Cost Minimization for Equilibrium Transition

Haoqiang Huang[1], Zihe Wang[2*], Zhide Wei[3], Jie Zhang[4]

[1]Hong Kong University of Science and Technology
[2]Renmin University of China
[3]Peking University
[4]University of Bath
hhuangbm@connect.ust.hk, wang.zihe@ruc.edu.cn, zhidewei@pku.edu.cn, jz2558@bath.ac.uk

Abstract

In this paper, we delve into the problem of using monetary incentives to encourage players to shift from an initial Nash equilibrium to a more favorable one within a game. Our main focus revolves around computing the minimum reward required to facilitate this equilibrium transition. The game involves a single row player who possesses m strategies and k column players, each endowed with n strategies. Our findings reveal that determining whether the minimum reward is zero is NP-complete, and computing the minimum reward becomes APX-hard. Nonetheless, we bring some positive news, as this problem can be efficiently handled if either k or n is a fixed constant. Furthermore, we have devised an approximation algorithm with an additive error that runs in polynomial time. Lastly, we explore a specific case wherein the utility functions exhibit single-peaked characteristics, and we successfully demonstrate that the optimal reward can be computed in polynomial time.

Introduction

Equilibrium analysis has remained a central focus in game theory research for several decades. The exploration of the fundamental Nash equilibrium and its various refinements has been the subject of extensive study. It is widely recognized that certain equilibria hold greater desirability than others, prompting investigations into the Price-of-Anarchy (Koutsoupias and Papadimitriou 2009) and Price-of-Stability (Anshelevich et al. 2008). In the dynamics of the game, players are driven to best respond to their counterparts' strategies (Hopkins 1999; Leslie, Perkins, and Xu 2020). Nevertheless, this approach may lead to suboptimal equilibria, as players overlook the potential benefits that could arise from deviating from the best response dynamics in response to successive changes made by other players.

In our research, we explore a scenario where a mediator aims to facilitate the transition from an initial equilibrium to a more favorable target equilibrium. The mediator achieves this by subsidizing the players to influence their behavior, encouraging them to progress towards the desirable equilibrium in gradual steps. Throughout this process, the mediator's primary objective is to minimize the overall cost involved. This study holds significant real-world applications. For instance, many governments worldwide are actively pursuing initiatives to boost the adoption of electric vehicles as part of their net-zero plans. Given that cars and vans contribute nearly a fifth of total emissions, expediting the transition to electric vehicles is crucial to accomplishing their environmental goals. One can liken the initial equilibrium to the prevailing usage of petrol and diesel cars. To drive the shift towards electric vehicles, the government may implement tax benefits and provide funds for charger installations, thus incentivizing drivers to make the switch. The more desirable target equilibrium, in this case, would involve completely phasing out the sale of new petrol and diesel cars.

In this intriguing game, we have a row player and k column players, each with specific payoffs represented by matrices R and C respectively. The game commences from an initial equilibrium and proceeds in rounds. As a mediator, the ultimate objective is to lead the players towards a desirable Nash equilibrium by offering rewards in each round. Even in the case of just one row player, the theoretical implications are already noteworthy. In practical terms, this single row player scenario effectively captures numerous monopolistic markets and the regulatory behavior of governments in markets. However, as we will demonstrate, the problem at hand becomes intractable in various setups. Thus, achieving positive outcomes in more general settings necessitates the application of additional constraints or the consideration of special cases.

Our Contribution

In this paper, we tackle the challenge of designing algorithms for computing the minimum reward needed to foster equilibrium transitions. Given a bimatrix game with payoff matrices $R_{m\times n}$ and $C_{m\times n}$, assuming that there are k column players playing against a row player, we show the following results:

- Determining whether the minimum reward is zero is NP-complete, and computing the minimum reward, in general, is APX-hard.
- However, computing the optimal reward scheme is slice-wise polynomial with respect to k and n, respectively.
- We design an approximation algorithm for this problem that runs in polynomial time. The additive approximation error is linear in the number of the row player's choices and the largest number of matrices R and C.

[*]Zihe Wang is the corresponding author.

- Last, we consider a special case where the utility functions are single-peaked and show that the optimal reward can be computed in polynomial time.

We show the hardness results by reductions from the EXACT COVER problem and a variant of the Knapsack problem. In the approximation algorithm, we construct a complete directed graph in which the weight of an edge corresponds to the solution of an integer linear programming (ILP). By approximating the solution of the ILP, the problem of finding the optimal transformation path from an initial equilibrium to a target equilibrium boils down to finding the shortest paths between vertices of the graph. Due to limited space, some proofs are skipped and can be found in the full version (Huang et al. 2023).

Related Work

Monderer and Tennenholtz (2004) consider the implementation of desirable outcomes by a reliable party who cannot modify game rules but can make non-negative payments to the players. They term this k-implementation problem an intermediate approach between algorithmic game theory (Papadimitriou 2011) and mechanism design (Nisan and Ronen 1999). They provide characterizations of k-implementation for the implementation of singletons in games with complete information and investigate several settings under which the problem is polynomial-time solvable or intractable. Deng, Tang, and Zheng (2016) follow up the study of k-implementation and prove that the problem is NP-complete for general games with respect to dominance by pure strategies. Furthermore, the authors study a variation of the k-implementation problem by characterizing its hardness and developing computationally efficient algorithms for supermodular games. Unlike k-implementation, which considers a single-round bi-matrix game, we aim to motivate a group of players to move to a target equilibrium step by step and minimize the total cost. Deng and Conitzer (2017) consider a disarmament game, in which players successively commit not to play certain strategies and thereby iteratively reduce their strategy spaces. Later on, in (Deng and Conitzer 2018), instead of removing a strategy in a game, they consider removing a resource which leads to ruling out all the strategies in which that resource is used simultaneously. They prove NP-completeness of several formulations of the problem of achieving desirable outcomes via disarmament.

The idea of allowing a mediator to influence the players' behavior and hence the outcome of a system has been widely studied. For example, Rozenfeld and Tennenholtz (2007) focus on the use of routing mediators in order to reach a correlated strong equilibrium. They show that a natural class of routing mediators allows the implementation of fair and efficient outcomes as a correlated super-strong equilibrium in a very wide class of games. Monderer and Tennenholtz (2009) propose to use mediators in order to enrich the set of situations where one can obtain stability against deviations by coalitions, in light of the understanding that strong equilibrium rarely exists. Augustine et al. (2015) address the question of whether a network designer can enforce particular equilibria or guarantee that efficient designs are consistent with users' selfishness by appropriately subsidizing some of the network links. They formulate this question as one of the optimization problems and present positive and negative results. Eidenbenz et al. (2007) consider the problem of a mechanism designer seeking to influence the outcome of a strategic game based on her creditability.

Studying the best response dynamics of players constitutes a fundamental aspect of game theory research. In a recent study, Amiet et al. (2021) explore games where the payoffs are drawn at random and demonstrate that a best-response dynamics approach leads to a pure Nash equilibrium with a high probability as the number of players increases. Feldman, Snappir, and Tamir (2017) delve into congestion games, where they investigate the inefficiency of various deviator rules. They find that the best response dynamics consistently converges to a pure Nash equilibrium in such games. Heinrich et al. (2022) analyze the performance of the best-response dynamic across all normal-form games using a random games approach. They show that the best-response dynamic converges to a pure Nash equilibrium in a vanishingly small fraction of all large games when players take turns according to a fixed cyclic order. By contrast, when the playing sequence is random, the dynamic converges to a pure Nash equilibrium if one exists in almost all large games.

Preliminary

We consider a game in which there is a row player and k column players. The row player and each of the k column players constitute a bimatrix game. Let $\mathscr{R}$ be the strategy set of the row player and $\mathscr{C}$ be the strategy set of a column player. Denote $\mathscr{C}^k$ the set of all possible strategy profiles of k column players. Let r^i and c^j be the row player's i-th strategy and a column player's j-th strategy, respectively, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Throughout the paper, the players only adopt pure strategies. Let $R_{m\times n}$ and $C_{m\times n}$ be the payoff matrices of the row player and the column players, respectively. Denote $(r(t), c_1(t), \ldots, c_k(t))$ the strategy profile of all players at time step t, where $r(t) \in \mathscr{R}$ and $c_i(t) \in \mathscr{C}, i = 1, \ldots, k$. For ease of notation, we denote $\mathcal{C}(t) = (c_1(t), \ldots, c_k(t)) \in \mathscr{C}^k$. This way, the row player's payoff at time t is $\sum_{1\leq i\leq k} R(r(t), c_i(t))$, i.e., the sum of its payoff in the k bimatrix games. The payoff of the column player i is $C(r(t), c_i(t))$, i.e., its payoff in the bimatrix game that it plays against the row player.

In the context of equilibrium transition, the game starts with an equilibrium. That is, the row player and the column players are at an equilibrium in each of the k bimatrix games. Assume that, from the mediator's perspective, there exists another more desirable equilibrium. The mediator is interested in designing an optimal reward scheme that motivates the players to move to the more desirable equilibrium over multiple rounds.

Specifically, consider a strategy profile $(r(t), \mathcal{C}(t))$ in round t, without any additional reward, the row player will best respond to the column players' strategy profile $\mathcal{C}(t)$ in the next round. Therefore, to incentivize the row player playing a specific strategy $r(t + 1) \in \mathscr{R}$ in the next round, we need to provide a reward of, denoted by $T_{\mathcal{C}(t)}(r(t+1))$,

$$\max_{r^i \in \mathscr{R}} \sum_{1 \le j \le k} R(r^i, c_j(t)) - \sum_{1 \le j \le k} R(r(t+1), c_j(t)),$$

where the first term is the maximum payoff that the row player can get by taking the best response strategy against column players' strategies $\mathcal{C}(t)$, and the second term is the row player's payoff when it takes the strategy $r(t+1) \in \mathscr{R}$.

Similarly, given the row player's strategy $r(t)$ in round t, to incentivize column players taking strategy profile $\mathcal{C}(t+1)$ in round $t+1$, the total reward needed is $T_{r(t)}(\mathcal{C}(t+1))$, which is

$$k \cdot \max_{c^i \in \mathscr{C}} C(r(t), c^i) - \sum_{1 \le j \le k} C(r(t), c_j(t+1)).$$

Given the initial equilibrium $(r(1), \mathcal{C}(1))$ and the target equilibrium $(r^*, \mathcal{C}^*)$, a reward scheme that incentivizes the players moving from the initial equilibrium to the target equilibrium consists of a *transformation path* $(r(1), \mathcal{C}(1)) \to (r(2), \mathcal{C}(2)) \to \cdots \to (r^*, \mathcal{C}^*)$. The total cost of this reward scheme is $T = \sum_t \{T_{r(t)}(\mathcal{C}(t+1)) + T_{\mathcal{C}(t)}(r(t+1))\}$.

We define the optimization problem OPT TRANSITION (k, m, n) as follows.

Problem 1: OPT TRANSITION (k, m, n).
Input: Payoff matrices $R_{m \times n}$ and $C_{m \times n}$. The initial equilibrium $(r(1), \mathcal{C}(1))$ and the target equilibrium $(r^*, \mathcal{C}^*)$,
Output: A transformation path from strategy profile $(r(1), \mathcal{C}(1))$ to $(r^*, \mathcal{C}^*)$ such that the total cost $T = \sum_t \{T_{r(t)}(\mathcal{C}(t+1)) + T_{\mathcal{C}(t)}(r(t+1))\}$ is minimized.

We also consider the following decision problem, called TRANSITION (k, m, n, T).

Problem 2: TRANSITION (k, m, n, T).
Input: Payoff matrices $R_{m \times n}$ and $C_{m \times n}$. The initial equilibrium $(r(1), \mathcal{C}(1))$ and the target equilibrium $(r^*, \mathcal{C}^*)$,
Output: YES, if a transformation path from $(r(1), \mathcal{C}(1))$ to $(r^*, \mathcal{C}^*)$ with the total cost no larger than T exists, and NO otherwise.

Complexity Results

This section investigates the complexity of the above two problems. We show that OPT TRANSITION (k, m, n) is APX-hard, which discourages us from designing efficient algorithms that can find a solution within some fixed multiplicative factor of the optimal cost T. In addition, even for the case that the row player has only two strategies, the decision problem TRANSITION $(k, 2, n, T)$ is NP-complete. However, OPT TRANSITION (k, m, n) becomes polynomial-time solvable when either k or n is a fixed constant.

General Values of k, m, and n

We show that the decision problem TRANSITION (k, m, n, T) is NP-complete when $T = 0$.

Theorem 1. *TRANSITION $(k, m, n, 0)$ is NP-complete.*

Proof. Given a transition path from the initial equilibrium to the target equilibrium, it is easy to verify whether it is a valid path and its cost is 0. For NP-hardness, we reduce from the EXACT COVER problem, which is known to be NP-complete (Karp 1972), to the decision problem TRANSITION $(k, m, n, 0)$. Recall the EXACT COVER problem:

Problem 3: EXACT COVER (s, w).
Input: A finite set $Z = \{1, 2, \ldots, 3s\}$ and a collection $X = \{X_1, X_2, \ldots, X_w\}$ of 3-element subsets of the set Z, where $s \le w$.
Output: YES, if there exists a collection $\{X_{i_1}, X_{i_2}, \ldots, X_{i_s}\} \subset X$ such that their union is Z, and NO otherwise.

Given an instance of EXACT COVER, we construct an instance of TRANSITION $(k, m, n, 0)$ as follows:

Let $k = s$. That is, there are s column players. Let $m = 3s + 2$ and $n = w + 2$. We design the column players' payoff matrix $C_{(3s+2)\times(w+2)}$ as below. Except for the last row and the last column, all elements in matrix C are 1. The intersection of the last row and last column is set to be 1 as well. Lastly, all the remaining elements are 0.

Column Players' Payoff Matrix C

	1	$\cdots$	$w+1$	$w+2$
1	1	$\cdots$	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$3s+1$	1	$\cdots$	1	0
$3s+2$	0	$\cdots$	0	1

Let $v_1, v_2, \ldots, v_w$ denote the characteristic vectors of $X_1, X_2, \ldots, X_w$. That is, for a vector $v_i = (v_{i,j})_{j=1,\ldots,3s}$, its elements $v_{i,j} = 1$ if $j \in X_i$ and $v_{i,j} = 0$ if $j \notin X_i$, $i = 1, \ldots, w$. We construct the row player's payoff matrix $R_{(3s+2)\times(w+2)}$ as follows. The entries in the first row are 0 except the last one being -1. The entries in the first column are 0 except the last one being $-s$. Then, for the last row, $R_{(3s+2),j} = 1/s, j = 2, \ldots, w+1$ and $R_{(3s+2),(w+2)} = 1$. For the last column, $R_{i,(w+2)} = 0, i = 2, \ldots, 3s+1$. The other entries of matrix R are filled by elements $v_{i,j}$ as shown below.

Row Player's Payoff Matrix R

	1	2	$\cdots$	$w+1$	$w+2$
1	0	0	$\cdots$	0	-1
2	0	$v_{1,1}$	$\cdots$	$v_{w,1}$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$3s+1$	0	$v_{1,3s}$	$\cdots$	$v_{w,3s}$	0
$3s+2$	$-s$	$1/s$	$\cdots$	$1/s$	1

We note that there are at least two pure Nash equilibria in the $(k+1)$-player game. They are $(r^1, c^1, \ldots, c^1)$ and $(r^{3s+2}, c^{w+2}, \ldots, c^{w+2})$; namely, all players choose their first strategy and all players choose their last strategy, respectively. In particular, let the initial Nash equilibrium $(r(1), \mathcal{C}(1))$ be $(r^1, c^1, \ldots, c^1)$ and the target equilibrium $(r^*, \mathcal{C}^*)$ be $(r^{3s+2}, c^{w+2}, \ldots, c^{w+2})$. Till now, we have constructed an instance of TRANSITION $(k = s, m = 3s+2, n = w+2, 0)$.

Reduction correctness. Given that there is a solution to TRANSITION $(s, 3s+2, w+2, 0)$, we can compute a solution to EXACT COVER (s, w). To this end, we disclose the features of a zero-cost transformation path of TRANSITION $(s, 3s+2, w+2, 0)$.

First, we observe from payoff matrix C, that a column player will not choose to play its last strategy c^{w+2} unless the row player has chosen its last strategy r^{3s+2}. Based on this, we observe from payoff matrix R, that the row player will not choose to play its last strategy r^{3s+2} as long as there is a single column player playing its first strategy c^1. This is because, in that case, the row player's payoff is at most $-s + \frac{s-1}{s}$ which is less than 0. In view of these observations, we conclude that the s column players must be playing some strategies within the set $\{c^2, \ldots, c^{w+1}\}$ on a zero-cost transformation path from the initial equilibrium $(r^1, c^1, \ldots, c^1)$ to the target equilibrium $(r^{3s+2}, c^{w+2}, \ldots, c^{w+2})$.

Second, for the second-last step on the transition path, we notice that these s column players must choose s different strategies amongst the set $\{c^2, \ldots, c^{w+1}\}$. Otherwise, suppose that two column players are choosing the same strategy, for example, c^2. Note that the characteristic vector $v_1 = (v_{1,j})_{j=1,\ldots,3s}$ has exactly three elements whose value are 1. Without loss of generality, assume $v_{1,1} = 1$. Then, by playing strategy r^2, the row player's payoff is 2, which is greater than its payoff 1 by playing strategy r^{3s+2}. Therefore, it contradicts the existence of a zero-cost transformation path on which the row player will transit to playing the last strategy.

Last, we notice that in these s columns of payoff matrix R that correspond to the s different strategies played by these column players, there should be only one 1-element in each row. Otherwise, by playing the strategy that corresponds to a row that has multiple 1-elements, the row player has a higher payoff and will not transit to playing the last strategy without a positive reward. The same contradiction occurs. So, the row player is indifferent between playing the last strategy r^{3s+2} and any one of the strategies in $\{r^2, \ldots, r^{3s+1}\}$, as its payoff is 1 in either case. Together with the fact that there are $3s$ rows (namely, row 2 to row $3s+1$) and each of these s columns has three 1-elements, we conclude that these s columns correspond to s characteristic vectors of 3-elements sets such that their union is the set $Z = \{1, 2, \ldots, 3s\}$.

To conclude, given a zero-cost transition path from the initial equilibrium to the target equilibrium, we can identify a set of values $v_{i,j}$ such that their corresponding 3-elements subsets $\{X_{i_1}, X_{i_2}, \ldots, X_{i_s}\}$ is a solution to the SET COVER problem. Also, if the SET COVER problem has a solution $\{X_{i_1}, \ldots, X_{i_s}\}$ such that their union is Z, then we can construct the transition path as $(r^1, c^1, \ldots, c^1) \to (r^1, C^{i_1}, C^{i_2}, \ldots, C^{i_s}) \to (r^{3s+2}, C^{i_1}, C^{i_2}, \ldots, C^{i_s}) \to (r^{3s+2}, C^{w+2}, C^{w+2}, \ldots, C^{w+2})$. It is easy to verify that the cost of this transition path is zero.

Hence, TRANSITION $(k, m, n, 0)$ is NP-complete. □

This result immediately implies that the problem OPT TRANSITION (k, m, n) is APX-hard.

Corollary 1. *Computing the optimal transformation path is APX-Hard and no multiplicative approximation is possible. That is, constructing a transformation path of cost $\alpha \cdot \mathrm{OPT}$ is NP-Hard for any $\alpha > 1$, where* OPT *is the cost of the optimal transformation path.*

When One of the Variables, k, m, n, Is a Constant.

Since we have shown that TRANSITION (k, m, n, T) is APX-hard, we then consider special cases when one of the variables, k, m, n, is a constant.

When $m = 2$**.** The following theorem shows that the problem TRANSITION (k, m, n, T) is hard to solve even if the row player has only two strategies.

Theorem 2. *TRANSITION $(k, 2, n, T)$ is NP-complete.*

The NP-hardness of the problem is shown by reducing from a variant of the Knapsack problem.

When k, the Number of Column Players, Is a Fixed Constant.

We show that the problem OPT TRANSITION (k, m, n) is polynomial-time solvable. This is done by reducing the problem of finding an optimal transformation path to the problem of finding the shortest path in a complete directed graph $G(V, E)$.

Given an instance of the problem OPT TRANSITION (k, m, n), the vertex $v \in V$ of graph G is a strategy profile $(r^v, \mathcal{C}^v)$ of the bi-matrix games, where $r^v \in \mathscr{R}$ and $\mathcal{C}^v = (c_1^v, \ldots, c_k^v), c_i^v \in \mathscr{C}$. For a directed edge $e_{vu} \in E$ from vertex v to vertex u, its weight w_{vu} is defined to be the reward needed to incentivize the players moving from strategy profile $(r^v, \mathcal{C}^v)$ to $(r^u, \mathcal{C}^u)$, which is $T_{\mathcal{C}^v}(r^u) + T_{r^v}(\mathcal{C}^u)$.

Let v^1 and v^* be the vertices corresponding to the initial equilibrium $(r(1), \mathcal{C}(1))$ and the target equilibrium $(r^*, \mathcal{C}^*)$, respectively. The shortest path from v^1 to v^* then corresponds to the optimal transformation path. Since G is a directed graph with non-negative edge weights, and the shortest path problem can be solved in $O(|E| + |V| \log\log |V|)$ time (Thorup 1999), we have the following result.

Theorem 3. *The optimal reward scheme can be computed in time $O(m^2 n^{2k})$. That is, the problem OPT TRANSITION (k, m, n) is slicewise polynomial with respect to k.*

Proof. Since each vertex of G corresponds to a strategy profile, the number of vertices $|V| = m \cdot n^k$. Given that G is a complete graph, the number of edges $|E|$ is $\Theta(|V|^2) = \Theta(m^2 n^{2k})$. Thus, the shortest path can be computed in $O(m^2 n^{2k})$ time. After getting the shortest path, we then construct the optimal transformation path by setting the strategy profile in round t to be the $(r^{v_t}, \mathcal{C}^{v_t})$, where v_t is the t-th vertex on the shortest path. The construction based on the shortest path takes $O(|V|)$ time. As a result, the theorem is proven. □

When n, the Number of Column Player Strategies, Is a Fixed Constant.

We show that the problem is slicewise polynomial.

Theorem 4. *When the number of a column player's strategies n is a fixed constant, we can compute the optimal reward scheme in time $O(m^2 k^{2n})$.*

Proof. Given the fact that the strategy space and utility of the k column players are the same, the k bimatrix games are identical. So, to incentivize the column players' transit to strategy profiles $\mathcal{C}^j$ and $\mathcal{C}^{j'}$ when the row player's strategy is r^i, respectively, the rewards $T_{r^i}(\mathcal{C}^j)$ and $T_{r^i}(\mathcal{C}^{j'})$ are the same, as long as the number of column players choosing each strategy in $\mathscr{C}$ is the same under both $\mathcal{C}^j$ and $\mathcal{C}^{j'}$.

Therefore, we can merge the vertices of the graph G that corresponds to $(r^i, \mathcal{C}^j)$ and $(r^i, \mathcal{C}^{j'})$ to form a new vertex. Since there are k column players and n different strategies, there are $\binom{n+k-1}{k-1}$ different ways for column players to choose strategies, which means the number of vertices $|V| = m\binom{n+k-1}{k-1} = O(mk^n)$ in the constructed graph. So, the number of edges $|E|$ is $\Theta(|V|^2) = \Theta(m^2k^{2n})$. □

Approximation Results

In light of the APX-Hardness of the problem OPT TRANSITION (k, m, n), in this section, we strive to design efficient algorithms that can find a solution within an additive factor of the optimal reward T. To this end, we define *alternating path* and use it to design an approximation algorithm.

Alternating Path

First, we define an alternating path as follows.

Definition 1. *The vertices of an alternating path are either a row player strategy $r(t)$ or a column players' strategy profile $\mathcal{C}(t)$. These vertices alternatingly appear on an alternating path as time epoch t varies.*

The edges of an alternating path are directed and weighted. The weight of edge $(r(t), \mathcal{C}(t+1))$ is $T_{r(t)}(\mathcal{C}(t+1))$. Namely, the weight is the reward needed to incentivize column players taking up strategy profile $\mathcal{C}(t+1)$ given that the row player's strategy is $r(t)$. Similarly, the weight of edge $(\mathcal{C}(t), r(t+1))$ is $T_{\mathcal{C}(t)}(r(t+1))$. The cost, $Cost(P)$, of an alternating path P is the sum of all its edges' weights. Then, given a transformation path from the initial equilibrium to the target equilibrium, we can derive two alternating paths with the same length as the transformation path as Figure 1 demonstrates. Notably, the total cost of these two alternating paths is equal to the cost, $T = \sum_t \{T_{r(t)}(\mathcal{C}(t+1)) + T_{\mathcal{C}(t)}(r(t+1))\}$, of the transformation path. As such, we derive the following lemma straightforwardly.

Algorithm 1: Transformation Path Construction

Input: An alternating path P with cost $Cost(P)$: $r(1) \to \cdots \to r(l) \to \mathcal{C}^*$
Output: A transformation path of length $l+1$ with cost $2 \cdot Cost(P)$
the 1st vertex $\leftarrow (r(1), \mathcal{C}(1))$;
the 2nd vertex $\leftarrow (r(1), \mathcal{C}(2))$;
while $3 \le t \le l$ **do**
  when t is odd, the t-th vertex $\leftarrow (r(t), \mathcal{C}(t-1))$;
  when t is even, the t-th vertex $\leftarrow (r(t-1), \mathcal{C}(t))$
end
the $(l+1)$-th vertex $\leftarrow (r(l), \mathcal{C}^*)$;
the $(l+2)$-th vertex $\leftarrow (r^*, \mathcal{C}^*)$

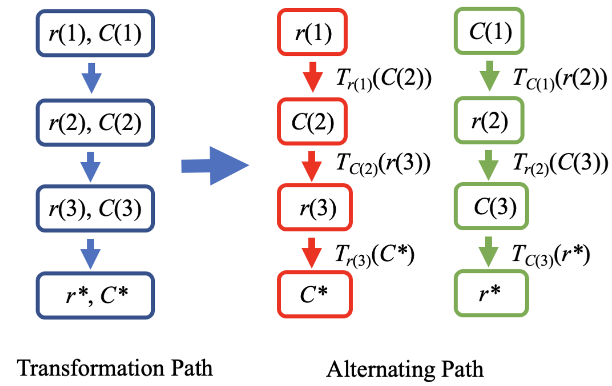

Figure 1: Two alternating paths decomposed from a transformation path.

Lemma 1. *Given the optimal transformation path from the initial equilibrium $(r(1), \mathcal{C}(1))$ to the target equilibrium $(r^*, \mathcal{C}^*)$, and the two alternating paths decomposed from this transformation path, the cost of the transformation path is at least two times the cost of the alternating path whichever is smaller.*

Now, fixing an initial equilibrium $(r(1), \mathcal{C}(1))$ and a target equilibrium $(r^*, \mathcal{C}^*)$, let us consider all possible alternating paths connecting $r(1)$ or $\mathcal{C}(1)$ and r^* or $\mathcal{C}^*$, respectively. Without loss of generality, assume the alternating path $r(1) \to \cdots \to r(l) \to \mathcal{C}^*$ has the smallest cost amongst all these alternating paths. Then, we can use Algorithm 1 to construct a transformation path from $(r(1), \mathcal{C}(1))$ to $(r^*, \mathcal{C}^*)$ whose cost is twice of the cost of the alternating path $r(1) \to \cdots \to r(l) \to \mathcal{C}^*$.

In the proof of the following theorem, we will present a constructive procedure for the optimal transformation path.

Theorem 5. *The cost of the transformation path constructed by Algorithm 1 is two times the cost of the alternating path $r(1) \to \cdots \to r(l) \to \mathcal{C}^*$. Moreover, the length of the output transformation path is one longer than the input alternating path.*

Directly following Lemma 1 and Theorem 5, we have a corollary as follows.

Corollary 2. *The cost of the optimal transformation path from the initial equilibrium $(r(1), \mathcal{C}(1))$ to the target equilibrium $(r^*, \mathcal{C}^*)$ is exactly two times the smallest cost of an alternating path connecting $r(1)$ or $\mathcal{C}(1)$ and r^* or $\mathcal{C}^*$.*

Moreover, if the alternating path with the smallest cost connecting $r(1)$ or $\mathcal{C}(1)$ and r^* or $\mathcal{C}^*$ is known, then we can construct the optimal transformation path from the initial equilibrium $(r(1), \mathcal{C}(1))$ to the target equilibrium $(r^*, \mathcal{C}^*)$. In addition, Algorithm 1 also allows us to upper bound the length of an optimal transformation path, i.e., the total number of rounds needed to move from the initial equilibrium to the target equilibrium.

Theorem 6. *Given any initial equilibrium $(r(1), \mathcal{C}(1))$ and target equilibrium $(r^*, \mathcal{C}^*)$, there exists an optimal transformation path whose length is at most $2m - 1$.*

Approximation Algorithm

In this subsection, we use the properties of the alternating path established above to design an approximation algorithm. We will construct a complete directed graph in which the vertices are row player strategies, and the weight of an edge is the cost of a length-2 alternating path connecting two adjacent vertices. Although the exact minimum cost of these length-2 alternating paths is hard to compute, as otherwise, the problem OPT TRANSITION (k, m, n) will become tractable following our construction, we approximate these costs by rounding the solutions of integer linear programmings. Building upon this complete directed graph, we can assemble an alternating path with nearly-optimal cost between any two row-player strategies. We can extend this approach by an additional handling technique to assemble a minimum-cost alternating path between any row or column player strategies. This way, we can approximate the cost of an alternating path connecting $r(1)$ or $\mathcal{C}(1)$ and r^* or $\mathcal{C}^*$. Building upon the alternating path with the smallest cost among these four paths, we can erect the optimal transformation path between the initial and target equilibria.

We start by constructing a weighted directed complete graph $G(V, E)$. In graph G, each vertex v_i corresponds to a strategy r^i of the row player. From each vertex v_i to a different vertex v_j, there is an edge e_{ij}. The weight w_{ij} of edge e_{ij} is the cost of an alternating path $r^i \to \mathcal{C}_{ij} \to r^j$, where $\mathcal{C}_{ij} \in \mathscr{C}^k$ is a column players' strategy profile. Clearly, the smaller w_{ij} is, the better we can approximate the solution to the problem OPT TRANSITION (k, m, n). In Algorithm 2, we employ a sub-routine **WEIGHT(r^i, r^j)** to compute w_{ij} and the corresponding column players' strategies $\mathcal{C}_{ij}$. Now that we have finished the construction of the graph G, we can compute the shortest path between any two vertices v_i and v_j. In the meantime, we insert the corresponding column player strategy profiles $\mathcal{C}_{ij}$ into every adjacent vertex v_i and v_j, to construct an alternating path (of the equilibrium transition problem) from the shortest path (of graph G). Clearly, the shortest path $P_{r^i r^j}$ between v_i and v_j corresponds to an approximately optimal alternating path from r^i to r^j. If both s_1 and s_2 are row player strategies, then no additional treatment on the shortest path is needed. So Algorithm 2 returns $P' = P_{s_1 s_2}$; otherwise, to obtain an approximately optimal alternating path from s_1 to s_2, we need to make an additional comparison to append one more column player strategy in front and at the end of the shortest path, so that the cost of the output alternating path is as small as possible.

In the following, we present an integer linear programming formulation used to approximate the cost of a length-2 alternating path in the sub-routine WEIGHT(r^i, r^j).

Sub-routine WEIGHT(r^i, r^j). Denote x_q the number of column players who play the strategy $c^q \in \mathscr{C}$. Then the column strategy profile $\mathcal{C}_{ij} \in \mathscr{C}^k$ is determined given a tuple $(x_1, \ldots, x_n)$. The hurdle to approximate $T_{r^i}(\mathcal{C}_{ij}) + T_{\mathcal{C}_{ij}}(r^j)$ is that $T_{\mathcal{C}_{ij}}(r^j)$ contains the max operator. To detour it, we introduce m integer linear programming (ILPs). For the z-th ILP, we eliminate the operator max by introducing a constraint that r^z is row player's best response strategy with respect to $\mathcal{C}_{ij}$. By allowing x_q to be a positive number, we further relax these ILPs to linear programmings (LPs). The z-th linear programming is shown as follows.

Algorithm 2: Construction of an Approximately Optimal Alternating Path $s_1 \to \cdots \to s_2$

```
Input: Strategy sets 𝓡 and 𝓒^k;
Payoff matrices R_{m×n} and C_{m×n};
Initial equilibrium (r(1), C(1)) and target equilibrium (r*, C*);
s_1 ∈ {r(1), C(1)}, s_2 ∈ {r*, C*}.
Output: An alternating path from s_1 to s_2
Initialize a complete directed graph G(V, E);
for i from 1 to m do
    v_i ← r^i;
    for j from 1 to m do
        e_ij ← an edge from v_i to v_j;
        (w_ij, C_ij) ← WEIGHT(r^i, r^j);
    end
end
for i from 1 to m do
    for j from 1 to m do
        P_{r^i r^j} ← the shortest path between v_i and v_j;
        for each edge e_pq along P_{r^i r^j} do
            Insert C_pq between v_p and v_q;
        end
    end
end
if s_1, s_2 ∈ 𝓡 then
    P' ← P_{s_1 s_2};
end
if s_1 ∈ 𝓡 and s_2 ∈ 𝓒^k then
    P' ← arg min_{P_{s_1 q}} (Cost(P_{s_1 q}) + T_q(s_2));
    Append s_2 to the end of P';
end
if s_1 ∈ 𝓒^k and s_2 ∈ 𝓡 then
    P' ← arg min_{P_{q s_2}} (Cost(P_{q s_2}) + T_{s_1}(q));
    Insert s_1 in the front of P';
end
if s_1, s_2 ∈ 𝓒^k then
    P' ←
      arg min_{P_pq} (T_{s_1}(p) + Cost(P_pq) + T_q(s_2));
    Append s_2 to the end of P';
    Insert s_1 in the front of P';
end
return P'
```

$$
\begin{aligned}
\min \quad & T_{r^i}(\mathcal{C}_{ij}) + \sum_{1 \le q \le n} x_q R(r^z, c^q) - \sum_{1 \le q \le n} x_q R(r^j, c^q) \\
\text{s.t.} \quad & \sum_{1 \le q \le n} x_q R(r^z, c^q) \ge \sum_{1 \le q \le n} x_q R(r^{z'}, c^q), \quad \forall z' \in [m] \\
& x_1 + \cdots + x_n = k \\
& x_q \ge 0, \quad \forall q \in [n]
\end{aligned}
$$

The objective function is the cost, $T_{r^i}(\mathcal{C}_{ij}) + T_{\mathcal{C}_{ij}}(r^j)$, of the length-2 alternating path $r^i \to \mathcal{C}_{ij} \to r^j$. It is linear in the variables x_q's. The first constraint ensures that r^z is the row player's best response strategy to column player strategy profile $\mathcal{C}_{ij}$. The remaining constraints state that the total number of column players equals precisely k and x_q is non-negative. After solving these m LPs, we take the one that gives the minimum value w^*_{ij} of the objective functions. Without loss of generality, assume that it is the z-th LP that achieves the minimum value. Hence, the objective function can be rewritten as $k \max_p C(r^i, c^p) - \sum_{1\le q\le n} x_q f(c^q)$, where $f(c^q)$ is $C(r^i, c^q) - R(r^z, c^q) + R(r^j, c^q)$. We can retrieve a strategy profile $\mathcal{C}_{ij}$ given the solution $(x_1, \dots, x_n)$. Albeit the solutions x_q's of the LPs are not necessarily integers, for now, we can interpret them as a mixed-strategy profile.

To round the fractional solution to an integral solution, we first round each x_q down to the nearest integer $\lfloor x_q \rfloor$, $q = 1, \dots, n$. Let $u = \arg\min_q \max_p R(r^p, c^q)$, we then increase x_u by $k - \sum_{1\le q\le n} \lfloor x_q \rfloor$. Till now, we have derived an integer solution $(\lfloor x_1 \rfloor, \dots, \lfloor x_n \rfloor)$, from which we can retrieve a valid strategy profile $\mathcal{C}_{ij}$. Based on $\mathcal{C}_{ij}$, we obtain an alternating path $r^i \to \mathcal{C}_{ij} \to r^j$ with cost w_{ij}.

Next, we bound the difference between the optimal objective value of the LPs and the transition cost w_{ij} that we derived from the alternating path.

Lemma 2. *$w_{ij} - w^*_{ij} \le 2\|R\|_{1,1} + \|C\|_{1,1}$, where $\|R\|_{1,1}$ and $\|C\|_{1,1}$ are the sum of all entries of R and C, respectively.*

Theorem 7. *Algorithm 2 returns an alternating path between s_1 and s_2 whose cost is at most $m(2\|R\|_{1,1} + \|C\|_{1,1})$ more than the cost of the optimal alternating path between s_1 and s_2.*

Finally, we bound the additive approximation error.

Theorem 8. *Denote* OPT *the cost of the optimal transformation path from $(r(1), \mathcal{C}(1))$ to $(r^*, \mathcal{C}^*)$. We can implement Algorithm 2 to find an approximately optimal alternating path and then Algorithm 1 to construct a transformation path with a cost* $\mathrm{OPT} + 2m(2\|R\|_{1,1} + \|C\|_{1,1})$ *in polynomial time.*

We note that the advantage of this additive approximation is that it is independent of the number of column players k. In many practical scenarios such as increasing the uptake of electric vehicles and retail store businesses, k is the number of drivers/customers which is significantly larger than the number of player strategies.

Tractable Cases

In this section, we turn our attention to the cases in which the optimal transformation path can be found in polynomial time. Recall the example in the Introduction: in the game, the row player is a service provider and the column players are the customers. We assume that all possible locations are distributed on a straight line. That is, the players' strategy space $\mathscr{R} = \mathscr{C}$. In particular, the row player is limited to choosing a location which is one of the column player strategies. The service provider seeks to minimize the sum of the Euclidean distance between its location and the locations of customers. A customer's payoff is negatively correlated with its distance to the service provider. With a slight abuse of notations, we denote r^i, c^j the axis of the locations on the line. W.l.o.g., we assume they are both sorted in increasing orders.

Definition 2. *In this one-dimensional domain, a player's payoff is single-peaked if it is maximized at a single point on the line, and the further its location to this point, the less its payoff. A player's payoff is exact-distance if it is equal to the negative value of the difference between its location to a single point on the line.*

In particular, we are interested in the case in which the row player's payoff is exact-distance and the column players' payoff is single-peaked. That is, given the strategy profile $(r^i, c_1^{j_1}, \dots, c_k^{j_k})$, $j_l \in [m]$, $l = 1, \dots, k$, the row player's payoff is $\sum_{l=1}^k R(r^i, c_l^{j_l}) = -\sum_{l=1}^k |r^i - c_l^{j_l}|$, and the column player l's payoff is $C(r^i, c_l^{j_l}) = g(|r^i - c_l^{j_l}|)$, where $g(\cdot)$ is monotone decreasing and $g(0) = 0$. In this case, we show that the problem is tractable.

Theorem 9. *When the row player's payoff is exact-distance and the column players' payoff is single-peaked, the optimal transformation path from an initial equilibrium to a target equilibrium can be found in polynomial time.*

Conclusions and Future Work

In this paper, we formulated an optimization problem to find the optimal transformation path from an initial equilibrium to a more desirable one. In this game, players move simultaneously in each round. We presented comprehensive complexity analyses of the problem. We proved that the problem is APX-hard when the number of the row player strategies, the number of the column player strategies, and k are input sizes. Furthermore, we showed that the problem is slicewise polynomial with respect to k and n, respectively. As for the parameter m, we proved that the problem is NP-Hard even when $m = 2$. Besides the hardness result, we designed a polynomial-time approximation algorithm with bounded additive error. Moreover, the approximation error is independent of the number of column players k. Finally, we considered cases where we can find the optimal transformation path in polynomial time.

It is important to acknowledge that the problem we have analyzed presents intractability in various setups, making any generalizations a formidable task. Viable solutions would likely require the application of additional restrictions or careful consideration of special cases. However, the realm of possibilities for further exploration is vast and captivating. Firstly, delving into the extension of our study to encompass scenarios where column players possess diverse strategies and payoff matrices holds great interest. Secondly, exploring the implications of incorporating randomized strategies and how they might alter the overall dynamics of the game poses an intriguing challenge. Thirdly, developing efficient algorithms for other cases, such as utilizing alternative distance measures to define player strategies corresponding to locations on general graphs and their associated payoffs, presents an enticing avenue for investigation. Additionally, broadening the scope of the study to incorporate multiple row players opens up exciting possibilities for research and discovery.

Acknowledgments

Zihe Wang was partially supported by the National Natural Science Foundation of China (Grant No. 62172422). Jie Zhang was partially supported by a Leverhulme Trust Research Project Grant (2021 – 2024) and the EPSRC grant (EP/W014912/1).

References

Amiet, B.; Collevecchio, A.; Scarsini, M.; and Zhong, Z. 2021. Pure Nash Equilibria and Best-Response Dynamics in Random Games. *Math. Oper. Res.*, 46(4): 1552–1572.

Anshelevich, E.; Dasgupta, A.; Kleinberg, J. M.; Tardos, É.; Wexler, T.; and Roughgarden, T. 2008. The Price of Stability for Network Design with Fair Cost Allocation. *SIAM J. Comput.*, 38(4): 1602–1623.

Augustine, J.; Caragiannis, I.; Fanelli, A.; and Kalaitzis, C. 2015. Enforcing Efficient Equilibria in Network Design Games via Subsidies. *Algorithmica*, 72(1): 44–82.

Deng, Y.; and Conitzer, V. 2017. Disarmament Games. In Singh, S. P.; and Markovitch, S., eds., *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*, 473–479. AAAI Press.

Deng, Y.; and Conitzer, V. 2018. Disarmament Games With Resource. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18)*, 981–988. AAAI Press.

Deng, Y.; Tang, P.; and Zheng, S. 2016. Complexity and Algorithms of K-implementation. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, Singapore, May 9-13, 2016*, 5–13. ACM.

Eidenbenz, R.; Oswald, Y. A.; Schmid, S.; and Wattenhofer, R. 2007. Mechanism Design by Creditability. In *Combinatorial Optimization and Applications, First International Conference, COCOA 2007*, volume 4616 of *Lecture Notes in Computer Science*, 208–219. Springer.

Feldman, M.; Snappir, Y.; and Tamir, T. 2017. The Efficiency of Best-Response Dynamics. In Bilò, V.; and Flammini, M., eds., *Algorithmic Game Theory - 10th International Symposium, SAGT 2017, L'Aquila, Italy, September 12-14, 2017, Proceedings*, volume 10504 of *Lecture Notes in Computer Science*, 186–198. Springer.

Heinrich, T.; Jang, Y.; Mungo, L.; Pangallo, M.; Scott, A.; Tarbush, B.; and Wiese, S. 2022. Best-response dynamics, playing sequences, and convergence to equilibrium in random games. arXiv:2101.04222.

Hopkins, E. 1999. A note on best response dynamics. *Games and Economic Behavior*, 29(1-2): 138–150.

Huang, H.; Wang, Z.; Wei, Z.; and Zhang, J. 2023. Cost Minimization for Equilibrium Transition. arXiv:2312.07603.

Karp, R. M. 1972. Reducibility Among Combinatorial Problems. In Miller, R. E.; and Thatcher, J. W., eds., *Proceedings of a symposium on the Complexity of Computer Computations, held March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, USA*, The IBM Research Symposia Series, 85–103. Plenum Press, New York.

Koutsoupias, E.; and Papadimitriou, C. H. 2009. Worst-case equilibria. *Comput. Sci. Rev.*, 3(2): 65–69.

Leslie, D. S.; Perkins, S.; and Xu, Z. 2020. Best-response dynamics in zero-sum stochastic games. *Journal of Economic Theory*, 189: 105095.

Monderer, D.; and Tennenholtz, M. 2004. K-Implementation. *J. Artif. Intell. Res.*, 21: 37–62.

Monderer, D.; and Tennenholtz, M. 2009. Strong mediated equilibrium. *Artif. Intell.*, 173(1): 180–195.

Nisan, N.; and Ronen, A. 1999. Algorithmic Mechanism Design (Extended Abstract). In Vitter, J. S.; Larmore, L. L.; and Leighton, F. T., eds., *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, May 1-4, 1999, Atlanta, Georgia, USA*, 129–140. ACM.

Papadimitriou, C. H. 2011. Games, algorithms, and the Internet. In Srinivasan, S.; Ramamritham, K.; Kumar, A.; Ravindra, M. P.; Bertino, E.; and Kumar, R., eds., *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, March 28 - April 1, 2011*, 5–6. ACM.

Rozenfeld, O.; and Tennenholtz, M. 2007. Routing Mediators. In *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 1488–1493.

Thorup, M. 1999. Undirected single-source shortest paths with positive integer weights in linear time. *Journal of the ACM (JACM)*, 46(3): 362–394.