Power of Correlation in Extensive-Form Games

Gabriele Farina, Chun Kai Ling, Fei Fang, Tuomas Sandholm
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
{gfarina,chunkail,feif,sandholm}@cs.cmu.edu

In this paper, we study—for the first time in non-toy examples—the benefits of extensive-form correlation in two-player sequential games with no chance moves. In this model, the strategic interaction of the two agents is complemented by an external device that privately recommends moves to the agents as the game progresses. Each agent is free to defect from the recommendation at any time, at the cost of stopping receiving future recommendations.

We propose two benchmark games; each game is naturally parametric, so that these games can scale in size as desired. The first game is a general-sum variant of the classic war game *Battleship*, where two competing fleets take turns in trying to sink their enemies. In this game, as the players' penalty for losing ships increases, we show that social-welfare-maximizing extensive-form correlated equilibria can achieve dramatically less destructive outcomes than any social-welfare-maximizing Nash equilibrium, while remaining incentive-compatible for both players. The second game is a simplified version of the *Sheriff of Nottingham* board game. In this game, a sheriff must decide whether or not to inspect a cargo that potentially carries illegal good, while the smuggler is trying to negotiate a suitable bribe for the sheriff. In both games, we show that the social-welfare-maximizing equilibria in this game are far from trivial, and can induce surprisingly subtle sequential behavior that so far has not received attention in the literature.

1 Introduction

Nash equilibrium (NE), introduced by Nash [1950], is the most seminal concept in non-cooperative game theory, and has transformed our understanding of strategic interaction. Nash equilibrium captures a multi-agent setting where each agent is selfishly motivated in maximizing their own payoff. In other words, the assumption underpinning Nash equilibrium is that the interaction is completely *decentralized*: the behavior of each agent is not regulated by any external orchestrator. Contrasted with the opposite—often utopian—extreme of a fully managed interaction, where an external dictator controls the behavior of each agent so that the whole system moves to a desired state, the social welfare that can be achieved by Nash equilibrium is generally lower, often dramatically so. Yet, in many realistic interactions, some intermediate form of centralized control can be achieved. In particular, in his landmark paper, Aumann [1974] proposed the concept of *correlated equilibrium* (CE), where a mediator (the *correlation device*) can *recommend* behavior, but not *enforce it*. In a CE, the designer of the interaction is responsible to make sure that the agents—which are still modeled as fully rational and selfish just like in an NE—have incentives to follow the private recommendation that is issued to them by the mediator. The idea of a mediator that is able to recommend behavior but not enforce it make CE a good candidate in multi-agent settings such as traffic control, and load balancing [Ashlagi *et al.*, 2008]. Crucially, correlated equilibria can engineered so that the whole system moves toward higher welfare, and is known to lead to much more favorable outcomes than Nash equilibrium [Aumann, 1974].

In this paper, we focus on the closely related notion of extensive-form correlated equilibrium (EFCE), introduced by Von Stengel and Forges [2008] in the context of extensive-form (that is, sequential) games.

Submitted to:
7th International Workshop on Strategic Reasoning at IJCAI'19. **Track (C), challenging open problems.**

Like CE, EFCE assumes that the strategic interaction is complemented by an external mediator; however, in an EFCE the mediator only privately reveals recommendation for the next move to each acting player, instead of issuing recommendations for the whole multi-stage strategy for each player. Furthermore, while each agent is free to defect from the recommendation at any time, this comes at the cost of stopping receiving future recommendations. Despite recent years have seen tremendous interest and progress around the problem of computing Nash equilibria in sequential partially-observable interactions, with significant milestones in the game of Poker [Brown and Sandholm, 2017; Bowling *et al.*, 2015; Moravčík *et al.*, 2017] and other large, real-world domains, not much has been done to increase our understanding of (extensive-form) correlated equilibria in these settings. At least for what concerns Nash equilibrium, sequential interactions are known to pose different challenges than their one-shot counterpart, especially in settings in which the agents retain private information. Conceptually, this is due to the fact that the players can strategically adjust to dynamic observations about the environment and their opponents as the game progresses. Even extremely small extensive-form games (such as Kuhn poker [Kuhn, 1950]) can exhibit rich behavior such as mandating that players bluff or use otherwise deceptive behavior as part of their optimal strategy. We show that EFCE is no exception.

Our primary objective with this paper is to spark more interest in the community towards a deeper understanding of the merits of correlation in sequential strategic interactions. As we will show later in the paper, despite the fact that the mediator in an EFCE cannot *enforce* behavior, it can nonetheless lead the players to drastically better social welfare than Nash equilibrium. Empirically, this appears to be usually achieved via a combination of the following behaviors:

- The EFCE mediator makes sure that the the threat of stopping receiving recommendations upon defection is a strong enough deterrent for each player; note that this deterrent is unique to *sequential* interactions and does not apply to correlated equilibria in one-shot interactions.
- In the case when the deviation of an agent A from the recommendation of the mediator can be detected by another agent B, the mediator can induce B to play some punitive actions against A. Again, this feature is only unique to sequential games.

In order to facilitate the study of extensive-form correlation, we propose two benchmark games; each game is naturally parametric, so that these games can scale in size as desired. The first game is a general-sum variant of the classic war game *Battleship*, where two competing fleets take turns in trying to sink their enemies. In this game, as the players' penalty for losing ships increases, we show that social-welfare-maximizing extensive-form correlated equilibria can achieve dramatically less destructive outcomes than any social-welfare-maximizing Nash equilibrium, while remaining incentive-compatible for both players. The second game is a simplified version of the *Sheriff of Nottingham* board game. In this game, a sheriff must decide whether or not to inspect a cargo that potentially carries illegal good, while the smuggler is trying to negotiate a suitable bribe for the sheriff. These games were chosen to cover two application domains that are natural for EFCEs: respectively conflict resolution via a mediator for the Battleship game, and bargaining and negotiation for the Sheriff game.

We plan to release the source code for these game generators, so that the research community can benefit from our implementation work.

2 Preliminaries

In this section, we recall several fundamental properties of extensive-form games, the primary model of strategic interaction that we will use in the rest of the paper. We also recall the definition and some properties of two solution concepts for extensive-form games: Nash equilibrium and extensive-form correlated equilibrium.

2.1 Extensive-Form Games

Extensive-form games (EFGs) are sequential, finite games that are played over a rooted game tree. Each vertex in the tree belongs to a player and corresponds to a decision point for that player. Edges leaving from a node v correspond to actions that can be taken by the player to which v belongs. Each terminal node in the game tree, also known as a *leaf* of the game, is associated with the tuple of payoffs that each player receives if the game ends in that state. To capture imperfect information, the set of vertices of each player is partitioned into *information sets*. Two vertices belong to the same information set when the player acting at them cannot distinguish between them. For example, in a game of Poker, a player cannot distinguish between certain states that only differ in opponent's private hand. Because vertices in a same information set are indistinguishable by the acting player, the strategy of the player must be the same for all of them. In particular, the distribution over the next action action to be taken must be the same for all vertices in a same information sets. Imperfect-information extensive-form games are

For the purposes of this paper, we only consider *perfect-recall* EFGs. This property means that each player does not forget any of their previous action, nor any private or public observation that the player has made. The perfect-recall property can be formalized by requiring that for any two vertices in a same information set, the paths from those vertices to the root of the game tree induce the same sequence of actions for the player that owns the information set.

A pure normal-form plan for player i in an EFG defines a choice of action for *every* information set belonging to i. An extensive-form game can be viewed as a normal-form game where each player chooses a distribution over their normal-form plans and plays according to a normal-form plan sampled from such distribution. However, this representation is of exponential size. Furthermore, it contains redundancies: some information sets for i may no longer be reachable after player i makes certain decisions higher up in the tree. Omitting these redundancies leads to the similar notion of *reduced-normal-form* strategy, which is known to be strategically equivalent to the normal form. The reduced-normal-form representation of the game is also of exponential size in the game tree. Fortunately, a classical result shows that in games with perfect recall, a third representation of the game, called the *sequence-form* representation, allows to capture a player's strategy in the sequential interaction using space polynomial in the game tree size [Von Stengel and Forges, 2008]. In the sequence-form representation, strategies for player i are expressed as *realization plans* $x_i \in \mathbb{R}^{m_i}$, where m_i is the number of possible actions for the player i across the entire game. Realization plans represent probabilities of performing a sequence of actions, in isolation from chance and other player's moves. Mathematically, this is represented by the linear constraints $Fx = f$, where F is a matrix with entries in $\{-1,0,1\}$, while f are vectors containing $\{0,1\}$. These specify 'flow' constraints and implicitly encode parent-child relationships and information sets. The space of valid strategies is described by the *sequence form polytope*, and may be seen as the extensive-form generalization of the requirement that x_i lies in the probability simplex.

Many solution concepts may be carried over from normal to extensive-form games while enjoying computational advantages offered by the sequence form. For example, the problem of finding a Nash equilibrium may be formulated as a Linear Complementarity Problem over the sequence form polytopes.

2.2 Extensive-Form Correlated Equilibrium

Extensive-form correlated equilibrium (EFCE) is a solution concept for extensive-form games introduced by Von Stengel and Forges [2008]. Like in the traditional correlated equilibrium (CE), introduced by Aumann [1974], a *correlation device* selects private signals for the players before the game starts. These signals are sampled from a joint probability distribution, and represent moves that the players are recommended to play. However, while in a CE the recommended moves for the whole game tree are privately revealed to the players when the game starts, in an EFCE the recommendations are revealed incrementally as the players progress in the game tree. In particular, a recommended move is only revealed when the player reaches the decision point in the game for which the recommendation is relevant. Furthermore, if a player ever deviates from the recommended move, he or she stops receiving recommendations from the device.[1]

In an EFCE, the players know less about the set of recommendations that were sampled by the correlation device. The benefits are twofold. First, the players can be more easily induced to play strategies that hurt them (but benefit the overall social welfare), as long as "on average" the players are indifferent as to whether or not to follow the recommendations: the set of EFCEs is a superset of that of CEs. Second, since the players observe less, the set of probability distributions for the correlation device for which no player has an incentive to deviate can be described succinctly in certain classes of games:

Theorem 1 (Von Stengel and Forges [2008], Theorem 1.1)**.** *In two-player, perfect-recall extensive-form games with no chance moves, the set of EFCEs can be described by a system of linear equations and inequalities of polynomial size.*

As a direct consequence of Theorem 1, in two-player, perfect-recall extensive-form games with no chance moves an EFCE that maximizes social welfare can be found as the solution of a linear program of polynomial size in the input game. However, the same result cannot hold in more general settings:

Theorem 2 (Von Stengel and Forges [2008], Section 3.7)**.** *In perfect-recall extensive-form games with three or more players, as well as in perfect-recall extensive-form games with two players and chance moves, the problem of determining the existence of an EFCE with social welfare greater than a given value is NP-hard.*

It is important to note that Theorem 2 only implies that the characterization of the set of EFCEs cannot be of polynomial size in general (unless P = NP). However, the problem of finding *one* EFCE can be solved in polynomial time: Huang [2011] and Huang and von Stengel [2008] show how to adapt the *Ellipsoid Against Hope* algorithm [Papadimitriou and Roughgarden, 2008; Jiang and Leyton-Brown, 2015] to compute an EFCE in polynomial time in games with more than two players and/or with chance moves.

3 Battleship

In this section and in the next, we illustrate the nature of welfare-maximizing EFCEs by means of several moderately complex games. We will see that even in well-structured 'real' games, EFCEs may lead to surprising and counter-intuitive behavior. We hope that by describing some of this behavior, the reader will gain an appreciation for the incentive constraints required by EFCE and find a use for it in their research. In the experiments which follow, the commercial software GUROBI was used to compute the equilibria.

The first game we consider is a non-zero sum variant of the classic game *Battleship*, which we describe in the next subsection. Broadly speaking, our variant is identical to the original game except that (a) the game has finite time horizon H and that (b) the payoff of each player is proportional to the sum of the

[1] An EFCE without the condition that deviating players stop receiving recommendations is customarily called *agent-form correlated equilibrium* [Von Stengel and Forges, 2008].

values of the opponent's ships that were destroyed, plus a penalty term for each ship that the plays has lost.

3.1 Description of the Game

A game of Battleship is parameterized by a tuple $(H, W, \mathscr{S}, r, \gamma)$, where

- the integers $H, W \geq 1$ define the height and width of the playing field for each player;
- $\mathscr{S}$ is an ordered list containing ship descriptions s_i for each player. Each description is a pair $s_i = (\ell_i, v_i)$, where ℓ_i is the length of the i-th ship and v_i is its value;
- $r \geq 1$ is the number of rounds in the game;
- $\gamma \geq 1$ is a *loss multiplier* that controls the relative value of a losing versus destroying ships.

The game proceeds in 2 phases: *ship placement* and *shooting*. During the ship placement phase, the players (starting with Player 1) take turns placing their ships on their playing field. The players must place all their ships, in the same order in which they appear in $\mathscr{S}$, on the playing field. The ship placement phase ends when all ships have been placed. We remark that the players' playing fields are separate: in other words, there are two playing fields of dimensions $H \times W$, one per player. The ships may be placed either horizontally or vertically on each player's grid (playing field); all ships must lie entirely within the playing field and may not overlap with other ships the player has already placed. Finally, the locations of a player's ships is private information for each player.

In the shooting phase, players take turns firing at each other; Player 1 starts first. This is done by selecting a pair of integer coordinates (x, y) that identify a cell within the playing field. After taking a shot, the player is told if the shot was a *hit*, that is, the selected cell (x, y) is occupied by a ship of the opponent, or if it is a *miss*, that is, (x, y) does not contain an opponent's ship. If all cells covered by a ship have been shot at, the ship is destroyed and this fact is announced. Note that the identity of the ship which was hit or sunk is not revealed; players only know that *some* ships was hit or sunk. The game ends when r shots have been made by each player, or if one player has lost all their ships, whichever comes first. At the end of the game, each player's payoff is computed as follows: for each opponent's ship that the player has destroyed, the player receives a payoff equal to the value v of that ships; for each ship that the player has lost to the opponent, the player incurs a negative payoff equal to $\gamma \cdot v$, that is the value of the ship times the loss multiplier γ. Note that when $\gamma > 1$ the game is general sum.

Since $\gamma \geq 1$, this asymmetric model describes situations where players are encouraged to destroy other ships, but are ultimately more protective of their own assets. The loss multiplier γ governs this gap; a higher value of γ makes so that each player values their ships more than destroying others. Note that when $\gamma = 1$, we obtain a zero-sum version of battleships (with varying scores for each ship).

For the remainder of the discussion, we define the *social welfare* (SW) of any outcome to be the sum of payoffs of each player. We will demonstrate that with the aid of a mediator (the correlation device), the social welfare of the optimal correlated equilibria are dramatically higher than the social welfare of even the best Nash equilibrium. In other words, the mediator leads to significantly less destructive outcomes, and leads to more frequent ties where the players sometimes agree to deliberately miss their opponents, while still retaining incentive-compatibility and rationality in the standard game-theoretic sense.

3.2 Social-Welfare-Maximizing Nash Equilibrium vs EFCE

For simplicity, consider the instance of Battleship with parameters $H = 3, W = 1, \mathscr{S} = [(1,1)], r = 2, \gamma = 2$; that is, the board is of size 3×1, each player commands just 1 ship of size 1, there are 2 rounds of shooting per player, and the loss multiplier is set to $\gamma = 2$. In this game, it can be shown that the social-welfare-maximizing Nash equilibrium is to have each player place and fire uniformly at random. The probability that Player 1 and 2 will end the game by destroying the opponent's ship is $5/9$ and $1/3$ respectively.[2] The probability that both players will end the game with their ships unharmed is a meagre $1/9$. Correspondingly, the maximum social welfare reached by any Nash equilibrium of the game is $-8/9$.

Below we will show that under the EFCE model, it is possible to induce the players to end the game with no damage to either ship with probability $5/18$, and obtain a corresponding SW of $-13/18$. In a nutshell, the correlation plan is constructed in a way that players are recommended to deliberately miss, and deviations from this are 'punished' by the mediator, who reveals the deviating player's ship location in future recommendations. This threat keeps players in line and encourages peace.[3] Figure 1 shows, as a function of γ, the probabilities with which Player 1 and 2 terminate the game by sinking their opponent's ship, if they play according to a SW-maximizing EFCE. For all values of γ, the probability that the games ends in a terminal state in which no ship is sunk, is higher than even the best Nash equilibrium, in which the game ends in a non-violent outcome only if both players miss their opponent twice, that is with probability $1/9$. Furthermore, as γ grows, the probability that each player sink the opponent tends to $1/3$: despite Player 1's inherent advantage for acting first, as γ increases each Player ends up sinking the opponent with equal probability.

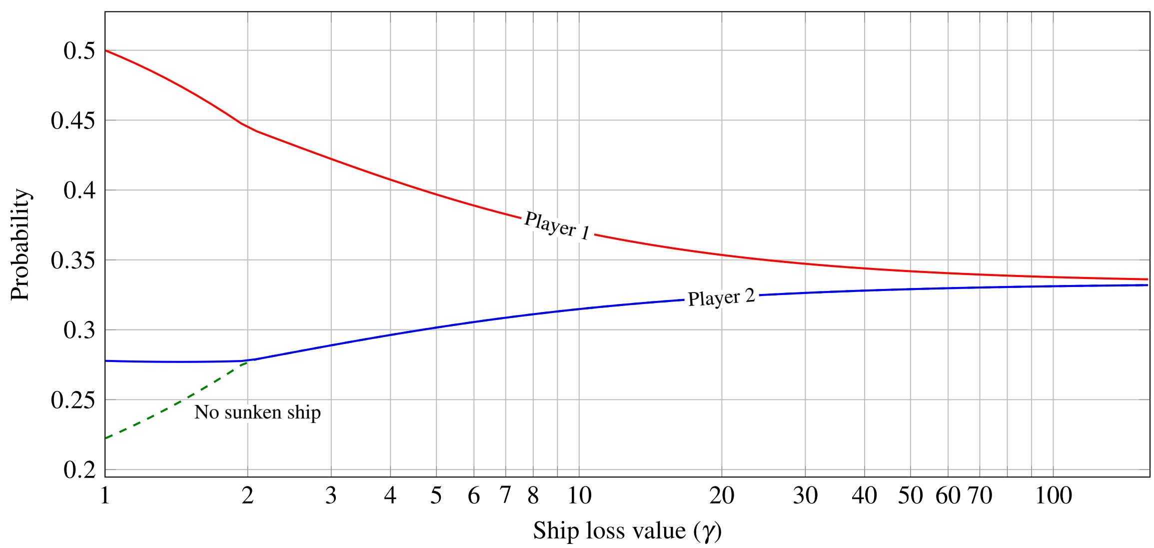

Figure 1: Probabilities of players sinking their opponent's ship when the players play according to a SW-maximizing EFCE in the Battleship game with parameters $H = 3, W = 1, \mathscr{S} = [(1,1)], r = 2, \gamma = 2$. For $\gamma \geq 2$, the probability of the game ending with no sunken ship and the probability of Player 2 sinking Player 1's ship coincide.

The next subsection analyzes how the SW-maximizing EFCE compels the players to play in a way that leads to a severely less destructive outcome than Nash equilibrium.

[2]Player 1 has an advantage since the game ends prematurely if he manages to hit Player 2's ship in the first turn.

[3]It is somewhat of a misnomer to describe the consequences of deviating to be a form of 'punishment' from the mediator. In an EFCE, the mediator does not play an 'active' role in the game—s/he does *not* observe player actions, nor does s/he adjust the recommended strategies based on to any deviations. The 'threats' and 'punishments' described so far are completely upfront, i.e., they are decided ex-ante. The behavior of the correlation is completely public to the players. In fact, the mediator simply samples (correlated) reduced-normal-form strategies; once this sampling is over, the mediator's role is effectively over (as long as we enforce that players only get recommendations at corresponding to their current information set). That said, it is often helpful to visualize the aforementioned situations are 'threats', and we continue to do so for the remainder of the paper.

3.3 Analysis of Social-Welfare-Maximizing EFCE

We analyze one social-welfare-maximizing EFCE in the same small instance of Battleship as the previous section. The mediator in this EFCE recommends the players a ship placement that is sampled uniformly at random and independently for each players. This results in 9 possible scenarios (one per possible ship placement) in the game, each occurring with probability $1/9$. Due to the symmetric nature of ship placements, only two scenarios are relevant: whether the two players are recommended to place their ship in the same spot, or in different spots. Figure 2 details the strategy of the the mediator in each of these two scenarios, assuming that the players do not deviate. Note that the game trees in Figure 2 are parametric on the recommended ship placements a and b; all 9 possible ship placements can be recovered from Figure 2 by setting a and b to appropriate values in $\{1,2,3\}$.

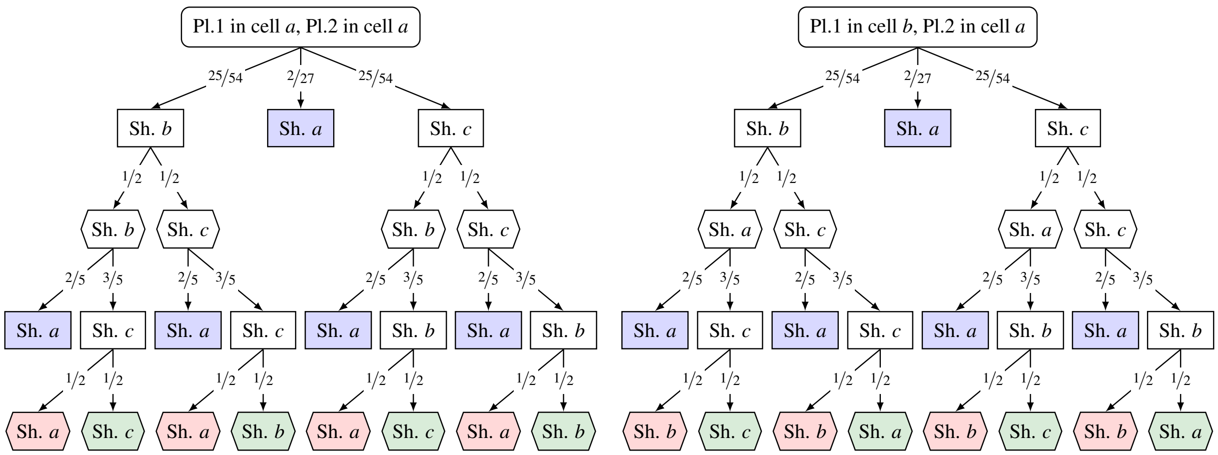

Figure 2: Example of a playthrough of Battleship assuming both players were recommended to place their ship in a (left), or that Player 1 and 2 were recommended to place their ships in a and b respectively (right). For both pictures, the numbers along each edge denote probabilities of each action being recommended; no edge is shown for actions recommended with zero probability. Squares and hexagons denote actions taken by Players 1 and 2 respectively. Similarly, blue and red nodes represent cases where Players 1 and 2 sink their opponent's ship, respectively. Green leaf nodes are where the game results in no ship loss. The *Shoot* action is abbreviated to 'Sh.'

For both game trees, note that the correlation device suggests that Player 1 shoot at the Player 2's ship with a low $2/27$ probability, and deliberately miss with high probability. As hinted in earlier sections, this type of recommendation is key to understanding why the EFCE succeeds in promoting less destructive outcomes. One may wonder why this behavior is incentive-compatible (that is, what are the incentives that compel Player 1 into not defecting), since the player may choose to randomly fire in any of the 2 locations that were *not* recommended, and get almost $1/2$ chance of winning the game immediately. The key is that if Player 1 does so and does not hit the opponent's ship, then the mediator can punish him by recommending that Player 2 shoot at the location of Player 1's ship. Since players value their ships more than destroying their opponents, the player is incentivized to avoid such a situation by accepting the recommendation to (most probably) miss.

A similar situation arises in the first move of Player 2. Here, Player 2 is recommended to *deliberately* miss, hitting each of the 2 empty spots with probability $1/2$. If he deviates and attempts to destroy Player 1's ship, then he risks the mediator revealing his location to his opponent if his shot misses; this risk is enough to keep Player 2 'in line'. The second move of Player 1 (third shot of the full game) bears a similar ideas. Here, Player 1 is recommended to hit Player 2's ship with probability $2/5$. Similar to his first shot, Player 1 may deviate and fire at the remaining location and enjoy 3/5 chance of winning the game out right. Yet, this behavior is discouraged, since in the 2/5 chance that he misses the shot (i.e., the recommendation was in fact, the correct location of Player 2's ship), then his location would be revealed by the mediator and he loses the next round. Again, this threat from the mediator encourages peaceful behavior, even though the recommendation to Player 1 reveals a more accurate 'posterior' of Player 2's ship location, as compared to the uniform distribution of 1/2. While making these recommendations, the mediator ensures that Player 2 has a uniform distribution of Player 1's ship location, meaning that even though Player 2 has the final move, he may not do better than guessing at uniform at this stage.

Remark. It is important to note that Figure 2 does not convey the full information of the correlated plans. Crucially, it does not show the consequences suffered if a player deviates from his recommended strategy—in this case, the deviating player stops receiving recommendations and risks having his ship's location revealed to the opponent. These 'counterfactual' scenarios may be counter-intuitive but are key to understanding how SW-maximizing EFCEs achieve their purpose.

4 Sheriff Game

We consider a simplified version of the "Sheriff of Nottingham" board game. The game comprises of two players, the *Smuggler* and the *Sheriff*. The game models the situation where a Smuggler is trying to smuggle illegal items in its cargo. The Sheriff can choose to inspect the cargo of the Smuggler. If the Sheriff chooses to inspect the cargo and finds illegal goods, the Smuggler must pay a fine to the Sheriff and have his goods confiscated. However, the Sheriff has to compensate the Smuggler if no illegal goods are found. Lastly, if the Sheriff decides not to inspect the cargo, the Smuggler's utility is proportional to the quantity of goods smuggled whereas the Sheriff's utility is zero. This simple game is made more interesting by including two additional features (also present in the board game): *Bribery* and *Bargaining*, which we describe in more detail later.

4.1 Formal Description

The Sheriff game is described by the the parameters $v, p, s \in \mathbb{R}^+, n_{\max}, b_{\max}, r \in \mathbb{N}$. The parameters $v, p, s \geq 0$ describe the value of *each* illegal item, the penalty that the Smuggler has to pay for *each* discovered illegal item, and the compensation that the Sheriff pays to the Smuggler in the case of a false alarm. At the beginning of the game, the Smuggler loads $n \in \{0, \ldots, n_{\max}\}$ items into his cargo. The amount of goods loaded is unknown to the Sheriff. The game then proceeds for $r \geq 1$ rounds of bargaining. Each round comprises two steps. First, the Smuggler offers a bribe $b_t \in \{0, \ldots, b_{\max}\}$ to the Sheriff, where $t \leq r$ is the round of bargaining. After that, the Sheriff responds with 'Yes' or 'No'.

All actions are public knowledge, except for the selection of cargo contents, which only the Smuggler knows. In the final step, we compute the payoffs to players. The outcome of the game is decided by the *last* step of bargaining. In particular, the first $r-1$ rounds of bargaining have no explicit bearing on the outcome of the game, except for purposes of coordination. The payoffs for each outcome are:

1. Sheriff accepts the bribe. The Smuggler's gets $n \cdot v - b_r$, and the Sheriff's gets the bribe offered b_r.
2. Sheriff inspects and discovers illegal items. The Smuggler is fined and gets a payoff of $-n \cdot p$ while the Sheriff gets a payoff of $n \cdot p$.
3. Sheriff chooses to inspect and does not find illegal items. The Smuggler receives a compensation of s, while the Sheriff gets $-s$.

The objective of the mediator is to maximize social welfare in the space EFCEs. Ideally, this will involve the Smuggler bringing in goods and the Sheriff accepting bribes – any other outcome would simply be zero-sum. On the other hand, a qualitative description of the welfare maximizing equilibrium is not obvious, since the game contains elements of both lying and bargaining.

Remark. The communication in the bargaining steps are similar to that in *cheap talk* [Crawford and Sobel, 1982], where costless and non-binding signals are transmitted between players. However, in our setting, the signals are transmitted in the middle of the game as opposed to just at the beginning. More importantly, the presence of the mediator during the phase of bargaining bestows more uses for the signals—in particular, the mediator may be able to take punitive measures against players who deviate from recommendations, since future recommendations will be withheld from players who deviate. This will be illustrated later.

4.2 Experiments and Discussion

Varying v, p **and** s For the purposes of quantitative discussion, we will focus on the case where $v = 5, p = 1, s = 1, n_{\max} = 10, b_{\max} = 2, r = 2$. We vary the item value v, item penalty p, and Sheriff compensation (penalty) s in isolation over a continuous range of values. The results are shown in Figure 3.[4]

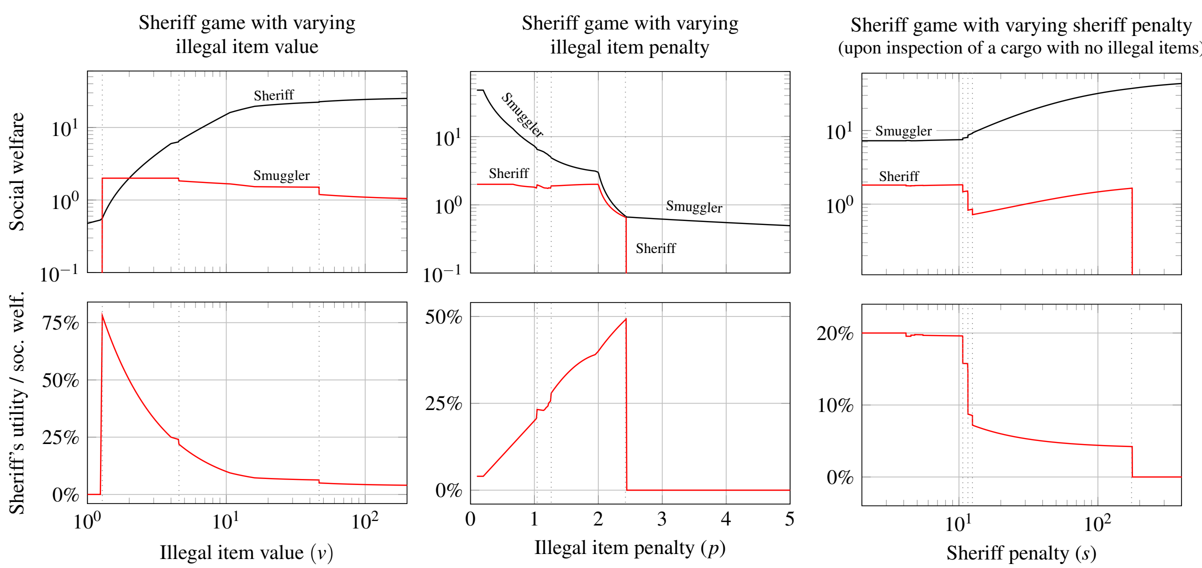

Figure 3: Payoffs to players with varying v, p and s for the SW maximizing equilibrium. Note the difference in scales for each of the 3 cases. Top: Payoffs in absolute terms. Bottom: Proportion of SW captured by the Sheriff.

Figure 3 illustrates several surprising trends. In all 3 situations, the effect of the parameter to the Smuggler is fairly consistent with intuition. However, the situation is more complicated for the Sheriff, whose source of utility comes from both making accurate inspections and accepting bribes. Consider the case where the case where s is gradually. Here, the Sheriff is discouraged from inspections and the Smuggler profits. As s increases, the Sheriff experiences a dip in payoff at around $s = 10$. However, the Sheriff's expected utility increases with $s \in [10, 100]$. In other words, harsher penalties for false alarms

[4]In order to verify that these plots are not the result of equilibrium selection issues, we reran experiments with a jittered objective and observed no qualitative change in behavior shown in the plots.

turns out to be beneficial to Sheriff (as well as the Smuggler) under an EFCE. We hypothesize that a higher s makes the Smuggler more willing to smuggle more, and in turn, more willing to hand out bribes to the Sheriff. However, as s increases even further, the Sheriff loses any leverage he once enjoyed.

Similar observations may be made in the other 2 settings. When p increases, the Sheriff initially benefits since he is able to collect more from each illegal item found. However, at around $p = 2$, the Sheriff begins to suffer from further increase in p, either because the Smuggler is less willing to bribe, or simply because he chooses to smuggle less. At around $p = 2.5$, the equilibrium undergoes yet another drastic change, and the Sheriff no longer extracts any utility from the entire setup.

Varying discrete variables $n_{\mathbf{max}}, b_{\mathbf{max}}, r$ Here, we try to empirically understand the impact of n and b on the SW maximizing equilibrium. As before we set $v = 5, p = 1, s = 1$ and vary n and r simultaneously while keeping b constant. The results are shown in Table 1.

$n_{\max}$	$r=1$	$r=2$	$r=3$	$r=4$
1	(4.00, 1.00)	(4.00, 1.00)	(4.00, 1.00)	(4.00, 1.00)
2	(1.24, 0.19)	(4.00, 1.00)	(4.00, 1.00)	(4.00, 1.00)
5	(0.89, 0.11)	(1.11, 1.00)	(4.00, 1.00)	(4.00, 1.00)
10	(0.82, 0.00)	(0.84, 1.00)	(3.62, 1.00)	(4.00, 1.00)
20	(0.83, 0.00)	(0.83, 0.86)	(2.84, 1.00)	(4.00, 1.00)

$n_{\max}$	$r=1$	$r=2$	$r=3$
1	(3.00, 2.00)	(3.00, 2.00)	(3.00, 2.00)
2	(8.00, 2.00)	(8.00, 2.00)	(8.00, 2.00)
5	(2.28, 1.26)	(8.00, 2.00)	(8.00, 2.00)
10	(1.76, 0.93)	(7.26, 1.82)	(8.00, 2.00)
20	(1.59, 0.77)	(6.65, 1.66)	(8.00, 2.00)

Table 1: SW maximizing payoffs for (Smuggler, Sheriff) with $b = 1$ (left) and $b = 2$ (right) and varying $n_{\max}$ and r.

A few observations are immediately apparent. First, increasing n while fixing r and b may *decrease* social welfare. For example, consider the case when $b = 2, n_{\max} = 2, r = 1$ (Table 1) where the payoffs are $(8.0, 2.0)$. Note that this is achieves the maximum attainable social welfare by smuggling $n_{\max} = 2$ items and having the Sheriff accept a bribe of 2. However, if we increase $n_{\max}$ to 5, the payoffs to *both* players drop significantly, and even more so when $n_{\max}$ increases further. This observation is counter-intuitive at first glance; since $n_{\max}$ only controls the maximum number of items smuggled, it is tempting to conclude that by 'ignoring' the possibility of smuggling more items, one should not do worse than before. However, this line if reasoning is incorrect; direct lifting of the equilibrium of $n_{\max} = 2$ to $n_{\max} = 3$ would not constitute an equilibrium, since the Smuggler could strictly benefit by smuggling 3 as opposed to 2 items. Consequently, the recommendation to the Sheriff would be adjusted to the *possibility* that 3 items were smuggled, ultimately leading to both parties suffering.

Secondly, the aforementioned 'paradox' may be alleviated by increasing r, the number of rounds of bargaining. Furthermore, there appears to be a maximum attainable SW determined by b, regardless of how large r is. When $b = 2$ (Table 1) this maximum appears to apportion payoffs of $(4, 1)$, with a SW of 5. Similarly, this maximum is $(8.0, 2.0)$ when $b = 2$. Interestingly, this appears to be a hard upper bound which is independent on $n_{\max}$, we believe that these boundary values are dependent on the relative settings of v, p and s. With sufficient bargaining steps, the Smuggler, with the aid of the mediator, is able to convince the Sheriff during the smuggling rounds that he is indeed bringing in the exact number of items recommended by the mediator. We explore this idea more in the following section.

4.3 Understanding the effect of additional rounds of bargaining r.

We illustrate the effect of the non-consequential bribes with two small settings, where $v = 5, p = 1, s = 1, n_{\max} = 3, b_{\max} = 2, r \in \{1, 2\}$. Examples of SW-maximizing equilibria are shown in Figure 4 and

Figure 5. [5]

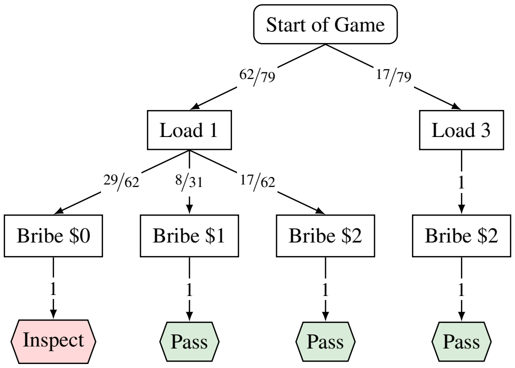

Figure 4: Example of a playthrough of the Sheriff game with $r = 1$. Edge labels correspond to action probabilities, edges with 0 probability are omitted. Squares and hexagons denote actions taken by Players 1 and 2 respectively, while green and red nodes denote the Sheriff choosing to pass or inspect.

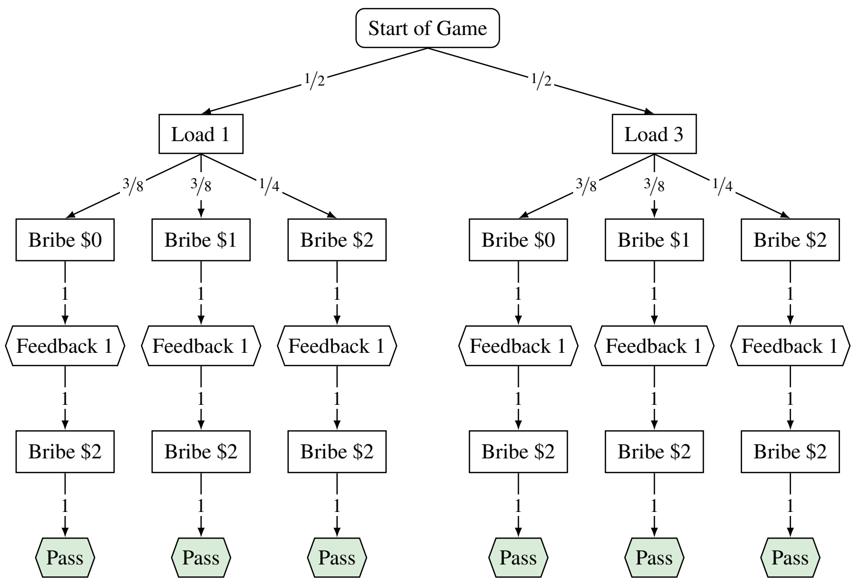

Figure 5: Example of a playthrough of the Sheriff game with $r = 2$. Edge labels correspond to action probabilities, edges with 0 probability are omitted. Squares and hexagons denote actions taken by Players 1 and 2 respectively, while green and red nodes denote the Sheriff choosing to pass or inspect.

The SW maximizing EFCE yields payoffs of $(3.89, 1.43)$ and $(8.0, 2.0)$ for $r = 1$ and $r = 2$ respectively. We will first consider the case where $r = 2$ (Figure 5. Here, what occurs happens along the equilibrium path is straightforward. The Smuggler loads in 1 or 3 items with equal probability. Next, he offers a (non-consequential) bribe of either 0, 1, or 2. Then, he receives some feedback of 1, and proceeds to offer a bribe of $2, which the Sheriff gladly accepts. The payoffs to players is $(13, 2)$ and $(3, 2)$ depending if the Smuggler was recommended to load 1 or 3 items, leading to an average payoff of $(8, 2)$.

[5] As before, note that this only shows interactions of players on the equilibrium path, that is, the graph omits what would happen if some player deviated.

The underlying mechanism is in fact fairly straightforward and mirrors the idea in the modified signalling game of Von Stengel and Forges [2008]. Assume that a random number is chosen uniformly from $\{0,\ldots,b\}$. This acts as a 'secret code' which the Sheriff expects from the Smuggler in the first round. This secret code part of the correlated plan, and will eventually be revealed to the Smuggler assuming he did not deviate when selecting the number of illegal items. [6] In other words, *the first (non-consequential) bribe may be used as a signal which hints to the Sheriff if the Smuggler has deviated*—if it is not equal to the secret code, the Smuggler *must* have deviated somewhere. On the other hand, a deviating Smuggler may guess the secret code with probability no greater than $1/(b+1)$; if the number of signals b is sufficiently large, then it is near impossible to guess the code. Using these tools, the mediator is able to engineer a 'deviation detector' which checks if the Smuggler ever deviated. Note, however, that unlike the Signaling game, the Sheriff is not able to glean exactly how many illegal items were loaded; he is only able to deduce if the player deviated from some recommendation, which could be to load either 1 or 3 items.

Issuing threats to the Smuggler becomes straightforward with this deviation detector. If the Sheriff knows the Smuggler is lying, he employs a 'grim trigger' for the rest of the game—in this case, the Sheriff opts to inspect all of the player's cargo, regardless of the bribe offered in the second round. The Smuggler could also be pretending to bring in illegal goods, i.e., by loading 0 items and hoping that he would guess the *incorrect* secret code, resulting in the Sheriff making a false accusation. However, because the Smuggler's payoff for deceiving the Sheriff in this manner is just 1, he remains incentivized to stick to the recommendations, which guarantees him a payoff of either 3 of 13.

We now make the following hypotheses. First, the effect of additional bargaining rounds r is that the chance of randomly guessing the secret code is reduced. If there are r rounds, then there are $(b+1)^{r-1}$ different possible signals that the Smuggler could have sent to the Sheriff through the first $r-1$ rounds. When $r=1$, this class of correlation plans fails since the bribe by the Smuggler serves both as the answer to the 'secret question' and as the actual bribe to be offered. This aliasing of roles is what leads to a lower payoff; the risk of sending an incorrect secret code is not sufficiently high to dissuade the Smuggler from deviating.

5 Conclusion

In this paper, we have proposed two parameterized benchmark games in which social-welfare-maximizing EFCE exhibits interesting behaviors. We have analyzed those behaviors both qualitatively and quantitatively, and isolated two ways through which a mediator is able to compel the agents to follow the recommendations. We hope that our analysis will bring attention to some of the potential practical uses of EFCEs, and that our benchmark games be used to evaluate EFCE solvers for large games.

References

Itai Ashlagi, Dov Monderer, and Moshe Tennenholtz. On the value of correlation. *Journal of Artificial Intelligence Research*, 33:575–613, 2008.

Robert Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

[6] Recall that the sequential nature of the EFCE means that the recommended amount to bribe is not revealed until the Smuggler loads the cargo with the recommended number of items.

Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Science*, 2015.

Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, Dec. 2017.

Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.

Wan Huang and Bernhard von Stengel. Computing an extensive-form correlated equilibrium in polynomial time. In *International Workshop On Internet And Network Economics (WINE)*, pages 506–513. Springer, 2008.

Wan Huang. *Equilibrium computation for extensive games*. PhD thesis, London School of Economics and Political Science, January 2011.

Albert Xin Jiang and Kevin Leyton-Brown. Polynomial-time computation of exact correlated equilibrium in compact games. *Games and Economic Behavior*, 91:347–359, 2015.

H. W. Kuhn. A simplified two-person poker. In *Contributions to the Theory of Games*. Princeton University Press, 1950.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 2017.

John Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36:48–49, 1950.

Christos H Papadimitriou and Tim Roughgarden. Computing correlated equilibria in multi-player games. *Journal of the ACM*, 55(3):14, 2008.

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