Multilinear Formulations for Computing a Nash Equilibrium of Multi-Player Games

Miriam Fischer
Department of Computing, Imperial College London, UK

Akshay Gupte[1]
School of Mathematics, The University of Edinburgh, UK

Abstract

We present multilinear and mixed-integer multilinear programs to find a Nash equilibrium in multi-player noncooperative games. We compare the formulations to common algorithms in Gambit, and conclude that a multilinear feasibility program finds a Nash equilibrium faster than any of the methods we compare it to, including the quantal response equilibrium method, which is recommended for large games. Hence, the multilinear feasibility program is an alternative method to find a Nash equilibrium in multi-player games, and outperforms many common algorithms. The mixed-integer formulations are generalisations of known mixed-integer programs for two-player games, however unlike two-player games, these mixed-integer programs do not give better performance than existing algorithms.

2012 ACM Subject Classification Theory of computation → Exact and approximate computation of equilibria; Theory of computation → Nonconvex optimization

Keywords and phrases Noncooperative n-person games, Nash equilibrium, Multilinear functions, Nonconvex problems, Mixed-integer optimization

Digital Object Identifier 10.4230/LIPIcs.SEA.2023.12

Supplementary Material
Dataset: https://github.com/economicsandcomputing/MultilinearNashEquilibria

Funding *Miriam Fischer*: Supported by a PhD scholarship from DeepMind.

Acknowledgements This research was initiated as part of the MSc dissertation of the first author in the School of Mathematics at the University of Edinburgh.

1 Introduction

A noncooperative game has n players, where $n \geq 2$ is finite, with each player having finitely many pure strategies which they do not discuss or reveal to each other. A mixed strategy for a player is a probability distribution over the player's pure strategies. Each player has a known payoff function which maps any combination of pure strategies of all the n players to a real number. Mixed strategies of all the players form a tuple whose payoff is calculated by taking expectation over the probability distributions. In his seminal work [11], Nash showed that every such game has a tuple of mixed strategies that is an equilibrium in the sense that no player increases their payoff if they were to change their mixed strategy while the others keep theirs fixed. Although existence of Nash equilibrium is guaranteed, uniqueness does not always hold, and there are also characterisations of when there exists an equilibrium formed solely by pure strategies.

[1] Corresponding author

This paper deals with the question of algorithmic and numerical computation of Nash equilibria. From a complexity perspective, computing an equilibrium was only somewhat recently formally settled to being PPAD-complete [1, 2] even for two-player games. There is a lot of literature for two-player bimatrix games, and the most well-known and established exact method to compute an equilibrium is the Lemke-Howson algorithm [8]. This gives a very good computational performance on many instances in practice, although its worst-case performance can take exponentially many pivoting steps [17].

However, for multi-player games, there do not seem to be commonly established approach for computing the equilibrium. Although there is a generalisation of the Lemke-Howson method to n-person games [14, 21], popular algorithmic approaches include a global Newton method [6], an iterated polymatrix approximation approach [7], a simplicial subdivision method [20], a simple search algorithm aiming to find an equilibrium with small support size [13], and a quantal response equilibrium method which gives an approximation to a Nash equilibrium [19]. Many of these methods are implemented in the game-theoretic library `Gambit` [10]. Experiments comparing different methods have been undertaken [16, 13], however it is rather unclear which of the methods is best for multi-player games. For example, the global Newton method gives solid performance for small games, however does not to scale well to larger games [5, 18]. Support enumeration algorithms are fast for games with pure equilibria but will be much slower for a game that only has equilibria of medium to large support size. There are also approximation algorithms, which tend to approximate a Nash equilibrium for large games [19, 5, 3].

We adopt the optimization approach, and propose different optimization formulations for computing a Nash equilibrium for n-person games for $n \geq 2$. Particularly, we present a multilinear polynomial continuous feasibility program of degree n (= number of players), which is a generalisation of the bilinear optimization problem for 2 players [9]. Further, we extend the two-player mixed-integer formulations of [16] to multi-player games, and give a large variety of mixed-integer formulations to find a Nash equilibrium in multi-player games. All our formulations find a Nash equilibrium of a $n \geq 2$ player game. We compare our programs to `gambit-gnm` (global Newton), `gambit-simpdiv` (simplicial subdivision), `gambit-logit` (quantal response equilibrium) algorithms in `Gambit`, focusing on random games and covariant games with negative covariance. We find that the mixed-integer formulations do not give better performance than existing algorithms, and our analysis of those is aimed to get an understanding of which mixed-integer formulations are most suited for finding a Nash equilibrium. We find that our multilinear continuous feasibility program is faster than all the methods in Gambit we compare it to, including the `gambit-logit` method, which is so far recommended for large games. Thus, we provide an alternative approach to computing Nash equilibrium in multi-player games.

The next section presents our continuous and mixed-integer multilinear optimization formulations. For each of them, their correctness, i.e., the fact that their optimal/feasible solutions correspond to Nash equilibria of the game, is proved in the Appendix.

2 Formulations

The multi-player multilinear formulation is an extension of a bilinear formulation for bimatrix games [9]. To motivate the multilinear formulation, we shortly recall the bilinear program that is equivalent to finding a Nash equilibrium in a bimatrix game. To do so, we introduce some notation. Let $A, B \in \mathbb{R}^{m \times n}$ be the payoff matrix of player 1 and player 2, with m pure strategies of player 1 and n pure strategies of player 2. Let $\boldsymbol{x} \in \mathbb{R}^m$ with $\boldsymbol{x} \geq 0$ and

$\sum_{i=1}^{m} x_i = 1$ be a (possibly mixed) strategy of player 1, with x_s being the probability placed on pure strategy s. Let $\boldsymbol{y} \in \mathbb{R}^n$ with $\boldsymbol{y} \geq 0$ and $\sum_{j=1}^{n} y_j = 1$ be a (possibly mixed) strategy of player 2. Let $\mathbf{1}_n$ and $\mathbf{1}_m$ denote vectors of all ones of dimension n and m. Any globally *optimal* solution (x, y, p, q) to the bilinear optimization problem in BLP is equivalent to a Nash equilibrium in a bimatrix game.

$$\text{(BLP)} \qquad \begin{aligned} \max_{x,y,p,q} \quad & x^\top Ay + x^\top By - p - q && \text{(1a)} \\ \text{s.t.} \quad & Ay \leq p\mathbf{1}_m, \quad B^\top x \leq q\mathbf{1}_n && \text{(1b)} \\ & \sum_{i=1}^{m} x_i = 1, \quad \sum_{j=1}^{n} y_j = 1, \quad x, y \geq \mathbf{0}. && \text{(1c)} \end{aligned}$$

It is easy to see that any feasible mixed strategies x, y will have objective function value less or equal to zero, as given player 2's (mixed) strategy, any pure strategy of player 1 can give payoff at most p, and given player 1's (mixed) strategy, any pure strategy of player 2 can give payoff at most q. This implies that any combination of pure strategies (i.e. any mixed strategy) of player 1 gives payoff at most p, and any combination of pure strategies of player 2 gives payoff at most q. Further, any Nash equilibrium x^*, y^* has objective function value equal to zero, thus maximises the objective function. This is because players play best responses, and thus $p^* = {x^*}^\top Ay^*$ and $q^* = {x^*}^\top By^*$. Importantly, *only* the Nash equilibria have objective function value of zero. This is because for any optimal solution (x^*, y^*, p^*, q^*) and any (x, y) with $x \geq 0$, $\sum_{i=1}^{m} x_i = 1$, $y \geq 0$, $\sum_{i=1}^{n} y_i = 1$, $x^\top Ay^* \leq p^*$, ${x^*}^\top By \leq q^*$, and thus ${x^*}^\top Ay^* \leq p^*$, ${x^*}^\top By^* \leq q^*$. As a Nash equilibrium has objective function value of zero and is guaranteed to exist, the optimal value of the bilinear formulation must be zero (as it is non-positive). Thus ${x^*}^\top Ay^* = p^*$, ${x^*}^\top By^* = q^*$. This implies $x^\top Ay^* \leq {x^*}^\top Ay^*$, ${x^*}^\top By \leq {x^*}^\top By^*$.

In this work, we propose a multilinear feasibility program whose every feasible solution is a Nash equilibrium to a corresponding multi-player game with $n \geq 2$ players. The formulation is based on an extension of the bilinear formulation to multi-player games. Although such an extension is straightforward, is has not been given much empirical analysis. We compare the multilinear feasibility formulation to established algorithms used to find an equilibrium in multi-player games. We find that our multilinear program is faster than a variety of algorithms in Gambit [10].

2.1 Multilinear formulation

Let us define some notation. Let $n \geq 2$ be the number of players, and $[n] = \{1, \ldots, n\}$ the set of players. Every player i comes with a finite set of pure strategies S_i, with $|S_i| = n_i$. Let $\mathcal{S} = S_1 \times S_2 \times \cdots \times S_n$ be the set of all n-tuples of pure strategy combinations of all players. We will further denote by $\mathcal{S}_{-i} = S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$ the set of all pure strategy tuples of all players except i. Let $\boldsymbol{s} = (s_1, \ldots, s_n) \in \mathcal{S}$ be a pure strategy tuple of all players and $\hat{\mathbf{s}} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n) \in \mathcal{S}_{-i}$ be a pure strategy tuple of all players other than i. We define payoff matrix $A_i : \mathbf{s} \in \mathcal{S} \to \mathbb{R}$ for player i. As an example, if we had three players, $A_1[s_1, s_2, s_3]$ denotes player 1's payoff when player 1 plays pure strategy s_1, player 2 plays pure strategy s_2 and player 3 plays pure strategy s_3. Likewise, $A_2[s_1, s_2, s_3]$ and $A_3[s_1, s_2, s_3]$ denote player 2 and 3's payoff for the strategy combination $(s_1, s_2, s_3) \in \mathcal{S}$. For pure strategy s of player i and pure strategies $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n) = \hat{\mathbf{s}} \in \mathcal{S}_{-i}$ of the other players, we write $A_i[s, \hat{\mathbf{s}}]$ to denote the payoff of player i when player i plays pure strategy $s \in \mathcal{S}_i$ and the other players play pure strategies $\hat{\mathbf{s}} \in \mathcal{S}_{-i}$. For every player i, we define strategy vector $\mathbf{x}^i \in \mathbb{R}^{n_i}$, with $\mathbf{x}^i \geq \mathbf{0}$ and $\sum_{s \in S_i} x^i_s = 1$. $\mathbf{x}^i$ is a probability distribution over player i's pure strategies, and thus a *mixed strategy.* Note that any pure strategy is also a mixed strategy[2]. Let $\mathbf{x} = (\mathbf{x}^1, \dots, \mathbf{x}^n)$ be a mixed strategy profile of all players, and $\mathbf{x}^{-i} = (\mathbf{x}^1, \dots, \mathbf{x}^{i-1}, \mathbf{x}^{i+1}, \mathbf{x}^n)$ be a mixed strategy profile of all players other than i. The product term $\prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j}$ for $\hat{\mathbf{s}} \in \mathcal{S}_{-i}$ denotes the combined probability of all players except i to play the pure strategy tuple $\hat{\mathbf{s}} \in \mathcal{S}_{-i}$. As an example, if we have three players, player 2 has pure strategies $s_{2,1}$ and $s_{2,2}$ and player 3 has pure strategies $s_{3,1}$ and $s_{3,2}$, then $\mathcal{S}_{-1} = \{(s_{2,1}, s_{3,1}), (s_{2,1}, s_{3,2}), (s_{2,2}, s_{3,1}), (s_{2,2}, s_{3,2})\}$. If player 2 plays $s_{2,1}$ with probability 1/2 and player 3 plays $s_{3,1}$ with probability 1/4, then $\prod_{s_j \in (s_{2,1}, s_{3,1})} = 1/2 * 1/4$, $\prod_{s_j \in (s_{2,1}, s_{3,2})} = 1/2 * 3/4$, $\prod_{s_j \in (s_{2,1}, s_{3,1})} = 1/2 * 1/4$, $\prod_{s_j \in (s_{2,2}, s_{3,2})} = 1/2 * 3/4$. Further, we define vector $\mathbf{p} \in \mathbb{R}^n$. p^i corresponds to player i's highest expected payoff.

▶ **Definition 1.** *Let $\Gamma = \langle \{1, \dots, n\}, (S_i), (A_i) \rangle$ be a game with n, S_i, A_i, $\mathbf{x}^i$ defined as above. Let $\mathbf{x}^i \geq 0$ with $\sum_{s \in S_i} x^i_s = 1$ be a mixed strategy of player i. Then, $\mathbf{x}^* = (\mathbf{x}^{*1}, \dots, \mathbf{x}^{*n})$ with $\mathbf{x}^{*i} \geq 0$ and $\sum_{s \in S_i} x^{*i}_s = 1$ for all players i is a (mixed) Nash equilibrium if for all players i and every mixed strategy $\mathbf{x}^i$, we have $\mathbb{E}[A_i[\mathbf{x}^*]] \geq \mathbb{E}\left[A_i[\mathbf{x}^i, \mathbf{x}^{*-i}]\right]$.*

We now present the multilinear optimization formulation.

$$\text{(MLP1)} \quad \begin{aligned} \max_{\mathbf{x}, \mathbf{p}} \quad & \sum_{i=1}^{n} \left(\sum_{\substack{(s, \hat{\mathbf{s}}) \\ \in S_i \times S_{-i}}} A_i[s, \hat{\mathbf{s}}] x^i_s \prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j} \right) - \sum_{i=1}^{n} p^i && \text{(2a)} \\ \text{s.t.} \quad & \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j} \leq p^i \qquad \forall i \in [n], s \in S_i && \text{(2b)} \\ & \sum_{s \in S_i} x^i_s = 1 \qquad \forall i \in [n] && \text{(2c)} \\ & 0 \leq x^i_s \leq 1 \qquad \forall i \in [n], s \in S_i && \text{(2d)} \end{aligned}$$

▶ **Theorem 2.** *A (mixed) strategy $(\mathbf{x}^1, \dots, \mathbf{x}^n)$ is a (mixed) Nash equilibrium of the n-player game $(A_1, \dots, A_n)$ if and only if there exist numbers $p^1, \dots, p^n$ such that $(\mathbf{x}^1, \dots, \mathbf{x}^n, p^1, \dots, p^n)$ is an optimal solution to the problem in MLP1.*

It is easy to see that for two players, MLP1 equals the bilinear formulation in BLP, with $\mathbf{x}^1, \mathbf{x}^2$ instead of $\mathbf{x}, \mathbf{y}$, p^1, p^2 instead of p, q, and A_1, A_2 instead of A, B. Computational experiments on small instances reveal that the solver takes significant time to solve MLP1 to optimality. However, further inspections reveal that it is more the verification of an optimal solution, rather than finding an optimal solution, that is the reason for this. Particularly, the solver finds a solution with objective function value 0 (i.e. a Nash equilibrium) relatively quickly, but spends a lot of time verifying that there is no feasible solution with objective function value larger than zero. However, as there cannot be a feasible solution with strictly positive objective value, it is sufficient for the solver to find a feasible solution with objective function value zero, instead of verifying that the upper bound to the optimisation program is zero. Thus, we reformulate the program into a feasibility program, for which the aim is to find a feasible solution for which the objective function (2a) of program MLP1 is non-negative. As a strictly positive solution is not possible, any feasible solution to MLP2 will have a value of zero, and thus be a Nash equilibrium.

[2] When we refer to (mixed) strategies or a (mixed) Nash equilibrium, this includes pure strategies or a pure Nash equilibrium.

▶ **Corollary 3.** *Every feasible solution of MLP2 is a Nash equilibrium.*

$$\text{(MLP2)} \quad \begin{aligned}
\max_{\mathbf{x},\mathbf{p}} \quad & 0 && && \text{(3a)}\\
\text{s.t.} \quad & \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j} \leq p^i && \forall i \in [n], s \in S_i && \text{(3b)}\\
& \sum_{i=1}^{n} \left(\sum_{\substack{(s,\hat{\mathbf{s}}) \\ \in S_i \times S_{-i}}} A_i[s, \hat{\mathbf{s}}] x^i_s \prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j} \right) - \sum_{i=1}^{n} p^i \geq 0 && && \text{(3c)}\\
& \sum_{s \in S_i} x^i_s = 1 && \forall i \in [n] && \text{(3d)}\\
& 0 \leq x^i_s \leq 1 && \forall i \in [n], s \in S_i && \text{(3e)}
\end{aligned}$$

2.2 Mixed-integer formulations

For a two-player game, four mixed-integer formulations whose solutions are equivalent to a Nash equilibrium in a two-player game were given in [16]. We generalize these formulations to multi-player games. The notation we use is similar to the notation introduced in the multilinear formulations. Further, we introduce $U^i = \max_{s^l, s^h \in S_i, \hat{\mathbf{s}}^{\mathbf{l}}, \hat{\mathbf{s}}^{\mathbf{h}} \in S_{-i}} A_i[s^h, \hat{\mathbf{s}}^{\mathbf{h}}] - A_i[s^l, \hat{\mathbf{s}}^{\mathbf{l}}]$ be the maximum difference of any two payoffs of player i for any pure strategies of all players.

We have four mixed-integer multilinear formulations, of which one is a feasibility program and three are optimisation programs. All programs have five sets of variables. $x^i_s, r^i_s, u^i_s, \overline{u}^i$ are real variables, and b^i_s is binary. The MIMLPs have the same interpretation and range of values for variables $x^i_s \geq 0$, $\overline{u}^i, u^i_s, r^i_s \geq 0$, further they also come with constraints (4b), (4c), (4d), (4e) (which are mostly such that variables $x^i_s, \overline{u}^i, u^i_s, r^i_s$ are defined as desired). x^i_s, for all players $i \in [n]$ and all pure strategies $s \in S_i$ of player i, denotes the probability with which player i plays pure strategy s. Hence, variables x^i_s give us the mixed strategy played by each player. In order to be valid strategies, all pure strategies of a player must be played with non-negative probability (Eq. 4h) and sum up to one (Eq. 4b), for all players. $\overline{u}^i$ denotes the highest utility player i can achieve by playing any strategy, given the other players mixed strategies. u^i_s is the expected utility of player i of playing pure strategy s, given the other players play their (mixed) strategies (Eq. 4c). Naturally, $\overline{u}^i \geq u^i_s$ (Eq. 4d). $r^i_s = \overline{u}^i - u^i_s$ (Eq. 4e) is the regret of player i of playing pure strategy s. It is defined as the difference of the highest utility of any strategy for the player to the utility of playing strategy s, given the other players' mixed strategies. By definition, the regret of any pure strategy must be non-negative (4h). Further, in any Nash equilibrium, every pure strategy that is played with strictly positive probability must have zero regret. If there was a pure strategy which the player plays and that has positive regret, the player can increase their payoff by putting more probability on a pure strategy with no regret and putting less probability on the pure strategy with regret. Hence, it would not be a Nash equilibrium.

The meaning of binary variables b^i_s is different in all formulations, with not all constraints of MIMLP 1 regarding this variable (Eq. 4f, 4g) present in MIMLP 2,3,4. In formulation 1, if b^i_s is 1, strategy s of player i is not played, hence $x^i_s = 0$. If $b^i_s = 0$, the probability on strategy s is allowed to be positive, however the regret of the strategy must be zero. (4f) ensures that b^i_s can only be set to 1 if zero probability is on s. Further, (4g) ensures that b^i_s can only be set to zero if the strategy's regret is zero (if $b^i_s = 1$, this constraint does not restrict any variable, as $r^i_s \leq U^i$ by definition).

▶ **Proposition 4.** *The set of feasible solutions to MIMLP1 is precisely the set of Nash equilibria for the corresponding multi-player game.*

$$
\text{(MIMLP1)} \quad
\begin{aligned}
\min \quad & 0 && && \text{(4a)}\\
\text{s.t.} \quad & \sum_{s\in S_i} x^i_s = 1 && \forall i \in [n] && \text{(4b)}\\
& u^i_s = \sum_{\hat{\mathbf{s}}\in S_{-i}} \prod_{s_j\in \hat{\mathbf{s}}} x^j_{s_j} A_i[s,\hat{\mathbf{s}}] && \forall i\in[n], \forall s\in S_i && \text{(4c)}\\
& \overline{u}^i \geq u^i_s && \forall i\in[n], \forall s\in S_i && \text{(4d)}\\
& r^i_s = \overline{u}^i - u^i_s && \forall i\in[n], \forall s\in S_i && \text{(4e)}\\
& x^i_s \leq 1 - b^i_s && \forall i\in[n], \forall s\in S_i && \text{(4f)}\\
& r^i_s \leq U^i b^i_s && \forall i\in[n], \forall s\in S_i && \text{(4g)}\\
& x^i_s, r^i_s \geq 0,\ u^i_s, \bar{u}_i \in \mathbb{R} && \forall i\in[n], \forall s\in S_i && \text{(4h)}\\
& b^i_s \in \{0,1\} && \forall i\in[n], \forall s\in S_i && \text{(4i)}
\end{aligned}
$$

MIMLP1 is a feasibility program, for which only Nash equilibria are feasible solutions. MIMLP2, MIMLP3, MIMLP4 have larger feasible regions, as pure strategies with positive probability are allowed to have positive regret, and pure strategies with positive regret are allowed to be played with positive probability. The formulations minimize a penalty, and it is only Nash equilibria for which the penalty is minimal. Thus, only Nash equilibria are optimal solutions. The advantage of these formulations is that, since finding a Nash equilibrium is assumed to be computationally intractable, these formulations can be used to stop the program before an equilibrium has been calculated, and thus give solutions which are close to an equilibrium, also called approximate equilibria. However, it is more difficult with these formulations to find a specific equilibrium among all equilibria, rather than just an arbitrary equilibrium.

MIMLP2 penalises the regret of a pure strategy that is played with positive probability in the objective function, and thus for optimal solutions, the regret of pure strategies with positive probability is zero. MIMLP3 penalises the probability placed on pure strategies with positive regret, and thus optimal solutions will have zero probability on pure strategies with positive regret. MIMLP4 combines the normalised regret and the probability as a penalty, and the solver can choose whether the regret or the probability should be minimized. As [16] noted, these formulations can be used to find approximate Nash equilibria.

MIMLP2 aims to minimize the regret of pure strategies that are played with positive probabilities. Particularly, the regret of a pure strategy played with positive probability serves as a penalty to the objective function. This is done by introducing variable f^i_s for all $i \in [n], s \in S_i$, which represents a pure strategy's regret if the strategy has positive probability and zero otherwise.

▶ **Proposition 5.** *The set of Nash equilibria minimizes the objective function of MIMLP2.*

$$
\text{(MIMLP2)} \quad
\begin{aligned}
\min \quad & \sum_{i=1}^{n}\sum_{s\in S_i} f^i_s - U^i b^i_s && && \text{(5a)}\\
\text{s.t.} \quad & (4b) - (4f), (4h), (4i) && && \\
& f^i_s \geq r^i_s && \forall i\in[n], \forall s\in S_i && \text{(5b)}\\
& f^i_s \geq U^i b^i_s && \forall i\in[n], \forall s\in S_i && \text{(5c)}
\end{aligned}
$$

MIMLP3 is similar to MIMLP 2, however instead of minimising the regret of pure strategies played with positive probability, the probabilities of pure strategies with positive regret is minimized. To do so, variables g_s^i are introduced, which are set such that a strategy's penalty in the objective function is zero if the strategy's regret is zero, and x_s^i (the probability with which it is played) otherwise. The set of Nash equilibria minimizes the objective, as strategies with positive regret are not played.

▶ **Proposition 6.** *The set of Nash equilibria minimizes the objective function of MIMLP3.*

$$
\text{(MIMLP3)} \quad
\begin{aligned}
\min \quad & \sum_{i=1}^{n} \sum_{s \in S_i} g_s^i - (1 - b_s^i) && \text{(6a)}\\
\text{s.t.} \quad & (4b) - (4e), (4g) - (4i) && \\
& g_s^i \geq x_s^i \qquad \forall i \in [n], \forall s \in S_i && \text{(6b)}\\
& g_s^i \geq 1 - b_s^i \qquad \forall i \in [n], \forall s \in S_i && \text{(6c)}
\end{aligned}
$$

MIMLP4 combines MIMLP 2 and MIMLP 3. Instead of penalising all pure strategies' regret (MIMLP 2) or penalising all pure strategies' probabilities if they have positive regret (MIMLP 3), this formulation lets the solver decide for each pure strategy whether to penalise the regret or the probability. The penalised regret is expressed with variables f_s^i, the penalised probabilities are expressed with variables g_s^i. When using both the regret and the probabilities, the regret must be normalised, as the probability of a pure strategy is between zero and one, but a pure strategy's regret can generally be larger than one. Hence, f_s^i uses normalised regret r_s^i/U^i, which is between zero and one.

▶ **Proposition 7.** *The set of Nash equilibria minimizes the objective function of MIMLP4.*

$$
\text{(MIMLP4)} \quad
\begin{aligned}
\min \quad & \sum_{i=1}^{n} \sum_{s \in S_i} f_s^i + g_s^i && \text{(7a)}\\
\text{s.t.} \quad & (4b) - (4e), (4h), (4i) && \\
& f_s^i \geq r_s^i/U^i \qquad \forall i \in [n], \forall s \in S_i && \text{(7b)}\\
& f_s^i \geq b_s^i \qquad \forall i \in [n], \forall s \in S_i && \text{(7c)}\\
& g_s^i \geq x_s^i \qquad \forall i \in [n], \forall s \in S_i && \text{(7d)}\\
& g_s^i \geq 1 - b_s^i \qquad \forall i \in [n], \forall s \in S_i && \text{(7e)}
\end{aligned}
$$

2.3 Continuous and feasibility formulations

For potential performance improvements of the mixed-integer multilinear programs, we further give continuous as well as feasibility formulations for the MIMLPs. Particularly, for all MIMLPs, we introduce continuous formulations[3] MIMLP1(C), MIMLP2(C), MIMLP3(C), MIMLP4(C), for which constraint 4i, i.e. constraints ($b_s^i \in \{0, 1\}$) is replaced by $b_s^i = (b_s^i)^2$ (which implies $0 \leq b_s^i \leq 1$ and $b_s^i = 0$ or $b_s^i = 1$). Thus, the continuous formulations are equivalent to the MIMLPs. The continuous formulation MIMLP1(C) for MIMLP1 is given in MIMLP1(C), likewise MIMLP2(C), MIMLP3(C), MIMLP4(C) are simply MIMLP2, MIMLP3, MIMLP4 but constraint (4i) replaced by (8a).

[3] We note that due to this, the formulations are no longer mixed-integer, however we will still refer to MIMLP(C), to make clear that they belong to the respective MIMLP

$$\text{(MIMLP1(C))} \qquad \begin{aligned} \min \quad & (4a) \\ \text{s.t.} \quad & (4b) - (4e), (4h) \\ & b_s^i = (b_s^i)^2 \qquad \forall i \in [n], \forall s \in S_i \end{aligned} \qquad (8a)$$

For MIMLP 2,3,4 we also introduce equivalent feasibility formulations MIMLP2(F), MIMLP3(F), MIMLP4(F), by introducing a constraint which requires the objective function of the respective MIMLP to be equal to the optimal value of the MIMLP. Particularly, MIMLP2 and MIMLP3 have optimal objective function of zero, and thus we introduce constraints (5a) = 0, i.e. $\sum_{i=1}^n \sum_{s\in S_i} f_s^i - U^i b_s^i = 0$ (MIMLP2) and (6a) = 0, i.e. $\sum_{i=1}^n \sum_{s\in S_i} g_s^i - (1-b_s^i) = 0$ for MIMLP3. MIMLP4 has optimal value $\sum_{i=1}^n |S_i|$, and thus we introduce constraint (7a) = $\sum_{i=1}^n |S_i|$, i.e. $\sum_{i=1}^n \sum_{s\in S_i} f_s^i + g_s^i = \sum_{i=1}^n |S_i|$. For all feasibility formulations, the objective function is changed to 0. MIMLP3(F) is given in MIMLP3(F).

Further, we introduce MIMLP2(C,F), MIMLP3(C,F), MIMLP4(C,F), which combine the continuous and feasibility formulations of MIMLP2,3,4, and are thus continuous multilinear formulations[4]. MIMLP3(C,F) is given in MIMLP3(C,F).

$$\text{(MIMLP3(F))} \qquad \begin{aligned} \min \quad & 0 \\ \text{s.t.} \quad & (4b) - (4e), (4g) - (4i), (6b), (6c) \\ & (6a) = 0 \end{aligned}$$

$$\text{(MIMLP3(C,F))} \qquad \begin{aligned} \min \quad & 0 \\ \text{s.t.} \quad & (4b) - (4e), (4g) - (4h), (6b), (6c), (8a) \\ & (6a) = 0 \end{aligned}$$

Table 1 Overview of all mixed-integer multilinear formulations.

MIMLP	1	2	3	4	1(C)	2(C)	3(C)	4(C)	2(F)	3(F)	4(F)	2(C,F)	3(C,F)	4(C,F)
Feasibility program	✓				✓				✓	✓	✓	✓	✓	✓
Optimality program		✓	✓	✓		✓	✓	✓						
Continuous					✓	✓	✓	✓				✓	✓	✓
Mixed-integer	✓	✓	✓	✓					✓	✓	✓			

3 Computational Experiments

All experiments are run on a MacBook Pro with 8GB RAM and Intel i5 CPU. Multilinear and mixed-integer formulations are implemented in AMPL [4]. We use BARON 21.1.13 [15] as the solver, which uses FilterSD and FilterSQP as non-linear subsolvers. As the multilinear formulation MLP2 in Equation (MLP2) is much faster than any of the MIMLPs (see Table 3), we decide to only compare MLP2 against common algorithms for multi-player games. The MIMLPs do not seem to give better performance than existing algorithms, and hence the analysis of those is focused on comparing the MIMLPs to each other, to get an understanding which MIMLP formulation is best. Thus, the experiments consist of two parts:

[4] MIMLP2(C,F), MIMLP3(C,F), MIMLP4(C,F) are thus no longer mixed-integer.

1. a comparison of MLP2 with common algorithms in `Gambit` [10] (results in Table 2),
2. a comparison of the different MIMLPs (results in Table 3).

All games are instanced in GAMUT [12] and have integer payoffs. We focus on *random games* and *covariance games* with negative covariance, as previous work [16, 13] indicates that covariance games with negative covariance are challenging to solve experimentally for a variety of algorithms as they tend to only have few equilibria with small support size. We refer to a covariance game with n players and $|S_i|$ actions per player and covariance ρ as CG(n,$|S_i|$,ρ), and to a random game with n players and $|S_i|$ actions per player as RG(n,$|S_i|$). For all games, we take the average of 10 randomly generated instances of that game, and if a method did not find a Nash equilibrium before the timeout (which, depending on the game, is 300 or 900 seconds), we add the timeout to the average.

Table 2 compares MLP2 to the simplicial subdivision method (SD), the global newton method (GN), and the quantal response equilibrium (QRE) in Gambit. The results can be summarised as follows: The simplicial subdivision algorithm is the slowest, and already small instances are sufficient for the algorithm to not find a Nash equilibrium in less than 15 minutes. The global Newton method, although fast on the instances for which it finds an equilibrium, in many instances terminates without giving an equilibrium back. This issues has been reported in different scenarios, see [18], and in these cases, we put the timeout towards the average. The logit algorithm and the multilinear formulation have similar runtime for smaller instances, but for larger games, our formulations seems to be faster. Thus, to conclude, our algorithm is faster than the algorithms in Gambit we test it with, and can be an alternative.

Table 3 presents the results for the MIMLPs and the reformulations. It should be pointed out than any of the MIMLPs takes much longer to find an equilibrium than MLP2, and thus none of the MIMLPs is suited to find an equilibrium for *large* multi-player games. This is different to the mixed-integer formulations for two-player games, for which [16] showed better performance on some instances than existing algorithms. Therefore, the analysis of the MIMLPs aims more to get an understanding what type of formulation is best to find a Nash equilibrium in a multi-player game, than to compare the MIMLPs to common algorithms.

First, the continuous formulations MIMLP2(C), MIMLP3(C), MIMLP4(C) of MIMLP2, MIMLP3, MIMLP4 don't give much performance improvement compared to MIMLP2,3,4. For MIMLP2 and MIMLP3, both the feasibility formulations MIMLP2(F) and MIMLP3(F) and the combined continuous and feasibility formulations MIMLP2(C,F) and MIMLP3(C,F) give better performance than MIMLP2 and MIMLP3, but whether MIMLP2(C,F) and MIMLP3(C,F) are better than MIMLP2(F) and MIMLP3(C,F) depends very much on the game. For MIMLP4, whether MIMLP4(F) or MIMLP4(C,F) are better than MIMLP4 depends on the game. Further, compared over all games, MIMLP1(C), i.e. the continuous formulation of the feasibility formulation MIMLP1 seems to give the best performance.

4 Future Work

Further questions include using different nonlinear solvers for the multilinear formulation. The solver we use finds a Nash equilibrium faster than any of the other algorithms we compare it to, other solvers should only improve the performance of the multilinear feasibility formulation. We also propose generating hard-to-solve instances. Even though GAMUT [12] offers many different types of games, many of these are easy to solve even for large multi-player games. Covariant games are among the few types of games that are (relatively) difficult to solve in the game generator GAMUT, and therefore we particularly use these instances. However, due to this, there is not much variety in the hard-to-solve instances we can use. Recent work has focused on hard-to-solve instances for polymatrix games (see [3] and `http://polymatrix-games.github.io`), and so more hard-to-solve instances is a direction to explore.

Table 2 Comparison of multilinear feasibility program to state-of-the-art algorithms.

Instance	MLP2	GN	SD	QRE	
CG(5,5,$\rho = -0.2$)	2.35	810.09	900	1.9	average (in seconds)
	100%	10%	0%	100%	percentage solved
	2.53	0.91	–	1.9	average on solved (in seconds)
CG(3,10,$\rho = -0.2$)	0.57	271.56	518.96	0.36	average (in seconds)
	100%	70%	50%	100%	percentage solved
	0.57	2.22	137.9	0.36	average on solved (in seconds)
RG(5,5)	2.23	540.57	632.98	2.08	average (in seconds)
	100%	40%	40%	100%	percentage solved
	2.23	1.43	232.45	2.08	average on solved (in seconds)
RG(3,10)	0.329	91.325	382.9	0.362	average (in seconds)
	100%	90%	70%	100%	percentage solved
	0.329	1.47	161.287	0.362	average on solved (in seconds)
CG(5,10,$\rho = -0.2$)	250.28	825.52	900	361.46	average (in seconds)
	100%	10%	0%	100%	percentage solved
	250.28	155.21	–	361.46	average on solved (in seconds)
RG(5,10)	208.79	900	900	564.32	average (in seconds)
	100%	0%	0%	90%	percentage solved
	208.79	–	–	527.02	average on solved (in seconds)
CG(5,10,$\rho = -0.1$)	220.72	813.47	900	415.64	average (in seconds)
	100%	10%	0%	90%	percentage solved
	220.72	34.77	–	361.82	average on solved (in seconds)

The time is the average over 10 instances of this game in seconds - if no solution is found after the timeout of 15 minutes, the timeout is evaluated as time for the instance.

Table 3 MIMLP results.

Method	RG(3,5)	RG(3,10)	CG(3,5,-0.2)	CG(5,3,-0.2)	RG(5,3)
	Time	Time	Time	Time	Time
MIMLP1	8.41	229.86	107.4	122.73	112.5
MIMLP1(C)	2.876	231.57	41.4	47.96	66.3
MIMLP2	19.11	202.38	16.16	538.17	660.5
MIMLP2(C)	55.93	272.15	165.5	469.5	453.14
MIMLP2(F)	14.4	150.33	29.72	410.94	345.45
MIMLP2(C,F)	14.3	279.03	68.516	198.7	218.16
MIMLP3	46.79	265.53	161.4	392.45	535.9
MIMLP3(C)	75.93	300	200.59	575.32	430.03
MIMLP3(F)	17.67	225.66	18.8	188.94	115.95
MIMLP3(C,F)	9.16	260.98	76.71	54.56	90.91
MIMLP4	5.84	220.969	79.4	359.54	49.75
MIMLP4(C)	110.3	300	69.6	479.0	462.62
MIMLP4(F)	56.5	221.26	129.78	129.59	56.88
MIMLP4(C,F)	59.19	270.69	65.12	248.85	84.52
MLP 2	0.03	0.36	0.035	0.12	0.09

The time is the average over 10 instances of this game in seconds - if no solution is found after the timeout, the timeout is evaluated as time for the instance
RG(3,5), RG(3,10), CG(3,5,-0.2): Timeout after 300 seconds [5 minutes]
CG(5,3,-0.2), RG(5,3): Timeout after 900 seconds [15 minutes]

References

1 Xi Chen and Xiaotie Deng. Settling the complexity of computing two-player Nash equilibrium. *Journal of the ACM*, 56(3):1–57, 2009. doi:10.1145/1516512.1516516.

2 Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009. doi:10.1137/070699652.

3 Argyrios Deligkas, John Fearnley, Tobenna Peter Igwe, and Rahul Savani. An empirical study on computing equilibria in polymatrix games. In *Proceedings of the 16st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '16, pages 186–195, 2016.

4 Robert Fourer, David M. Gay, and Brian W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Cengage Learning, 2nd edition, 2002.

5 Ian Gemp, Rahul Savani, Marc Lanctot, Yoram Bachrach, Thomas Anthony, Richard Everett, Andrea Tacchetti, Tom Eccles, and János Kramár. Sample-based approximation of Nash in large many-player games via gradient descent. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '22, pages 507–515. International Foundation for Autonomous Agents and Multiagent Systems, 2022. arXiv:2106.01285.

6 Srihari Govindan and Robert Wilson. A global newton method to compute Nash equilibria. *Journal of Economic Theory*, 110(1):65–86, 2003. doi:10.1016/S0022-0531(03)00005-X.

7 Srihari Govindan and Robert Wilson. Computing Nash equilibria by iterated polymatrix approximation. *Journal of Economic Dynamics and Control*, 28(7):1229–1241, 2004. doi:10.1016/S0165-1889(03)00108-8.

8 C. E. Lemke and J. T. Howson, Jr. Equilibrium points of bimatrix games. *SIAM Journal on Applied Mathematics*, 12(2):413–423, 1964. doi:10.1137/0112033.

9 O.L. Mangasarian and H. Stone. Two-person nonzero-sum games and quadratic programming. *Journal of Mathematical Analysis and Applications*, 9(3):348–355, 1964. doi:10.1016/0022-247X(64)90021-6.

10 Richard D. McKelvey, Andrew M. McLennan, and Theodore L. Turocy. Gambit: Software tools for game theory, version 16.0.2. URL: `http://www.gambit-project.org`.

11 John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951. `doi:10.2307/1969529`.

12 Eugene Nudelman, Jennifer Wortman, Yoav Shoham, and Kevin Leyton-Brown. Run the GAMUT: A comprehensive approach to evaluating game-theoretic algorithms. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '04, pages 880–887, USA, 2004. IEEE Computer Society.

13 Ryan Porter, Eugene Nudelman, and Yoav Shoham. Simple search methods for finding a Nash equilibrium. *Games and Economic Behavior*, 63(2):642–662, 2008. `doi:10.1016/j.geb.2006.03.015`.

14 Joachim Rosenmüller. On a generalization of the lemke–howson algorithm to noncooperative n-person games. *SIAM Journal on Applied Mathematics*, 21(1):73–79, 1971. `doi:10.1137/0121010`.

15 Nick V. Sahinidis. BARON 21.1.13: Global Optimization of Mixed-Integer Nonlinear Programs, *User's Manual*. `http://www.minlp.com/downloads/docs/baron%20manual.pdf`, 2017.

16 Thomas Sandholm, Andrew Gilpin, and Vincent Conitzer. Mixed-integer programming methods for finding Nash equilibria. In *Proceedings of the 20th National Conference on Artificial Intelligence - Volume 2*, AAAI'05, pages 495–501. AAAI Press, 2005.

17 Rahul Savani and Bernhard von Stengel. Hard-to-solve bimatrix games. *Econometrica*, 74(2):397–429, 2006. `doi:10.1111/j.1468-0262.2006.00667.x`.

18 Theodore L. Turocy. Answer to question regarding time limits of algorithms in gambit. URL: `https://github.com/gambitproject/gambit/issues/261#issuecomment-660894391`.

19 Theodore L. Turocy. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51(2):243–263, 2005. Special Issue in Honor of Richard D. McKelvey.

20 G. van der Laan, A. J. J. Talman, and L. van der Heyden. Simplicial variable dimension algorithms for solving the nonlinear complementarity problem on a product of unit simplices using a general labelling. *Mathematics of Operations Research*, 12(3):377–397, August 1987.

21 Robert Wilson. Computing equilibria of n-person games. *SIAM Journal on Applied Mathematics*, 21(1):80–87, 1971. `doi:10.1137/0121011`.

A Correctness of the Proposed Formulations

Here we present proofs for the claims made with regards to the formulations presented in this paper.

Proof of Theorem 2. We first show that if $(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium to $(A_1, \dots, A_m)$, then there exist numbers $\bar{p}^1, \dots, \bar{p}^n$ such that $(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n, \bar{p}^1, \dots, \bar{p}^n)$ is an optimal solution to the program in MLP1. Assume that $(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium. For any feasible solution $(\mathbf{x}^1, \dots, \mathbf{x}^n, p^1, \dots, p^n)$ of MLP1, constraints (2b), (2c) imply

$$\sum_{s \in S_i} \left(x^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} x^j_{s_j} \right) \le p^i$$

for all $i \in [n]$. This implies (2a) ≤ 0 for any feasible solution. Set

$$\bar{p}^i = \sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right)$$

for every $i \in [n]$. We show that $(\bar{\mathbf{x}}^1, \dots, \bar{\mathbf{x}}^n, \bar{p}^1, \dots, \bar{p}^n)$ is feasible and optimal to MLP1.

As $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium, we have

$$\sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) \geq \sum_{s \in S_i} \left(x^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right)$$

for all (mixed) strategies $\mathbf{x}^i \geq 0$ with $\sum_{s \in S_i} x^i_s = 1$. Choosing $\mathbf{x}^i = \mathbf{e}_k$, with $k \in \{1, \ldots, |S_i|\}$, hence the unit vector with all zeros except one in the k-th component, we have

$$\bar{p}^i = \sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) \geq \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[k, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \quad \forall k \in \{1, \ldots, |S_i|\},$$

satisfying constraint 2b. As we can apply this for all players $i \in [n]$ (and constraints 2c, 2d hold as $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium), it follows that $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^m, \bar{p}^1, \ldots, \bar{p}^n)$ is feasible. Further, the objective function value is zero at the point $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$. As the objective function value is at most zero and $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$ is feasible, it follows that it is optimal.

To show that if $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$ is an optimal solution to MLP1, $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium, we assume that $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$ indeed is optimal to MLP1. Since a Nash equilibrium exists in this game and has objective value of zero, and all feasible solutions have non-positive value, it follows that the objective value of $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$ must be zero. For any $\mathbf{x}^i \geq 0$ with $\sum_{s \in S_i} x^i_s = 1$, for all players $i \in [n]$, constraints 2b, 2c imply

$$\sum_{s \in S_i} \left(x^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) \leq \bar{p}^i \quad \forall i \in [n].$$

Particularly,

$$\sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) \leq \bar{p}^i \quad \forall i \in [n].$$

As further objective value of $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n, \bar{p}^1, \ldots, \bar{p}^n)$ is zero, we have

$$\sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) = \bar{p}^i \quad \forall i \in [n].$$

Hence,

$$\sum_{s \in S_i} \left(x^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) \leq \sum_{s \in S_i} \left(\bar{x}^i_s \sum_{\hat{\mathbf{s}} \in S_{-i}} A_i[s, \hat{\mathbf{s}}] \prod_{s_j \in \hat{\mathbf{s}}} \bar{x}^j_{s_j} \right) = \bar{p}^i$$

$\forall i \in [n], \forall \mathbf{x}^i \geq 0 : \sum_{s \in S_i} x^i_s = 1$. Therefore, $(\bar{\mathbf{x}}^1, \ldots, \bar{\mathbf{x}}^n)$ is a Nash equilibrium, as the constraint states that given the other players mixed strategies $\bar{\mathbf{x}}^j$, no strategy of player i can give higher payoff than strategy $\bar{\mathbf{x}}^i$, for all players. ◀

Proof of Proposition 4. For any player $i \in [n]$ and any pure strategy $s \in S_i$ of player i, x^i_s denotes the probability with which player i plays pure strategy s. Constraint 4b,4h guarantee $\boldsymbol{x}^i$ to be a valid mixed strategy for each player i, as all pure strategies are played with non-negative probability and sum up to one. Constraint 4c defines the expected payoff u^i_s of player i of playing pure strategy s (given the other players' mixed strategies), and 4d defines the highest possible expected payoff $\overline{u}^i$ of any (mixed) strategy of player i given the other players' (mixed) strategies. Constraint 4e,4h define the regret r^i_s of player i of playing pure strategy $s \in S_i$. The regret of a pure strategy is the difference of player i's highest possible expected payoff $\overline{u}^i$ and i's payoff of playing pure strategy s and is non-negative. Constraint 4i introduces binary variable b^i_s for any pure strategy s of any player i. Constraint 4f requires that b^i_s can only be set to one if player i puts zero probability on pure strategy s. Further, constraint 4g ensures that b^i_s can only be set to zero if the strategy's regret is zero (if $b^i_s = 1$, this constraint does not restrict any variable, as $r^i_s \leq U^i$ by definition). Thus, if b^i_s is 1, strategy s of player i is not played, hence $x^i_s = 0$. If $b^i_s = 0$, the probability on strategy s is allowed to be positive, however the regret of the strategy must be zero. Hence, only pure strategies with zero regret can be played with positive probability, which is precisely the definition of a Nash equilibrium. ◀

Proof of Proposition 5. Constraints 4b,4c,4d,4e,4h,4i guarantee that $\boldsymbol{x}^i, \boldsymbol{r}^i, \boldsymbol{u}^i, \overline{u}^i$ are correctly defined for all players. Due to constraint (4f), b^i_s can only be set to one if the probability on the pure strategy is zero. Then, due to minimising f^i_s in (5a) and Equation (5c), f^i_s must be set to U^i. In that case, f^i_s and $U^i b^i_s$ cancel out in the objective, and hence strategies with zero probability have no penalty. If $b^i_s = 0$, f^i_s equals r^i_s, due to minimising f^i_s and Equation (5b), and as $U^i b^i_s = 0$, the penalty of the pure strategy equals the regret of the strategy, and pure strategies that have no regret do not have a penalty. Thus, due to the objective function, it is encouraged to play pure strategies which have no regret, and to not play strategies with regret. Thus, any pure strategy will only contribute to the objective function if it has positive regret *and* probability. The Nash equilibria minimize the objective function, with optimal objective of zero. As any pure strategy in a Nash equilibrium will either have zero probability (hence no penalty) or zero regret (hence no penalty), the objective function will equal zero. Solutions which do not equal a Nash equilibrium have higher objective value, as for some strategies, $f^i_s > 0$ (as r^i_s and U^i are non-negative). ◀

Proof of Proposition 6. We recall that because of constraint (4g), b^i_s can only be set to zero if the strategy's regret r^i_s is zero. By constraint (6c) and minimising g^i_s, if $b^i_s = 0$, then $g^i_s = 1$. Thus, g^i_s and $1 - b^i_s$ cancel out in the objective function and the penalty of strategy s is zero. If $b^i_s = 1$, due to constraint (6b) and minimising g^i_s, $g^i_s = x^i_s$, and $1 - b^i_s = 0$. Hence, the penalty of strategy s equals x^i_s. Therefore, the probability a pure strategy is played with only contributes to the objective function if the strategy has positive regret. Nash equilibria minimize the objective function, and come with optimal value of zero. Constraint (4f) of MIMLP 1 (namely, $x^i_s \leq 1 - b^i_s$) is no longer in this formulation, and it is possible to set $b^i_s = 1$ even if some probability is placed on s. However, in a Nash equilibrium, b^i_s will only be set to 1 if the probability on s is indeed zero, as pure strategies with positive regret are not played. ◀

Proof of Proposition 7. Constraint (7b) demands that if $b^i_s = 0$, then $f^i_s = r^i_s / U^i$, which is at most 1. Further, due to (7e), $g^i_s = 1$. If $b^i_s = 1$, then $f^i_s = 1$ (constraint (7c)) and $g^i_s = x^i_s$ (constraint (7d)), which is at most 1. Hence, $f^i_s + g^i_s$ is at least 1 for every pure strategy s, and additional penalties (either the normalised regret or the probability of the strategy) contribute to the objective function if a strategy has positive probability and positive regret. Any feasible solution that is not a Nash equilibrium has $f^i_s + g^i_s > 1$ for some strategies, as not all strategies have either no regret or zero probability. Nash equilibria minimize the objective function, with value of $\sum_{i=1}^n |S_i|$, as the normalised regret is zero, or the probability of strategy is zero. ◀