Correlation in Extensive-Form Games: Saddle-Point Formulation and Benchmarks

Gabriele Farina[1] Chun Kai Ling[1] Fei Fang[2]
Tuomas Sandholm[1,3,4,5]

[1] Computer Science Department, Carnegie Mellon University
[2] Institute for Software Research, Carnegie Mellon University
[3] Strategic Machine, Inc.
[4] Strategy Robot, Inc.
[5] Optimized Markets, Inc.

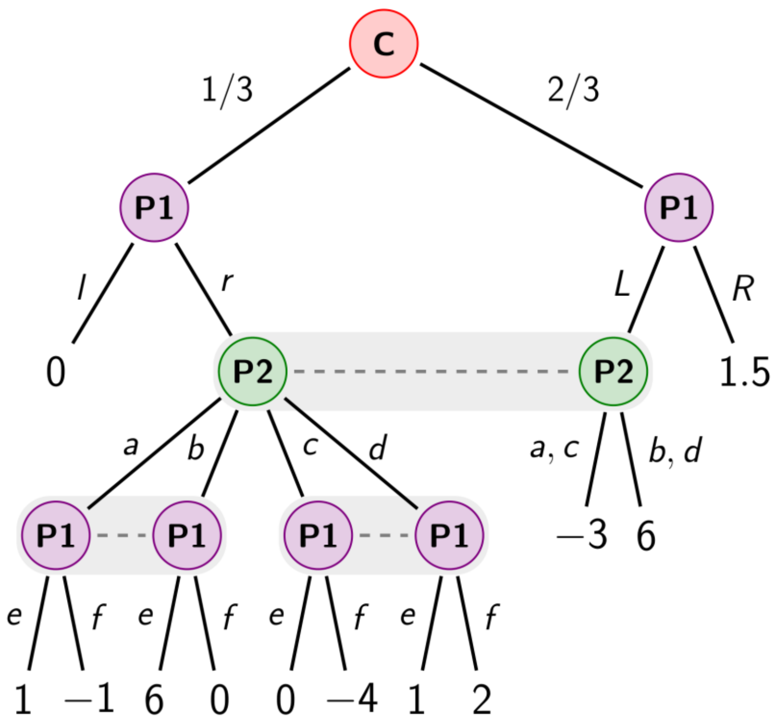
The concept of correlation

- Nash equilibrium assumes a fully decentralized interaction
  - Not the best solution concept in situations where some intermediate form of centralized control can be achieved
- Correlated equilibrium [Aumann 1974]: a mediator can recommend behavior but not enforce it
  - Well understood in normal-form games but not in extensive-form games

Summary of main contributions

- Primary objective: spark more interest in the community towards a deeper understanding of the behavioral and computational aspects of extensive-form correlation
- We propose two parametric benchmark games
  - Chosen to illustrate natural application domains of EFCE: conflict resolution and bargaining/negotiation
  - They can scale in size as desired
- We isolate two mechanisms through which a mediator is able to compel the agents to follow the recommendations
- We show that the problem of computing an optimal extensive-form correlated equilibrium is a saddle-point problem

Extensive-Form Games



- Can capture sequential and simultaneous moves
- Private information
- Each information set contains a set of "undistinguishable" tree nodes
- We assume perfect recall: no player forgets what the player knew earlier

Extensive-Form Correlated Equilibrium (EFCE)

- Introduced by von Stengel and Forges in 2008
- Correlation device selects private signals for the players before the game starts
  – The correlated distribution of signals is known to the players
- Recommendations are revealed incrementally as the players progress in the game tree
  – A recommended move is only revealed when the player reaches the decision point for which the recommendation is relevant
  – Players are free to defect, at the cost of future recommendations

Extensive-Form Correlated Equilibrium (EFCE)

- The players don’t know exactly what pair of strategies the correlation device is trying to induce the players to play
  - Bayesian reasoning: after observing each recommendation, the players update their posterior
- The players are free to defect, at the cost of future recommendations
  - The orchestrator cannot enforce behavior
  - The recommendations must be incentive-compatible
  - One of the orchestrator’s leverages: stop giving recommendations

Extensive-Form Correlated Equilibrium (EFCE)

- A social-welfare-maximizing orchestrator that is provably incentive-compatible can be constructed in polynomial time in two-player general-sum games with no chance moves [von Stengel and Forges, 2008]
  - Players can be induced to play strategies with significantly higher social welfare than Nash equilibrium…
  - …even despite the fact that each player to defect
  - Added benefit: players get told what to do---they do not need to come up with their own optimal strategy as in Nash equilibrium

Benchmark games

- EFCE can lead to better social welfare than Nash equilibrium
- EFCE is often highly nontrivial

First benchmark game: Battleship

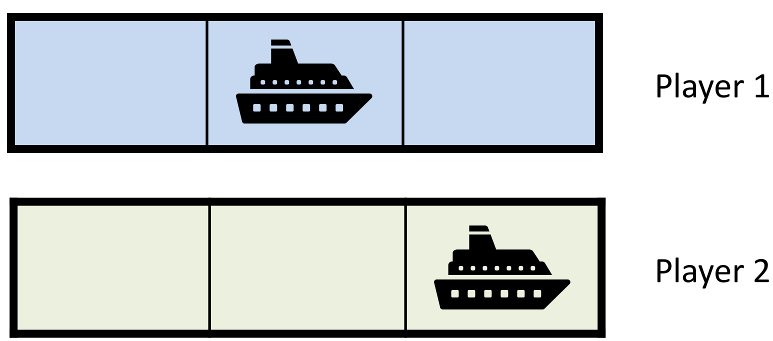
Conflict resolution via a mediator

Battleship

- Players take turns to secretly place a set of ships of varying sizes and value on separate grids of size $H \times W$
- After placements, players take turns firing at their opponent
- Ships which have been hit at all the tiles they lie on are considered destroyed
- The game continues until either one player has lost all of their ships, or each player has completed n shots
- Payoff: (value of opponent's ships that were destroyed) $-\gamma \cdot$ (value of own ships that were destroyed)

Toy example

- For now, let's focus on a specific instance of the game:
  - Board size: 3x1
  - Each player only has one ship: length 1, value 1
  - Max 2 rounds of shooting per player



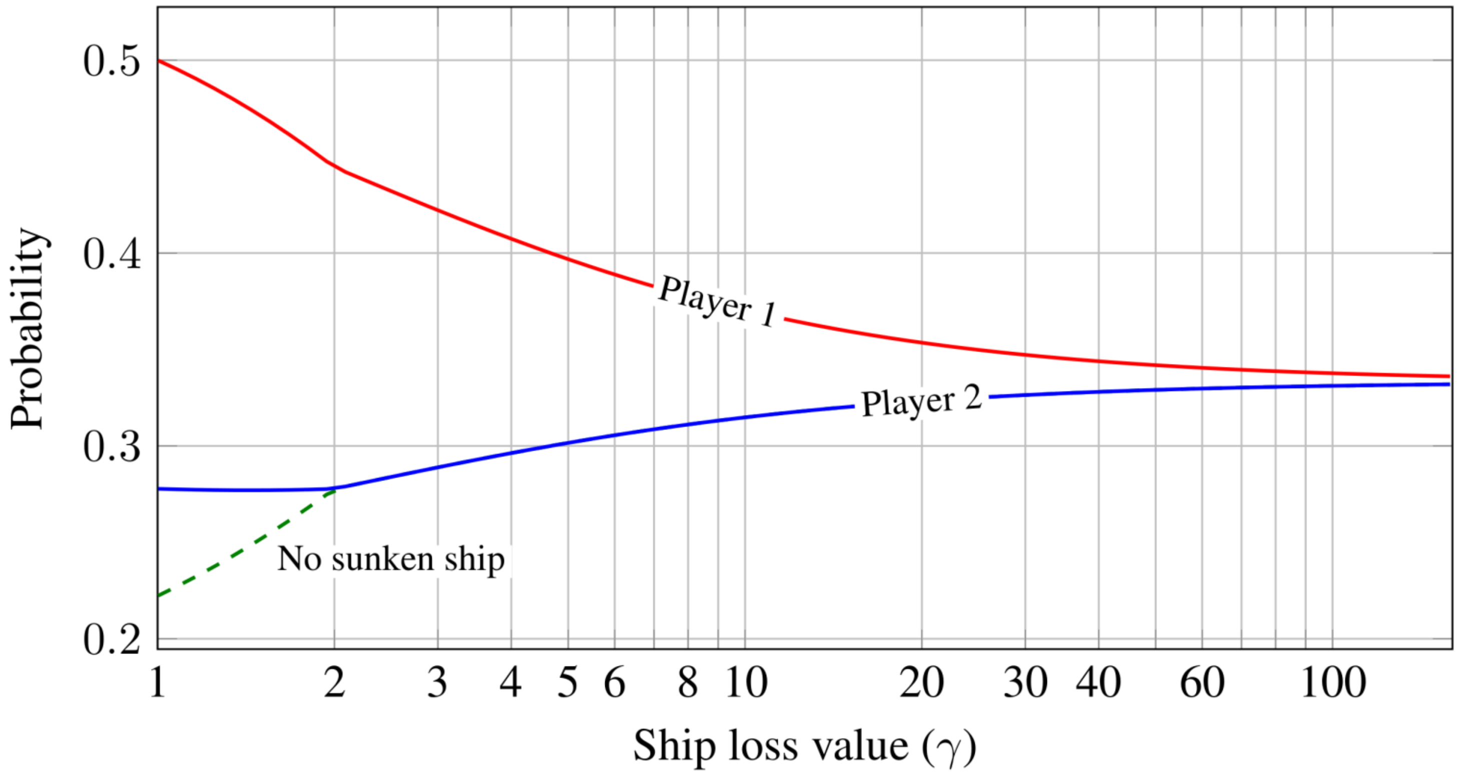
Nash vs EFCE

- The social-welfare-maximizing Nash equilibrium is to **place ships at random, and to shoot at random**
  - Player 1 wins with probability: 5/9
  - Player 2 wins with probability: 1/3
  - Probability of no ship destroyed: **1/9**
  - Social welfare of Nash equilibrium: -8/9 when $\gamma = 2$

Nash vs EFCE

- The social-welfare-maximizing Nash equilibrium is to **place ships at random, and to shoot at random**
  - Player 1 wins with probability: 5/9
  - Player 2 wins with probability: 1/3
  - Probability of no ship destroyed: **1/9**
  - Social welfare of Nash equilibrium: -8/9 when $\gamma = 2$
- The EFCE mediator is able to compel the players into not sinking any ship with probability **5/18** (when $\gamma = 2$)
  - 2.5x higher probability of peaceful outcome than Nash
  - Social welfare: -13/18 when $\gamma = 2$

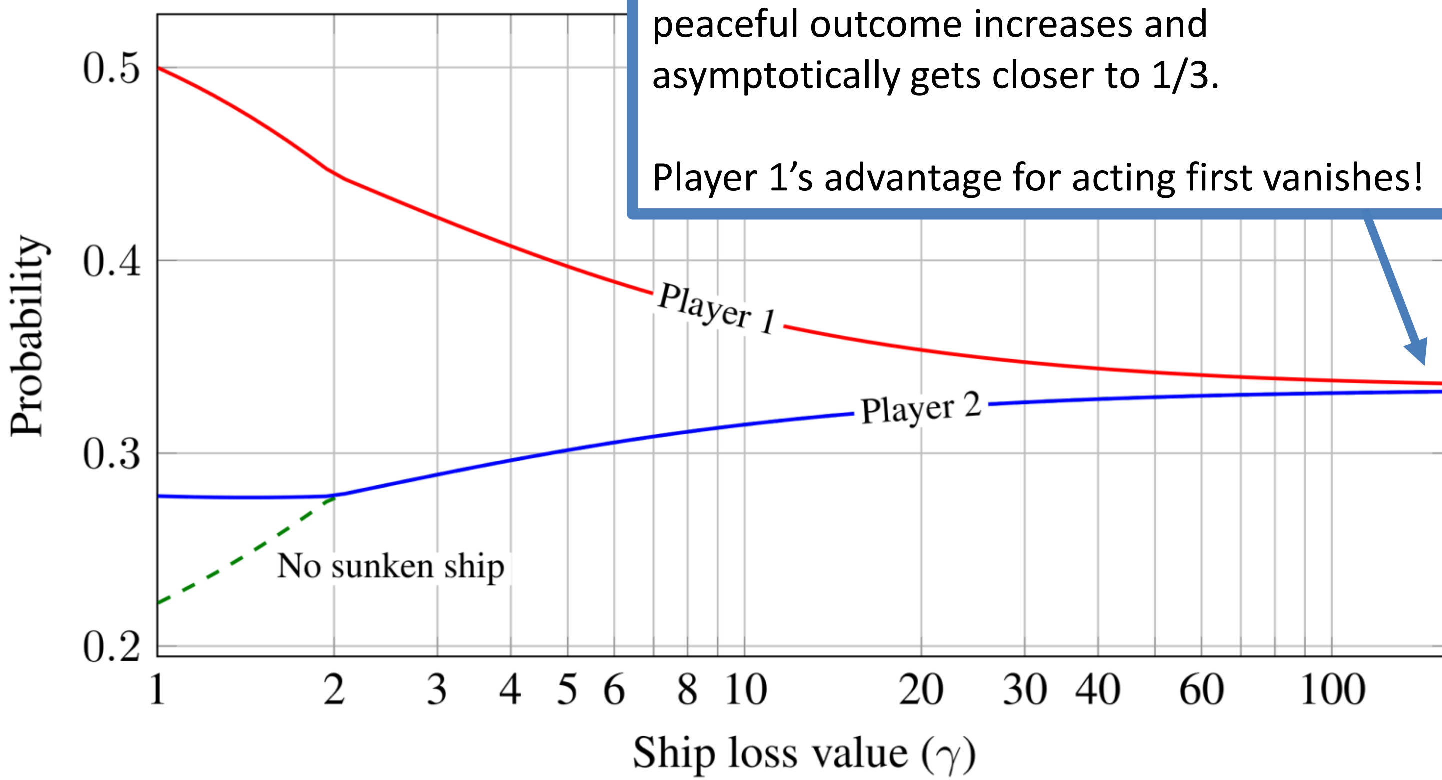
Probability of sinking ships



Probability of sinking ships

In the limit, the probability of reaching a peaceful outcome increases and asymptotically gets closer to 1/3.

Player 1’s advantage for acting first vanishes!

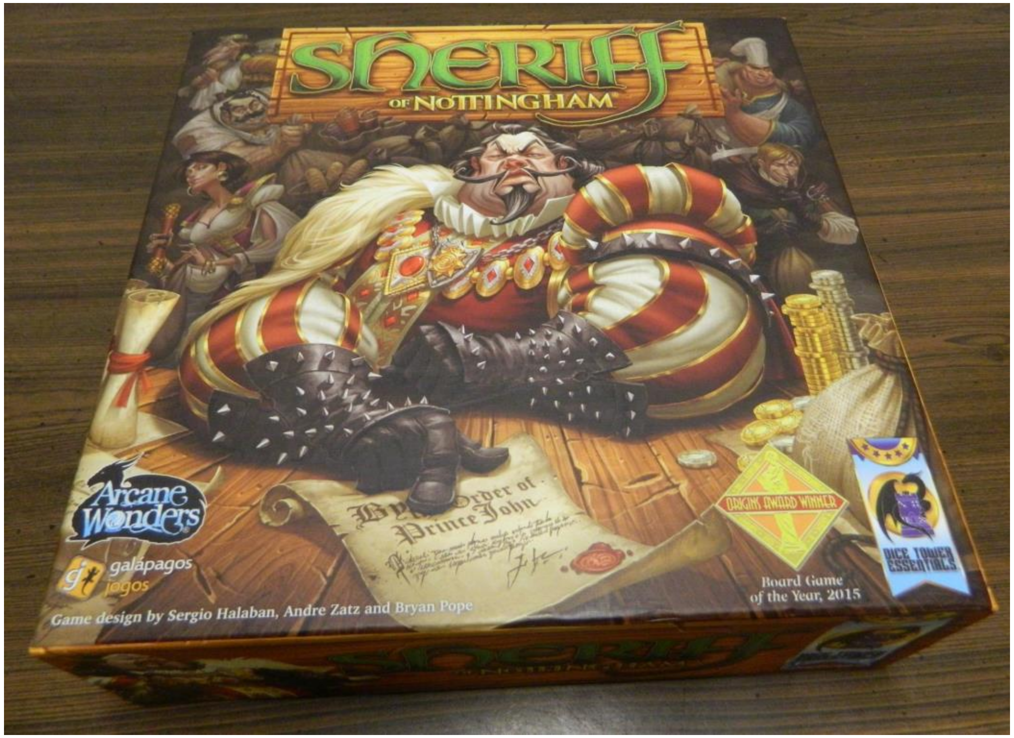


The strategy of the mediator

- In a nutshell:
  - Correlation plan is constructed so that players are recommended to deliberately miss
  - Incentive-compatibility: deviations are **punished by the mediator**, who **reveals to the opponent the ship location that was recommended to the deviating player**
- Details are complicated---see paper
  - Mediator must keep under check how much information is revealed with each recommendation, and account for the fact that players are free to defect at any point

Second Benchmark game: Sheriff

Bargaining and negotiation



Sheriff game

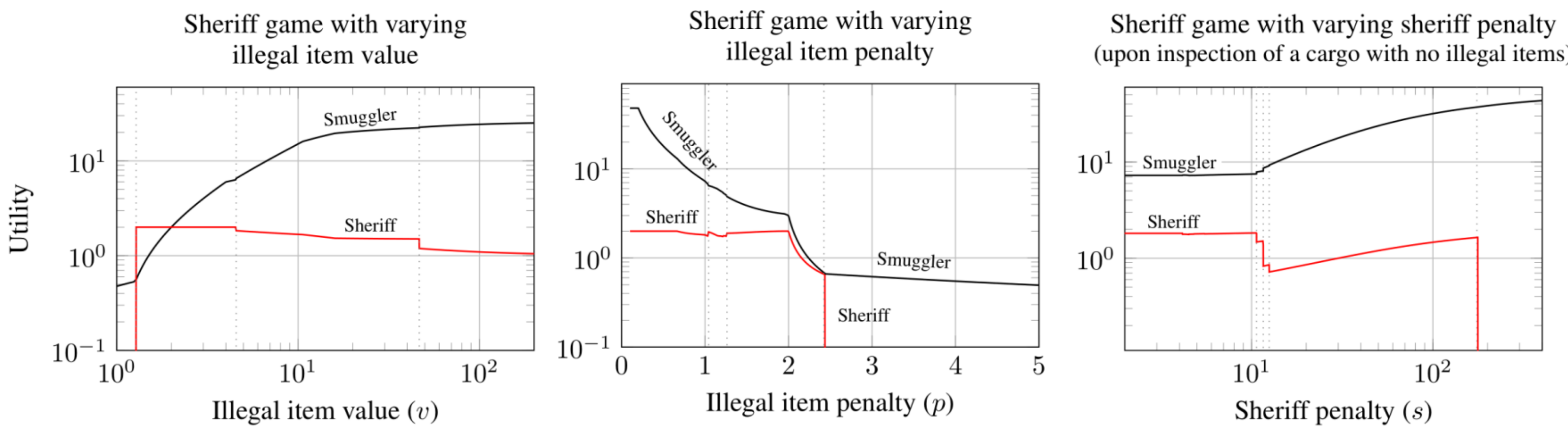
- The smuggler is trying to smuggle illegal items in their cargo
- The sheriff is trying to stop the Smuggler
- At the beginning of the game, the smuggler secretly loads his cargo with $n \in \{0, \ldots, n_{\max}\}$ illegal items
- At the end of the game, the sheriff decides whether to inspect the cargo or not
  - If yes, the smuggler must pay a fine $n \cdot p$ if $n > 0$, otherwise the sheriff must compensate the smuggler with a utility of s
  - If no, the smuggler utility is $n \cdot v$, and the sheriff's utility is 0

Sheriff game: bribery and bargaining rounds

- The game is made interesting by two additional elements (present in the original game too): bribery and bargaining
- After the smuggler loaded the cargo, the two players engage in r rounds of bargaining:
  - At each round $i = 1, \ldots, r$, the smuggler offers a bribe $b_i \in \{0, \ldots, b_{\max}\}$, and the sheriff responds whether or not he would accept the proposed bribe
  - This decision is non-consequential
  - If the sheriff accepts bribe b_r the smuggler gets a utility of $p \cdot n - b_r$ and the sheriff gets a utility of b_r

EFCEs in the Sheriff game

- Baseline instance: $v = 5,\ p = 1,\ s = 1,\ n_{\max} = 10,\ b_{\max} = 2,\ r = 2$



- Non-monotonic behavior
- Not even continuous!

EFCEs in the Sheriff game

- With sufficient bargaining steps, the smuggler, with the help of the mediator, is able to convince the sheriff that they have complied with the recommendation by the mediator
  - The mediator spends the first $r - 1$ bribes to **give a 'passcode' to the smuggler, so that the sheriff can verify compliance**
  - If an unexpected bribe is suggested, then the smuggler must have deviated, and the sheriff will inspect the cargo as punishment

Main takeaways

- EFCE is often nontrivial
- We offer the **first empirical observations** as to how EFCE is able achieve a better social welfare than Nash equilibrium while only recommending behavior without enforcing it
  - Mediator makes sure that the fact that players stop receiving recommendations upon defection is a deterrent
  - Furthermore, the mediator recommends punitive behavior to the opponent if the mediator detects deviations from the recommendations

Saddle-point formulation

- EFCE can be formulated as a bilinear min-max problem (just like Nash equilibrium)
- This enables the use of a wide array of tools beyond linear programming

Saddle-point formulation

- Finding an EFCE in a two-player game can be seen as a bilinear saddle-point problem
$$\min_{x \in X} \max_{y \in Y} x^T A y$$
where:
  - X, Y are convex polytopes
  - A is a real matrix
- This brings the problem of computing EFCE closer to several other concepts in game theory

Saddle-point formulation

- From a geometric angle, the saddle-point formulation better captures the combinatorial structure of the problem
  – Sets X and Y have well-defined meaning in terms of the input game tree
  – Algorithmic implications. For example, because of the structure of Y, the minimization problem can be performed via a single bottom-up game tree traversal

Saddle-point formulation

- From a computational point of view, the bilinear saddle-point formulation opens the way to the plethora of optimization algorithm that has been developed specifically for saddle-point problems
  - First-order methods (e.g., subgradient descent)
  - Regret minimization methods
- Our saddle-point formulation can be used to prove the correctness of the linear-programming-based approach of von Stengel and Forges (2008)

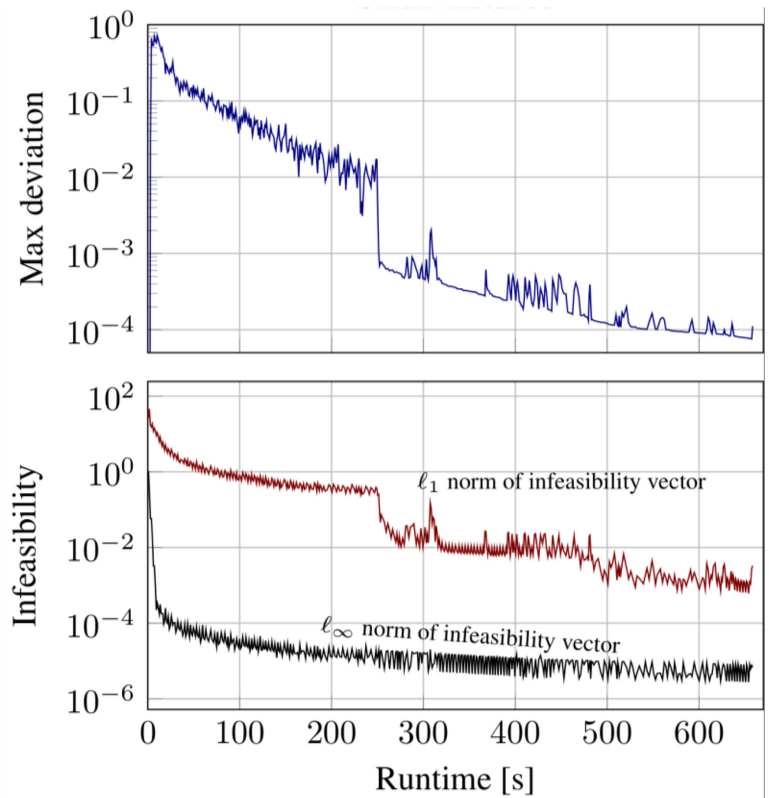
Projected subgradient method

- As a proof of concept, we implemented a recent method based on subgradient descent [Wang and Bertsekas, 2013] to solve the bilinear saddle-point problem
- Our method beats the commercial linear programming solver Gurobi in large Battleship games

(H, W)	r	Ship length	#Actions Pl 1	Pl 2	#Relevant seq. pairs	Time (LP) 10^{-1}	10^{-2}	10^{-3}	Time (ours) 10^{-1}	10^{-2}	10^{-3}
(2, 2)	3	1	741	917	35241	2s	2s	2s	1s	2s	3s
(3, 2)	3	1	15k	47k	3.89M	3m 6s	3m 17s	3m 24s	8s	34s	52s
(3, 2)	4	1	145k	306k	26.4M	42m 39s	42m 44s	43m	2m 48s	14m 1s	23m 24s
(3, 2)	4	2	970k	2.27M	111M	—- out of memory† —-			—- did not achieve ‡ —-		

Projected subgradient method

- Our method trades off feasibility of the iterates with their optimality
- Game instance in experiment to the right:
  - 15k unique actions for Pl. 1
  - 47k unique actions for Pl. 2



Regret minimization method

- We also designed the first efficient regret minimization method for computing EFCE
  - Designing such an algorithm is significantly more challenging than designing one for the Nash equilibrium counterpart: the constraints that define the space of correlation plans lack a hierarchical structure and might even form cycle
  - Our approach is based on a special convexity-preserving operation that we coin ‘scaled extension’
- Our regret-based approach is significantly faster than Gurobi in large games, and guaranteed to produce feasible iterates

Conclusions

- We introduced two benchmark games in which EFCE exhibits interesting behaviors
- We analyzed those behaviors both qualitatively and quantitatively
- We isolated two ways in which the mediator is able to compel the agents to follow the recommendations
- We showed that an EFCE can be computed via a bilinear saddle-point problem and demonstrated the merits of this formulation by designing algorithms that outperform standard LP-based methods