Proceedings of Machine Learning Research vol 331:1–12, 2026 8th Annual Conference on Learning for Dynamics and Control

Optimizing Coordination among Bounded Rational Agents

Zhewei Wang ZW23A@FSU.EDU
Department of Mechanical Engineering, FAMU-FSU College of Engineering, Florida State University, Tallahassee, FL 32306, USA

Marcos M. Vasconcelos M.VASCONCELOS@FSU.EDU
Department of Electrical and Computer Engineering, FAMU-FSU College of Engineering, Florida State University, Tallahassee, FL 32306, USA

Abstract

Coordination is a desirable feature in many multi-agent systems, such as robotic and socioeconomic networks. Traditionally, these agents are assumed to be perfectly rational, i.e., they are designed and trained to maximize their expected utilities. However, in many situations, perfectly rational behavior is not possible. We consider a binary networked coordination game over a weighted undirected regular graph with a sparsity constraint. Each agent exhibits bounded rationality and employs a distributed stochastic learning algorithm known as *Log Linear Learning* to update its action conditioned on the actions currently played by its neighbors. We optimize the probability that the multi-agent system will converge to a pure Nash equilibria of the game with respect to the graph weights. We provide analytical and numerical results for specific sparsity patterns considered in a classical behavioral economics experiment from Leavitt and Bavellas (1951).

Keywords: Distributed learning, bounded rationality, convex optimization.

1. Introduction

In multi-agent systems such as robotic and socioeconomic networks, agents must coordinate actions with neighbors to achieve collective goals. A key challenge is *bounded rationality*, where agents, limited by cognitive, informational, or computational constraints, make probabilistic suboptimal decisions (Simon, 1957; Tsiotras, 2021; Xianjia et al., 2023; Chasnov, 2024). This creates a spectrum in multi-agent systems: fully rational agents always choose best-responses, and irrational ones select uniformly at random. In practice, most artificial as well as natural agents operate between these extremes, preventing perfect coordination to emerge as a result of distributed learning processes. A classic illustration of how network structure and bounded rationality interact in collective decision-making was provided by *Leavitt's experiment* (Leavitt, 1951), which showed that network structure influences the efficiency and convergence rate of task completion in systems comprised of human agents (who are, of course, subject to bounded rationality). This important empirical observation motivates the analytical study of how rationality levels interact with network topology in achieving coordinated outcomes.

To study this interaction, we use the well-established framework for modeling bounded rationality in repeated games known as *log-linear learning* (LLL) (Blume, 1993; Marden and Shamma, 2012), which uses a single rationality parameter, β, for every agent. As $\beta \to 0$, agents behave uniformly at random; as $\beta \to \infty$, they approach best-response dynamics. This parameterization is related to both statistical mechanics (where rationality acts as the inverse of temperature) and behavioral game theory (through the notion of quantal response equilibrium) and has been widely validated in behavioral studies (McKelvey and Palfrey, 1995; Ghosal et al., 2023). We point out that other models bounded rationality exist: prospect theory (Tversky and Kahneman, 1992) and level-k reasoning (Fotiadis and Vamvoudakis, 2021; Guan et al., 2021) capture many interesting behavioral aspects. However, LLL is sufficiently rich, analytically tractable, and provides a behaviorally meaningful framework for network design in multi-agent systems due to the existence of a closed form stationary probability distribution that describes asymptotic bounded rational agent behavior.

Our focus is the design of weighted networks for a multi-agent system engaging in a network coordination game under bounded rationality using LLL. In our framework, agents choose binary actions to maximize local utilities, and the system designer seeks to maximize the stationary probability that the agents will converge to a Nash equilibrium. Additionally, we would like to minimize a cost on the network connectivity. We will introduce an objective function that trades-off: (1) the stationary probability mass on coordinated action profiles, and (2) a linear penalty on total graph weight. We first show that this optimization problem is convex for any fixed network sparsity pattern. Therefore, such problems are tractable. We then show that for complete graphs, the optimal solution uses uniform edge weights, which explicitly depends on the rationality β, agent number N, and connectivity cost ρ. We note that the optimal weights decrease with rationality β, establishing an inverse relationship between rationality and connectivity. To the best of our knowledge, this is the first work formalizing bounded-rational coordination as a network weight design convex optimization problem. Our contributions include: (1) formulating the problem under LLL and proving convexity; (2) deriving closed-form optimal weights for complete graphs and their dependence on β, N, and ρ; (3) showing weights decrease monotonically with β; (4) extending the same sparse networks considered in Leavitt (1951): line, cycle, star, and "Y" networks, whose optimal numerical solutions can obtained efficiently.

The interplay between network structure and agent behavior in coordination games has been extensively studied in economics and control theory (Jackson and Zenou, 2015; Li et al., 2022). Specifically when LLL is used by the agents, seminal results by Montanari and Saberi (2010) and Arieli et al. (2020) showed that graph structure is intimately related to the convergence rate to a Nash equilibrium as the rationality β increases without bound. Unlike the aforementioned works, our focus is on the asymptotic likelihood of converging to a coordinated strategy profile with finite rationality, where suboptimal decision-making is persistent. When the graphs are unweighted, our prior work (Zhang and Vasconcelos, 2024a,b), used LLL in coordination games to establish that increasing connectivity improves coordination likelihood under bounded rationality. However, optimizing over the space of unweighted graphs is a combinatorial problem. In this work, we take the first step to alleviate this complexity by relaxing the graph edges to be weighted, which leads to a convex optimization problem.

2. System Model

Consider a system with N agents connected by a weighted graph $\mathcal{G}([N], \mathcal{E})$, i.e., each link in the graph is associated with a weight $w_{ij} \in [0, 1]$. If $(i, j) \notin \mathcal{E}$, then $w_{ij} = 0$. Otherwise, $w_{ij} > 0$. Therefore, the network is characterized by a weight matrix $\mathbf{W} = [w_{ij}]$, whose entry w_{ij} represents the weight between agents i and j.

The agents in this system are playing a binary network coordination game defined on the connected weighted graph described by $\mathbf{W}$. The game is defined as follows. Consider a binary coordination game played between two agents in the system. Let $(i,j) \in \mathcal{E}$, and suppose that $a_i, a_j \in \{0,1\}$ are the actions played by agents i and j, respectively. Assume that the bimatrix game in Fig. 1 specifies the payoffs for the pairwise interaction between i and j with parameter parameter $\theta \in \mathbb{R}$, which represents the underlying difficulty of performing a collective task (c.f. Remark 1). This is a coordination game, because there are two Nash equilibria in pure strategies: $a_i = a_j = 1$ and $a_i = a_j = 0$, where both agents play the same action (they coordinate).

		a_j	
		1	0
a_i	1	$\left(1-\frac{\theta}{N}, 1-\frac{\theta}{N}\right)$	$\left(-\frac{\theta}{N}, 0\right)$
	0	$\left(0, -\frac{\theta}{N}\right)$	$\left(0, 0\right)$

Figure 1: A coordination game with parameter θ between two players.

Remark 1 (Payoff Interpretation) *The coordination game in Fig. 1 is an instance of a stag-hunt game (Skyrms, 2004). One of the applications of the framework proposed herein is in distributed multi-robot task allocation (Wei and Vasconcelos, 2023; Vasconcelos and Touri, 2023; Kanakia et al., 2016). In such application, there is a single task of difficulty θ, which is amortized among N agents in the system forming subtasks of difficulty θ/N. Each agent has one unit of power to use towards completing its subtask, but will only succeed if she works together with its neighbors.*

Extending the two-agent coordination game from Fig. 1 over a network, we obtain the following *network game* with N agents, where agent i simultaneously plays the same action with all of its neighbors $j \in \mathcal{N}_i$. Let $V_{ij} : \{0,1\}^2 \to \mathbb{R}$ be defined as

$$V_{ij}(a_i, a_j) \stackrel{\text{def}}{=} a_i\Big(a_j - \frac{\theta}{N}\Big). \tag{1}$$

In a network coordination game, every agent receives the weighted sum of all the payoffs of the bimatrix games $V_{ij}(a_i, a_j)$ played with each of its neighbors. Therefore, for the i-th agent, the utility is determined as follows

$$U_i(a_i, a_{-i}) \stackrel{\text{def}}{=} \sum_{j \in \mathcal{N}_i} w_{ij} V_{ij}(a_i, a_j). \tag{2}$$

Therefore, the payoff of the i-th agent in our game is

$$U_i(a_i, a_{-i}) = a_i\Big(\sum_{j \in \mathcal{N}_i} w_{ij} a_j - \frac{\theta}{N} \sum_{j \in \mathcal{N}_i} w_{ij}\Big). \tag{3}$$

Notice that the payoff in (3) reflects that more connected nodes face a larger difficulty due to the collaboration with other agents in the network. However, the effort is also aggregated, which leads to a potentially higher utility.

Definition 2 (Exact potential games) *Let $\mathcal{A}_i$ denote the action set of the i-th agent in a game with payoff functions $U_i(a_i, a_{-i})$, $i \in [N]$. Let $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$. A game is an exact potential game if there exists a potential function Φ: $\mathcal{A} \to \mathbb{R}$ such that*

$$U_i(a_i', a_{-i}) - U_i(a_i'', a_{-i}) = \Phi(a_i', a_{-i}) - \Phi(a_i'', a_{-i}), \tag{4}$$

for all $a_i', a_i'' \in \mathcal{A}_i$, $a_{-i} \in \mathcal{A}_{-i}$, $i \in [N]$.

Proposition 3 *The network coordination game defined above is an exact potential game if and only if the graph's weight matrix* $\mathbf{W}$ *is symmetric, i.e.,* $w_{ij} = w_{ji}$ *for all* $i, j \in [N]$.

Proof From Monderer and Shapley (1996), we know that a game with continuous action sets, which are intervals of real numbers, is an exact potential game if and only if the payoff functions are twice continuously differentiable, and

$$\frac{\partial^2 U_i}{\partial a_i \partial a_j} = \frac{\partial^2 U_j}{\partial a_i \partial a_j}, \quad i, j \in [N]. \tag{5}$$

Relaxing the action set $\mathcal{A}_i = \{0, 1\}$ to $\tilde{\mathcal{A}}_i = \mathbb{R}$, since the payoffs are twice continuously differentiable, the condition in (5) becomes

$$w_{ij} = w_{ji},\ i, j \in [N] \Rightarrow \mathbf{W} = \mathbf{W}^{\mathsf{T}}. \tag{6}$$

Since this condition must hold for the relaxed action set $\tilde{\mathcal{A}}_i$, it must also hold for $\mathcal{A}_i \subseteq \tilde{\mathcal{A}}_i$. ■

To obtain a potential function for our network game, define the following function

$$\phi_{ij}(a_i, a_j) \overset{\text{def}}{=} a_i a_j + (1 - a_i - a_j)\frac{\theta}{N}. \tag{7}$$

The function ϕ_{ij} is a potential function for the (weighted) two-agent coordination game in Fig 1, i.e.,

$$\phi_{ij}(1, a_j) - \phi_{ij}(0, a_j) = V_{ij}(1, a_j) - V_{ij}(0, a_j), \tag{8}$$

for all $i, j \in [N]$ and $a_j \in \{0, 1\}$. Let $\Phi_{\mathbf{W}} : \{0, 1\}^N \to \mathbb{R}$ be defined as follows

$$\Phi_{\mathbf{W}}(a) \overset{\text{def}}{=} \frac{1}{2} \sum_{i \in [N]} \sum_{j \neq i} w_{ij} \phi_{ij}(a_i, a_j). \tag{9}$$

Evaluating the summations, we obtain the potential function $\Phi_{\mathbf{W}}$ for the network coordination game with weight matrix $\mathbf{W}$ as

$$\Phi_{\mathbf{W}}(a) = \frac{1}{2} a^{\mathsf{T}} \mathbf{W} a - \frac{\theta}{N} \mathbb{1}^{\mathsf{T}} \mathbf{W} a + \frac{\theta}{2N} \mathbb{1}^{\mathsf{T}} \mathbf{W} \mathbb{1}. \tag{10}$$

Note that, while $\Phi_{\mathbf{W}}(a)$ is a binary quadratic form in a, the potential function in (10) is a linear function of $\mathbf{W}$. We state this property formally, without proof, in the following proposition.

Proposition 4 *For a fixed* $a \in \{0, 1\}^N$*, the potential function* $\Phi_{\mathbf{W}}(a)$ *is linear in* $\mathbf{W}$.

A fundamental result from (Monderer and Shapley, 1996), establishes a connection between the Nash Equilibria of the network coordination game and the maximizers of the corresponding potential function. Rewriting $\Phi_{\mathbf{W}}(a)$ in terms of the graph Laplacian, we obtain

$$\Phi_{\mathbf{W}}(a) = -\frac{1}{2} a^{\mathsf{T}} \mathbf{L}(\mathbf{W}) a + \left(\frac{1}{2} - \frac{\theta}{N}\right) \mathbb{1}^{\mathsf{T}} \mathbf{W} a + \frac{\theta}{2N} \mathbb{1}^{\mathsf{T}} \mathbf{W} \mathbb{1}, \tag{11}$$

where $\mathbf{L}(\mathbf{W}) \overset{\text{def}}{=} \text{diag}(\mathbf{W}\mathbb{1}) - \mathbf{W}$. We characterize the set of Nash equilibria of the weighted network coordination game in the following proposition.

Proposition 5 *Let $\mathcal{S}_{\mathrm{NE}}$ denote the set of pure strategy Nash equilibria of the weighted network coordination game with payoffs given by (3). For all connected weighted graphs with $\mathbf{W} \in [0,1]^{N \times N}$, the following holds: If $\theta \leq N/2$, the potential function is maximized at $a^\star = \mathbb{1}$; If $\theta > N/2$, the potential function is maximized at $a^\star = \mathbb{0}$. Therefore, the following holds*

$$\{\mathbb{0}, \mathbb{1}\} \subseteq \mathcal{S}_{\mathrm{NE}}. \tag{12}$$

Proof The proof follows from the potential function in (11) using the fact that $a^\mathsf{T}\mathbf{L}(\mathbf{W})a \geq 0$, with equality if and only if $a \in \{\mathbb{0}, \mathbb{1}\}$. Therefore, the following holds

$$\Phi_{\mathbf{W}}(a) \leq \left(\frac{1}{2} - \frac{\theta}{N}\right) \mathbb{1}^\mathsf{T}\mathbf{W}a + \frac{\theta}{2N}\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1}, \;\; a \in \{0,1\}^N. \tag{13}$$

Therefore, since the right hand side of (13) is affine in a, the global maximizer of the potential function is $a^\star = \mathbb{1}$ if $\theta \leq N/2$ and $a^\star = \mathbb{0}$ if $\theta > N/2$.

■

2.1. Log-linear learning

LLL is a learning algorithm that enables agents to converge to pure strategy Nash equilibria in potential games. The algorithm has a stochastic dynamics and is characterized by a parameter β, which controls the trade-off between exploration and exploitation, which can be interpreted as a measure of the agents' rationality. At time t, an agent i is randomly selected with uniform probability from $[N]$. Then, agent i updates its action given the actions of its neighbors played at time t, according to a softmax distribution, i.e.,

$$\mathbf{P}\big(A_i(t+1) = a_i \mid \{A_j(t) = a_j(t)\}_{j \in \mathcal{N}_i}\big) = \frac{e^{\beta U_i\left(a_i, a_{-i}(t)\right)}}{e^{\beta U_i\left(0, a_{-i}(t)\right)} + e^{\beta U_i\left(1, a_{-i}(t)\right)}}, \; a_i \in \{0,1\}. \tag{14}$$

When $\beta \to 0$, each agent selects an action uniformly at random from its action space. Conversely, as $\beta \to \infty$, the agent's behavior approaches that of a best-response policy. In the limit, the agents following the LLL dynamics converge to one of the game's pure-strategy Nash equilibria. However, under bounded rationality (i.e., $\beta < \infty$), the agents always have a nonzero probability of choosing a suboptimal action. LLL induces a Markov chain with a stationary distribution over the action space $\{0,1\}^N$ known as the Gibbs-Boltzmann distribution, given by

$$\mu_{\mathbf{W}}(a \mid \beta) \overset{\mathrm{def}}{=} \frac{e^{\beta\Phi_{\mathbf{W}}(a)}}{\sum_{a \in \{0,1\}^N} e^{\beta\Phi_{\mathbf{W}}(a)}}, \;\; a \in \{0,1\}^N. \tag{15}$$

Using Proposition 5, we can determine the probability of each of the possible equilibria for the game. If $a^\star = \mathbb{0}$ (when $\theta > N/2$) and $a^\star = \mathbb{1}$ (when $\theta < N/2$) , we have

$$\Phi_{\mathbf{W}}(\mathbb{0}) = \frac{\theta}{2N}\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1}, \quad \Phi_{\mathbf{W}}(\mathbb{1}) = \Big(\frac{1}{2} - \frac{\theta}{2N}\Big)\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1}, \tag{16}$$

which implies that

$$\Big(\mu_{\mathbf{W}}(\mathbb{0} \mid \beta)\Big)^{-1} = \sum_{a \in \{0,1\}^N} e^{\beta\left(\frac{1}{2}a^\mathsf{T}\mathbf{W}a - \frac{\theta}{N}\mathbb{1}^\mathsf{T}\mathbf{W}a\right)}, \tag{17}$$

and

$$\Big(\mu_{\mathbf{W}}(\mathbb{1} \mid \beta)\Big)^{-1} = \sum_{a \in \{0,1\}^N} e^{\beta\left(\frac{1}{2}a^\mathsf{T}\mathbf{W}a - \frac{\theta}{N}\mathbb{1}^\mathsf{T}\mathbf{W}a + \left(\frac{\theta}{N} - \frac{1}{2}\right)\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1}\right)}. \tag{18}$$

2.2. Optimization problem

We are interested in choosing the graph's weight matrix $\mathbf{W}$ such that the probabilities in (17) or (18) are maximized. Additionally, we impose a regularization term on the network connectivity. For $a^\star \in \{\mathbb{0}, \mathbb{1}\}$, $\rho > 0$ and network sparsity pattern $\mathcal{S}$, we would like to solve the following optimization problem

$$\begin{array}{ll} \text{minimize} & \left(\mu_{\mathbf{W}}(a^\star \mid \beta)\right)^{-1} + \frac{\rho}{2}\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1} \\ \text{subject to} & \operatorname{trace}(\mathbf{W}) = 0 \\ & \mathbf{W} = \mathbf{W}^\mathsf{T} \\ & 0 \leq \mathbf{W} \leq 1 \\ & \lambda_2\big(\mathbf{L}(\mathbf{W})\big) > 0 \\ & \mathbf{W} \in \mathcal{S} \end{array} \tag{19}$$

with variable $\mathbf{W} \in \mathbb{R}^{N\times N}$.

Remark 6 *The trace constraint in (19) ensures that the graph has no self-loops and is a linear equality constraint. The symmetry constraint guarantees that the weighted network game is potential and consists of element-wise equalities. The box inequality constraints are understood element-wise inequalities. Finally, the constraint on the algebraic connectivity $\lambda_2\big(\mathbf{L}(\mathbf{W})\big)$ ensures that the graph is connected, which in turn leads to the learning process having a unique stationary distribution $\mu_{\mathbf{W}}(a \mid \beta)$. This connectivity constraint is convex, as λ_2 is concave in $\mathbf{W}$ (Boyd et al., 2006).*

Remark 7 *To ensure at least one feasible solution exists, the sparsity pattern must be symmetric and allow at least one connected graph. Note that the sparsity pattern constraint is convex, and it differs from a constraint of the form "$\mathbf{W}$ corresponds to a star network," since the convex combination of two distinct star networks is generally not a star network. However, the set of all star networks with a fixed central node and peripheral nodes is convex.*

3. Convexity and a structural result

The optimization problem stated in the previous section is well-structured and admits a characterization of its optimal solution, presented in the following result.

Theorem 8 *The optimization problem in (19) is convex. When $\mathcal{S}$ is the set of all graphs, (19) admits an optimal solution of the form*

$$\mathbf{W}^\star = w^\star\big(\mathbb{1}\mathbb{1}^\mathsf{T} - \mathbf{I}\big), \tag{20}$$

where

$$w^\star = \arg\max_{w\in[0,1]} \sum_{d=0}^{N} \binom{N}{d} \exp\left(\beta w\Big(\frac{d^2-d}{2} - \frac{\theta}{N}(N-1)d\Big)\right) + \frac{\rho}{2}wN(N-1). \tag{21}$$

Proof Without loss of generality, assume $a^\star = \mathbb{0}$[1]. The objective function in (19) is

$$f_0(\mathbf{W}) \stackrel{\text{def}}{=} \sum_{a\in\{0,1\}^N} e^{\beta\left(\frac{1}{2}a^\mathsf{T}\mathbf{W}a - \frac{\theta}{N}\mathbb{1}^\mathsf{T}\mathbf{W}a\right)} + \frac{\rho}{2}\mathbb{1}^\mathsf{T}\mathbf{W}\mathbb{1}. \tag{22}$$

Each term in the sum over $\{0,1\}^N$ is a composition of an affine function of $\mathbf{W}$ with the convex and increasing function $\exp(\cdot)$, and is therefore convex. The first three constraints are linear in $\mathbf{W}$, and therefore convex. The last constraint involves the second smallest eigenvalue of the Laplacian matrix, which is a concave function of $\mathbf{W}$ (Boyd et al., 2006). Therefore, the graph connectedness is also a convex constraint. Finally, the sparsity pattern constraint is convex.

To prove the second part of the statement, let $\mathcal{S}$ be the set of all weighted graphs (an unconstrained sparsity pattern), notice that the feasible set is invariant to permutations in the following sense: Suppose $\mathbf{W}$ is feasible for the problem in (19). Let P be any permutation matrix in the set of all permutation matrices of dimension N, which we denote by $\mathcal{P}$. Let $\mathbf{W}' = P\mathbf{W}P^\mathsf{T}$. The two matrices $\mathbf{W}$ and $\mathbf{W}'$ and their corresponding Laplacian matrices are similar, and therefore, have the same spectrum (same eigenvalues). Therefore, the algebraic connectivity constraint is preserved. The other three constraints are trivially satisfied by $\mathbf{W}'$. Therefore, the feasible set is invariant to permutations.

Let $\mathcal{W}$ denote the feasible set for (19). Since $f_0(\mathbf{W})$ is continuous and the feasible set is compact due to the box constraints, there exists a $\mathbf{W}^\star \in \mathcal{W}$ that achieves the minimum of $f_0(\mathbf{W})$, denoted by $f_0^\star$. Suppose that $\mathbf{W}^\star$ is not of the form in (20). Based on $\mathbf{W}^\star$ we will construct an optimal solution to (19) with the desired form and the same objective value $f_0^\star$. Consider $P \in \mathcal{P}$ and let $\mathbf{W}_P \stackrel{\text{def}}{=} P\mathbf{W}^\star P^\mathsf{T}$. Since $\mathbf{W}_P$ is feasible, we evaluate

$$\begin{aligned}
f_0(\mathbf{W}_P) &= \sum_{a\in\{0,1\}^N} \exp\left(\beta\left(\frac{1}{2}a^\mathsf{T}\mathbf{W}_P a - \frac{\theta}{N}\mathbb{1}^\mathsf{T}\mathbf{W}_P a\right)\right)\\
&= \sum_{a\in\{0,1\}^N} \exp\left(\beta\left(\frac{1}{2}a^\mathsf{T}P\mathbf{W}^\star P^\mathsf{T}a - \frac{\theta}{N}\mathbb{1}^\mathsf{T}P\mathbf{W}^\star P^\mathsf{T}a\right)\right)\\
&\stackrel{(a)}{=} \sum_{\tilde{a}\in\{0,1\}^N} \exp\left(\beta\left(\frac{1}{2}\tilde{a}^\mathsf{T}\mathbf{W}^\star\tilde{a} - \frac{\theta}{N}\mathbb{1}^\mathsf{T}\mathbf{W}^\star\tilde{a}\right)\right)\\
&= f_0^\star,
\end{aligned}$$

where (a) follows from the change of variables $\tilde{a} = P^\mathsf{T}a$, and that the summation is over all possible vectors in $\{0,1\}^N$.

Define

$$\mathbf{W}_{\text{sym}} \stackrel{\text{def}}{=} \frac{1}{N!}\sum_{P\in\mathcal{P}} P\mathbf{W}^\star P^\mathsf{T}. \tag{23}$$

Since $\mathbf{W}_{\text{sym}}$ is feasible, and the objective function is convex, we have

$$f_0^\star \leq f_0(\mathbf{W}_{\text{sym}}) = f_0\Big(\frac{1}{N!}\sum_{P\in\mathcal{P}} P\mathbf{W}^\star P^\mathsf{T}\Big) \leq \frac{1}{N!}\sum_{P\in\mathcal{P}} \underbrace{f_0(\mathbf{W}_P)}_{=f_0^\star} = f_0^\star. \tag{24}$$

1. The same structural result holds for $a^\star = \mathbb{1}$ and is omitted for brevity.

Finally, we show that all nonzero entries in $\mathbf{W}_{\text{sym}}$ are equal. Let the entries below the diagonal of $\mathbf{W}^\star$ be denoted by

$$\{w_1^\star, w_2^\star, \dots, w_M^\star\}, \quad \text{where } M = \frac{N(N-1)}{2}. \tag{25}$$

Consider the set Π of all possible permutations of these M values. For each $\pi \in \Pi$, construct a matrix $\mathbf{W}_\pi^\star$ by assigning the permuted values symmetrically to the upper and lower triangular entries.

Let $\mathcal{P}$ denote the set of permutation matrices P such that, for any symmetric matrix $\mathbf{W}$ with zero diagonal, the matrix $P\mathbf{W}P^\mathsf{T}$ is also symmetric with zero diagonal. Then, for each permutation π, there exists a corresponding $P(\pi) \in \mathcal{P}$ such that

$$\mathbf{W}_\pi^\star = P(\pi)\mathbf{W}^\star P(\pi)^\mathsf{T}. \tag{26}$$

Therefore, the average of these matrices is

$$\mathbf{W}_{\text{sym}} = \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\mathbf{W}_\pi^\star = \frac{1}{|\mathcal{P}|}\sum_{P\in\mathcal{P}} P\mathbf{W}^\star P^\mathsf{T}. \tag{27}$$

Since each nonzero position in $\mathbf{W}_\pi^\star$, across all permutations in Π, takes on each of the M original subdiagonal values $w_1^\star, w_2^\star, \dots, w_M^\star$ the same number of times, each such entry in $\mathbf{W}_{\text{sym}}$ is equal to

$$w^\star \overset{\text{def}}{=} \frac{1}{M}\sum_{k=1}^{M} w_k^\star. \tag{28}$$

■

Remark 9 *The structural result in Theorem 8 has an intuitive interpretations: for agents with uniform rationality parameters β and unconstrained sparsity, the optimal graph is complete with a uniform weight, $w^\star$.*

4. Numerical results for complete and incomplete network sparsity patterns

To illustrate our theoretical results, consider the optimal design of the weighted graph for an unconstrained sparsity pattern. In other words, any pair of agents can be connected. Theorem 1 implies that there exists a fully connected graph with the uniform edge weight $w^\star$ that is optimal for the problem in (19). Considering a system with N agents and difficulty parameter $\theta = N/2 + 1$, the Nash equilibrium is $a^\star = \mathbb{0}$ (c.f. Prop 5). We find the optimal weight by computing $w^\star(\beta) \in (0, 1]$ that solves (21).

Figure 2 shows the numerical results[2] for $N = 20$ and connectivity cost $\rho = 1, 5, 10$ as a function of the agents' rationality, β. In Fig. 2 (left), we observe the resulting stationary probability of agents playing the Nash equilibrium $a^\star = \mathbb{0}$ as a function of their rationality β. In Fig. 2 (right), we observe how the optimal cost decreases as a function of the rationality. Finally, we see that the probability increases monotonically with β, but since the agents have bounded rationality, this

2. The code and datasets used in this work are available at https://github.com/MINDS-code/Teaming.git

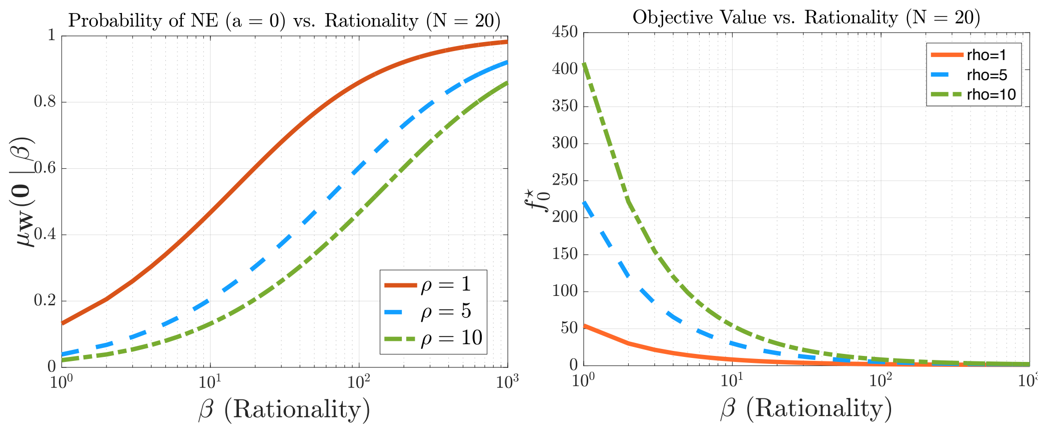

Figure 2: Numerical results for the optimal graph weight when $N = 20$, $\theta = N/2 + 1$, and $\rho = 1, 5, 10$ as a function of the rationality β: (left) shows the total optimal cost, and (right) shows the resulting probability of playing the Nash equilibrium $a^\star = \mathbb{0}$.

probability will always be bounded away from 1, i.e., perfect learning cannot be achieved with bounded rationality. However, our results establish the maximum probability of playing a Nash equilibrium for a given level of connectivity that the system designer can afford.

Our previous results hold for a possibly complete (i.e. unrestricted graph) where any two agents can be connected. However, the convexity of the optimization problem holds even when we constrain the graphs to have an arbitrary sparsity pattern, $\mathcal{S}$. The sparsity pattern of a graph specifies the connections/weights that must be equal to zero. As an example, the sparsity pattern for a weight matrix of a *line network* and *star network* with N nodes is given by

$$\mathcal{S}_{\text{line}} = \begin{bmatrix} 0 & * & 0 & \cdots & 0 \\ * & 0 & * & \cdots & 0 \\ 0 & * & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & * & 0 \end{bmatrix} \quad \mathcal{S}_{\text{star}} = \begin{bmatrix} 0 & * & * & \cdots & * \\ * & 0 & 0 & \cdots & 0 \\ * & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{29}$$

To obtain the optimal weight matrix for networks with incomplete sparsity patterns, we consider the three networks with $N = 5$ agents shown in Fig. 3: a line network (a), and star network (b) and the "Y" network (c). These are the same networks considered by Leavitt (1951). Notice that $(a), (b), (c)$ are networks with the same number of nodes, and the same number of edges in their respective sparsity patterns, which enables us to make a fair comparison between them.

The first observation is that for the star network (b), the optimal solution is always symmetric. For instance, when $\beta = 1$, $\theta = 3$ and $\rho = 5$, we have

$$\mathbf{W}^\star_{\text{star}} = \begin{bmatrix} 0 & 0.5174 & 0.5174 & 0.5174 & 0.5174 \\ 0.5174 & 0 & 0 & 0 & 0 \\ 0.5174 & 0 & 0 & 0 & 0 \\ 0.5174 & 0 & 0 & 0 & 0 \\ 0.5174 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{30}$$

with

$$f^\star_{0,\text{star}} = 26.5133. \tag{31}$$

However, when we optimize the line network in (a) for $\beta = 1$, $\theta = 3$ and $\rho = 5$, we obtain

$$\mathbf{W}^\star_{\text{line}} = \begin{bmatrix} 0 & 0.5363 & 0 & 0 & 0 \\ 0.5363 & 0 & 0.5056 & 0 & 0 \\ 0 & 0.5056 & 0 & 0.5056 & 0 \\ 0 & 0 & 0.5056 & 0 & 0.5363 \\ 0 & 0 & 0 & 0.5363 & 0 \end{bmatrix}, \tag{32}$$

whose weights are not equal and have objective equal to

$$f^\star_{0,\text{line}} = 26.4753. \tag{33}$$

Lastly, we investigate the performance of the three networks $(a), (b), (c)$ as a function on β for $\theta = 4$ and $\rho = 5$. Solving the optimization problem in (19) with the sparsity pattern constrain in each case, we obtain the following performance curves in Fig. 4, which shows the optimal objective value for each of the networks in Fig. 3 as a function of the agents' rationality β. Although the overall performance of the three networks are very close, we can clearly see that the star network performs worse than the other two. Intuitively, in a star network, having the central agent be both highly connected and having the same bounded rationality as peripheral agents can hurt the overall team performance.

It remains to compare the hybrid and line networks. Our numerical results in Fig. 4 suggest that being connected in a line network is uniformly better than the hybrid one. This provides evidence that even in small-scale settings, there are team structures that are better than others. We conjecture that such performance improvement comes from the fact that the sparsity pattern $\mathcal{S}_{\text{line}}$ is "closer" to the one of a regular graph. As established by Zhang and Vasconcelos (2024b), when the graphs are unweighted, regular graphs maximize the probability of learning to coordinate on the Nash equilibrium. A rigorous analysis of this observation is left for future work.

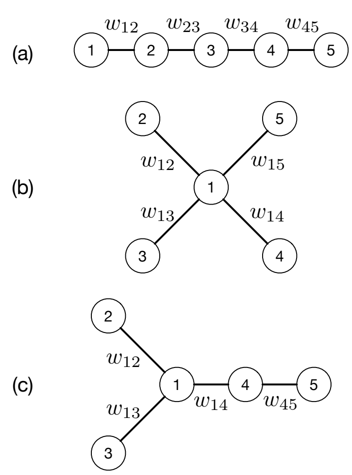

Figure 3: Line, Star and "Y" Networks.

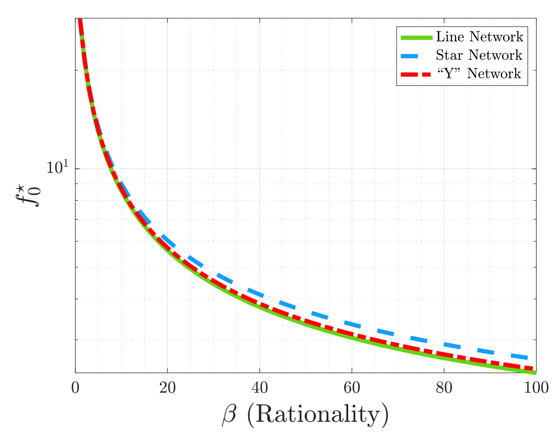

Figure 4: Performance of the Line, Star and "Y" networks as a function of β.

References

Itai Arieli, Yakov Babichenko, Ron Peretz, and H. Peyton Young. The speed of innovation diffusion in social networks. *Econometrica*, 88(2):569–594, 2020.

Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and economic behavior*, 5(3):387–424, 1993.

Stephen Boyd, Arpita Ghosh, Balaji Sun, and Persi Diaconis. Convex optimization of graph Laplacian eigenvalues. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 1311–1319, 2006.

Benjamin J. Chasnov. *Dynamics of Multi-Agent Learning Under Bounded Rationality: Theory and Empirical Evidence*. PhD thesis, University of Washington, 2024.

Filippos Fotiadis and Kyriakos G. Vamvoudakis. Recursive reasoning for bounded rationality in multi-agent non-equilibrium play learning systems. In *2021 IEEE Conference on Control Technology and Applications (CCTA)*, pages 741–746, 2021.

Gaurav R. Ghosal, Matthew Zurek, Daniel S. Brown, and Anca D. Dragan. The effect of modeling human rationality level on learning rewards from multiple feedback types. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)*, pages 671–681. AAAI Press, 2023.

Yue Guan, Dipankar Maity, Christopher M. Kroninger, and Panagiotis Tsiotras. Bounded-rational pursuit-evasion games. In *American Control Conference (ACC)*, pages 3216–3221. IEEE, 2021.

Matthew O. Jackson and Yves Zenou. Games on networks. In *Handbook of game theory with economic applications*, volume 4, pages 95–163. Elsevier, 2015.

Anshul Kanakia, Behrouz Touri, and Nikolaus Correll. Modeling multi-robot task allocation with limited information as global game. *Swarm Intelligence*, 10(2):147–160, 2016.

Harold J. Leavitt. Some effects of certain communication patterns on group performance. *The journal of abnormal and social psychology*, 46(1):38, 1951.

Tao Li, Guanze Peng, Quanyan Zhu, and Tamer Basar. The confluence of networks, games, and learning a game-theoretic framework for multiagent decision making over networks. *IEEE Control Systems Magazine*, 42(4):35–67, 2022.

Jason R. Marden and Jeff S. Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75(2):788–808, 2012.

Richard D. McKelvey and Thomas R. Palfrey. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, 10(1):6–38, July 1995.

Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1): 124–143, 1996.

Andrea Montanari and Amin Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196–20201, 2010.

Herbert A. Simon. *Model of Man; social and rational*, volume 1957b:198. Wiley, 1957.

Brian Skyrms. *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press, Cambridge, 2004.

Panagiotis Tsiotras. Bounded rationality in learning, perception, decision-making, and stochastic games. In *Handbook of Reinforcement Learning and Control*, pages 491–523. Springer, 2021.

Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

Marcos M. Vasconcelos and Behrouz Touri. On the coordination efficiency of strategic multi-agent robotic teams. In *62nd IEEE Conference on Decision and Control (CDC)*, pages 8130–8137, 2023.

Yi Wei and Marcos M. Vasconcelos. Strategic multi-task coordination over regular networks of robots with limited computation and communication capabilities. In *57th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6, 2023.

Wang Xianjia, Xue Linzhao, Zhipeng Yang, and Yang Liu. Rationality-bounded adaptive learning in multi-agent dynamic games. *Knowledge-Based Systems*, 268:110459, 2023.

Yifei Zhang and Marcos M. Vasconcelos. Rationality and connectivity in stochastic learning for networked coordination games. In *American Control Conference (ACC)*, pages 1622–1627. IEEE, 2024a.

Yifei Zhang and Marcos M. Vasconcelos. On the role of network structure in learning to coordinate with bounded rationality. In *63rd IEEE Conference on Decision and Control (CDC)*, pages 1684–1689, Milan, Italy, December 2024b. IEEE. doi: 10.1109/CDC56724.2024.10886163.