , indexed over  $\sigma \in \Sigma_i$ , defined as  $y[\sigma] := \sum_{\pi \in \Pi_i(\sigma)} \mu(\pi)$ . The set of sequence-form strategies that can be induced as  $\mu$  varies over  $\Delta(\Pi_i)$  is denoted by  $\mathcal{Y}_i$  and is known to be a convex polytope (called the *sequence-form polytope*) defined by a number of constraints equal to  $|\mathcal{I}_i|$  (Koller et al., 1996).

**TMECor as a Bilinear Saddle-Point Problem.** A TMECor strategy is a probability distribution  $\mu_{\mathsf{T}}$  over the set of randomized strategy profiles  $\mathcal{Y}_{\mathsf{T1}} \times \mathcal{Y}_{\mathsf{T2}}$  that guarantees maximum expected utility for the team against the best-responding opponent  $\mathsf{O}$ . Since each player has perfect recall, any randomized strategy for a player is equivalent to a distribution over reduced-normal-form pure strategies (Kuhn, 1953). Hence, any distribution over profiles of randomized strategies of the team members can be expressed in an equivalent way as a distribution over *deterministic* strategy profiles  $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$ . The benefit of this transformation is that  $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$  is a finite set, unlike  $\mathcal{Y}_{\mathsf{T1}} \times \mathcal{Y}_{\mathsf{T2}}$ . For this reason, TMECor is usually defined in the literature as a distribution over  $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$  without loss of generality. We will follow the same approach in our characterization.

For each leaf  $z$ , let  $\hat{u}_{\mathsf{T}}(z) := (u_{\mathsf{T1}}(z) + u_{\mathsf{T2}}(z))p_{\mathsf{C}}(z)$ . The expected utility of the team can be written as the following function of the distributions of play  $\mu_{\mathsf{T}} \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})$ ,  $\mu_{\mathsf{O}} \in \Delta(\Pi_{\mathsf{O}})$ :

$$u_{\mathsf{T}}(\mu_{\mathsf{T}}, \mu_{\mathsf{O}}) := \sum_{z \in Z} \hat{u}_{\mathsf{T}}(z) \left( \sum_{\substack{\pi_{\mathsf{T1}} \in \Pi_{\mathsf{T1}}(z) \\ \pi_{\mathsf{T2}} \in \Pi_{\mathsf{T2}}(z)}} \mu_{\mathsf{T}}(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}}) \right) \left( \sum_{\pi \in \Pi_{\mathsf{O}}(z)} \mu_{\mathsf{O}}(\pi) \right).$$

By definition, a *team-maxmin equilibrium with coordination device* (TMECor) is a Nash equilibrium of the game where the team plays according to the coordinated strategy  $\mu_{\mathsf{T}} \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})$ . In the zero-sum setting, this amounts to finding a solution of the optimization problem

$$\arg \max_{\mu_{\mathsf{T}} \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})} \min_{\mu_{\mathsf{O}} \in \Delta(\Pi_{\mathsf{O}})} u_{\mathsf{T}}(\mu_{\mathsf{T}}, \mu_{\mathsf{O}}). \quad (1)$$

<sup>1</sup> $\Delta(X)$  denotes the probability simplex over the finite set  $X$ .

The opponent's strategy  $\mu_{\mathsf{O}}$  can be compactly represented through its equivalent sequence-form representation. This is not the case for  $\mu_{\mathsf{T}}$ , which cannot be represented concisely through the sequence form as shown by Farina et al. (2018).

### 3. A New Formulation of TMECor Based on Extensive-Form Correlation Plans

We propose using a different representation of the correlated distribution of play  $\mu_{\mathsf{T}}$ , inspired by the growing body of literature on extensive-form correlated equilibria. Like the *realization form* by Farina et al. (2018), in our approach we represent  $\mu_{\mathsf{T}}$  as a vector with only a polynomial number of components. However, unlike the realization form, the number of components scales as the product of the number of sequences of the two players, which can be significantly larger than the number of leaves. This downside is amply outweighed by the following benefits. First, we show that in practice our proposed representation of  $\mu_{\mathsf{T}}$  enables us to compute best responses for the team significantly faster than the prior representations. Second, in certain classes of games, we even show that our proposed representation enables the computation of a TMECor in polynomial time. This is the case, for example, in Goofspiel, a popular benchmark game in computational game theory (Ross, 1971).

#### 3.1. Extensive-Form Correlation Plans

Our representation is based on the concept of *extensive-form correlation plans*, introduced by von Stengel & Forges (2008) in their seminal paper on extensive-form correlation. In particular, we map the correlated distribution of play  $\mu_{\mathsf{T}}$  of the team to the vector  $\xi_{\mathsf{T}}$  indexed over pairs of sequences  $(\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}) \in \Sigma_{\mathsf{T1}} \times \Sigma_{\mathsf{T2}}$ , where each entry is defined as

$$\xi_{\mathsf{T}}[(\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}})] := \sum_{\substack{\pi_{\mathsf{T1}} \in \Pi_{\mathsf{T1}}(\sigma_{\mathsf{T1}}) \\ \pi_{\mathsf{T2}} \in \Pi_{\mathsf{T2}}(\sigma_{\mathsf{T2}})}} \mu_{\mathsf{T}}[(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})]. \quad (2)$$

Here  $\xi_{\mathsf{T}}$  is not indexed over *all* pairs of sequences  $(\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}})$ —only *relevant* sequence pairs. While there are games in which this distinction is meaningless (that is, games in which all sequences pairs for the team members are relevant), in practice the number of *all* sequence pairs is usually significantly bigger than the number of relevant sequence pairs, as shown in Table 1(b).

The set of extensive-form correlation plans  $\xi_{\mathsf{T}}$  that can be induced as  $\mu_{\mathsf{T}}$  varies over the set of all correlated distributions of play for the team members is a convex polytope. We denote it as  $\Xi_{\mathsf{T}}$  and call it the *polytope of correlation plans*. We will recall existing results and provide new ones about the structure of  $\Xi_{\mathsf{T}}$  in Section 4.



## 4.2. Triangle-Freeness and Polynomial-Time Computation of TMECor

Proposition 2 shows that  $\Xi_T$  is a subset of the von Stengel-Forges polytope. There are games where the reverse inclusion does not hold. Farina & Sandholm (2020) gave a sufficient condition—called *triangle-freeness*—for the reverse inclusion to hold. We state the condition for our setting.

**Definition 2** (Farina & Sandholm (2020)). *The interaction of the team members T1 and T2 is triangle-free if, for any choice of distinct information sets  $I_1, I_2 \in \mathcal{I}_{T1}$  with  $\sigma_{T1}(I_1) = \sigma_{T1}(I_2)$  and any choice of distinct information sets  $J_1, J_2 \in \mathcal{I}_{T2}$  with  $\sigma_{T2}(J_1) = \sigma_{T2}(J_2)$ , it is never the case that  $(I_1 \rightleftharpoons J_1) \wedge (I_2 \rightleftharpoons J_2) \wedge (I_1 \rightleftharpoons J_2)$ .*

The above condition can be easily checked in polynomial-time by iterating over all possible quadruplets of information sets  $I_1, I_2, J_1, J_2$  and checking whether the condition (which itself can be checked by performing a standard traversal of the input game tree) holds. Farina & Sandholm (2020) show that when the information structure of correlating players (in our case, the team members) is triangle-free, then  $\Xi_T = \mathcal{V}_T$ . So, when the interaction of the team is triangle-free, a TMECor can be found in polynomial time by substituting constraint ④ in the LP in Proposition 1 with the von Stengel-Forges constraints of Definition 1. As far as we are aware, this positive complexity result has not been noted before in the literature. We show in Table 1(c) that Goofspiel is triangle free, since all chance outcomes are public (and that none of the other common benchmark games that we consider are).

## 4.3. Semi-Randomized Correlation Plans

We now give a third result about the structure of  $\Xi_T$ , which will enable us to replace Constraint ④ of Proposition 1 with something more practical. First, we introduce *semi-randomized correlation plans*, which are elements of a subset of the von Stengel-Forges polytope of the team, as we formalize shortly. A semi-randomized correlation plan represents a strategy profile in which one of the players plays a deterministic strategy, while the other player in the team independently plays a randomized strategy. Formally, we define the set of semi-randomized correlation plans for T1 and T2 as

$$\begin{aligned} \Xi_{T1}^* &= \{\xi \in \mathcal{V}_T : \xi[\emptyset, \sigma_{T2}] \in \{0, 1\} \quad \forall \sigma_{T2} \in \Sigma_{T2}\}, \\ \Xi_{T2}^* &= \{\xi \in \mathcal{V}_T : \xi[\sigma_{T1}, \emptyset] \in \{0, 1\} \quad \forall \sigma_{T1} \in \Sigma_{T1}\}, \end{aligned}$$

respectively. Crucially, a point  $\xi \in \Xi_i^*$  for  $i \in \{T1, T2\}$  can be expressed using real and binary variables, in addition to the linear constraints the define  $\mathcal{V}$  (Definition 1).

With that, we can show the following structural result for the polytope of extensive-form correlation plans  $\Xi_T$ .

**Proposition 3.** *In every game,  $\Xi_T$  is the convex hull of the*

set  $\Xi_{T1}^*$ , or equivalently of the set  $\Xi_{T2}^*$ . Formally,  $\Xi_T = \text{co } \Xi_{T1}^* = \text{co } \Xi_{T2}^* = \text{co}(\Xi_{T1}^* \cup \Xi_{T2}^*)$ .

Our notion of semi-randomized correlation plans is reminiscent of the *auxiliary game* construction of Farina et al. (2018), in which only one of the team members (the *pivot player*) is required to play a deterministic strategy. Our setting is very different, however, since we have a different representation of team strategies with many more variables and stronger combinatorial structure.

## 5. Computing TMECor with a Small Support of Semi-Randomized Plans of Fixed Size

From Proposition 3, it is known that  $\Xi_T$  is the convex hull of  $\Xi_{T1}^*$  and  $\Xi_{T2}^*$ . Furthermore, the polytopes  $\Xi_{T1}^*$  and  $\Xi_{T2}^*$  can be described via a number of linear constraints that is quadratic in the game size and a number of integer variables that is linear in the game size. So, we can replace Constraint ④ in Proposition 1 with the constraint that  $\xi_T$  be a convex combination of elements from  $\Xi_{T1}^*$  and  $\Xi_{T2}^*$ . We introduce variables  $\xi_T^{(1)}, \dots, \xi_T^{(n)} \in \Xi_{T1}^* \cup \Xi_{T2}^*$  and the corresponding convex combination coefficients  $\lambda^{(1)}, \dots, \lambda^{(n)}$ , and replace Constraint ④ with the linear constraint  $\xi_T = \sum_{i=1}^n \lambda^{(i)} \xi_T^{(i)}$ . Here,  $n$  is a parameter with which we can cap the number of semi-randomized correlation plans that can be included in the strategy. This gives the following mixed integer LP.

$$\left\{ \begin{array}{l} \arg \max_{\xi_T^{(1)}, \dots, \xi_T^{(n)}, \lambda^{(1)}, \dots, \lambda^{(n)}} v_\emptyset, \quad \text{subject to:} \\ \text{constraints ① ② ③ as in Proposition 1} \\ \text{④ } \xi_T = \sum_{i=1}^n \lambda^{(i)} \xi_T^{(i)} \\ \text{⑤ } \xi_T^{(1)} \in \Xi_{T1}^*, \xi_T^{(2)} \in \Xi_{T2}^*, \xi_T^{(3)} \in \Xi_{T1}^*, \xi_T^{(4)} \in \Xi_{T2}^*, \dots^\ddagger \\ \text{⑥ } \sum_{i=1}^n \lambda^{(i)} = 1, \lambda^{(i)} \geq 0 \quad \forall i \in \{1, \dots, n\}. \end{array} \right.$$

The larger  $n$  is, the higher the solution value obtained, but the slower the program. We can make this into an anytime algorithm by solving the integer program for increasing values of  $n$ . By Caratheodory's theorem, this program already yields an optimal solution to the LP in Proposition 1 when  $n \geq |\Sigma_1 \bowtie \Sigma_2| + 1$ . As we show in detail in Section 7, in practice we found that near-optimal coordination can be achieved through strategies with a significantly smaller value of  $n$ . Hence, oftentimes the team does not need a large number of complex profiles of randomized strategies to play optimally: a handful (often one or two) of carefully selected simple strategies often result in optimal coordination. That

<sup>‡</sup>In Constraint ⑤ we alternate the set of semi-randomized correlation plans (i.e., we alternate which player's turn it is to play a deterministic strategy). Empirically, this increases the diversity of the strategies of  $\Xi_T$  that can be represented with small values of  $n$  and leads to higher utilities for the team.



known. The above equality enables us to solve the problem because the right hand side is a well-defined mixed integer LP (MIP). We can use a commercial solver such as Gurobi to solve it. When the objective value of the pricing problem is non-positive, there is no variable that can be added to the master LP which would increase its value. Thus, the optimal solution to the master LP is guaranteed to be optimal for the LP in Proposition 1 and the main loop terminates.

## 6.2. Implementation Details

We further speed up the solution of the pricing problem in our implementation by the following techniques.

**Seeding phase.** To avoid having to go through many iterations of the main loop, each of which requires solving the pricing problem, we want to seed the master LP up front with a set of good candidate variables. While any seeding maintains optimality of the overall algorithm, seeding it with variables that are likely to be part of the optimal solution increases speed the most. We initialize the set of correlation plans  $S$  by running  $m$  iterations of a self-play no-external-regret algorithm. Specifically, we let each player run CFR+ (Tammelin et al., 2015; Bowling et al., 2015) and, at each iteration of that algorithm, we sample a pair of pure normal-form plans for the two team members according to the current strategies of the two players. At each iteration of that no-regret method, we set the utility of each team member to  $u_{T1} + u_{T2}$ . Finally, for each pair  $(\pi_{T1}, \pi_{T2}) \in \Pi_{T1} \times \Pi_{T2}$  of normal-form plans generated by that no-regret algorithm, we compute and add to  $S$  the correlation plan corresponding to the distribution  $\mu$  that assigns probability 1 to  $(\pi_{T1}, \pi_{T2})$  using Eq. (2). While self-play no-regret methods guarantee convergence to Nash equilibrium in two-player zero-sum game, no guarantee is available in our setting. However, we empirically find that this seeding strategy leads to a strong initial set of correlation plans.

**Linear relaxation.** Before solving the MIP formulation of the pricing problem, we first try to solve its linear relaxation  $\arg \max_{\hat{\xi}_T \in \mathcal{V}_T} c(\hat{\xi}_T)$ . We found that in many cases it outputs semi-randomized correlation plans, thus avoiding the overhead of having to solve a MIP.

**Solution pools.** Modern commercial MIP solvers such as Gurobi keep track of additional suboptimal feasible solutions (in addition to the optimal one) that were found during the process of solving a MIP. Since accessing those additional solutions is essentially free computationally, we add to  $S$  all the solutions (even suboptimal ones) that were produced in the process of solving the MIP. This can be viewed as a form of dynamic seeding and does not affect the optimality of the overall algorithm.

**Termination.** Because fast integer and LP solvers work with real-valued variables, near the end of the column-generation

loop the new variables that are generated in the pricing problem have reduced costs that are very close to zero. It is not clear whether they are actually positive or zero. Therefore, we set the numeric tolerance so that we stop the column-generation loop if the value of the pricing problem solution is less than  $10^{-6}$ .

**Dual values.** To obtain the dual values used in the pricing problem, we do not need to formulate and solve a dual LP as modern LP solvers already keep track of dual values.

## 7. Experimental Evaluation

We computationally evaluate the algorithms proposed in Section 5 and Section 6. We test on the common parametric games shown in Table 1. Appendix B provides additional detail about the games. We ran the experiments on a machine with a 16-core 2.80GHz CPU, and allow a maximum of four threads and 32GB of RAM to each experiment. We used Gurobi 9.1.1 to solve LPs and MIPs.

### 7.1. Small-Supported TMECor in Practice

Table 2 describes the maximum expected utility that the team can obtain by limiting the support of its distribution to  $n \in \{1, 2, 3\}$  semi-randomized correlation plans. Columns denoted by  $n = \infty$  show the optimal expected utility of the team at the TMECor (without any limit on the support size). We ran experiments with the opponent as the first ( $\circ = 1$ ), second ( $\circ = 2$ ), and third player ( $\circ = 3$ ) of each game. In all the games, distributions with as few as two or three semi-randomized coordination plans gave the team near-optimal expected utility. Moreover, in several games, one or two carefully selected semi-randomized coordination plans are enough to reach an optimal solution.

### 7.2. Column-Generation in Practice

We evaluate our column-generation algorithm against the two prior state-of-the-art algorithms for computing a TMECor: the column-generation technique by Celli & Gatti (2018) (henceforth CG-18), and the fictitious-team-play algorithm by Farina et al. (2018) (denoted FTP). Like our algorithm, CG-18 uses a column generation approach that lets  $\circ$  play sequence-form strategies, while the team’s strategy is directly represented as a distribution over joint normal-form plans  $\Pi_{T1} \times \Pi_{T2}$ . FTP is based on the bilinear saddle-point formulation of the problem and is essentially a variation of *fictitious play* (Brown, 1951). FTP operates on the bilinear formulation of TMECor (1): the team and the opponent are treated as two entities that converge to equilibrium in self-play. FTP only guarantees convergence in the limit to a TMECor, while our algorithm certifies optimality. So, the run-time comparison between our algorithm to FTP must be done with care, as the latter never stops, whereas our

Game	Ours		Fictitious Team Play (FTP)			CG-18	Pricers		Team utility after seeding			TMECor value
	Seeded	Not seed.	$\epsilon = 50\%$	$\epsilon = 10\%$	$\epsilon = 1\%$		Relax.	MIP	$m = 0$	100	1000	
[A]	1ms	2ms	2s <sup>†</sup>	10s <sup>†</sup>	1m 08s <sup>†</sup>	175ms	1	0	-0.556	0	0	0
[B]	1ms	14ms	3m 52s	37m 51s	> 6h	26.81s	2	0	-0.406	-0.042	-0.042	-0.042
[C]	4.96s	13.40s	4h 42m	> 6h	> 6h	> 6h	4	22	-0.343	-0.030	-0.021	-0.014
[D]	325ms	517ms	50s	9m 21s	> 6h	3m 09s	18	0	-1.000	0.247	0.252	0.252
[E]	1.18s	1.48s	4m 51s	2h 02m	> 6h	29m 38s	45	0	-2.933	0.239	0.248	0.253
[F]	1m 12s	4m 03s	> 6h	> 6h	> 6h	> 6h	40	8	0.000	0.276	0.284	0.284
[G]	> 6h	> 6h	> 6h	> 6h	> 6h	> 6h	—	—	-0.688	0.277	oom	—
[H]	2m 23s	3m 23s	> 6h	> 6h	> 6h	> 6h	20	171	-1.783	0.065	0.151	0.215
[I]	1h 07m	1h 19m	> 6h	> 6h	> 6h	> 6h	7	610	-1.216	-0.149	0.019	0.107
[J]	3m 29s	1m 50s	> 6h	> 6h	> 6h	> 6h	1346	18	-6.000	-0.222	0.387	0.516

(a) — Comparison of run times

(b)

(c)

Table 3: (a) Runtime comparison between our column generation algorithm, FTP, and CG-18. The seeded version of our algorithm runs  $m = 1000$  iterations of CFR+ (Section 6.2), while the non seeded version runs  $m = 0$ . ‘†’: since the TMECor value for the game is exactly zero, we measure how long it took the algorithm to find a distribution with expected value at least  $-\epsilon/10$  for the team. (b) Number of times the pricing problem for our column-generation algorithm was solved to optimality by the linear relaxation (‘Relax’) and by the MIP solver (‘MIP’) when using our column-generation algorithm. (c) Quality of the initial strategy of the team obtained for varying sizes of  $S$  compared to the expected utility of the team at the TMECor. ‘oom’: out of memory.

algorithm and CG-18 terminate after a finite number of iterations with an *exact* optimal strategy. We report the run time of FTP reaching solution quality that is  $\epsilon = 50\%$ ,  $10\%$ , and  $1\%$  off the optimal value (determined by the other two algorithms). Finally, the concurrent work by Zhang et al. (2020) proposes an alternative approach for computing a TMECor. According to the run times reported in their paper, our algorithm is significantly faster. In particular, our algorithm takes about 2m 00s to solve game [H], while their algorithm takes about 1h 21m, even though they conducted their experiments on a more powerful machine.<sup>3</sup>

We set a time limit of 6 hours, a memory limit of 32GB, and a cap of four threads for each algorithm. Table 3 shows the results with the opponent playing as the third player. According to Table 2, this is almost always the hardest setting. The results for the other two settings are in Appendix C.

Our column-generation algorithm dramatically outperforms FTP and CG-18. For example, in Liar’s dice instance [F], our algorithm finds an optimal TMECor in a few seconds while the prior algorithms exceed six hours. The last column of Table 3(c) shows the final team utility. Even when the opponent is playing as the third player, the team is able to reach positive expected utility. Finally, we identify Liar’s dice instance [G] as the current boundary of problem that

<sup>3</sup>The machine which they used has a 3.6GHz CPU, 32GB of memory, and they dedicated 12 threads to the algorithm. Moreover, we observe that the number of terminal nodes reported for 3L3 in their paper is inconsistent with the description of the game, which corresponds to our game [H]; in particular, their paper reports that the game has a larger number of terminal nodes than it actually has. This was confirmed by the authors in private communications.

just cannot be handled with current TMECor technology: it does not complete within six hours.

Using the linear relaxation of the pricing problem (“implementation details” in Section 6.2) often obviated the need to run the slower MIP pricing (see Table 3(b)). In all Goofspiel instances (games [D] and [E]) and in small Kuhn poker instances, the MIP pricing is never invoked.

Regret-based seeding further improves the performance of the algorithm. In the Liar’s dice instance [F], it reduced run time by roughly a factor of ten. The objective value of the master solution immediately after seeding (that is, before the first column generation step) increases significantly with the number of iterations of the no-regret algorithm that is used for seeding.

## 8. Conclusions

We studied finding an optimal strategy for a team with two members facing an opponent in an imperfect-information, zero-sum, extensive-form game. We focused on the setting where the team members cannot communicate during play but can coordinate before the game. We provided modeling results by drawing a connection to prior results on extensive-form correlation. Then, we developed an algorithm that computes an optimal joint distribution by just using profiles where only one of the team members gets to randomize in each profile. We can cap the number of such profiles we allow in the solution. This begets an anytime algorithm by increasing the cap. Moreover, we showed that often a handful of well-chosen such profiles suffice to reach optimal utility for the team. Inspired by this observation and lever-

---

aging theoretical concepts that we introduced, we developed an efficient column-generation algorithm for finding an optimal strategy for the team. We tested our algorithm on a suite of standard games, showing that it is three orders of magnitudes faster than the prior state of the art and also solves many games that were previously unsolvable.

## Acknowledgments

This material is based on work supported by the National Science Foundation under grants IIS-1718457, IIS-1901403, and CCF-1733556, the ARO under award W911NF2010081, and the Italian MIUR PRIN 2017 Project ALGADIMAR “Algorithms, Games, and Digital Market”. Gabriele Farina is supported by a Facebook fellowship.

## References

- Basilico, N., Celli, A., Nittis, G. D., and Gatti, N. Team-maxmin equilibrium: efficiency bounds and algorithms. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 356–362, 2017.
- Bowling, M., Burch, N., Johanson, M., and Tammelin, O. Heads-up limit hold’em poker is solved. *Science*, 347 (6218), January 2015.
- Brown, G. W. Iterative solutions of games by fictitious play. In Koopmans, T. C. (ed.), *Activity Analysis of Production and Allocation*, pp. 374–376. John Wiley & Sons, 1951.
- Brown, N. and Sandholm, T. Superhuman AI for multiplayer poker. *Science*, 365(6456):885–890, 2019.
- Celli, A. and Gatti, N. Computational results for extensive-form adversarial team games. In *AAAI Conference on Artificial Intelligence (AAAI)*, pp. 965–972, 2018.
- Farina, G. and Sandholm, T. Polynomial-time computation of optimal correlated equilibria in two-player extensive-form games with public chance moves and beyond. In *ArXiv preprint*, 2020.
- Farina, G., Celli, A., Gatti, N., and Sandholm, T. Ex ante coordination and collusion in zero-sum multi-player extensive-form games. In *Advances in Neural Information Processing Systems*, pp. 9638–9648, 2018.
- Ford, L. R. and Fulkerson, D. R. A suggested computation for maximal multi-commodity network flows. *Management Science*, 5(1):97–101, 1958. doi: 10.1287/mnsc.5.1.97.
- Koller, D., Megiddo, N., and von Stengel, B. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14(2), 1996.
- Kuhn, H. W. A simplified two-person poker. In Kuhn, H. W. and Tucker, A. W. (eds.), *Contributions to the Theory of Games*, volume 1 of *Annals of Mathematics Studies*, 24, pp. 97–103. Princeton University Press, Princeton, New Jersey, 1950.
- Kuhn, H. W. Extensive games and the problem of information. In Kuhn, H. W. and Tucker, A. W. (eds.), *Contributions to the Theory of Games*, volume 2 of *Annals of Mathematics Studies*, 28, pp. 193–216. Princeton University Press, Princeton, NJ, 1953.
- Lanctot, M., Waugh, K., Zinkevich, M., and Bowling, M. Monte Carlo sampling for regret minimization in extensive games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2009.
- Lisý, V., Lanctot, M., and Bowling, M. Online Monte Carlo counterfactual regret minimization for search in imperfect information games. In *Proceedings of the 2015 international conference on autonomous agents and multiagent systems*, pp. 27–36, 2015.
- Moravčík, M., Schmid, M., Burch, N., Lisý, V., Morrill, D., Bard, N., Davis, T., Waugh, K., Johanson, M., and Bowling, M. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, May 2017.
- Romanovskii, I. Reduction of a game with complete memory to a matrix game. *Soviet Mathematics*, 3, 1962.
- Ross, S. M. Goofspiel—the game of pure strategy. *Journal of Applied Probability*, 8(3):621–625, 1971.
- Southey, F., Bowling, M., Larson, B., Piccione, C., Burch, N., Billings, D., and Rayner, C. Bayes’ bluff: Opponent modelling in poker. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, July 2005.
- Tammelin, O., Burch, N., Johanson, M., and Bowling, M. Solving heads-up limit Texas hold’em. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.
- von Stengel, B. Efficient computation of behavior strategies. *Games and Economic Behavior*, 14(2):220–246, 1996.
- von Stengel, B. and Forges, F. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research*, 33(4):1002–1022, 2008.
- von Stengel, B. and Koller, D. Team-maxmin equilibria. *Games and Economic Behavior*, 21(1-2):309–321, 1997.
- Zhang, Y. and An, B. Computing team-maxmin equilibria in zero-sum multiplayer extensive-form games. In *AAAI*

---

*Conference on Artificial Intelligence (AAAI)*, pp. 2318–2325, 2020a.

Zhang, Y. and An, B. Converging to team-maxmin equilibria in zero-sum multiplayer games. In *International Conference on Machine Learning (ICML)*, 2020b.

Zhang, Y., An, B., and Černý, J. Computing ex ante coordinated team-maxmin equilibria in zero-sum multiplayer extensive-form games. In *ArXiv preprint:2009.12629*, 2020.

## A. Theoretical Details

### A.1. Representing Distributions of Play via Extensive-Form Correlation Plans

As mentioned in the body, every distribution over *randomized* strategy profiles for the team members is equivalent to a different distribution over *deterministic* strategy profiles by means of Kuhn's theorem (Kuhn, 1953), one of the most fundamental results about extensive-form game playing. Specifically, given two independent mixed strategies  $\mathbf{y}_{T1} \in \mathcal{Y}_{T1}$  and  $\mathbf{y}_{T2} \in \mathcal{Y}_{T2}$  for the team members, let  $\mu_{T1}$  and  $\mu_{T2}$  be the distributions over normal-form plans  $\Pi_{T1}, \Pi_{T2}$  equivalent to  $\mathbf{y}_{T1}$  and  $\mathbf{y}_{T2}$ , respectively. Then, the distribution over reandomized strategy profiles that assigns probability 1 to  $(\mathbf{y}_{T1}, \mathbf{y}_{T2})$  is equivalent to the product distribution of  $\mu_{T1}$  and  $\mu_{T2}$ , that is, the distribution over  $\Pi_{T1} \times \Pi_{T2}$  that picks a generic profile  $(\pi_{T1}, \pi_{T2})$  with probability  $\pi_{T1}(\pi_{T1}) \times \pi_{T2}(\pi_{T2})$ . The reverse is also true: a product distribution over  $\Pi_{T1} \times \Pi_{T2}$  is equivalent to a distribution over randomized profiles that picks exactly one profile with probability 1.

We now show that a similar result holds when the distribution over normal-form plans is represented as an extensive-form correlation plan. First, we introduce the notion of *product correlation plan*.

**Definition 3.** Let  $\xi_T \in \mathcal{V}$  be a vector in the von Stengel-Forges polytope. We say that  $\xi_T$  is a *product correlation plan* if

$$\xi_T[\sigma_{T1}, \sigma_{T2}] = \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma_{T2}]$$

for all  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \times \Sigma_{T2}$ .

**Lemma 1.** A *product correlation plan* is always an element of  $\Xi_T$ .

*Proof.* Let  $\xi_T$  be a product correlation plan. Since by definition,  $\xi_T \in \mathcal{V}$ , the vectors  $\mathbf{y}_{T1}, \mathbf{y}_{T2}$  indexed over  $\Sigma_{T1}$  and  $\Sigma_{T2}$ , respectively, and defined as

$$y[\sigma_{T1}] = \xi_T[\sigma_{T1}, \emptyset], y[\sigma_{T2}] = \xi_T[\emptyset, \sigma_{T2}]$$

are sequence-form strategies. By Kuhn's theorem, there exist distributions  $\mu_{T1}, \mu_{T2}$  over  $\Pi_{T1}$  and  $\Pi_{T2}$ , respectively, such that

$$y[\sigma_{T1}] = \sum_{\pi_{T1} \in \Pi_{T1}(\sigma_{T1})} \mu_{T1}[\pi_{T1}] \quad \forall \sigma_{T1} \in \Sigma_{T1}, \quad (3)$$

$$y[\sigma_{T2}] = \sum_{\pi_{T2} \in \Pi_{T2}(\sigma_{T2})} \mu_{T2}[\pi_{T2}] \quad \forall \sigma_{T2} \in \Sigma_{T2}. \quad (4)$$

Consider the distribution  $\mu_T$  over  $\Pi_{T1} \times \Pi_{T2}$  defined as the product distribution  $\mu_{T1} \otimes \mu_{T2}$ , that is,

$$\mu_T[\sigma_{T1}, \sigma_{T2}] := \mu_{T1}[\pi_{T1}] \cdot \mu_{T2}[\pi_{T2}]$$

for all  $(\pi_{T1}, \pi_{T2}) \in \Pi_{T1} \times \Pi_{T2}$ . We will show that is the extensive-form correlation plan corresponding to  $\mu_T$  according to (2), that is,

$$\xi_T[\sigma_{T1}, \sigma_{T2}] := \sum_{\substack{\pi_{T1} \in \Pi_{T1}(\sigma_{T1}) \\ \pi_{T2} \in \Pi_{T2}(\sigma_{T2})}} \mu_T[\pi_{T1}, \pi_{T2}]$$

for all  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \times \Sigma_{T2}$ . Indeed, using the fact that  $\xi_T$  is a *product correlation plan* together with (3) and (4):

$$\begin{aligned} \xi_T[\sigma_{T1}, \sigma_{T2}] &= \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma_{T2}] \\ &= y_{T1}[\sigma_{T1}] \cdot y_{T2}[\sigma_{T2}] \\ &= \left( \sum_{\pi_{T1} \in \Pi_{T1}(\sigma_{T1})} \mu_{T1}[\pi_{T1}] \right) \left( \sum_{\pi_{T2} \in \Pi_{T2}(\sigma_{T2})} \mu_{T2}[\pi_{T2}] \right) \\ &= \sum_{\substack{\pi_{T1} \in \Pi_{T1}(\sigma_{T1}) \\ \pi_{T2} \in \Pi_{T2}(\sigma_{T2})}} \mu_{T1}[\pi_{T1}] \cdot \mu_{T2}[\pi_{T2}] \\ &= \sum_{\substack{\pi_{T1} \in \Pi_{T1}(\sigma_{T1}) \\ \pi_{T2} \in \Pi_{T2}(\sigma_{T2})}} \mu_T[\pi_{T1}, \pi_{T2}]. \end{aligned}$$

This concludes the proof.  $\square$

**Lemma 2.** An *extensive-form correlation plan* is equivalent to a *distribution of play for the team* that picks one profile of randomized strategies  $(\mathbf{y}_{T1}, \mathbf{y}_{T2}) \in \mathcal{Y}_{T1} \times \mathcal{Y}_{T2}$  if and only if  $\xi_T$  is a *product correlation plan*. Furthermore, when that is the case,  $y_{T1}[\sigma_{T1}] = \xi_T[\sigma_{T1}, \emptyset]$ ,  $y_{T2}[\sigma_{T2}] = \xi_T[\emptyset, \sigma_{T2}]$  for all  $\sigma_{T1} \in \Sigma_{T1}, \sigma_{T2} \in \Sigma_{T2}$ .

*Proof.* The proof of Lemma 1 already shows that when  $\xi_T$  is a product correlation plan, it is equivalent to playing according to the distribution of play for the team with singleton support  $(\mathbf{y}_{T1}, \mathbf{y}_{T2})$ , where  $y_{T1}[\sigma_{T1}] = \xi_T[\sigma_{T1}, \emptyset]$ ,  $y_{T2}[\sigma_{T2}] = \xi_T[\emptyset, \sigma_{T2}]$  for all  $\sigma_{T1} \in \Sigma_{T1}, \sigma_{T2} \in \Sigma_{T2}$ . So, the only statement that remains to prove is that distributions  $\mu_T$  over randomized strategy profiles for the team members with a singleton support are mapped (Eq. (2)) to product correlation plans.

Let  $\{(\mathbf{y}_{T1}, \mathbf{y}_{T2})\} \subseteq \mathcal{Y}_{T1} \times \mathcal{Y}_{T2}$  be the (singleton) support of  $\mu_T$ , and let  $\mu_{T1}, \mu_{T2}$  be distributions over  $\Pi_{T1}$  and  $\Pi_{T2}$ , respectively, equivalent to  $\mathbf{y}_{T1}$  and  $\mathbf{y}_{T2}$ . Then,

$$y[\sigma_{T1}] = \sum_{\pi_{T1} \in \Pi_{T1}(\sigma_{T1})} \mu_{T1}[\pi_{T1}] \quad \forall \sigma_{T1} \in \Sigma_{T1}, \quad (5)$$

$$y[\sigma_{T2}] = \sum_{\pi_{T2} \in \Pi_{T2}(\sigma_{T2})} \mu_{T2}[\pi_{T2}] \quad \forall \sigma_{T2} \in \Sigma_{T2}. \quad (6)$$

Since by assumption the two team members sample strategies independently, their equivalent distribution of play over deterministic strategies is the product distribution  $\mu_T :=$

$\mu_{T1} \otimes \mu_{T2}$ . Using (2),  $\mu_T$  has a representation as extensive-form correlation plan given by

$$\begin{aligned} \xi_T[\sigma_{T1}, \sigma_{T2}] &= \sum_{\substack{\pi_{T1} \in \Pi_{T1}(\sigma_{T1}) \\ \pi_{T2} \in \Pi_{T2}(\sigma_{T2})}} \mu_T[\pi_{T1}, \pi_{T2}] \\ &= \sum_{\substack{\pi_{T1} \in \Pi_{T1}(\sigma_{T1}) \\ \pi_{T2} \in \Pi_{T2}(\sigma_{T2})}} \mu_{T1}[\pi_{T1}] \cdot \mu_{T2}[\pi_{T2}] \\ &= \left( \sum_{\pi_{T1} \in \Pi_{T1}(\sigma_{T1})} \mu_{T1}[\pi_{T1}] \right) \left( \sum_{\pi_{T2} \in \Pi_{T2}(\sigma_{T2})} \mu_{T2}[\pi_{T2}] \right) \\ &= y_{T1}[\sigma_{T1}] \cdot y_{T2}[\sigma_{T2}] \end{aligned} \quad (7)$$

for all  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \times \Sigma_{T2}$ . In particular, choosing  $\sigma_{T2} = \emptyset$  in (7), and using the fact that  $y_{T2}[\emptyset] = 1$ , we obtain

$$\xi_T[\sigma_{T1}, \emptyset] = y_{T1}[\sigma_{T1}] \quad \forall \sigma_{T1} \in \Sigma_{T1}.$$

Similarly,

$$\xi_T[\emptyset, \sigma_{T2}] = y_{T2}[\sigma_{T2}] \quad \forall \sigma_{T2} \in \Sigma_{T2}.$$

Substituting the last two equalities into (7) we can write

$$\xi_T[\sigma_{T1}, \sigma_{T2}] = \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma_{T2}]$$

for all  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \times \Sigma_{T2}$ . That, together with the inclusion  $\Xi_T \subseteq \mathcal{V}$ , shows that  $\xi_T$  is a product correlation plan.  $\square$

### Semi-randomized correlation plans are product plans

In the body we mentioned that semi-randomized correlation plans correspond to a distribution of play where one team member plays a deterministic strategy and the other team member plays a randomized strategy. We now give more formal grounding that that assertion.

**Lemma 3.** *Let  $\xi_T \in \Xi_{T1}^* \cup \Xi_{T2}^*$  be a semi-randomized plan. Then,  $\xi_T$  is a product plan.*

We reuse some ideas that already appeared in Farina & Sandholm (2020) to prove Lemma 3. In particular, in the proof we will make use of the following lemma.

**Lemma 4** (Farina & Sandholm (2020, Lemma 6)). *Let  $\xi_T \in \mathcal{V}$ . For all  $\sigma_{T1} \in \Sigma_{T1}$  such that  $\xi_T[\sigma_{T1}, \emptyset] = 0$ ,  $\xi_T[\sigma_{T1}, \sigma_{T2}] = 0$  for all  $\sigma_{T2} \in \Sigma_{T2} : \sigma_{T1} \bowtie \sigma_{T2}$ . Similarly, for all  $\sigma_{T2} \in \Sigma_{T2}$  such that  $\xi_T[\emptyset, \sigma_{T2}] = 0$ ,  $\xi_T[\sigma_{T1}, \sigma_{T2}] = 0$  for all  $\sigma_{T1} \in \Sigma_{T1} : \sigma_{T1} \bowtie \sigma_{T2}$ .*

*Proof of Lemma 3.* We will only show the proof for the case  $\xi_T \in \Xi_{T1}^*$ . The other case ( $\xi_T \in \Xi_{T2}^*$ ) is symmetric.

To show that

$$\xi_T[\sigma_{T1}, \sigma_{T2}] = \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma_{T2}]$$

for all  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \bowtie \Sigma_{T2}$ , we perform induction on the depth of the sequence  $\sigma_{T2}$ . The depth  $\text{depth}(\sigma_{T2})$  of a generic sequence  $\sigma_{T2} = (J, b) \in \Sigma_{T2}$  of Player  $i$  is defined as the number of actions that Player  $T2$  plays on the path from the root of the tree down to action  $b$  at information set  $J$ , included. Conventionally, we let the depth of the empty sequence be 0.

The base case for the induction proof corresponds to the case where  $\sigma_{T2}$  has depth 0, that is,  $\sigma_{T2} = \emptyset$ . In that case, the theorem is clearly true, because  $\xi_T[\emptyset, \emptyset] = 1$  as part of the von Stengel-Forges constraints (Definition 1).

Now, suppose that the statement holds as long as  $\text{depth}(\sigma_{T2}) \leq d$ . We will show that the statement will hold for any  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \bowtie \Sigma_{T2}$  such that  $\text{depth}(\sigma_{T2}) \leq d + 1$ . Indeed, consider  $(\sigma_{T1}, \sigma_{T2}) \in \Sigma_{T1} \bowtie \Sigma_{T2}$  such that  $\sigma_{T2} = (J, b)$  with  $\text{depth}(\sigma_{T2}) = d + 1$ .

There are only two possible cases:

- Case 1:  $\xi_T[\emptyset, \sigma_{T2}] = 0$ . From Lemma 4,  $\xi_T[\sigma_{T1}, \sigma_{T2}] = 0$  and the statement holds.
- Case 2:  $\xi_T[\emptyset, \sigma_{T2}] = 1$ . From the von Stengel-Forges constraints,  $\xi_T[\emptyset, \sigma(J)] = \sum_{b' \in A_J} \xi_T[\emptyset, (J, b')] = 1 + \sum_{b' \in A_J, b' \neq b} \xi_T[\emptyset, (J, b')] \geq 1$ . Hence, because all entries of  $\xi_T[\emptyset, \sigma_2]$  are in  $\{0, 1\}$  by definition of  $\Xi_{T1}^*$ , it must be  $\xi_T[\emptyset, \sigma(J)] = 1$  and  $\xi_T[\emptyset, (J, b')] = 0$  for all  $b' \in A_J, b' \neq b$ .

Using the inductive hypothesis, we have that

$$\xi_T[\sigma_{T1}, \sigma(J)] = \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma(J)] = \xi_T[\sigma_{T1}, \emptyset] \quad (8)$$

for all  $\sigma_{T1} \in \Sigma_{T1}, \sigma_{T1} \bowtie \sigma(J)$ . On the other hand, since  $\xi_T[\emptyset, (J, b')] = 0$  for all  $b' \in A_J, b' \neq b$ , from Lemma 4 we have that

$$\xi_T[\sigma_{T1}, (J, b')] = 0 \quad \forall \sigma_{T1} \bowtie J, b' \neq b. \quad (9)$$

Hence, summing over all  $b' \in A_J$  and using the von Stengel-Forges constraints, we get

$$\begin{aligned} \xi_T[\sigma_{T1}, \emptyset] \cdot \xi_T[\emptyset, \sigma_{T2}] &= \xi_T[\sigma_{T1}, \sigma(J)] \\ &= \sum_{b' \in A_J} \xi_T[\sigma_{T1}, (J, b')] \\ &= \xi_T[\sigma_{T1}, (J, b)] = \xi_T[\sigma_{T1}, \sigma_{T2}] \end{aligned}$$

for all  $\sigma_{T1} \bowtie (J, b)$ . This concludes the proof by induction.  $\square$

So, from Lemma 2 it follows that semi-randomized plans correspond to distributions of play over randomized profiles with the singleton support  $(y_{T1}, y_{T2}) \in \mathcal{Y}_{T1} \times \mathcal{Y}_{T2}$ . Furthermore, because of the second part of Lemma 2, when  $\xi_T \in \Xi_{T1}^*$ ,  $y_{T2}[\sigma_{T2}] \in \{0, 1\}$  for all  $\sigma_{T2} \in \Sigma_{T2}$ , which means that  $y_{T2}$  is a deterministic strategy for Player  $T2$  (a similar statement holds for  $\xi_T \in \Xi_{T2}^*$ ).



we have

$$\Xi_T = \text{co}\{f(\mathbb{1}_{(\pi_{T1}, \pi_{T2})}) : \pi_{T1} \in \Pi_{T1}, \pi_{T2} \in \Pi_{T2}\}.$$

Hence, to conclude the proof of this part, it will be enough to show that for each  $\pi_{T1} \in \Pi_{T1}, \pi_{T2} \in \Pi_{T2}$ , it holds that  $f(\mathbb{1}_{(\pi_{T1}, \pi_{T2})}) \in \Xi_{T1}^*$ . Since  $\mathbb{1}_{(\pi_{T1}, \pi_{T2})}$  assigns probability 1 to one profile and 0 to all other profiles,  $f(\mathbb{1}_{(\pi_{T1}, \pi_{T2})})$  is an extensive-form correlation plan whose entries are all in  $\{0, 1\}$ . So, in particular,  $f(\mathbb{1}_{(\pi_{T1}, \pi_{T2})}) \in \Xi_{T1}^*$ . This concludes the proof of the inclusion  $\Xi_T \subseteq \text{co} \Xi_{T1}^*$ .

Together, the two statements that we just prove show that  $\Xi_T = \text{co} \Xi_{T1}^*$ .

Finally, using the fact that unions and convex hulls commute, we have

$$\text{co}(\Xi_{T1}^* \cup \Xi_{T2}^*) = (\text{co} \Xi_{T1}^*) \cup (\text{co} \Xi_{T2}^*) = \Xi_T \cup \Xi_T = \Xi_T,$$

thereby concluding the proof.  $\square$

## B. Game Instances

The size of the parametric instances we use as benchmark is described in Table 1. In the following, we provide a detailed explanation of the rules of each game.

**Kuhn poker** Two-player Kuhn poker was originally proposed by Kuhn (1950). We employ the three-player variation described in Farina et al. (2018). In a three-player Kuhn poker game with rank  $r$  there are  $r$  possible cards. At the beginning of the game, each player pays one chip to the pot, and each player is dealt a single private card. The first player can check or bet, i.e., putting an additional chip in the pot. Then, the second player can check or bet after a first player’s check, or fold/call the first player’s bet. If no bet was previously made, the third player can either check or bet. Otherwise, the player has to fold or call. After a bet of the second player (resp., third player), the first player (resp., the first and the second players) still has to decide whether to fold or to call the bet. At the showdown, the player with the highest card who has not folded wins all the chips in the pot.

**Goofspiel** This bidding game was originally introduced by Ross (1971). We use a 3-rank variant, that is, each player has a hand of cards with values  $\{-1, 0, 1\}$ . A third stack of cards with values  $\{-1, 0, 1\}$  is shuffled and placed on the table. At each turn, a prize card is revealed, and each player privately chooses one of his/her cards to bid, with the highest card winning the current prize. In case of a tie, the prize is split evenly among the winners. After 3 turns, all the prizes have been dealt out and the payoff of each player

is computed as follows: each prize card’s value is equal to its face value and the players’ scores are computed as the sum of the values of the prize cards they have won.

**Goofspiel with limited information** This is a variant of Goofspiel introduced by Lanctot et al. (2009). In this variation, in each turn the players do not reveal the cards that they have played. Rather, they show their cards to a fair umpire, which determines which player has played the highest card and should therefore received the prize card. In case of tie, the umpire directs the players to split the prize evenly among the winners, just like in the Goofspiel game. This makes the game strategically more challenging as players have less information regarding previous opponents’ actions.

**Leduc poker** We use a three-player version of the classical Leduc hold’em poker introduced by Southey et al. (2005). We employ game instances of rank 3, in which the deck consists of three suits with 3 cards each. Our instances are parametric in the maximum number of bets, which in limit hold’em is not necessarily tied to the number of players. The maximum number of raise per betting round can be either 1, 2 or 3. As the game starts players pay one chip to the pot. There are two betting rounds. In the first one a single private card is dealt to each player while in the second round a single board card is revealed. The raise amount is set to 2 and 4 in the first and second round, respectively.

**Liar’s dice** Liar’s dice is another standard benchmark introduced by Lisỳ et al. (2015). In our three-player implementation, at the beginning of the game each of the three players privately rolls an unbiased  $k$ -face die. Then, the three players alternate in making (potentially false) claims about their toss. The first player begins bidding, announcing any face value up to  $k$  and the minimum number of dice that the player believes are showing that value among the dice of all the players. Then, each player has two choices during their turn: to make a higher bid, or to challenge the previous bid by declaring the previous bidder a ”liar”. A bid is higher than the previous one if either the face value is higher, or the number of dice is higher. If the current player challenges the previous bid, all dice are revealed. If the bid is valid, the last bidder wins and obtains a reward of +1 while the challenger obtains a negative payoff of -1. Otherwise, the challenger wins and gets reward +1, and the last bidder obtains reward of -1. All the other players obtain reward 0. We test our algorithms on Liar’s dice instances with  $k = 3$  and  $k = 4$ .

## C. Additional Experimental Results

All experiments were run 10 times, and the experimental tables show average run times. We always use the same random seed to sample no-regret strategies for the team

---

members in the seeding phase of our column-generation algorithm. The seed was never changed, and we don't treat it as a hyperparameter. So, all algorithms are deterministic, and the only source of randomness in the run time is due to system load. Consequently, we observed small standard deviations in the run times, less than 10% in all cases.

We used the same time limit for FTP that was found to be beneficial by the original authors (Farina et al., 2018), namely 15 seconds. For FTP and CG-18, we used the original implementations, with permission from the authors. In all algorithms, we observed that the majority of time is spent within Gurobi.

Table 4 and Table 5 show the comparison between our column-generation algorithm, FTP, and CG-18 when the opponent plays as the first and as the second player, respectively.

### **Comparison between the Algorithm of Section 5 and the Prior State of the Art**

Depending on the cap  $n$  on the number of semi-randomized correlation plans, the algorithm we describe in Section 5 might not reach the optimal TMEC<sub>or</sub> value for the team (although, as we argue in Section 7, a very small  $n$  already guarantees a large fraction of the optimal value empirically).

For completeness, we report the run time of the algorithm for a sample instance. We employ instance [H] with  $\bigcirc = 3$  as it has a good trade-off between dimensions and manageability. When  $n = 1$  the algorithm reaches an optimal solution in 9.21s. The optimal solution with  $n = 1$  achieves 15% of the optimal utility with no restrictions on the number of plans. With  $n = 2$  the run time is 12m 05s and the solution reaches 90% of the optimal value. With  $n = 3$  the run time is 1h 53m and the solution guarantees 92% of the optimal value.

The column-generation algorithm has better run time performances and guarantees to reach an optimal solution without having to pick the right support size. However, we observe that the algorithm of Section 5 already outperforms FTP and CG-18. Specifically, FTP cannot reach a strategy guaranteeing 50% of the optimal utility within the time limit, while our algorithm guarantees 90% of the optimal value within roughly 10 minutes. On the other hand, CG-18 cannot complete even a single iteration within the time limit. This confirms that our pricing formulation is significantly tighter than previous formulations.

Game	Ours		Fictitious Team Play (FTP)			CG-18	Pricers		Team utility after seeding			TMECor value
	Seeded	Not seed.	$\epsilon = 50\%$	$\epsilon = 10\%$	$\epsilon = 1\%$		Relax.	MIP	$m = 0$	100	1000	
[A]	2ms	2ms	0ms <sup>†</sup>	15.00s <sup>†</sup>	2m 35s <sup>†</sup>	66ms	5	0	-0.481	-0.133	-0.133	0
[B]	17ms	27ms	0ms	16m 39s	> 6h	1.01s	0	3	-0.307	0.037	0.037	0.038
[C]	4.67s	6.86s	7m 36s	> 6h	> 6h	> 6h	7	38	-0.381	0.055	0.058	0.066
[D]	302ms	654ms	2ms	> 6h	> 6h	1m 56s	17	0	-1.000	0.239	0.251	0.252
[E]	821ms	1.52s	6ms	> 6h	> 6h	23m 17s	31	0	-1.110	0.242	0.249	0.253
[F]	5.73s	15.34s	19m 25s	> 6h	> 6h	> 6h	1	0	-0.926	0	0	0
[G]	8.08s	1h 09m	> 6h	> 6h	> 6h	> 6h	0	2	-0.696	0.039	0.063	0.063
[H]	32.12s	40.16s	2h 49m	> 6h	> 6h	> 6h	1	90	-2.000	0.001	0.173	0.199
[I]	14m 51s	15m 12s	> 6h	> 6h	> 6h	> 6h	0	232	-2.000	-0.020	0.105	0.186
[J]	11m 23s	8m 24s	> 6h	> 6h	> 6h	> 6h	1042	123	-2.000	-0.433	0.395	0.549

(a) — Comparison of run times

(b)

(c)

Table 4: **Results for  $\alpha = 1$ .** (a) Runtime comparison between our column generation algorithm, FTP, and CG-18. The seeded version of our algorithm runs  $m = 1000$  iterations of CFR+ (Section 6.2), while the non seeded version runs  $m = 0$ . ‘†’: since the TMECor value for the game is exactly zero, we measure how long it took the algorithm to find a distribution with expected value at least  $-\epsilon/10$  for the team. (b) Number of times the pricing problem for our column-generation algorithm was solved to optimality by the linear relaxation (‘Relax’) and by the MIP solver (‘MIP’) when using our column-generation algorithm (seeded version with  $m = 1000$ ). (c) Quality of the initial strategy of the team obtained for varying sizes of  $S$  compared to the expected utility of the team at the TMECor.

Game	Ours		Fictitious Team Play (FTP)			CG-18	Pricers		Team utility after seeding			TMECor value
	Seeded	Not seed.	$\epsilon = 50\%$	$\epsilon = 10\%$	$\epsilon = 1\%$		Relax.	MIP	$m = 0$	100	1000	
[A]	0ms	3ms	0ms <sup>†</sup>	19s <sup>†</sup>	3m 09s <sup>†</sup>	147ms	1	0	-0.667	0	0	0
[B]	0ms	12ms	1m 39s	> 6h	> 6h	7.53s	1	0	-0.281	0.027	0.027	0.027
[C]	4.01s	4.22s	48m 08s	> 6h	> 6h	> 6h	6	26	-0.636	0.018	0.027	0.038
[D]	221ms	696ms	1ms	> 6h	> 6h	1m 46s	13	0	-1.000	0.247	0.252	0.252
[E]	1.11s	1.37s	1.39s	> 6h	> 6h	12m 30s	42	0	-1.110	0.241	0.252	0.253
[F]	20.72s	1m 03s	1h 30m	> 6h	> 6h	> 6h	38	2	-0.778	0.246	0.256	0.256
[G]	5h 48m	5h 44m	> 6h	> 6h	> 6h	> 6h	140	17	-0.969	0.260	0.264	—
[H]	2m 02s	2m 44s	> 6h	> 6h	> 6h	> 6h	24	168	-1.679	0.061	0.164	0.253
[I]	27m 17s	27m 50s	> 6h	> 6h	> 6h	> 6h	6	553	-1.911	-0.011	0.076	0.183
[J]	28m 05s	11m 21s	> 6h	> 6h	> 6h	> 6h	4600	254	-2.000	0.190	0.392	0.628

(a) — Comparison of run times

(b)

(c)

Table 5: **Results for  $\alpha = 2$ .** (a) Runtime comparison between our column generation algorithm, FTP, and CG-18. The seeded version of our algorithm runs  $m = 1000$  iterations of CFR+ (Section 6.2), while the non seeded version runs  $m = 0$ . ‘†’: since the TMECor value for the game is exactly zero, we measure how long it took the algorithm to find a distribution with expected value at least  $-\epsilon/10$  for the team. (b) Number of times the pricing problem for our column-generation algorithm was solved to optimality by the linear relaxation (‘Relax’) and by the MIP solver (‘MIP’) when using our column-generation algorithm (seeded version with  $m = 1000$ ). (c) Quality of the initial strategy of the team obtained for varying sizes of  $S$  compared to the expected utility of the team at the TMECor. ‘oom’: out of memory.