Faster Algorithms for Optimal Ex-Ante Coordinated Collusive Strategies in Extensive-Form Zero-Sum Games

Gabriele Farina,[1] **Andrea Celli,**[2] **Nicola Gatti,**[2] **Tuomas Sandholm**[1,3,4,5,]
[1]Computer Science Department, Carnegie Mellon University, [2]DEIB, Politecnico di Milano
[3]Strategic Machine, Inc., [4]Strategy Robot, Inc., [5]Optimized Markets, Inc.
gfarina@cs.cmu.edu, {andrea.celli, nicola.gatti}@polimi.it, sandholm@cs.cmu.edu

Abstract

We focus on the problem of finding an optimal strategy for a team of two 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 that 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 computes such an optimal distribution by only using profiles where only one of the team members gets to randomize in each profile. We can also cap the number of such profiles we allow 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 relatively 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 and 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. Also, all superhuman AI gaming milestones have been in two-player zero-sum games, with the exception of multi-player no-limit Texas hold'em recently (Brown and Sandholm 2019).

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 multi-player games (von Stengel and Koller 1997; Celli and 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 and 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. 2017; Zhang and 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 and 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 and Gatti 2018; Farina et al. 2018). Finding a TMECor strategy is NP-hard and inapproximable (Celli and Gatti 2018).

We propose a new formulation for the problem of finding a TMECor strategy. In doing so, we introduce the key notion of a *semi-randomized correlation plan* and draw connections with a particular strategy polytope defined by von Stengel and 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. 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 the theoretical concepts that we introduce, we develop an efficient optimal column-generation algorithm for finding a TMECor strategy. 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. It can also solve many games that were previously unsolvable.

2 Preliminaries: Extensive-Form Games

Extensive-form games (EFGs) are a standard model in game theory. They 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 $i \in \{\mathsf{T1}, \mathsf{T2}, \mathsf{O}\} \cup \{\mathrm{C}\}$ 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 $i \in \{\mathsf{T1}, \mathsf{T2}, \mathsf{O}\}$ 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 $p_{\mathrm{C}}(z)$.

Private information is represented via *information set* (infoset). In particular, the set of nodes belonging to $i \in \{\mathsf{T1}, \mathsf{T2}, \mathsf{O}\}$ 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 Σ_i, is defined as $\Sigma_i := \{(I, a) : I \in \mathcal{I}_i, a \in A_I\} \cup \{\varnothing\}$, where the special element $\varnothing$ 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 $\sigma(I) = \varnothing$ 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 the 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 $\sigma_i \bowtie \sigma_2$ 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 σ_i and $I_j \in \mathcal{I}_j$, we say that (σ_i, I_j) forms a relevant sequence-information set pair ($\sigma_i \bowtie I_j$), if $\sigma_i = \varnothing$ or if $\sigma_i = (I_i, a_i)$ and $I_i \rightleftharpoons I_j$.

Reduced-normal-form plans. A *reduced-normal-form* plan π_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 π itself. The set of reduced-normal-form plans of Player i is denoted Π_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 $\Pi_i(\varnothing) := \Pi_i$. Finally, given a leaf $z \in Z$, we denote with $\Pi_i(z) \subseteq \Pi_i$ the set of reduced-normal-form 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, Megiddo, and von Stengel 1996). Given a player $i \in \{\mathsf{T1}, \mathsf{T2}, \mathsf{O}\}$ and a normal-form strategy $\mu \in \Delta(\Pi_i)$,[1] the sequence-form strategy induced by μ is the real vector $\boldsymbol{y}$, 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 μ 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, Megiddo, and von Stengel 1996).

3 TMECor Formulation and Prior Work

A TMECor strategy is a probability distribution μ_{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 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.

[1] $\Delta(X)$ denotes the probability simplex over the finite set X.

TMECor as a Bilinear Saddle-Point Problem. For each leaf z, let $\hat{u}_{\mathsf{T}}(z) := (u_{\mathsf{T1}}(z) + u_{\mathsf{T2}}(z))p_{\mathrm{C}}(z)$. 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

$$\operatorname*{arg\,max}_{\mu_{\mathsf{T}} \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})} \min_{\mu_{\mathsf{O}} \in \Delta(\Pi_{\mathsf{O}})} u_{\mathsf{T}}(\mu_{\mathsf{T}}, \mu_{\mathsf{O}}). \tag{1}$$

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

Prior algorithms. Prior work on the computation of TMECor mainly differs in the way the team's distribution μ_{T} is represented. Celli and Gatti (2018) directly represent the strategy as a probability distribution over the set of joint reduced-normal-form plans $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$. The number of bits required to store such a distribution is exponential in the size of the game tree in the worst case. They propose a column-generation approach in which, at each iteration, a new pair of pure strategies is added to the support of the distribution μ_{T}. Farina et al. (2018) show that it suffices to employ $|Z|$-dimensional *vectors of realizations* where each z is mapped to its probability of being reached when the team follows μ_{T}. A *realization-form strategy* is a more concise representation than the original distribution μ_{T}. The authors propose a structural decomposition of the problem which is then used to prove convergence of a fictitious-play-like algorithm.

4 A Formulation of TMECor Based on Extensive-Form Correlation Plans

We propose using a different representation of the correlated distribution of play μ_{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 μ_{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 μ_{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).

Extensive-Form Correlation Plans

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

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

Here $\boldsymbol{\xi}_{\mathsf{T}}$ 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 relevant sequence pairs is only a tiny fraction of the total number of sequence pairs, as shown in Table 1(b).

The set of extensive-form correlation plans $\boldsymbol{\xi}_{\mathsf{T}}$ that can be induced as μ_{T} varies over the set of all correlated distributions of play for the team members is a convex polytope. We denote it as Ξ_{T} and call it the *polytope of correlation plans*. We will recall existing results and provide new ones about the structure of Ξ_{T} in Section 5.

Computing a TMECor using Correlation Plans

Extensive-form correlation plans encode a *superset* of the information encoded by realization plans. Indeed, for all z, $\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] = \rho_{\mathsf{T}}[z]$. Using the previous identity, we can rewrite the problem of computing a TMECor of a constant-sum game (1) as

$$\operatorname*{arg\,max}_{\boldsymbol{\xi}_{\mathsf{T}} \in \Xi_{\mathsf{T}}} \min_{\boldsymbol{y}_{\mathsf{O}} \in \mathcal{Y}_{\mathsf{O}}} \sum_{z \in Z} \hat{u}_{\mathsf{T}}(z) \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] y[\sigma_{\mathsf{O}}(z)].$$

By dualizing the inner linear minimization problem over $\boldsymbol{y}_{\mathsf{O}}$, we get the following proposition that shows that a TMECor can be found as the solution to a linear program (LP) with a polynomial number of variables. (All the proofs of this paper can be found in the appendix.)

Proposition 1. *An extensive-form correlation plan $\boldsymbol{\xi}_{\mathsf{T}}$ is a TMECor if and only if it is a solution to the LP*

$$\begin{cases} \operatorname*{arg\,max}_{\boldsymbol{\xi}_{\mathsf{T}}} \quad v_{\varnothing}, \quad \text{subject to:} \\ ① \; v_I - \displaystyle\sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I') = (I,a)}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z) = (I,a)}} \hat{u}_{\mathsf{T}}(z) \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \quad \forall\, (I,a) \in \Sigma_{\mathsf{O}} \setminus \{\varnothing\} \\ ② \; v_{\varnothing} - \displaystyle\sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I') = \varnothing}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z) = \varnothing}} \hat{u}_{\mathsf{T}}(z) \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \\ ③ \; v_{\varnothing} \text{ free}, v_I \text{ free} \quad \forall\, I \in \mathcal{I}_{\mathsf{O}} \\ ④ \; \boldsymbol{\xi}_{\mathsf{T}} \in \Xi_{\mathsf{T}}. \end{cases}$$

As a direct consequence of Proposition 1, a TMECor can be found in polynomial time whenever Ξ_{T} can be represented as the intersection of a set of polynomially many linear constraints. In Section 5, we recall when that is the case.

Game instance	Num. sequences $\lvert\Sigma_1\rvert$	$\lvert\Sigma_2\rvert$	$\lvert\Sigma_3\rvert$	Num. leaves $\lvert Z\rvert$	$\frac{\lvert\Sigma_{\mathsf{T1}} \bowtie \Sigma_{\mathsf{T2}}\rvert}{\lvert Z\rvert}$	$\frac{\lvert\Sigma_{\mathsf{T1}} \times \Sigma_{\mathsf{T2}}\rvert}{\lvert\Sigma_{\mathsf{T1}} \bowtie \Sigma_{\mathsf{T2}}\rvert}$	Triangle-free? O $= 1$	O $= 2$	O $= 3$
[A] Kuhn poker (3 ranks)	25	25	25	78	3.40	2.36	✗	✗	✗
[B] Kuhn poker (4 ranks)	33	33	33	312	1.59	2.19	✗	✗	✗
[C] Kuhn poker (12 ranks)	97	97	97	17 160	0.29	1.90	✗	✗	✗
[D] Goofspiel (3 ranks, limited info)	934	934	934	1296	9.54	70.59	✓	✓	✓
[E] Goofspiel (3 ranks)	1630	1630	1630	1296	15.54	131.96	✓	✓	✓
[F] Liar's dice (3 faces)	1021	1021	1021	13 797	5.27	14.43	✗	✗	✗
[G] Liar's dice (4 faces)	10 921	10 921	10 921	262 080	6.25	72.79	✗	✗	✗
[H] Leduc poker (3 ranks, 1 raise)	457	457	457	4500	2.62	17.70	✗	✗	✗
[I] Leduc poker (4 ranks, 1 raise)	801	801	801	16908	1.34	28.36	✗	✗	✗
[J] Leduc poker (2 ranks, 2 raises)	1443	1443	1443	3786	7.28	75.59	✗	✗	✗

(a) — Game instances and sizes; **(b)**; **(c)**

Table 1: **(a)** Size of the game instances used in our experiments, in terms of number of sequences $|\Sigma_i|$ for each player i, and number of leaves $|Z|$. **(b)** Ratio between the number of leaves $|Z|$, number of sequence pairs for the team members $|\Sigma_{\mathsf{T1}} \times \Sigma_{\mathsf{T2}}|$, and number of *relevant* sequence pairs for the team members $|\Sigma_{\mathsf{T1}} \bowtie \Sigma_{\mathsf{T2}}|$ in various benchmark games. For all games reported in the subtable, we chose the first two players to act as the team members. **(c)** The subtable reports whether the interaction of the team members is triangle-free (Farina and Sandholm 2020), given the opponent player O.

5 Semi-Randomized Correlation Plans and the Structure of Ξ_{T}

Even though Ξ_{T} is a convex polytope, the set of (potentially exponentially many) linear constraints that define it is not known in general. So, alternative characterizations of the set Ξ_{T} are needed before the LP in Proposition 1 can be solved. In this section, we recall two known results about the structure of Ξ_{T}, and propose a new one (Proposition 3). We will use our result to arrive at two different approaches to tackle the LP of Proposition 1 in Section 6 and 7, respectively.

Containment in the von Stengel-Forges Polytope

The first result about the structure of Ξ_{T} has to do with a particular polytope that was introduced by von Stengel and Forges (2008).

Definition 1. *The* von Stengel-Forges polytope *of the team, denoted $\mathcal{V}_{\mathsf{T}}$, is the polytope of all vectors $\boldsymbol{\xi} \in \mathbb{R}_{\geq 0}^{|\Sigma_{\mathsf{T1}} \bowtie \Sigma_{\mathsf{T2}}|}$ indexed over relevant sequence pairs that satisfy the following polynomially-sized set of linear constraints.*

① $\xi[\varnothing, \varnothing] = 1$

② $\displaystyle\sum_{a_{\mathsf{T1}} \in A_{I_{\mathsf{T1}}}} \xi[(I_{\mathsf{T1}}, a_{\mathsf{T1}}), \sigma_{\mathsf{T2}}] = \xi[\sigma(I_{\mathsf{T1}}), \sigma_{\mathsf{T2}}] \quad \forall I_{\mathsf{T1}} \bowtie \sigma_{\mathsf{T2}}$

③ $\displaystyle\sum_{a_{\mathsf{T2}} \in A_{I_{\mathsf{T2}}}} \xi[\sigma_{\mathsf{T1}}, (I_{\mathsf{T2}}, a_{\mathsf{T2}})] = \xi[\sigma_{\mathsf{T1}}, \sigma(I_{\mathsf{T2}})] \quad \forall \sigma_{\mathsf{T1}} \bowtie I_{\mathsf{T2}}.$

These can be interpreted as "probability mass conservation" constraints. They are interlaced sequence-form constraints.

The following result by von Stengel and Forges (2008) is immediate from the definition of ξ_{T} in (2).

Proposition 2 (von Stengel and Forges (2008))**.** *The set of extensive-form correlation plans is a subset of the von Stengel-Forges polytope. Formally, $\Xi_{\mathsf{T}} \subseteq \mathcal{V}_{\mathsf{T}}$.*

Triangle-Freeness and Polynomial-Time Computation of TMECor

Proposition 2 shows that Ξ_{T} is a subset of the von Stengel-Forges polytope. There are games where the reverse inclusion does not hold. Farina and 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 and Sandholm (2020))**.** *The interaction of the team members* T1 *and* T2 *is* triangle-free *if, for any choice of distinct information sets $I_1, I_2 \in \mathcal{I}_{\mathsf{T1}}$ with $\sigma_{\mathsf{T1}}(I_1) = \sigma_{\mathsf{T1}}(I_2)$ and any choice of distinct information sets $J_1, J_2 \in \mathcal{I}_{\mathsf{T2}}$ with $\sigma_{\mathsf{T2}}(J_1) = \sigma_{\mathsf{T2}}(J_2)$, it is never the case that $(I_1 \rightleftharpoons J_1) \wedge (I_2 \rightleftharpoons J_2) \wedge (I_1 \rightleftharpoons J_2)$.*

Farina and Sandholm (2020) show that when the information structure of correlating players (in our case, the team members) is triangle-free, then $\Xi_{\mathsf{T}} = \mathcal{V}_{\mathsf{T}}$. 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 (and that none of the other common benchmark games that we consider are).

Semi-Randomized Correlation Plans

We now give a third result about the structure of Ξ_{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 subsets of the von Stengel-Forges polytope of the team. They represent strategy profiles 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

$$\Xi^*_{\mathsf{T1}} = \{\boldsymbol{\xi} \in \mathcal{V}_{\mathsf{T}} : \xi[\varnothing, \sigma_{\mathsf{T2}}] \in \{0,1\} \quad \forall\, \sigma_{\mathsf{T2}} \in \Sigma_{\mathsf{T2}}\},$$
$$\Xi^*_{\mathsf{T2}} = \{\boldsymbol{\xi} \in \mathcal{V}_{\mathsf{T}} : \xi[\sigma_{\mathsf{T1}}, \varnothing] \in \{0,1\} \quad \forall\, \sigma_{\mathsf{T1}} \in \Sigma_{\mathsf{T1}}\},$$

respectively. Crucially, a point $\boldsymbol{\xi} \in \Xi^*_i$ for $i \in \{\mathsf{T1}, \mathsf{T2}\}$ 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 Ξ_{T}.

Proposition 3. *In every game, Ξ_{T} is the convex hull of the set $\Xi^*_{\mathsf{T}1}$, or equivalently of the set $\Xi^*_{\mathsf{T}2}$. Formally, $\Xi_{\mathsf{T}} = \operatorname{co}\Xi^*_{\mathsf{T}1} = \operatorname{co}\Xi^*_{\mathsf{T}2} = \operatorname{co}(\Xi^*_{\mathsf{T}1} \cup \Xi^*_{\mathsf{T}2})$.*

6 Computing TMECor with a Small Support of Semi-Randomized Plans of Fixed Size

From Proposition 3, it is known that Ξ_{T} is the convex hull of $\Xi^*_{\mathsf{T}1}$ and $\Xi^*_{\mathsf{T}2}$. Furthermore, the polytopes $\Xi^*_{\mathsf{T}1}$ and $\Xi^*_{\mathsf{T}2}$ 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 $\boldsymbol{\xi}_{\mathsf{T}}$ be a convex combination of elements from $\Xi^*_{\mathsf{T}1}$ and $\Xi^*_{\mathsf{T}2}$. We introduce variables $\boldsymbol{\xi}^{(1)}_{\mathsf{T}}, \dots, \boldsymbol{\xi}^{(n)}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}1} \cup \Xi^*_{\mathsf{T}2}$ and the corresponding convex combination coefficients $\lambda^{(1)}, \dots, \lambda^{(n)}$, and replace Constraint ④ with the linear constraint $\boldsymbol{\xi}_{\mathsf{T}} = \sum_{i=1}^{n} \lambda^{(i)} \boldsymbol{\xi}^{(i)}_{\mathsf{T}}$. 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.

$$\begin{cases} \displaystyle\mathop{\arg\max}_{\boldsymbol{\xi}^{(1)}_{\mathsf{T}},\dots,\boldsymbol{\xi}^{(n)}_{\mathsf{T}},\lambda^{(1)},\dots,\lambda^{(n)}} \quad v_\varnothing, \quad \text{subject to:} \\ \text{constraints } \textcircled{1}\,\textcircled{2}\,\textcircled{3} \text{ as in Proposition 1} \\ \textcircled{4}\ \boldsymbol{\xi}_{\mathsf{T}} = \sum_{i=1}^{n} \lambda^{(i)} \boldsymbol{\xi}^{(i)}_{\mathsf{T}} \\ \textcircled{5}\ \boldsymbol{\xi}^{(1)}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}1}, \boldsymbol{\xi}^{(2)}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}2}, \boldsymbol{\xi}^{(3)}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}1}, \boldsymbol{\xi}^{(4)}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}2}, \dots^{\ddagger} \\ \textcircled{6}\ \sum_{i=1}^{n} \lambda^{(i)} = 1,\ \lambda^{(i)} \ge 0 \quad \forall i \in \{1,\dots,n\}. \end{cases}$$

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 $n \ge |\Sigma_1 \bowtie \Sigma_2| + 1$. As we show in detail in Section 8, 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 of carefully selected simple strategies often result in optimal coordination.

7 A Fast Column Generation Approach

In this section, we show a different approach to solving the LP in Proposition 1—using column generation (Ford and Fulkerson 1958). First, we proceed with a *seeding* phase. We pick a set S containing one or more points $\boldsymbol{\xi}^{(1)}_{\mathsf{T}}, \boldsymbol{\xi}^{(2)}_{\mathsf{T}}, \dots, \boldsymbol{\xi}^{(m)}_{\mathsf{T}}$ that are known to belong to Ξ_{T}. Then, the main loop starts. First, for $i \in \{1, \dots, |S|\}$, let

$$\beta^{(i)}(\sigma_{\mathsf{O}}) := \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z) = \sigma_{\mathsf{O}}}} \hat{u}_{\mathsf{T}}(z) \xi^{(i)}_{\mathsf{T}}[\sigma_{\mathsf{T}1}(z), \sigma_{\mathsf{T}2}(z)] \quad \forall\, \sigma_{\mathsf{O}} \in \Sigma_{\mathsf{O}}.$$

Then we solve the LP of Proposition 1 where Constraint ④ has been substituted with $\boldsymbol{\xi}_{\mathsf{T}} \in \operatorname{co} S$:

$$(*):\begin{cases} \displaystyle\mathop{\arg\max}_{\lambda^{(1)},\dots,\lambda^{(|S|)}} \quad v_\varnothing, \quad \text{subject to:} \\ \textcircled{1}\ v_I - \displaystyle\sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I') = \sigma_{\mathsf{O}}}} v_{I'} - \sum_{i=1}^{|S|} \beta^{(i)}(\sigma_{\mathsf{O}})\, \lambda^{(i)} \le 0 \\ \qquad\qquad\qquad\qquad\qquad \forall\, \sigma_{\mathsf{O}} \in \Sigma_{\mathsf{O}} \setminus \{\varnothing\} \\ \textcircled{2}\ v_\varnothing - \displaystyle\sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I') = \varnothing}} v_{I'} - \sum_{i=1}^{|S|} \beta^{(i)}(\varnothing)\, \lambda^{(i)} \le 0 \\ \textcircled{3}\ \sum_{i=1}^{|S|} \lambda^{(i)} = 1 \\ \textcircled{4}\ \lambda^{(i)} \ge 0 \quad \forall\, i \in \{1, \dots, |S|\} \\ \textcircled{5}\ v_\varnothing \text{ free}, v_I \text{ free} \quad \forall\, I \in \mathcal{I}_{\mathsf{O}}. \end{cases}$$

This is called the *master LP*.[2]

Given the solution to the master LP, a *pricing problem* is created. The goal of the pricing problem is to generate a new element $\boldsymbol{\xi}^{|S|+1}_{\mathsf{T}}$ to be added to S so as to increase the team utility in the next iteration, that is, the next solve of the master LP that then has an additional variable. This main loop of solving the larger and larger master LP keeps repeating until termination (discussed later).

The Pricing Problem

The pricing problem consist of finding a correlation plan $\hat{\boldsymbol{\xi}}_{\mathsf{T}} \in \Xi_{\mathsf{T}}$ which, if included in the convex combination computed by $(*)$, would lead to the maximum gradient of the objective (that is, the maximum *reduced cost*). By exploiting the theory of linear programming duality, such a correlation plan can be computed starting from the solution of the dual of $(*)$. In particular, let $\boldsymbol{\gamma}$ be the $|\Sigma_{\mathsf{O}}|$-dimensional vector of dual variables corresponding to Constraints ① and ② of $(*)$, and $\gamma' \in \mathbb{R}$ be the dual variable corresponding to Constraint ③. Then, the reduced cost of any candidate $\hat{\boldsymbol{\xi}}_{\mathsf{T}}$ is

$$c(\hat{\boldsymbol{\xi}}_{\mathsf{T}}) := -\gamma' + \sum_{z \in Z} \hat{u}_{\mathsf{T}}(z) \hat{\xi}_{\mathsf{T}}[\sigma_{\mathsf{T}1}(z), \sigma_{\mathsf{T}2}(z)] \gamma[\sigma_{\mathsf{O}}(z)].$$

Now comes our crucial observation. Since $c(\hat{\boldsymbol{\xi}}_{\mathsf{T}})$ is a linear function, and since from Proposition 3 we know that $\Xi_{\mathsf{T}} = \operatorname{co}\Xi^*_{\mathsf{T}1}$, by convexity

$$\max_{\hat{\boldsymbol{\xi}}_{\mathsf{T}} \in \Xi_{\mathsf{T}}} c(\hat{\boldsymbol{\xi}}_{\mathsf{T}}) = \max_{\hat{\boldsymbol{\xi}}_{\mathsf{T}} \in \Xi^*_{\mathsf{T}1}} c(\hat{\boldsymbol{\xi}}_{\mathsf{T}}).$$

We want to solve the LP on the left hand side, but—as discussed in Section 5—the constraints defining Ξ_{T} are not

[‡] 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 Ξ_{T} that can be represented with small values of n and leads to higher utilities for the team.

[2] In $(*)$ the convex combination is among *given* correlation plans, while in the MIP of Section 6, the elements to combine are themselves variables.

Game		Opponent player O = 1				Opponent player O = 2				Opponent player O = 3			
		$n=1$	$n=2$	$n=3$	$n=\infty$	$n=1$	$n=2$	$n=3$	$n=\infty$	$n=1$	$n=2$	$n=3$	$n=\infty$
Kuhn poker	**[A]**	0	★	★	0	0	★	★	0	0	★	★	0
	[B]	0.0208	0.0379	★	0.0379	0.0018	0.0246	0.0265	0.0265	−0.0417	★	★	−0.0417
	[C]	0.0470	0.0655	0.0663	0.0664	0.0128	0.0367	0.0376	0.0380	−0.0227	−0.0153	−0.0141	−0.0140
Goofspiel	**[D]**	0.2389	0.2524	★	0.2524	0.2389	0.2524	★	0.2524	0.2389	0.2524	★	0.2524
	[E]	0.2389	0.2534	★	0.2534	0.2389	0.2534	★	0.2534	0.2389	0.2534	★	0.2534
Liar's dice	**[F]**	0	★	★	0	0.2099	0.2554	0.2562	0.2562	0.2716	0.2840	★	0.2840
	[G]	0.0625	★	★	0.0625	0.2500	0.2656	0.2656	—	0.2656	—	—	—
Leduc poker	**[H]**	0.1453	0.2246	0.2466	0.2765	0.2107	0.2863	0.3143	0.3450	0.1840	0.2448	0.2815	0.2926
	[I]	—	—	—	0.1422	—	—	—	0.1420	—	—	—	0.0850
	[J]	0.2449	0.7037	0.7975	0.8359	0.2101	0.9222	0.9695	0.9709	0.2449	0.7037	0.7975	0.8359

Table 2: Expected utility of the team for varying support sizes (n). All values for $n \in \{1, 2, 3\}$ were computed using the MIP of Section 6, while the values corresponding to $n = \infty$ were computed using our column generation approach (Section 7). '★': A provably optimal utility has already been obtained with a lower value of the support size n. '—': We were unable to compute the exact value, because the corresponding algorithm hit the time limit.

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.

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 $u_{\mathsf{T}1} + u_{\mathsf{T}2}$. Finally, for each pair $(\pi_{\mathsf{T}1}, \pi_{\mathsf{T}2}) \in \Pi_{\mathsf{T}1} \times \Pi_{\mathsf{T}2}$ of normal-form plans generated by that no-regret algorithm, we compute and add to S the correlation plan corresponding to the distribution μ that assigns probability 1 to $(\pi_{\mathsf{T}1}, \pi_{\mathsf{T}2})$ 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 $\arg\max_{\hat{\boldsymbol{\xi}}_\mathsf{T} \in \mathcal{V}_\mathsf{T}} c(\hat{\boldsymbol{\xi}}_\mathsf{T})$. 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.

8 Experimental Evaluation

We computationally evaluate the algorithms proposed in Section 6 and Section 7. 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 2.40GHz CPU and 32GB of RAM. We used Gurobi 9.0.3 to solve LPs and MIPs.

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 (O = 1), second (O = 2), and third player (O = 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

Game	Ours		Fictitious Team Play (FTP)			CG-18	Pricers		Team utility after seeding			TMECor
	Seeded	Not seed.	$\epsilon = 50\%$	$\epsilon = 10\%$	$\epsilon = 1\%$		Relax.	MIP	$m = 1$	1000	10 000	value
[A]	1ms	1ms	2s†	10s†	1m 08s†	175ms	1	0	−0.500	0	0	0
[B]	1ms	34ms	3m 52s	37m 51s	> 6h	26.81s	2	0	−0.365	−0.021	−0.020	−0.042
[C]	17.20s	18.61s	4h 42m	> 6h	> 6h	> 6h	2	25	−0.155	−0.020	−0.020	−0.014
[D]	267ms	682ms	50s	9m 21s	> 6h	3m 09s	14	0	−0.436	0.252	0.252	0.252
[E]	1.34s	1.77s	4m 51s	2h 02m	> 6h	29m 38s	48	0	−0.830	0.248	0.250	0.253
[F]	1m 41s	11m 22s	> 6h	> 6h	> 6h	> 6h	20	7	−0.481	0.252	0.252	0.284
[G]	> 6h	> 6h	> 6h	> 6h	> 6h	> 6h	—	—	−0.688	0.277	oom	—
[H]	5m 20s	5m 53s	> 6h	> 6h	> 6h	> 6h	23	204	−2.354	0.087	0.125	0.293
[I]	1h 30m	1h 44m	> 6h	> 6h	> 6h	> 6h	5	638	−1.827	0.013	0.036	0.085
[J]	11m 08s	14m 49s	> 6h	> 6h	> 6h	> 6h	1232	48	−3.333	0.646	0.668	0.836

(a) — Comparison of run times **(b)** **(c)**

Table 3: **(a)** Runtime comparison between our algorithm, FTP, and CG-18. The seeded version of our algorithm runs $m = 1000$ iterations of CFR+ (Section 7), while the non seeded version runs $m = 1$. '†': 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.

carefully selected semi-randomized coordination plans are enough to reach an optimal solution.

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 and Gatti (2018) (henceforth CG-18), and the fictitious-team-play algorithm by Farina et al. (2018) (denoted FTP). Like our algorithm, CG-18 uses column generation approach which lets ○ play sequence-form strategies, while the team's strategy is directly represented as a distribution over joint normal-form plans. On the other hand, 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 an approximate 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 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). We set a time limit of 6 hours and a cap of at most four threads for each algorithm. Table 3 shows the results with the opponent playing as the third player. By 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. There are settings, such as Liar's dice instance [F], where we our algorithm needs just a few seconds to compute an optimal TMECor, while previous algorithms exceed 6 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 just cannot be handled with current TMECor technology.

Using the linear relaxation of the pricing problem ("implementation details" in Section 7) 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 ameliorates the performance of the algorithm. In the Liar's dice instance [F], it reduced run time by roughly a factor of ten. The value of the initial master solution (that is, before the first pricing) increases significantly with the number of iterations of the no-regret algorithm used for seeding.

9 Conclusions

We studied the problem of 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 scenario in which team members are not allowed to communicate during play but can coordinate before the game. First, we provided modeling results by drawing a connection to previous results on extensive-form correlation. Then, we developed an algorithm that computes an optimal joint distribution by only 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 suffices to reach optimal utility for the team. Inspired by this observation and leveraging 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 order of magnitudes faster than the state of the art and also solves many games that were previously intractable.

Broader Impact

Enabling the computation of strong, game-theoretic strategies for imperfect-information adversarial team games has complex effects. Such technology could be used by a team of malicious players to exploit an interaction or a specific opponent. On the other hand, the technology could also be used defensively, to play in such a way as to minimize the value that can be extracted from the agent herself. Whether the technology has a positive or negative societal impact (or none) varies depending on the nature of the imperfect-information interaction and the way the technology is implemented. We believe that publishing the algorithm increases its dissemination, thereby helping even the playing field between educated expert players and ones who might be less privileged and could thus benefit more from algorithmic strategy support.

References

Basilico, N.; Celli, A.; Nittis, G. D.; and Gatti, N. 2017. Team-maxmin equilibrium: efficiency bounds and algorithms. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 356–362.

Bowling, M.; Burch, N.; Johanson, M.; and Tammelin, O. 2015. Heads-up Limit Hold'em Poker is Solved. *Science* 347(6218).

Brown, G. W. 1951. Iterative Solutions of Games by Fictitious Play. In Koopmans, T. C., ed., *Activity Analysis of Production and Allocation*, 374–376. John Wiley & Sons.

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

Celli, A.; and Gatti, N. 2018. Computational Results for Extensive-Form Adversarial Team Games. In *AAAI Conference on Artificial Intelligence (AAAI)*, 965–972.

Farina, G.; Celli, A.; Gatti, N.; and Sandholm, T. 2018. Ex ante coordination and collusion in zero-sum multi-player extensive-form games. In *Advances in Neural Information Processing Systems*, 9638–9648.

Farina, G.; and Sandholm, T. 2020. Polynomial-Time Computation of Optimal Correlated Equilibria in Two-Player Extensive-Form Games with Public Chance Moves and Beyond. In *ArXiv preprint*.

Ford, L. R.; and Fulkerson, D. R. 1958. A Suggested Computation for Maximal Multi-Commodity Network Flows. *Management Science* 5(1): 97–101. doi:10.1287/mnsc.5.1.97.

Koller, D.; Megiddo, N.; and von Stengel, B. 1996. Efficient Computation of Equilibria for Extensive Two-Person Games. *Games and Economic Behavior* 14(2).

Kuhn, H. W. 1950. 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*, 97–103. Princeton, New Jersey: Princeton University Press.

Kuhn, H. W. 1953. 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*, 193–216. Princeton, NJ: Princeton University Press.

Lanctot, M.; Waugh, K.; Zinkevich, M.; and Bowling, M. 2009. Monte Carlo Sampling for Regret Minimization in Extensive Games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*.

Lisỳ, V.; Lanctot, M.; and Bowling, M. 2015. 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*, 27–36.

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

Romanovskii, I. 1962. Reduction of a Game with Complete Memory to a Matrix Game. *Soviet Mathematics* 3.

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

Southey, F.; Bowling, M.; Larson, B.; Piccione, C.; Burch, N.; Billings, D.; and Rayner, C. 2005. Bayes' Bluff: Opponent Modelling in Poker. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*.

Tammelin, O.; Burch, N.; Johanson, M.; and Bowling, M. 2015. Solving Heads-up Limit Texas Hold'em. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*.

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.

von Stengel, B.; and Koller, D. 1997. Team-maxmin equilibria. *Games and Economic Behavior* 21(1-2): 309–321.

Zhang, Y.; and An, B. 2020a. Computing Team-Maxmin Equilibria in Zero-Sum Multiplayer Extensive-Form Games. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2318–2325.

Zhang, Y.; and An, B. 2020b. Converging to Team-Maxmin Equilibria in Zero-Sum Multiplayer Games. In *International Conference on Machine Learning (ICML)*.

A Theoretical Details

Representing Distributions of Play via Extensive-Form Correlation Plans

As mentioned in the body, every distribution over *randomized* stratregy 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 $\boldsymbol{y}_{\text{T1}} \in \mathcal{Y}_{\text{T1}}$ and $\boldsymbol{y}_{\text{T2}} \in \mathcal{Y}_{\text{T2}}$ for the team members, let μ_{T1} and μ_{T2} be the distributions over normal-form plans $\Pi_{\text{T1}}, \Pi_{\text{T2}}$ equivalent to $\boldsymbol{y}_{\text{T1}}$ and $\boldsymbol{y}_{\text{T2}}$, respectively. Then, the distribution over reandomized strategy profiles that assignes probability 1 to $(\boldsymbol{y}_{\text{T1}}, \boldsymbol{y}_{\text{T2}})$ is equivalent to the product distribution of μ_{T1} and μ_{T2}, that is, the distirbution over $\Pi_{\text{T1}} \times \Pi_{\text{T2}}$ that picks a generic profile $(\pi_{\text{T1}}, \pi_{\text{T2}})$ with probability $\pi_{\text{T1}}(\pi_{\text{T1}}) \times \pi_{\text{T2}}(\pi_{\text{T2}})$. The reverse is also true: a product distribution over $\Pi_{\text{T1}} \times \Pi_{\text{T2}}$ 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 $\boldsymbol{\xi}_{\text{T}} \in \mathcal{V}$ be a vector in the von Stengel-Forges polytope. We say that $\boldsymbol{\xi}_{\text{T}}$ is a* product *correlation plan if*

$$\xi_{\text{T}}[\sigma_{\text{T1}}, \sigma_{\text{T2}}] = \xi_{\text{T}}[\sigma_{\text{T1}}, \varnothing] \cdot \xi_{\text{T}}[\varnothing, \sigma_{\text{T2}}]$$

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

Lemma 1. *A product correlation plan is always an element of Ξ_{T}.*

Proof. Let $\boldsymbol{\xi}_{\text{T}}$ be a product correlation plan. Since by definition, $\boldsymbol{\xi}_{\text{T}} \in \mathcal{V}$, the vectors $\boldsymbol{y}_{\text{T1}}, \boldsymbol{y}_{\text{T2}}$ indexed over Σ_{T1} and Σ_{T2}, repsectively, and defined as

$$y[\sigma_{\text{T1}}] = \xi_{\text{T}}[\sigma_{\text{T1}}, \varnothing], y[\sigma_{\text{T2}}] = \xi_{\text{T}}[\varnothing, \sigma_{\text{T2}}]$$

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

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

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

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

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

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

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

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

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

This concludes the proof. □

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

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

Let $\{(\boldsymbol{y}_{\text{T1}}, \boldsymbol{y}_{\text{T2}})\} \subseteq \mathcal{Y}_{\text{T1}} \times \mathcal{Y}_{\text{T2}}$ be the (singleton) support of μ_{T}, and let $\mu_{\text{T1}}, \mu_{\text{T2}}$ be distributions over Π_{T1} and Π_{T2}, respectively, equivalent to $\boldsymbol{y}_{\text{T1}}$ and $\boldsymbol{y}_{\text{T2}}$. Then,

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

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

Since by assumption the two team members sample strategies independently, their equivalent distribution of play over determinitic strategies is the product distribution $\mu_{\text{T}} := \mu_{\text{T1}} \otimes \mu_{\text{T2}}$. Using (2), μ_{T} has a representation as extensive-form correlation plan given by

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

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

$$\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] = y_{\mathsf{T1}}[\sigma_{\mathsf{T1}}] \qquad \forall\, \sigma_{\mathsf{T1}} \in \Sigma_{\mathsf{T1}}.$$

Similarly,

$$\xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}] = y_{\mathsf{T2}}[\sigma_{\mathsf{T2}}] \qquad \forall\, \sigma_{\mathsf{T2}} \in \Sigma_{\mathsf{T2}}.$$

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

$$\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] = \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] \cdot \xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}]$$

for all $(\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}) \in \Sigma_{\mathsf{T1}} \times \Sigma_{\mathsf{T2}}$. That, together with the inclusion $\Xi_{\mathsf{T}} \subseteq \mathcal{V}$, shows that $\boldsymbol{\xi}_{\mathsf{T}}$ is a product correlation plan. □

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 $\boldsymbol{\xi}_{\mathsf{T}} \in \Xi^*_{\mathsf{T1}} \cup \Xi^*_{\mathsf{T2}}$ be a semi-randomized plan. Then, $\boldsymbol{\xi}_{\mathsf{T}}$ is a product plan.*

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

Lemma 4 (Farina and Sandholm (2020, Lemma 6))**.** *Let $\boldsymbol{\xi}_{\mathsf{T}} \in \mathcal{V}$. For all $\sigma_{\mathsf{T1}} \in \Sigma_{\mathsf{T1}}$ such that $\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] = 0$, $\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] = 0$ for all $\sigma_{\mathsf{T2}} \in \Sigma_{\mathsf{T2}} : \sigma_{\mathsf{T1}} \bowtie \sigma_{\mathsf{T2}}$. Similarly, for all $\sigma_{\mathsf{T2}} \in \Sigma_{\mathsf{T2}}$ such that $\xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}] = 0$, $\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] = 0$ for all $\sigma_{\mathsf{T1}} \in \Sigma_{\mathsf{T1}} : \sigma_{\mathsf{T1}} \bowtie \sigma_{\mathsf{T2}}$.*

Proof of Lemma 3. We will only show the proof for the case $\boldsymbol{\xi}_{\mathsf{T}} \in \Xi^*_{\mathsf{T1}}$. The other case ($\boldsymbol{\xi}_{\mathsf{T}} \in \Xi^*_{\mathsf{T2}}$) is symmetric.

To show that

$$\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] = \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] \cdot \xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}]$$

for all $(\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}) \in \Sigma_{\mathsf{T1}} \bowtie \Sigma_{\mathsf{T2}}$, we perform induction on the depth of the sequence σ_{T2}. The depth $\mathrm{depth}(\sigma_{\mathsf{T2}})$ of a generic sequence $\sigma_{\mathsf{T2}} = (J, b) \in \Sigma_{\mathsf{T2}}$ 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 σ_{T2} has depth 0, that is, $\sigma_{\mathsf{T2}} = \varnothing$. In that case, the theorem is clearly true, because $\xi_{\mathsf{T}}[\varnothing, \varnothing] = 1$ as part of the von Stengel-Forges constraints (Definition 1).

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

There are only two possible cases:

- Case 1: $\xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}] = 0$. From Lemma 4, $\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] = 0$ and the statement holds.
- Case 2: $\xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}] = 1$. From the von Stengel-Forges constraints, $\xi_{\mathsf{T}}[\varnothing, \sigma(J)] = \sum_{b' \in A_J} \xi_{\mathsf{T}}[\varnothing, (J, b')] = 1 + \sum_{b' \in A_J, b' \neq b} \xi_{\mathsf{T}}[\varnothing, (J, b')] \ge 1$. Hence, because all entries of $\xi_{\mathsf{T}}[\varnothing, \sigma_2]$ are in $\{0, 1\}$ by definition of Ξ^*_{T1}, it must be $\xi_{\mathsf{T}}[\varnothing, \sigma(J)] = 1$ and $\xi_{\mathsf{T}}[\varnothing, (J, b')] = 0$ for all $b' \in A_J, b' \neq b$.

  Using the inductive hypothesis, we have that

$$\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma(J)] = \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] \cdot \xi_{\mathsf{T}}[\varnothing, \sigma(J)] = \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] \tag{8}$$

  for all $\sigma_{\mathsf{T1}} \in \Sigma_{\mathsf{T1}}, \sigma_{\mathsf{T1}} \bowtie \sigma(J)$. On the other hand, since $\xi_{\mathsf{T}}[\varnothing, (J, b')] = 0$ for all $b' \in A_J, b' \neq b$, from Lemma 4 we have that

$$\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, (J, b')] = 0 \quad \forall \sigma_{\mathsf{T1}} \bowtie J, b' \neq b. \tag{9}$$

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

$$\begin{aligned} \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \varnothing] \cdot \xi_{\mathsf{T}}[\varnothing, \sigma_{\mathsf{T2}}] &= \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma(J)] \\ &= \sum_{b' \in A_J} \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, (J, b')] \\ &= \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, (J, b)] = \xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}, \sigma_{\mathsf{T2}}] \end{aligned}$$

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

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

Convex combinations of product plans Both of the algorithms we presented in the paper ultimately produce an extensive-form correlation plan $\boldsymbol{\xi}_{\mathsf{T}}$ that is a convex combination of semi-randomized plans $\boldsymbol{\xi}^{(1)}_{\mathsf{T}}, \dots, \boldsymbol{\xi}^{(n)}_{\mathsf{T}}$, that is, of the form

$$\boldsymbol{\xi}_{\mathsf{T}} = \lambda^{(1)} \boldsymbol{\xi}^{(1)}_{\mathsf{T}} + \cdot + \lambda^{(n)} \boldsymbol{\xi}^{(n)}_{\mathsf{T}}$$

for $\lambda^{(i)} \ge 0$ such that $\lambda^{(1)} + \cdots + \lambda^{(n)} = 1$. Since semi-randomized correlation plans are product correlation plans (Lemma 3), from Lemma 2 each $\boldsymbol{\xi}^{(i)}_{\mathsf{T}}$ is equivalent to the team playing a single profile of randomized strategies $(\boldsymbol{y}^{(i)}_{\mathsf{T1}}, \boldsymbol{y}^{(i)}_{\mathsf{T2}}) \in \mathcal{Y}_{\mathsf{T1}} \times \mathcal{Y}_{\mathsf{T2}}$ with probability 1. By linearity, it is immediate to show that $\boldsymbol{\xi}_{\mathsf{T}}$ is equivalent to playing according to the distribution over randomized strategies for the team that picks $(\boldsymbol{y}^{(i)}_{\mathsf{T1}}, \boldsymbol{y}^{(i)}_{\mathsf{T2}})$ with probability $\lambda^{(i)}$.

TMECor Formulation Based on Extensive-Form Correlation Plans

Proposition 1. *An extensive-form correlation plan* $\boldsymbol{\xi}_{\mathsf{T}}$ *is a TMECor if and only if it is a solution to the LP*

$$\begin{cases} \arg\max_{\boldsymbol{\xi}_{\mathsf{T}}} \quad v_{\varnothing}, \quad \text{subject to:} \\ ① \; v_I - \sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I')=(I,a)}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z)=(I,a)}} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \qquad \forall\, (I,a) \in \Sigma_{\mathsf{O}} \setminus \{\varnothing\} \\ ② \; v_{\varnothing} - \sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I')=\varnothing}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z)=\varnothing}} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \\ ③ \; v_{\varnothing} \text{ free}, v_I \text{ free} \quad \forall\, I \in \mathcal{I}_{\mathsf{O}} \\ ④ \; \boldsymbol{\xi}_{\mathsf{T}} \in \Xi_{\mathsf{T}}. \end{cases}$$

Proof. We follow the steps mentioned in the body, starting from the bilinear saddle point problem formulation of the problem of computing a TMECor strategy for the team:

$$\arg\max_{\boldsymbol{\xi}_{\mathsf{T}} \in \Xi_{\mathsf{T}}} \min_{\boldsymbol{y}_{\mathsf{O}} \in \mathcal{Y}_{\mathsf{O}}} \sum_{z \in Z} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)]y[\sigma_{\mathsf{O}}(z)].$$

Expanding the constraint $\boldsymbol{y}_{\mathsf{O}} \in \mathcal{Y}_{\mathsf{O}}$ using the *sequence-form constraints* (Koller, Megiddo, and von Stengel 1996; von Stengel 1996), the inner minimization problem is

$$(P): \begin{cases} \min_{\boldsymbol{y}_{\mathsf{O}}} \quad \sum_{z \in Z} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)]y[\sigma_{\mathsf{O}}(z)] \\ ① \; -y[\sigma(I)] + \sum_{a \in A_I} y_{\mathsf{O}}[(I,a)] = 0 \qquad \forall I \in \mathcal{I}_{\mathsf{O}} \\ ② \; y_{\mathsf{O}}[\varnothing] = 1 \\ ③ \; y_{\mathsf{O}}[\sigma_{\mathsf{O}}] \ge 0 \qquad \forall\, \sigma_{\mathsf{O}} \in \Sigma_{\mathsf{O}}. \end{cases}$$

Introducing the free dual variables $\{v_I : I \in \mathcal{I}_{\mathsf{O}}\}$ for Constraint ①, and the free dual variable $v_{\varnothing}$ for Constraint ②, we obtain the dual linear program

$$(D): \begin{cases} \max_{v_I, v_{\varnothing}} \quad v_{\varnothing}, \quad \text{subject to:} \\ ① \; v_I - \sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I')=(I,a)}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z)=(I,a)}} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \qquad \forall\, (I,a) \in \Sigma_{\mathsf{O}} \setminus \{\varnothing\} \\ ② \; v_{\varnothing} - \sum_{\substack{I' \in \mathcal{I}_{\mathsf{O}} \\ \sigma_{\mathsf{O}}(I')=\varnothing}} v_{I'} \le \sum_{\substack{z \in Z \\ \sigma_{\mathsf{O}}(z)=\varnothing}} \hat{u}_{\mathsf{T}}(z)\xi_{\mathsf{T}}[\sigma_{\mathsf{T1}}(z), \sigma_{\mathsf{T2}}(z)] \\ ③ \; v_{\varnothing} \text{ free}, v_I \text{ free} \quad \forall\, I \in \mathcal{I}_{\mathsf{O}}. \end{cases}$$

So, $\boldsymbol{\xi}_{\mathsf{T}}$ is a TMECor if and only if it is a solution of $\arg\max_{\boldsymbol{\xi}_{\mathsf{T}} \in \Xi_{\mathsf{T}}} (D)$, which is exactly the statement. □

Semi-Randomized Correlation Plans

Proposition 3. *In every game,* Ξ_{T} *is the convex hull of the set* Ξ^*_{T1}*, or equivalently of the set* Ξ^*_{T2}*. Formally,* $\Xi_{\mathsf{T}} = \operatorname{co} \Xi^*_{\mathsf{T1}} = \operatorname{co} \Xi^*_{\mathsf{T2}} = \operatorname{co}(\Xi^*_{\mathsf{T1}} \cup \Xi^*_{\mathsf{T2}})$.

Proof. We will show that $\Xi_{\mathsf{T}} = \operatorname{co} \Xi^*_{\mathsf{T1}}$. The proof that $\Xi_{\mathsf{T}} = \operatorname{co} \Xi^*_{\mathsf{T2}}$ is symmetric.

We will break the proof of $\Xi_{\mathsf{T}} = \operatorname{co} \Xi^*_{\mathsf{T1}}$ into two parts:

($\subseteq$) In the first part of the proof, we argue that $\Xi^*_{\mathsf{T1}} \subseteq \Xi_{\mathsf{T}}$. This is straightforward: from Lemma 3 we know that all elements of Ξ^*_{T1} are product correlation plans (Definition 3), which implies that $\Xi^*_{\mathsf{T1}} \subseteq \Xi_{\mathsf{T}}$ by Lemma 1. Since convex hulls preserve inclusions, we have

$$\operatorname{co} \Xi^*_{\mathsf{T1}} \subseteq \operatorname{co} \Xi_{\mathsf{T}},$$

which is exactly the statement $\Xi^*_{\mathsf{T1}} \subseteq \Xi_{\mathsf{T}}$ upon using the known fact that Ξ_{T} is a convex polytope and therefore $\operatorname{co} \Xi_{\mathsf{T}} = \Xi_{\mathsf{T}}$.

($\supseteq$) To complete the proof, we now argue that the reverse inclusion, namely $\Xi_{\mathsf{T}} \subseteq \operatorname{co} \Xi^*_{\mathsf{T1}}$, also holds. Let $f : \mu_{\mathsf{T}} \mapsto \xi_{\mathsf{T}}$ be the mapping from the distribution of play $\mu_{\mathsf{T}} \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})$ to its corresponding extensive-form correlation plan defined in Eq. (2). By definition, $\Xi_{\mathsf{T}} = f(\Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}))$. Let $\mathbb{1}_{(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})}$ denote the distribution of play with singleton support $(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})$, that is, the distribution of play that assigns the deterministic strategy profile $(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})$ for the team with probability 1. Since f is linear, and since

$$\Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}) = \operatorname{co}\{\mathbb{1}_{(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})} : \pi_{\mathsf{T1}} \in \Pi_{\mathsf{T1}}, \pi_{\mathsf{T2}} \in \Pi_{\mathsf{T2}}\},$$

we have

$$\Xi_{\mathsf{T}} = \operatorname{co}\{f(\mathbb{1}_{(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}})}) : \pi_{\mathsf{T1}} \in \Pi_{\mathsf{T1}}, \pi_{\mathsf{T2}} \in \Pi_{\mathsf{T2}}\}.$$

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

Together, the two statements that we just prove show that $\Xi_{\mathsf{T}} = \operatorname{co} \Xi^*_{\mathsf{T1}}$.

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

$$\operatorname{co}(\Xi^*_{\mathsf{T1}} \cup \Xi^*_{\mathsf{T2}}) = (\operatorname{co} \Xi^*_{\mathsf{T1}}) \cup (\operatorname{co} \Xi^*_{\mathsf{T2}}) = \Xi_{\mathsf{T}} \cup \Xi_{\mathsf{T}} = \Xi_{\mathsf{T}},$$

thereby concluding the proof. □

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 necessarely 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, Lanctot, and Bowling (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 6 and the Prior State of the Art

Depending on the cap n on the number or semi-randomized correlation plans, the algorithm we describe in Section 6 might not reach the optimal TMECor value for the team (although, as we argue in Section 8, 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 is has a good trade-off between dimensions and manageability. When $n = 1$ the algorithm reaches an optimal solution in 9.74s. The optimal solution with $n = 1$ achieves 63% of the optimal utility with no restrictions on the number of plans. With $n = 2$ the run time is 5m38s and the solution reaches 84% 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 6 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 84% of the optimal value within roughly 5 minutes. On the other hand, CG-18 cannot complete even a single iteration within the time limit. This confirms the 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	$m = 1$	1000	10 000	
[A]	2ms	1ms	0ms†	15.00s†	2m 35s†	66ms	5	0	–0.567	–0.133	–0.133	0
[B]	21ms	3ms	0ms	16m 39s	$>$ 6h	1.01s	0	3	–0.375	0.037	0.038	0.038
[C]	5.69s	5.79s	7m 36s	$>$ 6h	$>$ 6h	$>$ 6h	8	41	–0.166	0.058	0.058	0.066
[D]	186ms	304ms	2ms	$>$ 6h	$>$ 6h	1m 56s	19	0	–0.492	0.251	0.252	0.252
[E]	464ms	860ms	6ms	$>$ 6h	$>$ 6h	23m 17s	33	0	–1.000	0.249	0.253	0.253
[F]	2.14s	4.21s	19m 25s	$>$ 6h	$>$ 6h	$>$ 6h	1	0	–0.748	0	0	0
[G]	1m 11s	41m 23s	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	0	2	–0.721	0.063	oom	0.063
[H]	43.72s	1m 24s	2h 49m	$>$ 6h	$>$ 6h	$>$ 6h	9	79	–3.142	0.210	0.228	0.277
[I]	43m 58s	46m 08s	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	0	614	–3.091	0.111	0.122	0.142
[J]	3m 48s	11m 36s	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	1612	37	–4.000	0.627	0.651	0.836

(a) — Comparison of run times **(b)** **(c)**

Table 4: **Results for O** $= 1$**.** **(a)** Runtime comparison between our algorithm, FTP, and CG-18. The seeded version of our algorithm runs $m = 1000$ iterations of CFR+ (Section 7), while the non seeded version runs $m = 1$. '$\dagger$': 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.

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	$m = 1$	1000	10 000	
[A]	0ms	1ms	0ms†	19s†	3m 09s†	147ms	1	0	–0.633	0.000	0.000	0
[B]	0ms	11ms	1m 39s	$>$ 6h	$>$ 6h	7.53s	1	0	–0.250	0.027	0.027	0.027
[C]	6.47s	5.64s	48m 08s	$>$ 6h	$>$ 6h	$>$ 6h	6	33	–0.126	0.027	0.033	0.038
[D]	144ms	368ms	1ms	$>$ 6h	$>$ 6h	1m 46s	14	0	–0.384	0.252	0.252	0.252
[E]	641ms	904ms	1.39s	$>$ 6h	$>$ 6h	12m 30s	40	0	–3.000	0.252	0.252	0.253
[F]	55.00s	8m 59s	1h 30m	$>$ 6h	$>$ 6h	$>$ 6h	21	0	–0.630	0.256	0.256	0.256
[G]	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	—	—	–0.766	0.264	oom	—
[H]	7m 30s	8m 05s	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	25	335	–2.002	0.177	0.201	0.345
[I]	57m 32s	1h 09m	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	1	492	–2.505	0.096	0.110	0.142
[J]	7m 11s	5m 16s	$>$ 6h	$>$ 6h	$>$ 6h	$>$ 6h	2508	37	–7.500	0.630	0.819	0.971

(a) — Comparison of run times **(b)** **(c)**

Table 5: **Results for O** $= 2$**.** **(a)** Runtime comparison between our algorithm, FTP, and CG-18. The seeded version of our algorithm runs $m = 1000$ iterations of CFR+ (Section 7), while the non seeded version runs $m = 1$. '$\dagger$': 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.