Predicting deliberative outcomes

Vikas K. Garg [1] **Tommi Jaakkola** [1]

Abstract

We extend structured prediction to deliberative outcomes. Specifically, we learn parameterized games that can map any inputs to equilibria as the outcomes. Standard structured prediction models rely heavily on global scoring functions and are therefore unable to model individual player preferences or how they respond to others asymmetrically. Our games take as input, e.g., UN resolution to be voted on, and map such contexts to initial strategies, player utilities, and interactions. Players are then thought to repeatedly update their strategies in response to weighted aggregates of other players' choices towards maximizing their individual utilities. The output from the game is a sample from the resulting (near) equilibrium mixed strategy profile. We characterize conditions under which players' strategies converge to an equilibrium in such games and when the game parameters can be provably recovered from observations. Empirically, we demonstrate on two real voting datasets that our games can recover interpretable strategic interactions, and predict strategies for players in new settings.

1. Introduction

Structured prediction methods (Lafferty et al., 2001; Taskar et al., 2003; Tsochantaridis et al., 2005; Nowozin and Lampert, 2011) typically operate on parametric scoring functions whose maximizing assignment is used as the predicted configuration. Since the parameters can be learned directly to maximize prediction accuracy, often via surrogate losses, the methods have been successful across areas (Chen et al., 2015; Globerson et al., 2015; London et al., 2016; Cortes et al., 2016; Osokin et al., 2017; Pan and Srikumar, 2018).

Not all structured observations can be naturally modeled as extrema of scoring functions. For instance, votes cast in response to a bill in the US Congress likely involve rounds of negotiations prior to casting the final votes. As a result, the votes reflect individual utilities of representatives as well as their interactions with subsets of others. Absent some unifying scoring function, contentious votes are more naturally modeled in game theoretic terms as an equilibrium rather than a maximizing assignment (Waugh et al., 2011; Garg and Jaakkola, 2017). For example, one could consider the voting outcome directly as a *pure strategy* Nash equilibrium (PSNE) of some fixed underlying *graphical game* (Irfan and Ortiz, 2014; Honorio and Ortiz, 2015; Garg and Jaakkola, 2016; 2017). The observed votes would be in this case directly actions in the game, and complete observations help estimation. However, most games do not permit pure strategy equilibria (Tardos and Vazirani, 2007). The voting outcome is therefore best viewed as a sample from a product distribution that represents a mixed strategy equilibrium (guaranteed to exist for any game).

Our goal is to use games in a predictive sense. We must therefore explicitly model the game dynamics, i.e., how an equilibrium outcome is reached from the initial conditions. Previous methods for estimating game parameters from data often side-step the issue of game dynamics. It is therefore not possible in such approaches to directly output an equilibrium as the predicted outcome. For games to be useful in structured prediction, the game parameters as well as the initial strategies for the players, prior to deliberation, must be also conditioned on a common context such as a bill to be voted on.

Strategic prediction. In this paper, we turn games into structured prediction methods and call such approaches strategic prediction methods. At a high level, our model takes the available context such as the written bill – the input – and maps it to a mixed strategy equilibrium – the output. The input-output mapping is obtained via a deliberative process akin to multiple rounds of negotiations between the countries. The observed actions by players, e.g., how countries voted on the bill, are then viewed as samples from this mixed strategy equilibrium. We model the impact of context by parametrically mapping it to each player as player types. These types are simply learned vectors quantifying what each player derives from the context so as to guide their behavior in the game. The types also specify the utility func-

[1]CSAIL, MIT. Correspondence to: Vikas Garg <vgarg@csail.mit.edu>, Tommi Jaakkola <tommi@csail.mit.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020. Copyright 2020 by the author(s).

tions and initial strategies for players. A key part of strategic prediction pertains to how players interact with each other. In this paper, we limit our focus to so called *aggregative games* where the players respond to weighted aggregates of other players' actions or strategies. The interactions are weighted and asymmetric, and the weights can be positive or negative representing cooperative or competing relationships between the players. We follow a best-response game dynamics to evolve players' initial strategies towards the predicted output, a (near) mixed strategy equilibrium. We call our parametrized games *strategic aggregative prediction* (SAP) models.

Our games can be viewed as conditional versions of directed graphical games (Kearns et al., 2001; Kearns and Mansour, 2002), restricted to a rich subclass of aggregative games that subsumes Cournot oligopoly, mean field, public goods, and population games (Garg and Jaakkola, 2017). These games shield each player from specific information about any neighbor since players respond to aggregate (weighted sum) of their neighbors' strategies.

Modeling game dynamics. A key novelty of our approach is to explicitly incorporate game dynamics. In our implementation, we follow a k-step best response dynamics, seeded with predicted initial strategies. As we operate on continuous strategies and our best response updates are differentiable, we can back-propagate through the k-step strategy updates, and thus learn parameters efficiently, unlike, e.g. (Garg and Jaakkola, 2016; 2017). Since game dynamics plays a critical role in predictive use of games, we also introduce and provide a deeper analysis of more general dynamics and types of aggregation strategies. Our work yields exact conditions under which strategies converge to different types of equilibrium.

Identifiability. We provide identifiability guarantees for the game parameters. For the analysis, we adopt a simpler *one-shot* setting, where the observed outcome is sampled from player strategies after just one round of communication instead of following k-steps. We characterize conditions under which one-shot interaction weights are identifiable in the sense that we can recover the neighbors of any player as well as the correct sign (positive or negative) of their interaction. Prior guarantees exist for the recovery of equilibria, not game parameters, showing conditions for the recovery of a set of pure strategy equilibria under various noise assumptions (Ghoshal and Honorio, 2017).

The rest of this paper is structured as follows. The basic setup is introduced in section 2. We describe parameter estimation in section 3, and extension to new players in section 4. Identifiability is discussed in section 5, and convergence of dynamics in section 6. We present detailed experiments on two real datasets to illustrate the benefits of our approach in section 7. We defer the details of our proofs to the Supplementary.

2. Basic strategic prediction model

We first introduce our basic strategic prediction model. To this end, we need to define several components of the model. These include (a) the graphical layout of the game, and how players influence each other; (b) the player types and how these are derived from the context; (c) individual utilities for the players; (d) initial strategies for the players before witnessing the play of others; (e) and the game dynamics, i.e., how players respond to others. Later, in the *transferable* setting, we will no longer individuate players through their identities but instead introduce feature vectors to predict the strategies of new players in new contexts.

Let $G = (V, E)$ be a connected digraph such that vertex i identifies player $i \in [n] \triangleq \{1, 2, \ldots, n\}$, $n = |V|$. Let A be the finite discrete set of actions, the same for all the players. Each player i plays a randomized (mixed) strategy, i.e., a distribution over actions $\sigma_i \in \triangle(A)$, from the simplex

$$\triangle(A) \triangleq \left\{ \sigma_i \;\middle|\; \sum_{a_i \in A} \sigma_i(a_i) = 1, \sigma_i(a_i) \geq 0 \;\; \forall a_i \in A \right\}.$$

We will denote a joint strategy profile of all the players by (σ_i, σ_{-i}) to emphasize the distinction between player i and the other players.

We model the influence of players on others through a weighted aggregation of the strategies. The weights $w_{ij} \in \mathbb{R}$ denote the strength of influence of player j on player i. We will call players $j \in [n] \setminus \{i\}$ that have $w_{ij} \neq 0$ the *neighbors* of players i. We define a weight matrix $W \in \mathbb{R}^{n \times n}$ such that $W(i,i) = 0$ and $W(i,j) = w_{ij}$. Player i communicates with other players only through aggregator $\mathcal{A}_i$ that maps the strategies of other players, i.e., σ_{-i} to the weighted sum $\sum_{j \neq i} w_{ij} \sigma_j$, the effective influence of others.

The context influences the game through the types. The type of a player is her valuation that quantifies what she derives from the context. The mapping from context to types could be defined in multiple ways. For simplicity, we parameterize the private type z_i of each player i via a matrix θ_i that maps any context $x \in \mathcal{X} \subseteq \mathbb{R}^d$ to vector $z_i(x) = \theta_i x$. Henceforth, we will keep the dependence on context implicit for simplicity, and write $z_i(x)$ as z_i when the context is clear.

We model the utility or payoff of player $i \in [n]$, of type z_i, under strategy profile (σ_i, σ_{-i}) as:

$$U_i(\sigma_i, \sigma_{-i}, z_i) = \sigma_i^\top \left(\mathcal{A}_i(\sigma_{-i}) - z_i \right) + \tau \mathbb{H}(\sigma_i) \,, \quad (1)$$

where $\mathbb{H}(\sigma_i)$ is the entropy associated with σ_i and $\tau \geq 0$. The entropy encourages *completely mixed* strategy choices, in the interior of simplex $\Delta(A)$. The type of a player is by design hidden from the other players. Thus by subtracting the player type, the payoff function mitigates the (indirect) role of player's valuation in guiding the aggregate response of the others. Our games are flexible to accommodate a wide variety of other payoff functions.

Our choice of payoff function (1) maintains a linear dependence on the strategies of others for two reasons. First, our games are related to *linear influence games* that have been studied previously, e.g., as models of how to vaccinate against a disease, install antivirus software, or get home insurance (Irfan and Ortiz, 2014; Honorio and Ortiz, 2015; Ghoshal and Honorio, 2017). We extend these games by allowing multi-way actions and types. Second, the linear dependence also helps establish provable recovery guarantees for the game parameters.

The utilities naturally define the *best response* of player i when it observes the aggregate input $\mathcal{A}_i(\sigma_{-i})$:

$$\beta_i^{\tau}(\mathcal{A}_i(\sigma_{-i}), z_i) \in \arg\max_{\sigma_i \in \triangle(A_i)} U_i(\sigma_i, \sigma_{-i}, z_i) \,. \tag{2}$$

We say that $(\sigma_i^*, \sigma_{-i}^*, z_i, z_{-i})$ form a *mixed strategy Nash equilibrium MSNE* (MSNE) or simply NE iff

$$U_i(\sigma_i^*, \sigma_{-i}^*, z_i) \;\geq\; U_i(\sigma_i, \sigma_{-i}^*, z_i) \quad \forall i \in [n], \sigma_i \in \triangle(A) \,. \tag{3}$$

It is easy to show that our game has at least one MSNE. For $\tau > 0$, our best response function is continuous, and maps each strategy profile from product simplex (compact space) to a unique point in itself. The Nash equilibrium is therefore guaranteed to exist by Brouwer's fixed point theorem. For $\tau = 0$, the mapping may be set-valued, but the Nash equilibrium is still guaranteed to exist by Nash's theorem (Nash, 1951) via Kakutani's fixed point argument.

We say that MSNE is *strict* (SNE) when (3) is strict for all $i \in [n], \sigma_i \in \Delta(A) \backslash \{\sigma_i^*\}$, *completely mixed* (CMNE) when $\sigma_i^*(a_i) > 0$ for all $i \in [n], a_i \in A$, and *pure* (PSNE) when for all $i \in [n]$ there exists an $a_i \in A$ such that $\sigma_i^*(a_i) = 1$.

It remains to specify how an equilibrium is reached, i.e., the game dynamics. To begin with, players observe context x, and evaluate types z_i. The types, in turn, give rise to initial strategies $\sigma_i^0 = \psi(z_i)$, where $\psi : T_i \to \Delta(A)$ (e.g. softmax). The best response dynamics from this point on depends on the details of the aggregator and whether the dynamics is defined over strategies or actions directly. We provide details on general game dynamics for two different aggregators along with associated convergence guarantees in section 6. In our empirical analysis, we adopt a k-step dynamics as described in the following section. Once a (near) equilibrium is reached, a sample action profile $y \in \mathcal{Y} \subseteq A^n$ is drawn from player strategies.

We emphasize that our setting dispenses with the restrictive assumption made by Bayesian games (Harsanyi, 1967; Kalai, 2004; Jiang and Leyton-Brown, 2010) that the conditional distribution $P(z_{-i}|z_i)$ is known to player i.

3. Parameter estimation

We learn our games from data as structured prediction methods. Specifically, given a dataset $D = \{(x^{(m)}, y^{(m)})\} \in \mathcal{X} \times \mathcal{Y}, m \in [M]\}$ linking contexts to sampled action profiles, our objective is to estimate the type parameters θ_i and the influences of neighbors $w_i \triangleq (w_{ij})_{j \neq i}, i \in [n]$.

Each pair $(x^{(m)}, y^{(m)})$ is treated as follows. A linear transformation $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$ maps the context $x^{(m)}$ to the types $\hat{z}(x^{(m)}) \triangleq (\hat{z}_1^{(m)}, \ldots, \hat{z}_n^{(m)})$ that result in initial strategies $\hat{\sigma}_i^0(x^{(m)}) = \zeta(\hat{z}_i^{(m)})$ of the players $i \in [n]$, where ζ is the softmax nonlinearity. The aggregators $\hat{\mathcal{A}}_i$ evaluate weighted sums, and are parametrized by weights $\hat{w}_i = (\hat{w}_{ij})_{j \neq i}$. A sequence of k update steps

$$\hat{\sigma}_i^{t+1}(x^{(m)}) = \zeta(\nu(\hat{\sigma}_i^t) + \alpha(\hat{\mathcal{A}}_i(\hat{\sigma}_{-i}^t) - \hat{z}_i^{(m)})), \tag{4}$$

$t \in \{0, 1, \ldots, k-1\}$ is then followed: k and α are hyperparameters, and ν defines the type of update. Several choices of ν are possible; e.g., $\nu(\hat{\sigma}_i^t) = 0$ pertains to best response $\beta_i^{1/\alpha}(\hat{\mathcal{A}}_i(\hat{\sigma}_{-i}^t), \hat{z}_i^{(m)})$ defined in (2), the identity mapping $\nu(\hat{\sigma}_i^t) = \hat{\sigma}_i^t$ defines a gradient step, and $\nu(\hat{\sigma}_i^t) = \log \hat{\sigma}_i^t$ corresponds to a proximal update based on KL-divergence: we define $\hat{\sigma}_i^{t+1}$ as

$$\arg\max_{\sigma_i} \;\; \sigma_i^\top(\hat{A}_i(\hat{\sigma}_{-i}^t) - \hat{z}_i^{(m)}) + (1/\alpha)\mathrm{KL}(\sigma_i || \hat{\sigma}_i^t) \,.$$

Note that we can use functions other than softmax for our strategy updates. For instance, we can take a gradient step and then project on the probability simplex. One attractive feature of softmax is that it is differentiable and its gradient can be computed in a closed form, so it can be readily backpropagated through in a neural model. On the other hand, projection on the probability simplex typically requires specialized optimization algorithms such as (Duchi et al., 2008) or (Condat, 2016).

Our estimation criterion for the game is to minimize expected cross-entropy loss $\mathbb{E}[\ell(\hat{\sigma}^k(x^{(m)}), y^{(m)}]$ between the predicted mixed strategies and the observed profiles, where the expectation is with respect to the empirical distribution over pairs $(x^{(m)}, y^{(m)})$, ℓ is the cross entropy loss, and $\hat{\sigma}^k(x^{(m)}) \triangleq (\hat{\sigma}_i^k(x^{(m)}))_{i \in [n]}$. We use standard backpropagation to evaluate gradients through the k-step strategy updates efficiently.

Note that since we do not assume that the observed configurations $y^{(m)}$ correspond to pure strategy Nash equilibria, we naturally avoid having to enforce specific consistency constraints. Instead, we can measure a degree of agreement that is amenable to more complex parameterizations such as neural networks. Our estimation procedure also scales better than methods that rely on PSNEs, e.g., (Garg and Jaakkola, 2016; 2017). The overall dependence of our approach is linear in number of observed game outcomes and the number of update steps, and quadratic in number of players.

4. Transferable strategic prediction

Here we generalize SAP models to permit different players from one game to another. Unlike in section 2, we can no longer assume a fixed interaction structure across games. Instead, the neighbor influences are determined by context and player *feature vectors*. Specifically, we construct a feature vector $b_i \in \mathcal{B}$ for each player $i \in [n]$. Such information is often available, e.g., education and gender of judges; human development indicators of countries, etc. This enables us to predict the behavior of new players in new contexts.

Each game is played with a different subset of players $\mathcal{I} \subseteq [n]$, and is unrolled as follows. A context $x \in \mathcal{X}$ is mapped to latent (more general) player types using a parametric function $f_z : \mathcal{X} \times \mathcal{B} \to \mathcal{Z}$, taking each pair $(x, b_v), v \in \mathcal{I}$ as input, and mapping it to $z_{x,v} \in \mathcal{Z}$. These types define initial strategies as before $\sigma^0_{x,v} = \phi(\Gamma z_{x,v})$, where Γ is a transformation matrix that yields a vector in $\mathbb{R}^{|A|}$, and ϕ (e.g., softmax) maps the result to a distribution in the simplex $\Delta(A)$. Unlike before, the (asymmetric) influences between players are now calculated parametrically from the types: $w_{x,v,v'} = f_w(z_{x,v}, z_{x,v'})$ using a parametric mapping $f_w : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$. Each player v still responds to other players $v' \in \mathcal{I} \setminus \{v\}$ through its aggregator

$$\mathcal{A}_{x,v,\mathcal{I}}(\sigma_{x,-v}) \triangleq \sum_{v' \in \mathcal{I}: v' \neq v} w_{x,v,v'} \sigma_{x,v'} \ .$$

We can extend the definition of the payoffs slightly to incorporate the more general types:

$$U_{v,\mathcal{I}}(\sigma_{x,v}, \sigma_{x,-v}, z_{x,v}) = \sigma_v^\top \left(\mathcal{A}_{x,v,\mathcal{I}}(\sigma_{x,-v})\right) - \Gamma z_{x,v}) \ .$$

where Γ is an additional parameter matrix to be learned. The game dynamics dictates the course of play in the same fashion as the basic strategic setting. We learn the model, now parameterized by f_z, f_w, and Γ, by minimizing the loss between predicted k-step strategies and observed action profiles.

5. Identifiability of the games

We now characterize the conditions under which the strategic interactions in our games become identifiable. We focus on the *one-shot* setting (i.e., $k = 1$) with the gradient update so that the observed outcome is sampled from player strategies after one round of communication. We also use binary actions to simplify the exposition. Specifically, we estimate from data D the *support* S_i or the set of neighbors of i defined with respect to the unknown influences $w^*_{ij} \neq 0$. We also recover the correct sign of these influences.

Our recovery procedure is a novel adaptation of the primal-dual witness method (Wainwright, 2009) for structure estimation in games. The method has previously been applied in several non-strategic settings such as Lasso (Wainwright, 2009) and Ising models (Ravikumar et al., 2010). In contrast to Lasso and Ising models, our setting poses some new challenges. Unlike these models, the nodes of our games are players that actively communicate with others, and refine their strategies toward maximizing their individual utilities. Moreover, context and dynamics play no part in Ising models. In our setting, each outcome is sampled from a separate joint strategy profile following one step of dynamics initiated under a different context.

(Ghoshal and Honorio, 2017) employed the witness method to recover the entire set of PSNE from data consisting of a subset of PSNE, and a small fraction of non-equilibrium outcomes assumed to be sampled under their noise models in the setting of linear influence games. However, the problem of structure recovery is significantly harder: it is known (Honorio and Ortiz, 2015; Ghoshal and Honorio, 2017) that the problem becomes non-identifiable in the setting of PSNE, since multiple game structures may pertain to the same of PSNE. We leverage the dynamics to circumvent these issues, and characterize the conditions when we can provably recover the structure of our games.

Our analysis on recovering the influences makes the simplifying assumption that the type parameters θ are known. We denote by $\phi^{(m)}_j$ the probability assigned by the initial strategy $\sigma^0_j(x^{(m)})$ to action 1 for player j on example m. Let $\ell_i(w_i; D)$ be the average cross-entropy loss between one-step strategies under candidate weights $w_i \triangleq (w_{ij})_{j \neq i}$ and the observed actions for player i, and $\lambda > 0$ be a regularization parameter. We solve the following problem for each player $i \in [n]$:

$$\arg \min_{w_i \in \mathbb{R}^{n-1}} \ \ell_i(w_i; D) \ + \ \lambda ||w_i||_1 \ , \tag{5}$$

Let H^M_i be the sample Hessian $\nabla^2 \ell_i(w_i; D)$ under w_i, and H^{*M}_i pertain to true weights w^*_i. Let $\Lambda_{min}(\cdot)$ and $\Lambda_{max}(\cdot)$ denote the minimum and the maximum eigenvalues. The following assumptions serve as our analogues of the conditions for support recovery in Lasso (Wainwright, 2009) and Ising models (Ravikumar et al., 2010):

$$\Lambda_{\min}\left(H^{*M}_{i,SS}\right) \quad \geq \quad \alpha^2 C_{min} \tag{6}$$

$$\Lambda_{\max}\left(\frac{1}{M}\sum_{m=1}^{M} \Phi^{(m)}_{-i} {\Phi^{(m)}_{-i}}^\top\right) \quad \leq \quad C_{max} \tag{7}$$

$$|||H^{*M}_{i,S^cS}(H^{*M}_{i,SS})^{-1}|||_\infty \quad \leq \quad 1 - \gamma \ , \tag{8}$$

where $\Phi^{(m)}_{-i} \triangleq (\phi^{(m)}_j)_{j \neq i}$, $C_{\min} > 0$, $C_{\max} < \infty$,

$\gamma \in (0, 1]$; $|||A|||_\infty$ is the maximum over L_1-norm of rows in A; $H^{*M}_{i,SS}$ is the submatrix obtained by restricting H^{*M}_i to rows and columns corresponding to neighbors, i.e., players in S_i, and H^{*M}_{i,SS^c} is restricted to rows pertaining to S_i and columns to S_i^c (non-neighbors). Let the number of neighbors for any player be at most $d \leq n-1$. We have the following result.

Theorem 1. *Let* $M > \dfrac{80^2 C^2_{\max}}{C^4_{\min}} \left(\dfrac{2-\gamma}{\gamma}\right)^4 d^2 \log(n)$*, and* $\lambda_M \geq \dfrac{8\alpha(2-\gamma)}{\gamma}\sqrt{\dfrac{\log(n)}{M}}$*. Suppose the sample satisfies assumptions* (6), (7)*, and* (8).

Define $C_{\alpha,\gamma} = \dfrac{\gamma^2}{32\alpha^2(2-\gamma)^2}$ *. Consider any player* $i \in [n]$*. The following results hold with probability at least* $1 - 2\exp(-C_{\alpha,\gamma}\lambda_M^2 M) \to 1$ *for* i*.*

1. *The corresponding ℓ_1-regularized optimization problem has a unique solution, i.e., a unique set of neighbors for i.*

2. *The set of predicted neighbors of i is a subset of the true neighbors. Additionally, the predicted set contains all true neighbors j for which $|w^*_{ij}| \geq \dfrac{10}{\alpha^2 C_{\min}}\sqrt{d}\lambda_M$. In particular, the set of true neighbors of i is exactly recovered if*

$$\min_{j \in S_i} |w^*_{ij}| \geq \frac{10}{\alpha^2 C_{\min}}\sqrt{d}\lambda_M \ .$$

Note that taking a union bound over players, our results imply that we recover the true signed neighborhoods for all players in our game with high probability. Our analysis can possibly be extended to include unknown type parameters θ by computing the gradient and the Hessian of loss with respect to θ, similarly to influences. However, additional assumptions may be needed to establish recovery.

6. General game dynamics and convergence

We now give an overview of general game dynamics along with associated convergence guarantees. The aggregator acts as a privacy preserving component, hiding specific neighbor actions or strategies, only offering aggregate statistics. We design dynamics under two different kinds of feedback from the aggregator. In our *active aggregator* (AA) setting, the players get a prediction about the anticipated aggregate of their neighbors. In contrast, a *passive aggregator* (PA) only provides the aggregate of *empirical frequencies* used by the neighbors, and changes in the aggregate are estimated by the player. Intuitively, AA reveals more information about the neighbors' strategy evolution.

Note that the payoff received by player i when she samples a pure action $a_i \in A$ according to σ_i (denoted by $a_i \sim \sigma_i$), and others sample actions a_{-i} according to σ_{-i} is

$$U_i(e_{a_i}, e_{a_{-i}}, z_i) = e_{a_i}^\top \left(\mathcal{A}_i(e_{a_{-i}}) - z_i\right) + \tau \mathbb{H}(\sigma_i),$$

where $e_{a_i} \in \triangle(A)$ is the strategy corresponding to pure action a_i, and $e_{a_{-i}} \triangleq \{e_{a_j} : j \neq i\}$.

We devise two new protocols that may be viewed as derivative action adaptations (Shamma and Arslan, 2005) of smooth fictitious play (FP) and gradient play (GP) (Brown, 1951; Robinson, 1951; Fudenberg and Kreps, 1993; Shapley, 1964) in an aggregative game setting. The protocols differ by how players respond to the (predicted) aggregate: one can play the best response or adapt the strategy via a gradient update. Formally, in our protocols, player i samples an action $a_i^k \sim \sigma_i^k$ at time $k > 0$ based on

$$\begin{aligned} q_i^k &= q_i^{k-1} + (e_{a_i^{k-1}} - q_i^{k-1})/k \in \Delta(A) \\ \sigma_i^k &= g_i(\mathcal{A}_i(h_{-i}(q_{-i}^k)), z_i) \ , \end{aligned} \tag{9}$$

where q_i^k is the empirical frequency of actions played by i till time k, and $g_i : \mathbb{R}^{|A|} \times \mathbb{R}^{|A|} \to \Delta(A)$ and $h_i : \mathbb{R}^{|A|} \to \mathbb{R}^{|A|}$ are appropriately defined Lipschitz mappings possibly involving small input noise. We let $q_{-i}^{k-1} \triangleq \{q_j^{k-1} | j \neq i\}$, and $h_{-i}(q_{-i}^{k-1}) \triangleq \{h_j(q_j^{k-1}) | j \neq i\}$. We also define the base case $q_i^0 = \sigma_i^0 = \phi(z_i)$. Note that player i communicates only with $\mathcal{A}_i$. We define a *passive* aggregator (PA) by letting h_i be the identity mapping, i.e. $h_i(v_i) = v_i$. Alternatively, when $h_i(v_i) = v_i + \gamma \nabla \tilde{v}_i$ for some $\gamma > 0$ and a difference approximation $\nabla \tilde{v}_i$ of a temporal derivative ∇v_i, we obtain an *active* aggregator (AA). Intuitively, AA views each q_j as discretization of a continuous signal $q_j(t)$ so that when $\nabla \tilde{q}_j(t) \approx \nabla q_j(t)$, for neighbors j of i, we have

$$\begin{aligned} h_j(q_j(t)) &\approx q_j(t) + \gamma \nabla q_j(t) \approx q_j(t+\gamma) \\ &\implies \mathcal{A}_i(h_{-i}(q_{-i}(t))) \approx \mathcal{A}_i(q_{-i}(t+\gamma)), \end{aligned}$$

and therefore $\mathcal{A}_i$ offers a predicted aggregate to player i. We consider two forms of best response dynamics encoded in g_i, SAP-FP and SAP-GP, based on derivative FP and derivative GP, respectively. In SAP-FP we set $\tau > 0$ in the utility functions. This lets us have a unique best response:

$$\begin{aligned} &(\textbf{SAP-FP}) \\ \text{AA yields } \overline{u}^k:\ g_i(\overline{u}^k, z_i) &= \beta_i^\tau(\overline{u}^k, z_i), \\ \text{PA yields } u^k:\ g_i(u^k, z_i) &= \beta_i^\tau(u^k + \gamma \nabla \hat{u}^k, z_i), \end{aligned}$$

where the AA case differs from PA in terms of where the difference approximation happens. In AA, it happens prior to aggregation thus g_i is defined directly in terms of the output of AA or $\bar{u}^k$ which absorbs any temporal approximation error. In PA case, the player constructs a temporal prediction of the aggregate, and the approximation is $(\nabla u^k - \nabla \hat{u}^k)$.

$$
\begin{aligned}
&(\textbf{SAP-FP/AA}) & \dot{q}_i &= \beta_i^{\tau}(\mathcal{A}_i(q_{-i}+\gamma\dot{r}_{-i}), z_i) - q_i, & \dot{r}_i &= \lambda(q_i - r_i) && (10)\\
&(\textbf{SAP-FP/ PA}) & \dot{q}_i &= \beta_i^{\tau}(\mathcal{A}_i(q_{-i})+\gamma\dot{r}_i, z_i) - q_i, & \dot{r}_i &= \lambda(\mathcal{A}_i(q_{-i}) - r_i) && (11)\\
&(\textbf{SAP-GP/AA}) & \dot{q}_i &= \Pi_{\Delta}[q_i + \mathcal{A}_i(q_{-i}+\gamma\dot{r}_{-i}) - z_i] - q_i, & \dot{r}_i &= \lambda(q_i - r_i) && (12)\\
&(\textbf{SAP-GP/PA}) & \dot{q}_i &= \Pi_{\Delta}[q_i + \mathcal{A}_i(q_{-i}) + \gamma\dot{r}_i - z_i] - q_i, & \dot{r}_i &= \lambda(\mathcal{A}_i(q_{-i}) - r_i) && (13)
\end{aligned}
$$

In SAP-GP, we set $\tau = 0$, and player i takes a gradient step to maximize the anticipated payoff followed by a Euclidean projection to get a unique mapping g_i (since any such projection on a closed convex set is unique):

$$
\begin{aligned}
&(\textbf{SAP-GP})\\
\text{AA yields } \overline{u}^k: \quad & g_i(\overline{u}^k, z_i) = \Pi_{\Delta}(q_i + \overline{u}^k - z_i),\\
\text{PA yields } u^k: \quad & g_i(u^k, z_i) = \Pi_{\Delta}(q_i^k + u^k + \gamma\nabla\hat{u}^k - z_i),\\
\text{where} \quad & \Pi_{\Delta}(q) \triangleq \text{argmin}_{\tilde{q}\in\triangle(A)} ||\tilde{q} - q||_2 .
\end{aligned}
$$

Thus under both protocols, players take actions stochastically according to σ_i^k and the best response mapping g_i is unique for each k. For analysis, we associate an ODE system to characterize the evolution of player strategies, and specify conditions when the fixed points of this ODE are *locally asymptotic stable* (*l.a.s.*). An equilibrium point s is said to be *l.a.s.* if every ODE trajectory that starts at a point in a small neighborhood of s remains forever in that neighborhood and eventually converges to s. As a consequence, our discrete updates would converge to a Nash equilibrium with positive probability (Shamma and Arslan, 2005). Our updates in (9) lead to the implicit ODEs (39)-(42) for SAP-FP and SAP-GP under AA and PA settings, where $\lambda > 0$, $\dot{r}_i$ is an estimate for $\dot{q}_i$, and $\dot{r}_{-i} \triangleq \{\dot{r}_j | j \neq i, w_{ij} \neq 0\}$. We will call a matrix *stable* if all its eigenvalues have strictly negative real parts. Let I denote the identity matrix. We now state results that characterize conditions under which different dynamics lead to asymptotically stable equilibria.

We state and prove our convergence results as theorems in the Supplementary (see Table 1 for a summary). Specifically, we prove the convergence of player strategies to SNE via carefully crafted *Lyapunov functions* that are locally positive definite and have a locally negative semidefinite time derivative. The other proofs track the evolution of game dynamics around an equilibrium: we analyze conditions under which the Jacobian matrix of the linearization of ODE is *Hurwitz stable*, i.e., all the eigenvalues have negative real roots, and exploit the fact that the behavior of the ODE near equilibrium is same as its linear approximation when all eigenvalues have non-zero real parts.

Our results have several important implications. (Garg and Jaakkola, 2016; 2017) enforced margin constraints on the payoff functions in their discriminative PSNE setup, without establishing SNE that guarantees the players would be *strictly* worse off by a unilateral switch to another strategy. In contrast, Theorems 5 and 7 specify exactly the conditions for SNE, which under additional assumptions, may be sampled to get a PSNE (Azrieli and Shmaya, 2013) if desired. It is known that classic FP fails to converge in some simple games, e.g. (Shapley, 1964), that have a unique CMNE. Theorems 4 and 6 specify conditions that circumvent such negative results. Our stability conditions can be simplified further when λ is sufficiently large, whence the behavior may be understood solely in terms of γ.

Table 1: Convergence results (details in the Supplementary)

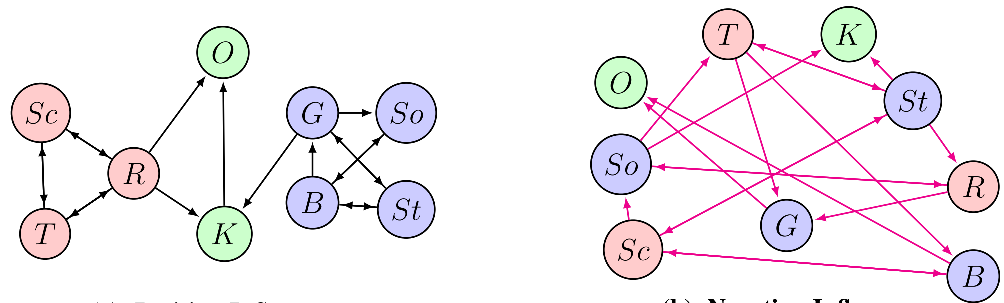
Setting	Converges to	Conditions
SAP-FP/AA	NE	**Theorem 2**
SAP-FP/PA	NE	**Theorem 3**
SAP-GP/AA	CMNE	**Theorem 4**
SAP-GP/AA	SNE	**Theorem 5**
SAP-GP/PA	CMNE	**Theorem 6**
SAP-GP/PA	SNE	**Theorem 7**

7. Experiments

We now describe the results of our experiments on two real datasets that provide insights into some important aspects of our games. We first show that our SAP model qualitatively recovers the known strategic behavior of the Justices from US Supreme Court data. We also provide quantitative evidence in terms of two measures: (a) the fraction of correct edges recovered, and (b) coherence of groups in terms of cut-size, to show that SAP game outperforms prior methods on two different measures. We then show that the structure estimated by SAP game on the UN General Assembly data (Voeten et al., 2009) is meaningful, and helps unravel the subtle behavior of member countries. Finally, we demonstrate that SAP games incur a lower cross-entropy loss than our baseline, so can be effectively transferred to predict strategies in new settings with different sets of players.

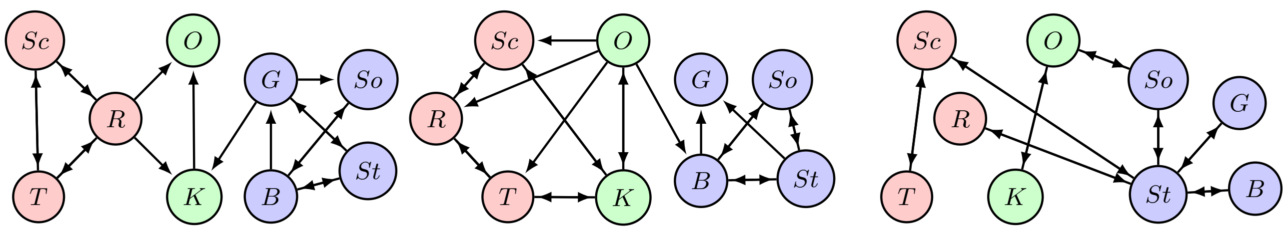
Experimental setup

We found that SAP games performed well over a wide range of hyperparameters. We implemented the models in sections 3 and 4 with L_1 regularization for structure estimation as described in section 5. Our models performed well for a wide range of α, λ, and k. We report the results with $k = 5$,



(a) Positive Influence **(b) Negative Influence**

Figure 1: **Supreme Court structure recovery with our SAP model**: (a) and (b) show, respectively, justices with the most positive and the most negative influence, quantified by $\hat{w}_{ij}$, on each Justice i. The estimated connections are consistent with the known jurisprudence of the Court. In particular, (a) shows coherence between the conservatives (red), that between the liberals (blue), and the separation of these ideologies from the moderates (green). Likewise, (b) shows all the negative connections are between the blocs. Determining influences by heuristics such as ordering by pairwise vote agreements does not work; e.g, that would imply K had a strong positive influence on R, since R agreed with K more than with anyone else.



(a) SAP (Ours) **(b) Local AG** **(c) Tree Structured Potential Game**

Figure 2: **Quantitative comparison with prior methods**: (a), (b), and (c) show the structures estimated by SAP, local AG, and tree structured potential game on US Supreme Court data. We evaluate the different methods using size of the *cut*, i.e., minimum number of edges to be removed to decompose the structure into its three components (reds, blues, and greens). A low value of cut quantifies high coherence within each component, and thus pertains to a good structure. The cut size for SAP game (3) is much lower than other methods (6 each).

$\alpha = 0.1$, and $\lambda = 0.1$ for all our experiments, except the transferable setting where we set $\alpha = 0.01$ and $\lambda = 0$. We set ν to be the identity function in (4). We trained our models in batches of size 200, with default settings of the RMSprop optimizer in PyTorch. To account for effect of randomness in neural training toward structure estimation, we averaged the parameters of each model across 5 independent runs. We characterize the influence of other players j on player i in terms of the ordering of the corresponding estimated average weight values from most positive to most negative.

We did not impose L_1 penalty in the transferable setting since the interactions are learned for new individuals, i.e., they are not specific to any fixed set of players, unlike the basic setting.

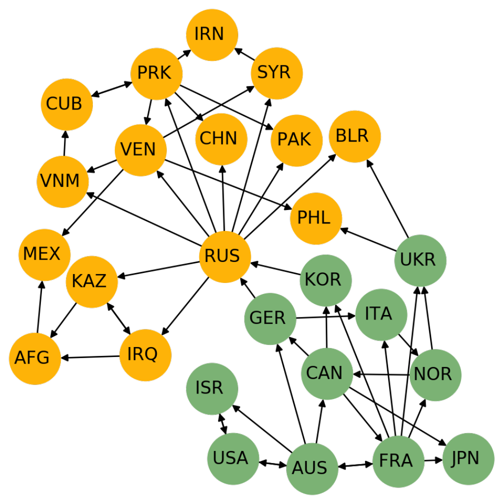
7.1. US Supreme Court Data

We included all the cases from the *Rehnquist Court*, during the period 1994-2005 that had votes documented for all the 9 Justices.[1] Justices Rehnquist (R), Scalia (Sc), and Thomas (T) represented the conservative side; Justices Stevens (St), Souter (So), Ginsburg (G), and Breyer (B) formed the liberal bloc; and Justices Kennedy (K) and O'Connor (O), often called *swing votes*, followed a moderate ideology. Each Justice is treated as a player in our game. Our contexts comprise of 32 binary attributes about the specifics of the appeal, e.g., the disposition of lower court. Our observation outcome $y^{(m)}$ for each context $x^{(m)}$ pertains to the corresponding votes of the Justices. The votes belong to one of the three categories: yes, no, or complex.

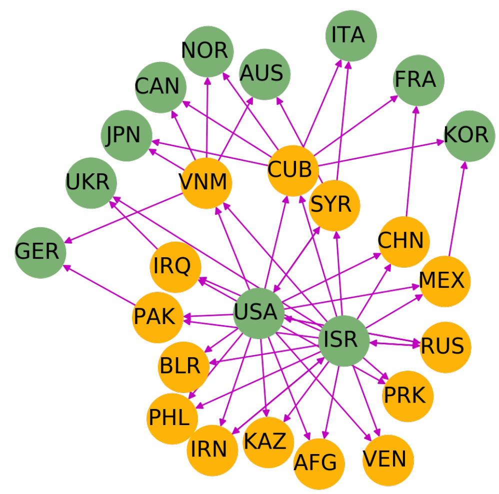
Structure recovery from Supreme Court data

Fig. 1 describes in detail how our method yields a structure that is qualitatively consistent with the known jurisprudence of the Rehnquist Court. Specifically, the conservatives and the liberals form two separate coherent, strongly-connected blocs that are well-segregated from each other.

[1] data available at: http://scdb.wustl.edu/data.php



(a) **Positive influence**



(b) **Negative influence**

Figure 3: **UNGA structure recovery with SAP games**: Incoming arrows show (a) 2 countries with the most positive influence (black edges), and (b) 2 countries with the most negative influence (magenta edges), quantified by $\hat{w}_{ij}$, on each country i. The estimated links are largely consistent with the expected alignments. In particular, (a) shows the two blocs (in yellow and green) are well segregated from each other. More interesting alignments are revealed, e.g., (1) strong affinity between NATO members on one side, and Syria, Iran, Venezuela etc. on the other, (2) link from Germany and Korea to Russia hinting at the trade influence despite their differences, and (3) geographical influence of Russia and Ukraine on Belarus. Additionally, (b) reveals that a significant fraction of negative connections emanate from or end at Israel and USA on one side, and some yellow node on the other.

Comparison with previous works

In order to position our work with respect to prior work on structure recovery, we also quantitatively compare SAP games with the PSNE based Potential Game (Garg and Jaakkola, 2016) and Local Aggregative Game (Garg and Jaakkola, 2017) methods.

Fig. 2 provides a detailed quantitative comparison that reveals how SAP compared favorably with these measures in terms of size of the *cut*, i.e., minimum number of edges to be removed in order to decompose the recovered structure into the constituent blocs. Note that a low value of cut indicates coherence within each component, and thus quantifies a good structure.

7.2. United Nations General Assembly (UNGA) Data

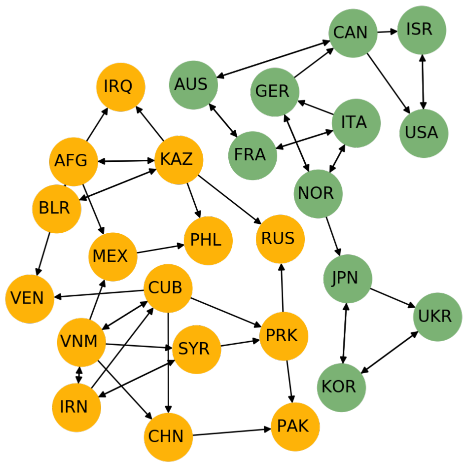
Our second dataset consists of the roll call votes of the member countries on the resolutions considered in the UN General Assembly. Each resolution is a textual description that provides a context while the votes of the countries on the resolution pertain to the observed outcome. We compiled data on all resolutions in UNGA during 1992-2017 to understand the interactions of member nations since the dissolution of Soviet Union.

We considered 25 countries that have dominated the United States (USA) politics, and are generally known to belong to one of the two blocs: pro-USA, namely, Australia (AUS), Canada (CAN), France (FRA), Germany (GER), Israel (ISR), Italy (ITA), Japan (JPN), South Korea (KOR), Norway (NOR), Ukraine (UKR); and others, namely, Afghanistan (AFG), Belarus (BLR), China (CHN), Cuba (CUB), Iran (IRN), Iraq (IRQ), Mexico (MEX), Pakistan (PAK), Philippines (PHL), North Korea (PRK), Kazakhstan (KAZ), Russia (RUS), Syria (SYR), Venezuela (VEN) and Vietnam (VNM). Each country is viewed as a player in our game.

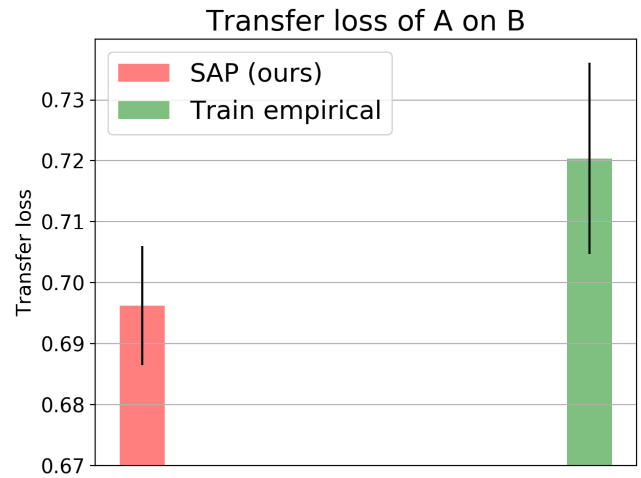
We used pretrained GLoVe embeddings to represent each resolution as a 50-dimensional context vector $x^{(m)}$ obtained by taking the mean embedding of the words in its resolution. Each vote was interpreted to take one of the three values: 1 (yes), 2 (absent/abstain), or 3 (no), and we represented $y^{(m)}$ as a 26-dimensional vector.

Structure and type recovery from UNGA data

Fig. 3 shows the structure estimated by our method, i.e., the weights $\hat{w}_i$ learned for each country i. The weights $\hat{w}_{ij}, j \neq i$ have a natural interpretation in terms of influence: the more positive w_{ij} is, the more positive the influence of j on i. A similar connotation holds for the negative weights. To aid visualization, we disentangle the positive



(a) **Type similarity**



(b) **Transfer performance**

Figure 4: **(Left: UNGA type recovery)** Incoming arrows show 2 countries with the highest cosine similarity $\hat{\theta}_i^\top \hat{\theta}_j / (||\hat{\theta}_i||\ ||\hat{\theta}_j||)$ for each country i. Type vectors were reasonably well aligned for members in the same bloc. **(Right: transfer performance)** SAP models were more effective in predicting strategies for new players as quantified by a lower transfer loss than that incurred by the baseline strategy.

and the negative connections, and depict only the most influential connections for either case. Fig. 3 describes how our method estimated a meaningful structure, and unraveled subtle influences beyond the prominent two-bloc structure. The learned type parameters $\hat{\theta}_i$ were also found to be similar for members in the same bloc (Fig. 4).

Prediction and transfer results

Our final set of experiments focused on transferring SAP games. Besides UNGA data, we compiled a 73-dimensional feature vector for each country from its HDI Indicators (UN2, 2007) that span various spheres including health, education, trade, and social-economic sustainability.

We kept aside three-fourths of the data to set up games over small sets of players by randomly sampling 5 countries independently for each context. The left-out players from these contexts were then sampled independently to get another set of 5 countries per context. We call these two sets A and B respectively. We formed a third set C of 5 countries per game from the untouched data (i.e. the one-fourth fraction). Thus, A and B were defined on the same contexts, that were disjoint from C. We averaged our results over 10 such independent triplets (A, B, C) to mitigate sampling effects. We trained a model for A using the procedure in section 4, and computed the loss on B. Our baseline, *train empirical*, used the empirical distribution of actions for each player i over the games it participated in A as its predicted strategy for the games in B and C.

The cross-entropy loss of SAP games (0.852) turned out to be lower than the baseline (0.865) on C. Moreover, as Fig. 4 shows, the loss of SAP games (0.696) was found to be significantly lower than the baseline (0.720) on B even though HDI does not fully reflect complex country characteristics. Thus, our results clearly underscore the benefits of using SAP games for strategic prediction.

Conclusion

We extended structured prediction to strategic settings, and specified conditions under which our games are identifiable from data, and when strategies converge to an equilibrium. Empirical results demonstrate effectiveness of our models in uncovering meaningful strategic interactions from data, predicting player strategies on new contexts, and transferring knowledge to predict strategies for new players.

We sidestepped the issue of rate of convergence to equilibrium in our analysis. Since there might be multiple mixed strategy equilibria in a game, convergence rate to an equilibrium depends on the proximity of the initial joint strategy profile, and the spectrum of Jacobian matrix resulting from linearization of the associated ODE.

Acknowledgments

TJ and VG acknowledge support from ONR and the MIT-IBM collaboration.

References

Human Development Reports. Available at *http://hdr.undp.org/en/countries/profiles/*. United Nations Development Programme, 2007.

Y. Azrieli and E. Shmaya. Lipschitz games. *Mathematics of Operations Research*, 38(2):350–357, 2013.

M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44:328–348, 2005.

M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions, part ii: Applications. *Mathematics of Operations Research*, 31(4):673–695, 2006.

V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

G. W. Brown. Iterative solutions of games by fictitious play. In *Activity Analysis of Production and Allocation*, 1951.

L.-C. Chen, A. Schwing, A. Yuille, and R. Urtasun. Learning deep structured models. In *International Conference on Machine Learning (ICML)*, 2015.

L. Condat. Fast projection onto the simplex and the l_1 ball. *Mathematical Programming*, 158(1–2):575–585, 2016.

C. Cortes, V. Kuznetsov, M. Mohri, and S. Yang. Structured prediction theory based on factor graph complexity. In *Neural Information Processing Systems (NIPS)*, 2016.

J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra. Efficient projections onto the l1-ball for learning in high dimensions. In *International Conference on Machine Learning (ICML)*, 2008.

D. Fudenberg and D. M. Kreps. Learning mixed equilibria. *Games and Economic Behavior*, 5(3):320–367, 1993.

V. K. Garg and T. Jaakkola. Learning tree structured potential games. In *Neural Information Processing Systems (NIPS)*, 2016.

V. K. Garg and T. Jaakkola. Local aggregative games. In *Neural Information Processing Systems (NIPS)*, 2017.

A. Ghoshal and J. Honorio. Learning graphical games from behavioral data: Sufficient and necessary conditions. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017.

A. Globerson, T. Roughgarden, D. Sontag, and C. Yildirim. How hard is inference for structured prediction? In *International Conference on Machine Learning (ICML)*, 2015.

J. C. Harsanyi. Games with incomplete information played by "bayesian" players (part i, the basic model). *Management Science*, 14(3):159–182, 1967.

J. Honorio and L. Ortiz. Learning the structure and parameters of large-population graphical games from behavioral data. *Journal of Machine Learning Research (JMLR)*, 16: 1157–1210, 2015.

M. T. Irfan and L. E. Ortiz. On influence, stable behavior, and the most influential individuals in networks: A game-theoretic approach. *Artificial Intelligence*, 215:79–119, 2014.

A. X. Jiang and K. Leyton-Brown. Bayesian action-graph games. In *Neural Information Processing Systems (NIPS)*, 2010.

E. Kalai. Large robust games. *Econometrica*, 72(6):1631–1665, 2004.

M. Kearns and Y. Mansour. Efficient nash computation in large population games with bounded influence. In *Uncertainty in Artificial Intelligence (UAI)*, 2002.

M. Kearns, M. L. Littman, and S. Singh. Graphical models for game theory. In *Uncertainty in Artificial Intelligence (UAI)*, 2001.

J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *International Conference on Machine Learning (ICML)*, 2001.

B. London, B. Huang, and L. Getoor. Stability and generalization in structured prediction. *Journal of Machine Learning Research (JMLR)*, 2016.

J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

S. Nowozin and C. H. Lampert. Structured learning and prediction in computer vision, 2011.

A. Osokin, F. Bach, and S. Lacoste-Julien. On structured prediction theory with calibrated convex surrogate losses. In *Neural Information Processing Systems (NIPS)*, 2017.

X. Pan and V. Srikumar. Learning to speed up structured output prediction. In *International Conference on Machine Learning (ICML)*, pages 3996–4005, 2018.

P. Ravikumar, M. J. Wainwright, and J. D. Lafferty. High-dimensional ising model selection using ℓ^1-regularized logistic regression. *The Annals of Statistics*, 38(3):1287–1319, 2010.

J. Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54(2), 1951.

J. S. Shamma and G. Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *IEEE Transactions on Automatic Control*, 50 (3):312–327, 2005.

L. S. Shapley. Some topics in two-person games. *Advances in Game Theory*, pages 1–29, 1964.

E. Tardos and V. Vazirani. Introduction to game theory: Basic solution concepts and computational issues, 2007.

B. Taskar, C. Guestrin, and D. Koller. Max-margin markov networks. In *Neural Information Processing Systems (NIPS)*, 2003.

I. Tsochantaridis, T. Joachims, T. Hofmann, and Y. Altun. Large margin methods for structured and interdependent output variables. *Journal of Machine Learning Research (JMLR)*, 6(2):1453–1484, 2005.

E. Voeten, A. Strezhnev, and M. Bailey. United nations general assembly voting data, 2009. URL `https://hdl.handle.net/1902.1/12379`.

M. J. Wainwright. Sharp thresholds for high-dimensional and noisy sparsity recovery using l1-constrained quadratic programming (lasso). *IEEE Transactions on Information Theory*, 55(5):2183–2202, 2009.

K. Waugh, B. D. Ziebart, and J. A. Bagnell. Computational rationalization: The inverse equilibrium problem. In *International Conference on Machine Learning (ICML)*, 2011.

Supplementary material

We now provide a detailed analysis on two fundamental aspects of our games: convergence and identifiability. That is, we characterize conditions under which players converge to an equilibrium, and when the game parameters can be provably recovered from observed outcomes.

A. Identifiability of our games

We begin with the results on provably recovering the structure of our one-shot games from data. Specifically, we characterize the conditions under which our games with one step of dynamics become identifiable, and provide an algorithm to recover the structure of the game, i.e., the neighbors for each player $i \in [n]$ with the signs (positive or negative) of their respective influences.

Our recovery procedure adapts the primal-dual witness method (Wainwright, 2009) for structure estimation in games. The method has previously been applied in several non-strategic settings such as Lasso (Wainwright, 2009) and Ising models (Ravikumar et al., 2010). Recently, (Ghoshal and Honorio, 2017) employed this method to recover a set of pure strategy Nash equilibria (PSNE) from data consisting of a subset of PSNE, and a small fraction of non-equilibrium outcomes assumed to be sampled under their noise models in the setting of linear influence games. However, the problem of structure recovery is significantly harder: it is known (Honorio and Ortiz, 2015; Ghoshal and Honorio, 2017) that the problem becomes non-identifiable in the setting of PSNE, since multiple game structures may pertain to the same of PSNE. We leverage dynamics to fill this gap by characterizing conditions under which our one-shot games become identifiable.

Our approach follows the general proof structure of primal-dual witness method in the context of model selection for Ising models (Ravikumar et al., 2010). However, our setting is significantly different from the setting in (Ravikumar et al., 2010) where context and dynamics play no part, and all the observed data is assumed to be sampled from a common (global) distribution expressible in a closed form. In contrast, each observed outcome in our setting is sampled from a separate joint strategy profile following one-step of dynamics initiated under a different context.

Specifically, in the one-shot setting, consider a dataset $D = \{(x^{(m)}, a^{(m)}) \in \mathcal{X} \times \mathcal{Y}, m \in [M]\}$ where $a^{(m)}$ is the action profile (i.e. observed outcome) sampled from the joint player strategies after one round of communication. Assume that the type parameters $\theta = (\theta_1, \ldots, \theta_n)$ are known. Then, since types for any context are determined by the parameters θ, we have access to the player types $z^{(m)}(x^{(m)}) = (z_1^{(m)}, \ldots, z_n^{(m)})$, which in turn determine determine the initial strategies for all the examples $m \in [M]$. We focus on binary actions here since they let us simplify the exposition while conveying the essential ideas. Specifically, each player $i \in [n]$ initially plays action 1 with probability

$$\phi_i^{(m)} = \xi(z_i^{(m)}) \triangleq \frac{1}{1 + \exp(-z_i^{(m)})} \ ,$$

and the action 0 with probability $1 - \phi_i^{(m)}$. We define $\phi^{(m)} = (\phi_1^{(m)}, \ldots, \phi_n^{(m)})$, and $\Phi_{-i}^{(m)} = (\phi_j^{(m)})_{j \neq i}$. We focus on the gradient update setting where after one round of communication, player i responds to its neighbors with its updated strategy $(\sigma_i^{*(m)}, 1 - \sigma_i^{*(m)})$, where

$$\sigma_i^{*(m)} \triangleq \xi \left(\phi_i^{(m)} + \alpha (\sum_{j \neq i} w_{ij}^* \phi_j^{(m)} - z_i^{(m)}) \right) \ ,$$

such that $\alpha > 0$, and $w_{ij}^* \in \mathbb{R}$ is the true influence (i.e. interaction weight) of player $j \in [n] \setminus \{i\}$ on i. Recall that we call player j a neighbor of i if $|w_{ij}^*| > 0$. Finally, action $a_i^{(m)}$ is sampled from the updated strategy, and we obtain the joint profile $a^{(m)} = \{a_i^{(m)}, i \in [n]\}$ as the observed outcome. Our goal is to estimate, from D and α, the *support* S_i, or the set of neighbors j for i, i.e., the players that have influence $w_{ij}^* \neq 0$. We can thus separate the influence of neighbors of i from the non-neighbors by defining the set of non-zero weights $w_{i,S}^* = \{w_{ij}^* | j \in S_i\}$. We denote the complement of a set A by A^c. Thus, $w_{ij}^* = 0$ for $j \in S_i^c$. We equivalently write $w_{i,S^c}^* = \mathbf{0}$. We are interested in recovering not only the support of each player i, but also the correct sign of influence (i.e. positive or negative) of each neighbor j on i.

We consider the average cross-entropy loss between the strategy under w_i and the observed outcome.

$$\ell_i(w_i; D) \; = \; \frac{1}{M}\sum_{m=1}^{M} -\left(a_i^{(m)}\log(\sigma_i^{(m)}) + (1 - a_i^{(m)})\log(1 - \sigma_i^{(m)})\right) \; . \tag{14}$$

We compute the gradient and the Hessian of the sample loss:

$$\nabla \ell_i(w_i; D) \; = \; \frac{\alpha}{M}\sum_{m=1}^{M}(\sigma_i^{(m)} - a_i^{(m)}) \; \Phi_{-i}^{(m)} \; , \tag{15}$$

$$H_i^M \; \triangleq \; \nabla^2 \ell_i(w_i; D) \; = \; \frac{\alpha^2}{M}\sum_{m=1}^{M}\sigma_i^{(m)}(1 - \sigma_i^{(m)}) \; \Phi_{-i}^{(m)}\Phi_{-i}^{(m)^\top} \; . \tag{16}$$

We will often use the variance function $\eta_i(w_i; m) \triangleq \alpha^2\sigma_i^{(m)}(1 - \sigma_i^{(m)})$ as a shorthand, and write

$$H_i^M = \frac{1}{M}\sum_{m=1}^{M}\eta_i(w_i; m) \; \Phi_{-i}^{(m)}\Phi_{-i}^{(m)^\top} \; . \tag{17}$$

We denote by $H_{i,SS}^{*M}$ the submatrix obtained by restricting the Hessian H_i^{*M}, pertaining to true weights, to rows and columns corresponding to neighbors, i.e., players in S_i. Likewise, H_{i,SS^c}^{*M} denotes the submatrix restricted to rows pertaining to S_i (neighbors) and columns to S_i^c (non-neighbors).

We will provide detailed analysis under sample Fisher matrix assumptions. We will omit the analysis for the population setting that can be derived by imposing analogous assumptions directly on the population matrices, and making concentration arguments that show these assumptions hold in the sampled setting with high probability. Recall from the main text that we make the following assumptions that are reminiscent of those for support recovery under Lasso (Wainwright, 2009), and model selection in Ising models (Ravikumar et al., 2010). We first recall our assumptions from the main text.

Assumptions.

$$\Lambda_{\min}\left(H_{i,SS}^{*M}\right) \; \geq \; \alpha^2 C_{min} \; . \tag{18}$$

$$\Lambda_{\max}\left(\frac{1}{M}\sum_{m=1}^{M}\Phi_{-i}^{(m)}\Phi_{-i}^{(m)^\top}\right) \; \leq \; C_{max} \; . \tag{19}$$

$$|||H_{i,S^cS}^{*M}(H_{i,SS}^{*M})^{-1}|||_\infty \; \leq \; 1 - \gamma \; , \tag{20}$$

such that $C_{\min} > 0$, $C_{\max} < \infty$, and $\gamma \in (0, 1]$. In our notation, $|||A|||_\infty$ denotes the maximum ℓ_1 norm across rows of matrix A, and $|||A|||_2$ denotes the spectral norm (i.e. maximum singular value) of A. $\Lambda_{min}(A)$ and $\Lambda_{max}(A)$ refer respectively to the minimum and the maximum eigenvalue of a square matrix A.

Analysis. We propose to solve the following regularized problem for each player $i \in [n]$ separately.

$$\arg\min_{w_i \in \mathbb{R}^{n-1}} \ell_i(w_i; D) \; + \; \lambda_{M,n,d}||w_i||_1 \; , \tag{21}$$

where $\lambda_{M,n,d} > 0$ is a regularization parameter that depends on the sample size M, the number of players n, and the maximum *degree* (i.e. number of neighbors) d of any player. For brevity, we will omit the dependence of this parameter on n and d, and simply write λ_M. This problem is convex but not differentiable everywhere because of the ℓ_1 penalty. Note that since the problem is not strictly convex, it might have multiple minimizing solutions. For any such optimal solution $\hat{w}_i$, we must have by KKT conditions,

$$\nabla \ell_i(\hat{w}_i; D) \; + \; \lambda_M \hat{\kappa}_i \; = \; \mathbf{0} \; , \tag{22}$$

where the subgradient $\hat{\kappa}_i \in \mathbb{R}^{n-1}$ is such that

$$\hat{\kappa}_{ij} = \text{sign}(\hat{w}_{ij}) \in \{\pm 1\} \text{ if } \hat{w}_{ij} \neq 0, \text{ and } |\hat{\kappa}_{ij}| \leq 1 \text{ otherwise.} \tag{23}$$

We would like to ensure the following conditions in order to recover the signed neighborhood for i.

$$\begin{aligned} \text{sign}(\hat{\kappa}_{ij}) &= \text{sign}(w^*_{ij}),\ \forall j \in S_i & (24)\\ \hat{w}_{ij} &= 0,\ \forall j \in S_i^c\ . & (25)\end{aligned}$$

Our analysis is built on the *primal-dual witness* (PDW) method (Wainwright, 2009). This method has the following steps. First, only for the sake of analysis, we presuppose that some Oracle provides the true neighbors S_i. Therefore, we solve the following problem to recover the signs of true neighbors.

$$\hat{w}_{i,S} = \arg \min_{(w_{i,S},\mathbf{0}) \in \mathbb{R}^{n-1}} \ell_i(w_i; D) \ + \ \lambda_M ||w_{i,S}||_1\ , \tag{26}$$

We then set the components of the dual vector κ_i that pertain to neighbors of i to the sign of corresponding components in $\hat{w}_{i,S}$. That is, $\hat{\kappa}_{i,j} \ = \ \text{sign}(\hat{w}_{i,j}),\ \forall j \in S_i$. We next set $\hat{w}_{i,S^c} = \mathbf{0}$, and thus (25) is satisfied. We then solve for $\hat{\kappa}_{i,S^c}$ by plugging $\hat{w}_{i,S}$, $\hat{\kappa}_{i,S}$, and $\hat{w}_{i,S^c}$ in (22). Thus, we are left to show that (23) and (24) are satisfied. We impose conditions on M, n, and d under which these conditions are satisfied with high probability. In fact, we prove a stronger result for (23), namely, strict dual feasibility for non-neighbors, i.e., $|\hat{\kappa}_{i,j}| < 1$ for all $j \in S_i^c$.

We argue that our construction yields a unique optimal primal solution $\hat{w}_i$. Specifically, we invoke Lemma 1 from (Ravikumar et al., 2010) that states that so long as $||\hat{\kappa}_{i,S^c}||_\infty < 1$, any optimal primal solution $\tilde{w}_i$ satisfies $\tilde{w}_{i,S^c} = \mathbf{0}$. This is established by our construction above. Moreover, Lemma 1 asserts that $\hat{w}_i$ is the unique solution to (21) if $\Lambda_{\min}(\hat{H}^M_{i,SS}) > 0$, i.e., if the sample Hessian under $\hat{w}_i$ is positive definite when restricted to the rows and columns in the true support S_i. We show that assumption (18) implies $\Lambda_{\min}\left(\hat{H}^M_{i,SS}\right) \ \geq \ \alpha^2 \dfrac{C_{min}}{2} > 0$, and this guarantees that we correctly recover the signed neighborhood of i.

To proceed, we define $G_i^M = -\nabla \ell_i(w_i^*; D)$ and rewrite (22) as

$$\nabla \ell_i(\hat{w}_i; D) - \nabla \ell_i(w_i^*; D) \ = \ G_i^M \ - \ \lambda_M \hat{\kappa}_i\ . \tag{27}$$

Applying the mean value theorem component-wise, we can write (27) as

$$\nabla^2 \ell_i(w_i^*; D)(\hat{w}_i - w_i^*) \ = \ G_i^M \ - \ \lambda_M \hat{\kappa}_i \ - \ R_i^M\ , \tag{28}$$

where

$$R^M_{i,j} \ = \ \left(\nabla^2 \ell_i(\overline{w}_i^{(j)}; D) - \nabla^2 \ell_i(w_i^*; D)\right)_j^\top (\hat{w}_i - w_i^*)\ ,$$

for some vector $\overline{w}_i^{(j)} = t_j \hat{w}_i + (1 - t_j) w_i^*$, $t_j \in [0, 1]$. Here, $(A)_j^\top$ denotes row j of matrix A.

We will now prove some auxiliary results that we will use in the proof of Theorem 2.

Lemma 1. *We have that*

$$\mathbb{P}\left(||G_i^M||_\infty \geq \frac{\lambda_M}{4} \frac{\gamma}{2-\gamma}\right) \ \leq \ 2 \exp\left(-\frac{\gamma^2 \lambda_M^2}{32\alpha^2(2-\gamma)^2} M + \log(n)\right)\ ,$$

which converges to zero at rate $\exp(-C_{\alpha,\gamma} \lambda_M^2 M)$ *(where constant* $C_{\alpha,\gamma}$ *depends on* α *and* γ*) whenever*

$$\lambda_M \geq \frac{8\alpha(2-\gamma)}{\gamma} \sqrt{\frac{\log(n)}{M}}\ .$$

Proof. We note that

$$G_i^M \ = \ -\nabla \ell_i(w_i^*; D) \ = \ \frac{1}{M} \sum_{m=1}^M \underbrace{-\alpha(\sigma_i^{*(m)} - a_i^{*(m)})\, \Phi_{-i}^{(m)}}_{Z_{i,m}}\ ,$$

where $|Z^u_{i,m}| \leq \alpha$ for each component $Z^m_{i,u}$ of random vector Z^m_i. Moreover, $\mathbb{E}(Z^m_{i,u}) = 0$ under w_i^*, and $Z^1_{i,u}, \ldots, Z^M_{i,u}$ are independent. Invoking the Hoeffding's inequality, we have that for any $\delta > 0$,

$$\mathbb{P}(|G^M_{i,u}| \geq \delta) \ \leq \ 2 \exp\left(-\frac{M\delta^2}{2\alpha^2}\right)\ ,$$

where $G_{i,u}^M$ denotes the component at index u of vector G_i^M. Setting $\delta = \dfrac{\gamma\lambda_M}{4(2-\gamma)}$, we get

$$\mathbb{P}\left(|G_{i,u}^M| \geq \frac{\gamma\lambda_M}{4(2-\gamma)}\right) \;\leq\; 2\exp\left(-\frac{M}{2\alpha^2}\frac{\gamma^2\lambda_M^2}{16(2-\gamma)^2}\right) \,.$$

Then, applying a union bound over indices $u \in [n-1]$, we get

$$\begin{aligned}\mathbb{P}\left(||G_i^M||_\infty \geq \frac{\gamma\lambda_M}{4(2-\gamma)}\right) \quad &\leq \quad 2(n-1)\exp\left(-\frac{M}{2\alpha^2}\frac{\gamma^2\lambda_M^2}{16(2-\gamma)^2}\right) \\ &< \quad 2\exp\left(-\frac{M}{2\alpha^2}\frac{\gamma^2\lambda_M^2}{16(2-\gamma)^2} + \log(n)\right) \,.\end{aligned}$$

□

Lemma 2. *Let $\lambda_M d \;\leq\; \dfrac{\alpha C_{\min}^2}{10 C_{\max}}$ and $||G_i^M||_\infty \;\leq\; \dfrac{\lambda_M}{4}$. Then,*

$$||\hat{w}_{i,S} - w_{i,S}^*||_2 \;\leq\; \frac{5}{\alpha^2 C_{\min}}\lambda_M\sqrt{d} \,.$$

Proof. We define a function $F : \mathbb{R}^d \to \mathbb{R}$ that quantifies the change in optimization objective at a distance $\Delta_{i,S}$ from the true parameters $w_{i,S}^*$. Specifically,

$$F(\Delta_{i,S}) \;\triangleq\; \ell_i(w_{i,S}^* + \Delta_{i,S}; D) - \ell_i(w_{i,S}^*; D) + \lambda_M(||w_{i,S}^* + \Delta_{i,S}||_1 - ||w_{i,S}^*||_1) \,.$$

Note that F is convex and $F(\mathbf{0}) \;=\; 0$. Moreover, F is minimized for $\hat{\Delta}_{i,S} = \hat{w}_{i,S} - w_{i,S}^*$. Therefore, $F(\hat{\Delta}_{i,S}) \leq 0$. We show that the function F is strictly positive on the surface of a Euclidean ball of radius B for some $B > 0$. Then, the vector $\hat{\Delta}_{i,S}$ lies inside the ball, i.e.,

$$||\hat{w}_{i,S} - w_{i,S}^*||_2 \leq B \,.$$

This follows since otherwise, the convex combination $t\hat{\Delta}_{i,S} + (1-t)\mathbf{0}$ would lie on boundary of the ball for some $t \in (0,1)$, which would imply the contradiction

$$F(t\hat{\Delta}_{i,S} + (1-t)\mathbf{0}) \leq tF(\hat{\Delta}_{i,S}) + (1-t)F(\mathbf{0}) \leq 0.$$

Therefore, let $\Delta \in \mathbb{R}^d$ be an arbitrary vector such that $||\Delta||_2 = B$. We then have from Taylor's series

$$F(\Delta) = \nabla\ell_i(w_{i,S}^*; D)^\top\Delta + \Delta^\top\nabla^2\ell(w_{i,S}^* + \theta\Delta; D)\Delta + \lambda_M(||w_{i,S}^* + \Delta||_1 - ||w_{i,S}^*||_1) \,, \tag{29}$$

for some $\theta \in [0,1]$. We lower bound $F(\Delta)$ by bounding each term on the right side of (29).

We let $B = O\lambda_M\sqrt{d}$ where we will choose $O > 0$ later. From Cauchy-Schwartz inequality,

$$\begin{aligned}\nabla\ell_i(w_{i,S}^*; D)^\top\Delta \quad &\geq \quad -||\nabla\ell_i(w_{i,S}^*; D)||_\infty||\Delta||_1 &(30)\\ &\geq \quad -||\nabla\ell_i(w_{i,S}^*; D)||_\infty\sqrt{d}||\Delta||_2 &(31)\\ &\geq \quad -(\lambda_M\sqrt{d})^2\frac{O}{4} \,, &(32)\end{aligned}$$

where in the last inequality we have used $||\Delta||_2 = B = O\lambda_M\sqrt{d}$, and

$$-||\nabla\ell_i(w_{i,S}^*; D)||_\infty \geq -||\nabla\ell_i(w_i^*; D)||_\infty \;=\; -||-\nabla\ell_i(w_i^*; D)||_\infty \;=\; -||G_i^M||_\infty \;\geq\; -\frac{\lambda_M}{4}$$

by our assumption on $||G_i^M||_\infty$ in the lemma statement. Next, by triangle inequality, we have

$$\lambda_M(||w_{i,S}^* + \Delta||_1 - ||w_{i,S}^*||_1) \;\geq\; -\lambda_M||\Delta||_1 \;\geq\; -\lambda_M\sqrt{d}||\Delta||_2 \;\geq\; -(\lambda_M\sqrt{d})^2 O \,. \tag{33}$$

We now bound the quantity $\Delta^\top \nabla^2 \ell(w^*_{i,S} + \theta\Delta; D)\Delta$. We note that

$$
\begin{aligned}
\Delta^\top \nabla^2 \ell(w^*_{i,S} + \theta\Delta; D)\Delta &\geq \min_{||\tilde{\Delta}||_2 = B} \tilde{\Delta}^\top \nabla^2 \ell(w^*_{i,S} + \theta\Delta; D)\tilde{\Delta} \\
&\geq \min_{\tilde{\theta}\in[0,1]} B^2 \Lambda_{\min}(\nabla^2 \ell(w^*_{i,S} + \tilde{\theta}\Delta; D)) \\
&= B^2 \min_{\tilde{\theta}\in[0,1]} \Lambda_{\min}\left(\frac{1}{M}\sum_{m=1}^{M} \eta_i(w^*_{i,S} + \tilde{\theta}\Delta; m)\, \Phi^{(m)}_{-i}\Phi^{(m)^\top}_{-i}\right) .
\end{aligned}
$$

Applying Taylor's series expansion, we note that $\Delta^\top \nabla^2 \ell(w^*_{i,S} + \theta\Delta; D)\Delta$

$$
\begin{aligned}
&\geq B^2 \Lambda_{\min}\left(\frac{1}{M}\sum_{m=1}^{M} \eta_i(w^*_{i,S}; m)\Phi^{(m)}_{-i}\Phi^{(m)^\top}_{-i}\right) \\
&\quad - B^2 \max_{\tilde{\theta}\in[0,1]} \left|\left|\left|\frac{1}{M}\sum_{m=1}^{M} \eta'_i(w^*_{i,S} + \overline{\theta}\Delta; m)(\Phi^{(m)^\top}_{-i}\tilde{\theta}\Delta)\, \Phi^{(m)}_{-i}\Phi^{(m)^\top}_{-i}\right|\right|\right|_2 \\
&= B^2 \Lambda_{\min}(H^{*M}_{i,SS}) - B^2 \max_{\tilde{\theta}\in[0,1]} \left|\left|\left|\frac{1}{M}\sum_{m=1}^{M} \eta'_i(w^*_{i,S} + \overline{\theta}\Delta; m)(\Phi^{(m)^\top}_{-i}\tilde{\theta}\Delta)\, \Phi^{(m)}_{-i}\Phi^{(m)^\top}_{-i}\right|\right|\right|_2 \\
&= B^2 \alpha^2 C_{\min} - B^2 \max_{\tilde{\theta}\in[0,1]} \left|\left|\left|\frac{1}{M}\sum_{m=1}^{M} \eta'_i(w^*_{i,S} + \overline{\theta}\Delta; m)(\Phi^{(m)^\top}_{-i}\tilde{\theta}\Delta)\, \Phi^{(m)}_{-i}\Phi^{(m)^\top}_{-i}\right|\right|\right|_2 .
\end{aligned}
$$

Now, a simple calculation shows that $|\eta'_i(\cdot)| \leq \alpha^3$. Moreover, we note for $\tilde{\theta} \in [0,1]$,

$$
|\Phi^{(m)^\top}_{-i}\tilde{\theta}\Delta| \leq ||\Phi^{(m)}_{-i}||_\infty ||\tilde{\theta}\Delta||_1 \leq ||\Phi^{(m)}_{-i}||_\infty ||\Delta||_1 \leq ||\Delta||_1 \leq \sqrt{d}||\Delta||_2 = O\lambda_M d .
$$

Putting all these facts together, along with our assumption (19), we get

$$
\Delta^\top \nabla^2 \ell(w^*_{i,S} + \theta\Delta; D)\Delta \geq B^2\alpha^2 C_{\min} - B^2\alpha^3(O\lambda_M d)C_{\max} \geq B^2\alpha^2\frac{C_{\min}}{2} \tag{34}
$$

when $\lambda_M \leq \dfrac{C_{\min}}{2\alpha C_{\max} O d}$. Therefore, plugging the lower bounds from (30), (33), and (34) in (29),

$$
F(\Delta) \geq \lambda_M^2 d\left(-\frac{O}{4} - O + \frac{O^2\alpha^2 C_{\min}}{2}\right) > 0 ,
$$

for $O = \dfrac{5}{\alpha^2 C_{\min}}$. Thus, for $\lambda_M \leq \dfrac{C_{\min}}{2\alpha C_{\max} O d} = \dfrac{\alpha C^2_{\min}}{10 C_{\max} d}$, we must have

$$
||\hat{w}_{i,S} - w^*_{i,S}||_2 \leq B = O\lambda_M\sqrt{d} = \frac{5}{\alpha^2 C_{\min}}\lambda_M\sqrt{d} .
$$

□

Lemma 3. *Let* $\lambda_M d \leq \dfrac{\alpha C^2_{\min}}{100 C_{\max}}\dfrac{\gamma}{2-\gamma}$ *and* $||G^M_i||_\infty \leq \dfrac{\lambda_M}{4}$. *Then,*

$$
\frac{||R^M_i||_\infty}{\lambda_M} \leq \frac{25 C_{\max}}{\alpha C^2_{\min}}\lambda_M d \leq \frac{1}{4}\left(\frac{\gamma}{2-\gamma}\right) \leq \frac{\gamma}{4} .
$$

Proof. We have for $j \in [n] \setminus \{i\}$ and some $\overline{w}_i^{(j)} = t_j \hat{w}_i + (1-t_j) w_i^*$, $t_j \in [0,1]$,

$$\begin{aligned}
R_{i,j}^M &= \left(\nabla^2 \ell_i(\overline{w}_i^{(j)}; D) - \nabla^2 \ell_i(w_i^*; D)\right)_j^\top (\hat{w}_i - w_i^*) \\
&= \frac{1}{M} \sum_{m=1}^{M} \left(\left(\eta_i(\overline{w}_i^{(j)}; m) - \eta_i(w_i^*; m)\right) \Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top}\right)_j^\top (\hat{w}_i - w_i^*) \\
&= \frac{1}{M} \sum_{m=1}^{M} \left(\eta_i'(\overline{\overline{w}}_i^{(j)}; m) \left(\Phi_{-i}^{(m)^\top} (\overline{w}_i^{(j)} - w_i^*)\right) \Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top}\right)_j^\top (\hat{w}_i - w_i^*) \,,
\end{aligned}$$

where $\overline{\overline{w}}_i^{(j)}$ is a point on the line between $\overline{w}_i^{(j)}$ and w_i^*, by the mean value theorem. We note that

$$\left(\Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top}\right)_j^\top = \phi_j^{(m)} \Phi_{-i}^{(m)^\top} \,.$$

We thus write

$$\begin{aligned}
R_{i,j}^M &= \frac{1}{M} \sum_{m=1}^{M} \eta_i'(\overline{\overline{w}}_i^{(j)}; m) \phi_j^{(m)} \left((\overline{w}_i^{(j)} - w_i^*)^\top \Phi_{-i}^{(m)}\right) \Phi_{-i}^{(m)^\top} (\hat{w}_i - w_i^*) \\
&= \frac{1}{M} \sum_{m=1}^{M} \eta_i'(\overline{\overline{w}}_i^{(j)}; m) \phi_j^{(m)} \left((\overline{w}_i^{(j)} - w_i^*)^\top \Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top} (\hat{w}_i - w_i^*)\right) \\
&= \frac{1}{M} \sum_{m=1}^{M} \underbrace{\eta_i'(\overline{\overline{w}}_i^{(j)}; m) \phi_j^{(m)}}_{p^{(m)}} \underbrace{\left(t_j (\hat{w}_i - w_i^*)^\top \Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top} (\hat{w}_i - w_i^*)\right)}_{q^{(m)}} \,,
\end{aligned}$$

which is of the form $\dfrac{1}{M} p^\top q$, where $p, q \in \mathbb{R}^M$. Thus, we have by Cauchy-Schwartz inequality,

$$|R_{i,j}^M| = \frac{1}{M} |p^\top q| \leq \frac{1}{M} ||p||_\infty ||q||_1 \,.$$

It can be shown that $p^{(m)} = \alpha^3 \overline{\overline{\sigma}}_i^{(m)} (1 - \overline{\overline{\sigma}}_i^{(m)})(1 - 2\overline{\overline{\sigma}}_i^{(m)})$, whereby $||p||_\infty \leq \alpha^3$.

Finally, we see that $q^{(m)} = t_j \left|\left|\Phi_{-i}^{(m)^\top} (\hat{w}_i - w_i^*)\right|\right|_2^2 \geq 0$ since $t_j \in [0,1]$. Therefore $||q||_1 = q^\top \mathbf{1}$, where $\mathbf{1} \in \mathbb{R}^M$ is a vector of all ones. Moreover, since $\hat{w}_{i,S^c} = w_{i,S^c}^* = \mathbf{0}$, we note that

$$\begin{aligned}
\frac{1}{M} ||q||_1 &= t_j (\hat{w}_i - w_i^*)^\top \left(\frac{1}{M} \sum_{m=1}^{M} \Phi_{-i}^{(m)} \Phi_{-i}^{(m)^\top}\right) (\hat{w}_i - w_i^*) \\
&= t_j (\hat{w}_{i,S} - w_{i,S}^*)^\top \left(\frac{1}{M} \sum_{m=1}^{M} \Phi_{-i,S}^{(m)} \Phi_{-i,S}^{(m)^\top}\right) (\hat{w}_{i,S} - w_{i,S}^*) \\
&\leq C_{\max} \left|\left|\hat{w}_{i,S} - w_{i,S}^*\right|\right|_2^2 \,.
\end{aligned}$$

Since $\gamma \in (0,1]$, so

$$\lambda_M d \leq \frac{\alpha C_{\min}^2}{100 C_{\max}} \frac{\gamma}{2-\gamma} \leq \frac{\alpha C_{\min}^2}{100 C_{\max}} \leq \frac{\alpha C_{\min}^2}{10 C_{\max}} \,.$$

Therefore, we can invoke Lemma 2 when $||G_i^M||_\infty \leq \dfrac{\lambda_M}{4}$. Specifically, we then have for each j,

$$|R_{i,j}^M| \leq \alpha^3 C_{\max} \left|\left|\hat{w}_{i,S} - w_{i,S}^*\right|\right|_2^2 \leq \alpha^3 C_{\max} \left(\frac{5}{\alpha^2 C_{\min}} \lambda_M \sqrt{d}\right)^2 = \frac{25 C_{\max}}{\alpha C_{\min}^2} \lambda_M^2 d \,.$$

This immediately yields $\dfrac{||R_i^M||_\infty}{\lambda_M} \leq \dfrac{25 C_{\max}}{\alpha C_{\min}^2} \lambda_M d \,.$ $\square$

We are now ready to prove our main result.

Theorem 1. *Let* $M > \dfrac{80^2 C_{\max}^2}{C_{\min}^4}\left(\dfrac{2-\gamma}{\gamma}\right)^4 d^2\log(n)$*, and* $\lambda_M \geq \dfrac{8\alpha(2-\gamma)}{\gamma}\sqrt{\dfrac{\log(n)}{M}}$*. Suppose the sample satisfies assumptions* (18)*,* (19)*, and* (20)*. Define* $C_{\alpha,\gamma} = \dfrac{\gamma^2}{32\alpha^2(2-\gamma)^2}$ *. Consider any player* $i \in [n]$*. The following results hold with probability at least* $1 - 2\exp(-C_{\alpha,\gamma}\lambda_M^2 M) \to 1$ *for* i*.*

1. *The corresponding* ℓ_1*-regularized optimization problem has a unique solution, i.e., a unique set of neighbors for* i*.*
2. *The set of predicted neighbors of* i *is a subset of the true neighbors. Additionally, the predicted set contains all true neighbors* j *for which* $|w_{ij}^*| \geq \dfrac{10}{\alpha^2 C_{\min}}\sqrt{d}\lambda_M$*. In particular, the set of true neighbors of* i *is exactly recovered if*

$$\min_{j\in S_i}|w_{ij}^*| \geq \frac{10}{\alpha^2 C_{\min}}\sqrt{d}\lambda_M \ .$$

Taking a union bound over players, our results imply that we recover the true signed neighborhoods for all players in the game with probability at least $1 - 2n\exp(-C_{\alpha,\gamma}\lambda_M^2 M)$ *.*

Proof. Since $\lambda_M \geq \dfrac{8\alpha(2-\gamma)}{\gamma}\sqrt{\dfrac{\log(n)}{M}}$, Lemma 1 holds. Thus, with high probability (as stated in the theorem statement), we obtain

$$||G_i^M||_\infty \ \leq\ \frac{\lambda_M}{4}\frac{\gamma}{2-\gamma} \ \leq\ \frac{\gamma\lambda_M}{4} \ \leq\ \frac{\lambda}{4}\,, \tag{35}$$

since $\gamma \in (0,1]$. Moreover, for the specified lower bound on sample size M, a simple computation shows

$$\lambda_M d \ \leq\ \frac{\alpha C_{\min}^2}{10 C_{\max}}\frac{\gamma}{2-\gamma}\ . \tag{36}$$

Thus the conditions required for both Lemma 2 and Lemma 3 are satisfied. By our primal-dual construction, $\hat{w}_{i,S^c} = \mathbf{0}$. Furthermore, using (18), $\Lambda_{\min}(H_{i,SS}^{*M}) > 0$, and so $H_{i,SS}^{*M}$ is invertible. Separating the rows in the support of i and others, we write (28) as

$$\begin{aligned} H_{i,S^cS}^{*M}(\hat{w}_{iS} - w_{iS}^*) &= G_{i,S^c}^M - \lambda_M\hat{\kappa}_{i,S^c} - R_{i,S^c}^M \\ H_{i,SS}^{*M}(\hat{w}_{iS} - w_{iS}^*) &= G_{i,S}^M - \lambda_M\hat{\kappa}_{i,S} - R_{i,S}^M\ . \end{aligned}$$

These two equations can be combined into one as

$$H_{i,S^cS}^{*M}(H_{i,SS}^{*M})^{-1}\left(G_{i,S}^M - \lambda_M\hat{\kappa}_{i,S} - R_{i,S}^M\right) \ =\ G_{i,S^c}^M - \lambda_M\hat{\kappa}_{i,S^c} - R_{i,S^c}^M\ .$$

Recalling that $||\hat{\kappa}_{i,S}||_\infty < 1$, we immediately get that $\lambda_M||\hat{\kappa}_{i,S^c}||_\infty$

$$\begin{aligned} &\leq\ \left|\left|\left|H_{i,S^cS}^{*M}(H_{i,SS}^{*M})^{-1}\right|\right|\right|_\infty\left(||G_{i,S}^M||_\infty + ||R_{i,S}^M||_\infty + \lambda_M\right) + ||G_{i,S^c}^M||_\infty + ||R_{i,S^c}^M||_\infty \\ &\leq\ (1-\gamma)\left(||G_{i,S}^M||_\infty + ||R_{i,S}^M||_\infty + \lambda_M\right) + ||G_{i,S^c}^M||_\infty + ||R_{i,S^c}^M||_\infty \\ &\leq\ (1-\gamma)\lambda_M + ||G_i^M||_\infty + ||R_i^M||_\infty \\ &\leq\ \lambda_M\left(1-\gamma+\frac{\gamma}{4}+\frac{\gamma}{4}\right) \\ &=\ \lambda_M\left(1-\frac{\gamma}{2}\right)\ . \end{aligned}$$

Since $\gamma \in (0,1]$ and $\lambda_M > 0$, we immediately get $||\hat{\kappa}_{i,S^c}||_\infty < 1$. Therefore, strict dual feasibility is established and (23) is verified. Then, using Lemma 1 of (Ravikumar et al., 2010), we note that any optimal solution $\tilde{w}_i$ of (21) must have $\tilde{w}_{i,S^c} = \mathbf{0}$. In particular, we have $\hat{w}_{i,S^c} = \mathbf{0}$ as desired. Thus, we can focus on $\hat{w}_{i,S}$. We now prove uniqueness of $\hat{w}_i$ by showing that $\Lambda_{\min}\left(\hat{H}_{i,SS}^M\right) > 0$. Let $\Delta = \hat{w}_{i,S} - w_{i,S}^* \in \mathbb{R}^d$. Then, using Lemma 2, we have

$$||\Delta||_2 \ \leq\ \frac{5}{\alpha^2 C_{\min}}\lambda_M\sqrt{d}\,.$$

Note that

$$\begin{aligned}
\Lambda_{\min}\left(\hat{H}^M_{i,SS}\right) &= \Lambda_{\min}\left(\frac{1}{M}\sum_{m=1}^{M}\eta_i(\hat{w}_i;m)\ \Phi^{(m)}_{-i,S}\Phi^{(m)^\top}_{-i,S}\right)\\
&= \Lambda_{\min}\left(\frac{1}{M}\sum_{m=1}^{M}\eta_i(\hat{w}_{i,S};m)\ \Phi^{(m)}_{-i,S}\Phi^{(m)^\top}_{-i,S}\right)\\
&= \Lambda_{\min}\left(\frac{1}{M}\sum_{m=1}^{M}\eta_i(w^*_{i,S}+\Delta;m)\ \Phi^{(m)}_{-i,S}\Phi^{(m)^\top}_{-i,S}\right)\ .
\end{aligned}$$

Performing a Taylor expansion around $w^*_{i,S}$, and making arguments similar to the proof segment between (33) and (34) in Lemma 2, we can show that

$$\begin{aligned}
\Lambda_{\min}\left(\hat{H}^M_{i,SS}\right) &\geq \alpha^2 C_{\min} - \alpha^3\sqrt{d}||\Delta||_2 C_{\max}\\
&\geq \alpha^2 C_{\min} - \left(\frac{5\alpha C_{\max}}{C_{\min}}\right)\lambda_M d\\
&\geq \alpha^2 C_{\min} - \alpha^2\frac{C_{\min}}{2}\frac{\gamma}{2-\gamma}\\
&\geq \alpha^2\frac{C_{\min}}{2}\ ,
\end{aligned}$$

which is greater than 0. Therefore, $\hat{H}^M_{i,SS}$ is positive definite, and Lemma 1 of (Ravikumar et al., 2010) guarantees that $\hat{w}_i$ is the unique optimal primal solution for (21).

We finally argue about the only remaining condition (24). In order for neighbor j to be correctly recovered with sign, i.e., $\text{sign}(\hat{w}_{ij}) = \text{sign}(w^*_{ij})$, it suffices to have

$$|\hat{w}_{ij} - w^*_{ij}| \ \leq\ \frac{|w^*_{ij}|}{2}\ . \tag{37}$$

Moreover to recover the neighborhood of i exactly, it is sufficient to show

$$\min_{j\in S_i}|w^*_{ij}| \geq 2||\hat{w}_{i,S} - w^*_{i,S}||_\infty\ , \tag{38}$$

which implies (37). We note that

$$||\hat{w}_{i,S} - w^*_{i,S}||_\infty \ \leq\ ||\hat{w}_{i,S} - w^*_{i,S}||_2 \ \leq\ \frac{5}{\alpha^2 C_{\min}}\lambda_M\sqrt{d}\ .$$

Using (38), it immediately follows that the neighborhood of i is recovered with correct sign if

$$\min_{j\in S_i}|w^*_{ij}| \geq \frac{10}{\alpha^2 C_{\min}}\lambda_M\sqrt{d}\ .$$

□

B. General game dynamics and convergence

In this section we provide an in-depth look at the game dynamics along with associated convergence guarantees.

Recall that in of our protocols, players take actions stochastically according to σ^k_i and the best response mapping g_i is unique for each k. Assuming that the error sequence in updating $\{\sigma^k_i\}$ is a martingale, our updates satisfy the conditions outlined in (section 2.1 of (Borkar, 2008)) and we can analyze the stochastic evolution of each setting as a noisy discretization of a limiting ordinary differential equation (ODE). In particular, Lipschitz condition is satisfied since g_i and h_i are both Lipschitz continuous, step size condition is fulfilled since the sequence $b^{k-1} = 1/k$ satisfies $\sum_k b^{k-1} = \infty$ and $\sum_k (b^{k-1})^2 < \infty$,

$$
\begin{aligned}
&(\textbf{SAP-FP/AA}) \quad \dot{q}_i = \beta_i^\tau(\mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}), z_i) - q_i, && \dot{r}_i = \lambda(q_i - r_i) && (39)\\
&(\textbf{SAP-FP/ PA}) \quad \dot{q}_i = \beta_i^\tau(\mathcal{A}_i(q_{-i}) + \gamma \dot{r}_i, z_i) - q_i, && \dot{r}_i = \lambda(\mathcal{A}_i(q_{-i}) - r_i) && (40)\\
&(\textbf{SAP-GP/AA}) \quad \dot{q}_i = \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i] - q_i, && \dot{r}_i = \lambda(q_i - r_i) && (41)\\
&(\textbf{SAP-GP/PA}) \quad \dot{q}_i = \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i}) + \gamma \dot{r}_i - z_i] - q_i, && \dot{r}_i = \lambda(\mathcal{A}_i(q_{-i}) - r_i) && (42)
\end{aligned}
$$

and our iterates remain bounded since they remain confined to $\Delta(A)$. Thus, we can investigate the conditions under which the fixed points of the limiting ODE are *locally asymptotically stable* (*l.a.s.*), and as a consequence, our discrete updates would converge to a Nash equilibrium with positive probability (Shamma and Arslan, 2005). An equilibrium point s is said to be *l.a.s.* if every ODE trajectory that starts at a point in a small neighborhood of s remains forever in that neighborhood and eventually converges to s.

Our updates in (9) lead to the implicit ODEs (39)-(42) for SAP-FP and SAP-GP under AA and PA settings, where $\lambda > 0$, $\dot{r}_i$ is an estimate for $\dot{q}_i$, and $\dot{r}_{-i} \triangleq \{\dot{r}_j | j \neq i, w_{ij} \neq 0\}$. We will call a matrix *stable* if all its eigenvalues have strictly negative real parts. Let I denote the identity matrix. We now state results that characterize conditions under which different dynamics lead to asymptotically stable equilibria.

Theorem 2. (SAP-FP/AA convergence to NE) *Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a NE under the dynamics in* (39)*. There exists a matrix $\mathcal{D}$ such that the linearization of* (39) *with $\gamma > 0$ is l.a.s. for $\lambda > 0$ if and only if the following matrix is stable*

$$
\begin{bmatrix} -I + (1+\gamma\lambda)\mathcal{D} & -\gamma\lambda\mathcal{D} \\ \lambda I & -\lambda I \end{bmatrix}.
$$

Theorem 3. (SAP-FP/PA convergence to NE) *Let the weight matrix W be stochastic. Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a NE under the dynamics in* (40)*. There exists a matrix $\mathcal{D}_1$ with zero diagonal, and a block diagonal matrix $\mathcal{D}_2$ such that the linearization of* (40) *with $\gamma > 0$ is l.a.s. for $\lambda > 0$ if and only if if the following matrix is stable*

$$
\begin{bmatrix} -I + (1+\gamma\lambda)\mathcal{D}_1 & -\gamma\lambda\mathcal{D}_2 \\ \lambda W & -\lambda I \end{bmatrix}.
$$

Theorem 4. (SAP-GP/AA convergence to CMNE) *Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a completely mixed NE under the dynamics in* (41)*. Then the linearization of* (41) *with $\gamma > 0$ is l.a.s. for $\lambda > 0$ if and only if the following matrix is stable*

$$
\begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda I & -\lambda I \end{bmatrix}.
$$

Theorem 5. (SAP-GP/AA convergence to SNE) *Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a strict NE under the dynamics in* (41)*. The associated equilibrium point $(q_i = q_i^*, q_{-i} = q_{-i}^*, r_i = q_i^*, r_{-i} = q_{-i}^*)$ is l.a.s. for any $\gamma > 0$ and $\lambda > 0$.*

Theorem 6. (SAP-GP/PA convergence to CMNE) *Let the weight matrix W be stochastic. Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a completely mixed NE under the dynamics in* (42)*. Then the linearization of* (42) *with $\gamma > 0$ is l.a.s. for $\lambda > 0$ if and only if the following matrix is stable*

$$
\begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda W & -\lambda I \end{bmatrix}.
$$

Theorem 7. (SAP-GP/PA convergence to SNE) *Let the weight matrix W be doubly stochastic. Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a strict NE under the dynamics in* (42)*. The equilibrium point $(q_i = q_i^*, r_i = A_i(q_{-i}^*))_{i \in [n]}$ is l.a.s. for sufficiently small $\gamma\lambda$, where $\gamma > 0$ and $\lambda > 0$.*

We now provide some insight into our proof techniques. We follow the general proof structure of (Shamma and Arslan, 2005). Specifically, we prove convergence to SNE via carefully crafted *Lyapunov functions* $\mathcal{V}$ that are locally positive definite and have a locally negative semidefinite time derivative, and thus satisfy the *Lyapunov stability* criterion. The other proofs track the evolution of game dynamics around an equilibrium, where $\dot{q}_i = 0$ and $\dot{r}_i = 0$. Specifically, we analyze conditions under which the Jacobian matrix of the linearization is *Hurwitz stable*, i.e., all the eigenvalues have negative real roots, and exploit the fact that the behavior of the ODE near equilibrium is same as its linear approximation when the real parts of all eigenvalues are non-zero. Our discrete updates would then converge to a Nash equilibrium with positive probability (Shamma and Arslan, 2005).

Recall that AA reveals more information about the evolution of neighbors' strategy. As a result, the PA settings, i.e. (40) and (42), require additional subtle reasoning since at equilibrium r_i^* converges only to $\mathcal{A}_i(q_{-i}^*)$ and not to q_i^*. Since q_i evolves within $\Delta(A)$, stochasticity assumptions are required to ensure r_i stays within the probability simplex as well. Note that the SAP-FP updates to strategies are smooth due to the entropy term (since $\tau > 0$), unlike SAP-GP. Consequently, the results for SAP-GP require a separate treatment of completely mixed NE and strict NE, unlike SAP-FP where they can be analyzed without distinction. Note that $\tau > 0$ ensures that best response is a singleton set and therefore we could leverage the ODE formulations. Differential inclusions (Benaïm et al., 2005; 2006) could possibly be used to handle $\tau = 0$.

We now provide detailed proofs on convergence of dynamics. We restructure the theorem statements to have the results for the active aggregator setting precede those for the passive aggregator setting. We use AA1, AA2 etc. to indicate that the result pertains to convergence in an active aggregator setting. Likewise, we will use PA1 etc. for the passive aggregator setting. We start with the active aggregator.

Theorem AA1. (Convergence under SAP-FP/AA to NE) *Let* $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ *be a NE under the dynamics in* (39)*. There exists a matrix* $\mathcal{D}$ *such that the linearization of* (39) *with* $\gamma > 0$ *is locally asymptotically stable for* $\lambda > 0$ *if and only if the following matrix is stable*

$$\begin{bmatrix} -I + (1+\gamma\lambda)\mathcal{D} & -\gamma\lambda\mathcal{D} \\ \lambda I & -\lambda I \end{bmatrix}.$$

Proof. Since $\tau > 0$, best response is a singleton set, and the unique best response σ_i^* can be obtained by setting the gradients of the payoff functions to 0. In particular, we have the best response

$$\beta_i^\tau(\mathcal{A}_i(\sigma_{-i}), z_i) = \zeta\left(\frac{\sum_{j\neq i} w_{ij}\sigma_j - z_i}{\tau}\right) = \zeta\left(\frac{\mathcal{A}_i(\sigma_{-i}) - z_i}{\tau}\right), \tag{43}$$

where ζ is the softmax function with output coordinate ℓ given by

$$(\zeta(x))_\ell = \exp(x_\ell) \Big/ \sum_k \exp(x_k).$$

Now recall from (39) that we have the following ODE:

$$\dot{q}_i = \underbrace{\beta_i^\tau(\mathcal{A}_i(q_{-i} + \gamma\dot{r}_{-i}), z_i^*) - q_i}_{\triangleq F_i(q_i, q_{-i}, r_{-i})} \tag{44}$$

$$\dot{r}_i = \lambda(q_i - r_i). \tag{45}$$

Since β_i^τmaps it input to the simplex $\Delta(A)$, we note that the right side of (44) is a difference between two probability distributions. Therefore this difference must sum to zero. Moreover, since $|A| = m$, we have $m-1$ degrees of freedom that can be used to express this difference. Therefore, we can investigate the evolution of q_i via a matrix $N \in \mathbb{R}^{m\times(m-1)}$ of $(m-1)$ orthonormal columns such that

$$N^\top N = I_{m-1}, \text{ and } 1_m^\top N = 0_{m-1},$$

where I_{m-1} is the identity matrix of order $m-1$, and 1_m and 0_m are m-dimensional vectors with all coordinates set to 1 and 0 respectively. We will sometimes omit the subscripts for I_m, 1_m, and 0_m when the size will be clear from the context. The equilibrium (q_i^*, q_{-i}^*) corresponds to a point $(q_i(t) = q_i^*, q_{-i}(t) = q_{-i}^*, r_i(t) = q_i^*, r_{-i}(t) = q_{-i}^*)$ of the dynamics. It will be convenient to investigate the dynamics as the evolution of deviations around this point. Since q_i is confined to $\Delta(A)$, we can express

$$q_i(t) = q_i^* + N\delta x_{q_i}(t),$$

where $\delta x_{q_i}(t) \in \mathbb{R}^{m-1}$ is uniquely specified, and likewise $r_i = q_i^* + \delta x_{r_i}(t)$ for some $\delta x_{r_i}(t)$. Thus, we can define a block diagonal matrix $\mathcal{N} \in \mathbb{R}^{2nm\times 2n(m-1)}$, with each diagonal block set to N and all other elements set to 0, such that

$$(q_1(t) - q_1^*, \ldots, q_n(t) - q_n^*, \quad r_1(t) - q_1^*, \ldots, r_n(t) - q_n^*)^\top = \mathcal{N}\delta x(t)\,, \tag{46}$$

where

$$\delta x(t) = (\delta x_{q_1}(t), \ldots, \delta x_{q_n}(t), \delta x_{r_1}(t), \ldots, \delta x_{r_n}(t))^\top \in \mathbb{R}^{2n(m-1)}$$

is formed by stacking together the deviations at time t in a column vector. Then, the following is immediate from (46):

$$\mathcal{N}^\top (q_1(t) - q_1^*, \ldots, q_n(t) - q_n^*, r_1(t) - q_1^*, \ldots, r_n(t) - q_n^*)^\top \;=\; \mathcal{N}^\top \mathcal{N} \delta x(t) \;=\; \delta x(t). \tag{47}$$

Denote the Jacobian matrix obtained by taking derivatives of vector y with respect to vector x by $J_x y$. We will linearize $\dot{q}_i = F_i(q_i, q_{-i}, r_{-i})$ in (44) around $\triangleq (q_1^*, q_{-1}^*, q_1^*, q_{-1}^*)$ using first order Taylor series. Then, since $\dot{q}_i^* = 0$, we note from (44) and (47) that

$$\dot{\delta} x_{q_i} = N^\top (\dot{q}_i - \dot{q}_i^*) = N^\top \dot{q}_i(t) = N^\top F_i(q_i, q_{-i}, r_{-i}). \tag{48}$$

Now, at equilibrium, we have $\dot{q}_i = 0$ for all $i \in [n]$, and therefore we have from (44) that

$$F_i(q_i^*, q_{-i}^*, r_{-i}^*) = 0_m.$$

Let $\mathrm{diag}(b)$ be a diagonal matrix with vector b on the diagonal and all other elements set to 0. Ignoring the second order and higher terms, we therefore have by the Taylor series approximation that

$$\begin{aligned}
& F_i(q_i, q_{-i}, r_{-i}) \\
\approx\; & \sum_{k=1}^{n} J_{q_k} F_i(q_k, q_{-k}^*, q_{-k}^*)\Big|_{q_k = q_k^*} (q_k - q_k^*) \;+\; \sum_{k \neq i} J_{r_k} F_i(q_k^*, q_{-k}^*, q_{-ki}^*, r_k)\Big|_{r_k = q_k^*} (r_k - q_k^*) \\
=\; & \sum_{k=1}^{n} J_{q_k} F_i(q_k, q_{-k}^*, q_{-k}^*)\Big|_{q_k = q_k^*} N \delta x_{q_k} \;+\; \sum_{k \neq i} J_{r_k} F_i(q_k^*, q_{-k}^*, q_{-ki}^*, r_k)\Big|_{r_k = q_k^*} N \delta x_{r_k} \\
=\; & -N \delta x_{q_i} + \sum_{k \neq i} J_{q_k} F_i(q_k, q_{-k}^*, q_{-k}^*)\Big|_{q_k = q_k^*} N \delta x_{q_k} \;+\; \sum_{k \neq i} J_{r_k} F_i(q_k^*, q_{-k}^*, q_{-ki}^*, r_k)\Big|_{r_k = q_k^*} N \delta x_{r_k} \\
=\; & -N \delta x_{q_i} \;+\; (1 + \gamma\lambda) \sum_{k \neq i} \tilde{D}_{ik} N \delta x_{q_k} \;-\; \gamma\lambda \sum_{k \neq i} \tilde{D}_{ik} N \delta x_{r_k},
\end{aligned}$$

where $\tilde{D}_{ik} \triangleq \dfrac{w_{ik}}{\tau} \nabla\zeta \left(\dfrac{A_i(q_{-i}^*) - z_i}{\tau} \right)$, and $\nabla\zeta(b) \triangleq \mathrm{diag}(\zeta(b)) - \zeta(b)\zeta^\top(b)$.

Define $D_{ik} = N^\top \tilde{D}_{ik} N$. Since $N^\top N = I_{m-1}$, it follows immediately from (48) that

$$\dot{\delta} x_{q_i} = -\delta x_{q_i} \;+\; (1 + \gamma\lambda) \sum_{k \neq i} D_{ik} \delta x_{q_k} \;-\; \gamma\lambda \sum_{k \neq i} D_{ik} \delta x_{r_k}. \tag{49}$$

Linearizing (45), we see that the Taylor approximation results in

$$\dot{\delta} x_{r_i} = \lambda(\delta x_{q_i} - \delta x_{r_i}). \tag{50}$$

We define

$$\mathcal{D} = \begin{bmatrix} 0 & D_{12} & D_{13} & \ldots & D_{1n} \\ D_{21} & 0 & D_{23} & \ldots & D_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & D_{n3} & \ldots & 0 \end{bmatrix}.$$

Combining (49) and (50) together, we can write

$$\dot{\delta} x = \begin{bmatrix} -I + (1 + \gamma\lambda)\mathcal{D} & -\gamma\lambda\mathcal{D} \\ \lambda I & -\lambda I \end{bmatrix} \delta x.$$

The statement of the theorem now follows immediately from the Hurwitz stability criterion. $\square$

Theorem AA2. (Convergence under SAP-GP/AA to CMNE) *Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a completely mixed NE under the dynamics in* (41)*. Then the linearization of* (41) *with $\gamma > 0$ is locally asymptotically stable for $\lambda > 0$ if and only if the following matrix is stable*

$$\begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda I & -\lambda I \end{bmatrix}.$$

Proof. Recall the ODE from (41):

$$\begin{aligned} \dot{q}_i &= \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i] - q_i && (51) \\ \dot{r}_i &= \lambda(q_i - r_i). && (52) \end{aligned}$$

At equilibrium $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$, $\dot{q}_i = 0$ and $\dot{r}_i = 0$. Therefore, using (56), we have:

$$q_i^* = \Pi_\Delta[q_i^* + \mathcal{A}_i(q_{-i}^*) - z_i].$$

Since the equilibrium is completely mixed, q_i^* is in the interior of $\Delta(A)$. We invoke Lemma 4.1 in (Shamma and Arslan, 2005) to get the following:

$$\begin{aligned} NN^\top(\mathcal{A}_i(q_{-i}^*) - z_i) &= 0 && (53) \\ \Pi_\Delta[q_i^* + \mathcal{A}_i(q_{-i}^*) - z_i + \delta y] - q_i^* &= NN^\top\Big(\mathcal{A}_i(q_{-i}^*) \;-\; z_i + \delta y\Big), \end{aligned}$$

for δy sufficiently small, and N as defined in the proof of Theorem AA1. Then, for a sufficiently small deviation δx, where δx is as defined in Theorem 1, we get the following dynamics:

$$\begin{aligned} \dot{q}_i &= NN^\top[\mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i] && (54) \\ \dot{r}_i &= \lambda(q_i - r_i). && (55) \end{aligned}$$

Linearizing these equations and noting that $N^\top N = I$, we get

$$\begin{aligned} \dot{\delta x}_{q_i} &= N^\top \left(NN^\top (1+\gamma\lambda) \sum_{k \neq i} w_{ik} N \delta x_{q_k} \right) \;-\; N^\top \left(NN^\top \gamma\lambda \sum_{k \neq i} w_{ik} N \delta x_{r_k} \right) \\ &= (1+\gamma\lambda) N^\top \sum_{k \neq i} w_{ik} N \delta x_{q_k} \;-\; \gamma\lambda N^\top \sum_{k \neq i} w_{ik} N \delta x_{r_k} \\ &= (1+\gamma\lambda) \sum_{k \neq i} w_{ik} \delta x_{q_k} \;-\; \gamma\lambda \sum_{k \neq i} w_{ik} \delta x_{r_k}, \end{aligned}$$

and

$$\dot{\delta x}_{r_i} = \lambda(\delta x_{q_i} - \delta x_{r_i}).$$

It follows immediately that

$$\dot{\delta x} = \begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda I & -\lambda I \end{bmatrix} \delta x,$$

where the weight matrix

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} & \ldots & w_{1n} \\ w_{21} & 0 & w_{23} & \ldots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \ldots & 0 \end{bmatrix}.$$

□

Theorem AA3. (Convergence under SAP-GP/AA to SNE) *Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a strict NE under the dynamics in* (41)*. The associated equilibrium point $(q_i = q_i^*, q_{-i} = q_{-i}^*, r_i = q_i^*, r_{-i} = q_{-i}^*)$ is locally asymptotically stable for any $\gamma > 0$ and $\lambda > 0$.*

Proof. Recall the ODE from (41):

$$\dot{q}_i = \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i] - q_i \tag{56}$$
$$\dot{r}_i = \lambda(q_i - r_i). \tag{57}$$

To prove the local asymptotic stability of the ODE dynamics, we will define a Lyapunov function $\mathcal{V}$ that is locally positive definite and has locally negative semi-definite time derivative. Consider

$$\mathcal{V}(q_i, q_{-i}, r_i, r_{-i}) \triangleq \frac{1}{2}\sum_{i=1}^{n} \left((q_i - q_i^*)^\top (q_i - q_i^*) + \lambda (r_i - q_i)^\top (r_i - q_i)\right). \tag{58}$$

We define the shorthand $d_i \triangleq q_i + \mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i$. Applying the chain rule, we see that the time derivative of $\mathcal{V}$,

$$\begin{aligned}
\dot{\mathcal{V}} &= \sum_{i=1}^{n} \left(\frac{\partial \mathcal{V}}{\partial q_i}\right)^\top \dot{q}_i + \sum_{i=1}^{n} \left(\frac{\partial \mathcal{V}}{\partial r_i}\right)^\top \dot{r}_i \\
&= \sum_{i=1}^{n} \left[(q_i - q_i^*) + \lambda(q_i - r_i)\right]^\top \dot{q}_i - \lambda^2 \sum_{i=1}^{n} (r_i - q_i)^\top (r_i - q_i) \\
&= \sum_{i=1}^{n} (q_i - q_i^*)^\top \dot{q}_i + \lambda \sum_{i=1}^{n} (q_i - r_i)^\top \dot{q}_i \; - \; \lambda^2 \sum_{i=1}^{n} ||r_i - q_i||^2 \\
&= \sum_{i=1}^{n} (q_i - q_i^*)^\top \Pi_\Delta(d_i) - \sum_{i=1}^{n} (q_i - q_i^*)^\top q_i \; + \; \lambda \sum_{i=1}^{n} (q_i - r_i)^\top \dot{q}_i - \lambda^2 \sum_{i=1}^{n} ||r_i - q_i||^2.
\end{aligned}$$

Also, we note that

$$\begin{aligned}
\sum_{i=1}^{n} ||\dot{q}_i||^2 &= \sum_{i=1}^{n} ||\Pi_\Delta(d_i) - q_i||^2 \\
&= \sum_{i=1}^{n} ||\Pi_\Delta(d_i)||^2 + \sum_{i=1}^{n} q_i^\top q_i - 2 \sum_{i=1}^{n} q_i^\top \Pi_\Delta(d_i).
\end{aligned}$$

This immediately implies

$$\begin{aligned}
\dot{\mathcal{V}} \; + \; \sum_{i=1}^{n} ||\dot{q}_i||^2 &= \sum_{i=1}^{n} \underbrace{\left(\Pi_\Delta(d_i) - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i\right)}_{(B)} \\
&\quad + \; \lambda \sum_{i=1}^{n} (q_i - r_i)^\top \dot{q}_i - \lambda^2 \sum_{i=1}^{n} ||r_i - q_i||^2.
\end{aligned} \tag{59}$$

Consider $(B) = \left(\Pi_\Delta(d_i) - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i\right)$. Since $\Delta(A)$ is a convex set, the projection property implies

$$\left[\Pi_\Delta(d_i)\right]^\top \left(\Pi_\Delta(d_i) - q_i\right) \leq d_i^\top \left(\Pi_\Delta(d_i) - q_i\right),$$

whence we note

$$\begin{aligned}
(B) &= \left(\Pi_\Delta(d_i) - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i\right) \\
&= \left[\Pi_\Delta(d_i)\right]^\top \left(\Pi_\Delta(d_i) - q_i\right) - \left(\Pi_\Delta(d_i) - q_i\right)^\top q_i^* \\
&\leq d_i^\top \left(\Pi_\Delta(d_i) - q_i\right) - \left(\Pi_\Delta(d_i) - q_i\right)^\top q_i^* \\
&= \left(d_i - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i\right) \\
&= \left(q_i + \mathcal{A}_i(q_{-i} + \gamma \dot{r}_{-i}) - z_i - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i\right).
\end{aligned}$$

Now, we note from the definition of $\mathcal{V}$ in (58) that by decreasing the distances $(q_i - r_i)$ and $(q_i - q_i^*)$, we can make $\mathcal{V}(q_i, q_{-i}, r_i, r_{-i})$ arbitrarily close to 0 from the right. In other words, we can consider a sufficiently small neighborhood around the equilibrium such that as $r_i, q_i \to q_i^*$, (B) tends to

$$\begin{aligned} & \left(q_i^* + \mathcal{A}_i(q_{-i}^* + \delta y) - z_i - q_i^*\right)^\top \left(\Pi_\Delta(d_i) - q_i^*\right) \\ = & \quad \left(\mathcal{A}_i(q_{-i}^* + \delta y) - z_i\right)^\top \left(\Pi_\Delta(d_i) - q_i^*\right), \\ = & \quad \left(\Pi_\Delta(d_i) - q_i^*\right)^\top \left.\frac{\partial U_i(q_i, q_{-i}^* + \delta y, z_i)}{\partial q_i}\right|_{q_i = q_i^*} < 0 \end{aligned}$$

for some sufficiently small δy and $\Pi_\Delta(d_i) \neq q_i^*$. The last inequality follows since $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ is a strict Nash equilibrium, whereby (a) (q_i^*, q_{-i}^*) is a pure strategy Nash equilibrium (since $\mathcal{A}_i(\cdot)$ is a linear transformation and the payoff maximization happens at the vertex), and (b) q_i^* is a (strictly) best response to q_{-i}^* and nearby strategies. Therefore, we see from (59) that for a sufficiently small neighborhood around the equilibrium point,

$$\begin{aligned} \dot{\mathcal{V}} & \leq -\sum_{i=1}^n ||\dot{q}_i||^2 + \lambda \sum_{i=1}^n (q_i - r_i)^\top \dot{q}_i - \lambda^2 \sum_{i=1}^n ||r_i - q_i||^2 \\ & = -\sum_{i=1}^n \left(||\dot{q}_i||^2 + \lambda^2 ||r_i - q_i||^2\right) + \lambda \sum_{i=1}^n (q_i - r_i)^\top \dot{q}_i \\ & \leq -\sum_{i=1}^n \left(||\dot{q}_i||^2 + \lambda^2 ||r_i - q_i||^2\right) + \frac{1}{2} \sum_{i=1}^n \left(||\dot{q}_i||^2 + \lambda^2 ||r_i - q_i||^2\right) \\ & = -\frac{1}{2} \sum_{i=1}^n \left(||\dot{q}_i||^2 + \lambda^2 ||r_i - q_i||^2\right), \end{aligned}$$

where we have invoked the Cauchy-Schwarz inequality in the penultimate line. Since this quantity is non-positive, we see that $\dot{\mathcal{V}}$ is locally negative semi-definite. Finally, it is clear from (58) that $\mathcal{V}(q_i, q_{-i}, r_i, r_{-i}) > 0$ in the neighborhood $(q_i, q_{-i}, r_i, r_{-i})$ of the equilibrium point $(q_i = q_i^*, q_{-i} = q_{-i}^*, r_i = q_i^*, r_{-i} = q_{-i}^*)$, and $\mathcal{V}(q_i^*, q_{-i}^*, q_i^*, q_{-i}^*) = 0$. Thus, $\mathcal{V}$ is locally positive definite, and the statement of the theorem follows. □

We will now characterize conditions for convergence in the passive aggregator setting.

Theorem PA1. (Convergence under SAP-FP/PA to NE) *Let the weight matrix W be stochastic. Let $(q_1^*, \ldots, q_n^*, z_1, \ldots, z_n)$ be a NE under the dynamics in* (40)*. There exists a matrix $\mathcal{D}_1$ with zero diagonal, and a block diagonal matrix $\mathcal{D}_2$ such that the linearization of* (40) *with $\gamma > 0$ is locally asymptotically stable for $\lambda > 0$ if and only if if the following matrix is stable*

$$\begin{bmatrix} -I + (1 + \gamma\lambda)\mathcal{D}_1 & -\gamma\lambda\mathcal{D}_2 \\ \lambda W & -\lambda I \end{bmatrix}.$$

Proof. We reproduce the ODE from (40):

$$\begin{aligned} \dot{q}_i &= \beta_i^\tau(\mathcal{A}_i(q_{-i}) + \gamma\dot{r}_i), z_i) - q_i && (60) \\ \dot{r}_i &= \lambda(\mathcal{A}_i(q_{-i}) - r_i). && (61) \end{aligned}$$

Note that at equilibrium $\dot{r}_i = 0$, but unlike Theorem AA1, r_i does not converge to q_i^*. Specifically, we note that the equilibrium (q_i^*, q_{-i}^*) corresponds to a point $(q_i(t) = q_i^*, q_{-i}(t) = q_{-i}^*, r_i(t) = \mathcal{A}_i(q_{-i}^*)), i \in [n]$, of the dynamics. Therefore, we will instead linearize around this point. Since the weight matrix W is stochastic, we must have $\mathcal{A}_i(q_{-i}^*) \in \Delta(A)$. Therefore, we can investigate the deviation of r_i around $\mathcal{A}_i(q_{-i}^*)$ with the help of matrix N defined in Theorem AA1. In particular, we can express the deviation vector $\delta x = (\delta x_{q_1}, \ldots, \delta x_{q_n}, \delta x_{r_1}, \ldots, \delta x_{r_n})^\top$ as:

$$\Big(q_1(t) - q_1^*, \ldots, q_n(t) - q_n^*,\ r_1(t) - \mathcal{A}_1(q_{-1}^*), \ldots, r_n(t) - \mathcal{A}_n(q_{-n}^*)\Big)^\top = \mathcal{N}\delta x(t), \tag{62}$$

where the block diagonal matrix $\mathcal{N}$ is as defined in Theorem AA1. Linearizing around our equilibrium point and proceeding similarly to Theorem AA1, we get

$$\dot{\delta x}_{q_i} = -\delta x_{q_i} + (1 + \gamma\lambda) \sum_{k \neq i} D_{ik} \delta x_{q_k} - \gamma\lambda C_i \delta x_{r_i}. \tag{63}$$

where

$$D_{ik} \triangleq \frac{w_{ik}}{\tau} N^\top \nabla\zeta\left(\frac{A_i(q^*_{-i}) - z_i}{\tau}\right) N,$$
$$C_i \triangleq \frac{1}{\tau} N^\top \nabla\zeta\left(\frac{A_i(q^*_{-i}) - z_i}{\tau}\right) N,$$

and

$$\nabla\zeta(b) \triangleq \mathrm{diag}(\zeta(b)) - \zeta(b)\zeta^\top(b),$$

with $\zeta(b)$ the same as in Theorem AA1. Additionally, we have

$$\dot{\delta} x_{r_i} = \lambda \sum_{k \neq i} w_{ik} \delta x_{q_i} - \lambda \delta x_{r_i}. \tag{64}$$

Recall that the weight matrix

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & 0 & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & 0 \end{bmatrix}.$$

Define

$$\mathcal{D}_1 \triangleq \begin{bmatrix} 0 & D_{12} & D_{13} & \dots & D_{1n} \\ D_{21} & 0 & D_{23} & \dots & D_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & D_{n3} & \dots & 0 \end{bmatrix}, \quad \text{and}$$

$$\mathcal{D}_2 \triangleq \begin{bmatrix} C_1 & 0 & 0 & \dots & 0 \\ 0 & C_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_n \end{bmatrix}.$$

Then, the proof follows by combining (63) and (64), since we can express the deviations as

$$\dot{\delta} x = \begin{bmatrix} -I + (1+\gamma\lambda)\mathcal{D}_1 & -\gamma\lambda\mathcal{D}_2 \\ \lambda W & -\lambda I \end{bmatrix} \delta x.$$

□

Theorem PA2. (Convergence under SAP-GP/PA to CMNE) *Let the weight matrix W be stochastic. Let $(q_1^*, \dots, q_n^*, z_1, \dots, z_n)$ be a completely mixed NE under the dynamics in* (42). *Then the linearization of* (42) *with $\gamma > 0$ is locally asymptotically stable for $\lambda > 0$ if and only if the following matrix is stable*

$$\begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda W & -\lambda I \end{bmatrix}.$$

Proof. Recall the ODE from (42):

$$\dot{q}_i = \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i}) + \gamma\dot{r}_i - z_i] - q_i \tag{65}$$
$$\dot{r}_i = \lambda(\mathcal{A}_i(q_{-i}) - r_i). \tag{66}$$

At equilibrium $(q_1^*, \dots, q_n^*, z_1, \dots, z_n)$, $\dot{q}_i = 0$ and $\dot{r}_i = 0$. Therefore, using (65), we have:

$$q_i^* = \Pi_\Delta[q_i^* + \mathcal{A}_i(q_{-i}^*) - z_i].$$

Proceeding along the lines of proof of Theorem AA2, for a sufficiently small deviation δx as defined in Theorem PA1, we can equivalently analyze the following dynamics:

$$\dot{q}_i = NN^\top[\mathcal{A}_i(q_{-i}) + \gamma\dot{r}_{-i} - z_i]$$
$$\dot{r}_i = \lambda(q_i - r_i).$$

Linearizing these equations and noting $N^\top N = I$, we get

$$\begin{aligned}
\dot{\delta x}_{q_i} &= N^\top \left(NN^\top (1+\gamma\lambda) \sum_{k\neq i} w_{ik} N \delta x_{q_k} \right) \; - \; N^\top \left(NN^\top \gamma\lambda \sum_{k\neq i} N \delta x_{r_k} \right) \\
&= (1+\gamma\lambda) N^\top \sum_{k\neq i} w_{ik} N \delta x_{q_k} - \gamma\lambda N^\top \sum_{k\neq i} w_{ik} N \delta x_{r_k} \\
&= (1+\gamma\lambda) \sum_{k\neq i} w_{ik} \delta x_{q_k} - \gamma\lambda \sum_{k\neq i} w_{ik} \delta x_{r_k},
\end{aligned}$$

and

$$\dot{\delta x}_{r_i} = \lambda \sum_{k\neq i} w_{ik} \delta x_{q_i} - \lambda \delta x_{r_i}.$$

It follows immediately that

$$\dot{\delta x} = \begin{bmatrix} (1+\gamma\lambda)W & -\gamma\lambda W \\ \lambda W & -\lambda I \end{bmatrix} \delta x,$$

where the weight matrix

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & 0 & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & 0 \end{bmatrix}.$$

□

Theorem PA3. (Convergence under SAP-GP/PA to SNE) *Let the weight matrix W be doubly stochastic. Let $(q_1^*, \dots, q_n^*, z_1, \dots, z_n)$ be a strict NE under the dynamics in (42). The associated equilibrium point $(q_i = q_i^*, r_i = A_i(q_{-i}^*))_{i\in[n]}$ is locally asymptotically stable for sufficiently small $\gamma\lambda$, where $\gamma > 0$ and $\lambda > 0$.*

Proof. Recall the ODE from (42):

$$\begin{aligned}
\dot{q}_i &= \Pi_\Delta[q_i + \mathcal{A}_i(q_{-i}) + \gamma \dot{r}_i - z_i] - q_i \\
\dot{r}_i &= \lambda(\mathcal{A}_i(q_{-i}) - r_i).
\end{aligned}$$

We will prove local asymptotic stability via a Lyapunov function $\mathcal{V}$ that is locally positive definite and has locally negative semi-definite time derivative. Consider

$$\mathcal{V}(q_i, q_{-i}, r_i, r_{-i}) \triangleq \frac{1}{2} \sum_{i=1}^n \Big((q_i - q_i^*)^\top (q_i - q_i^*) \; + \; \lambda \left(r_i - \mathcal{A}_i(q_{-i}) \right)^\top \left(r_i - \mathcal{A}_i(q_{-i}) \right) \Big). \tag{67}$$

We define the shorthand $\tilde{d}_i \triangleq q_i + \mathcal{A}_i(q_{-i}) + \gamma \dot{r}_i - z_i$. Applying the chain rule, we see that the time derivative of $\mathcal{V}$,

$$\begin{aligned}
\dot{\mathcal{V}} &= \sum_{i=1}^n \left(\frac{\partial \mathcal{V}}{\partial q_i} \right)^\top \dot{q}_i + \sum_{i=1}^n \left(\frac{\partial \mathcal{V}}{\partial r_i} \right)^\top \dot{r}_i \\
&= \sum_{i=1}^n \left[(q_i - q_i^*) - \lambda \sum_{k\neq i} w_{ki} (r_k - \mathcal{A}_k(q_{-k})) \right]^\top \dot{q}_i \; - \; \lambda^2 \sum_{i=1}^n \left(r_i - \mathcal{A}_i(q_{-i}) \right)^\top \left(r_i - \mathcal{A}_i(q_{-i}) \right)) \\
&= \sum_{i=1}^n (q_i - q_i^*)^\top \dot{q}_i - \lambda^2 \sum_{i=1}^n ||r_i - \mathcal{A}_i(q_{-i})||^2 \; - \; \lambda \sum_{i=1}^n \left(\sum_{k\neq i} w_{ki} (r_k - \mathcal{A}_k(q_{-k})) \right)^\top \dot{q}_i \; .
\end{aligned}$$

Also, we note that

$$\begin{aligned}\sum_{i=1}^{n}||\dot{q}_i||^2 &= \sum_{i=1}^{n}||\Pi_\Delta(\tilde{d}_i)-q_i||^2\\ &= \sum_{i=1}^{n}||\Pi_\Delta(\tilde{d}_i)||^2+\sum_{i=1}^{n}q_i^\top q_i-2\sum_{i=1}^{n}q_i^\top\Pi_\Delta(\tilde{d}_i).\end{aligned}$$

It can be shown that

$$\begin{aligned}(q_i-q_i^*)^\top\dot{q}_i+||\dot{q}_i||^2 &= (\Pi_\Delta(\tilde{d}_i)-q_i^*)^\top(\Pi_\Delta(\tilde{d}_i)-q_i)\\ &\le (\tilde{d}_i-q_i^*)^\top(\Pi_\Delta(\tilde{d}_i)-q_i)\\ &= (q_i+\mathcal{A}_i(q_{-i})+\gamma\dot{r}_i-z_i-q_i^*)^\top(\Pi_\Delta(\tilde{d}_i)-q_i)\ .\end{aligned}$$

As $r_i, q_i \to q_i^*$, this quantity tends to

$$\begin{aligned}&\left(q_i^*+\mathcal{A}_i(q_{-i}^*)+\gamma\lambda(\mathcal{A}_i(q_{-i}^*)-q_i^*)-z_i-q_i^*\right)^\top\left(\Pi_\Delta(\tilde{d}_i)-q_i^*\right)\\ =\;&\left(\mathcal{A}_i(q_{-i}^*)+\gamma\lambda(\mathcal{A}_i(q_{-i}^*)-q_i^*)-z_i\right)^\top\left(\Pi_\Delta(\tilde{d}_i)-q_i^*\right)\\ =\;&\left((1+\gamma\lambda)\mathcal{A}_i(q_{-i}^*)-\gamma\lambda q_i^*-z_i\right)^\top\left(\Pi_\Delta(\tilde{d}_i)-q_i^*\right)\\ =\;&\left(\mathcal{A}_i((1+\gamma\lambda)q_{-i}^*)-\gamma\lambda q_i^*-z_i\right)^\top\left(\Pi_\Delta(\tilde{d}_i)-q_i^*\right)\end{aligned}$$

which can be expressed in the form

$$\left(\Pi_\Delta(\tilde{d}_i)-q_i^*\right)^\top\left.\frac{\partial U_i(q_i,q_{-i}^*+\delta y,z_i)}{\partial q_i}\right|_{q_i=q_i^*} < 0$$

when $\gamma\lambda$ is sufficiently small and $\Pi_\Delta(\tilde{d}_i)\neq q_i^*$, by arguing along the lines of proof for Theorem AA3. Therefore,

$$\begin{aligned}\dot{\mathcal{V}} &\le -\sum_{i=1}^{n}||\dot{q}_i||^2+\lambda\sum_{i=1}^{n}\left(\sum_{k\neq i}w_{ki}(\mathcal{A}_k\left(q_{-k}\right)-r_k)\right)^\top\dot{q}_i\ -\ \lambda^2\sum_{i=1}^{n}||r_i-\mathcal{A}_i(q_{-i})||^2\\ &= -\sum_{i=1}^{n}\left(||\dot{q}_i||^2+\lambda^2||r_i-\mathcal{A}_i(q_{-i})||^2\right)\ +\ \sum_{i=1}^{n}\sum_{k\neq i}w_{ki}\left(\lambda(\mathcal{A}_k\left(q_{-k}\right)-r_k)^\top\dot{q}_i\right)\\ &\le -\sum_{i=1}^{n}\left(||\dot{q}_i||^2+\lambda^2||r_i-\mathcal{A}_i(q_{-i})||^2\right)\ +\ \frac{1}{2}\sum_{i=1}^{n}\sum_{k\neq i}w_{ki}\left(\lambda^2||r_k-\mathcal{A}_k(q_{-k})||^2+||\dot{q}_i||^2\right)\end{aligned}$$

by noting that $w_{ki}\ge 0$ for all $i\in[n]$, $k\neq i$ and invoking Cauchy-Schwarz. Furthermore, since W is doubly stochastic, we have $\sum_{k\neq i}w_{ki}=1$ and $\sum_{k\neq i}w_{ik}=1$ for all $i\in[n]$. Thus, we may decompose the second term on the right in the last equation as

$$\begin{aligned}&\frac{1}{2}\sum_{i=1}^{n}\sum_{k\neq i}w_{ki}\left(\lambda^2||r_k-\mathcal{A}_k(q_{-k})||^2+||\dot{q}_i||^2\right)\\ =\;&\frac{\lambda^2}{2}\sum_{i=1}^{n}\sum_{k\neq i}w_{ki}||r_k-\mathcal{A}_k(q_{-k})||^2+\frac{1}{2}\sum_{i=1}^{n}||\dot{q}_i||^2\sum_{k\neq i}w_{ki}\\ =\;&\frac{\lambda^2}{2}\sum_{i=1}^{n}\sum_{k\neq i}w_{ki}||r_k-\mathcal{A}_k(q_{-k})||^2+\frac{1}{2}\sum_{i=1}^{n}||\dot{q}_i||^2.\end{aligned}$$

The first term in the last equation may be interpreted as a weighted outgoing flow from player i to player $k \neq i$. Now viewing this from the equivalent perspective of total incoming flow, we have

$$\begin{aligned}
\dot{\mathcal{V}} &\leq -\sum_{i=1}^{n}\left(||\dot{q}_i||^2 + \lambda^2||r_i - \mathcal{A}_i(q_{-i})||^2\right) \;+\; \frac{\lambda^2}{2}\sum_{i=1}^{n}\sum_{k\neq i} w_{ik}||r_i - \mathcal{A}_i(q_{-i})||^2 + \frac{1}{2}\sum_{i=1}^{n}||\dot{q}_i||^2 \\
&= -\sum_{i=1}^{n}\left(||\dot{q}_i||^2 + \lambda^2||r_i - \mathcal{A}_i(q_{-i})||^2\right) \;+\; \frac{\lambda^2}{2}\sum_{i=1}^{n}||r_i - \mathcal{A}_i(q_{-i})||^2\sum_{k\neq i} w_{ik} + \frac{1}{2}\sum_{i=1}^{n}||\dot{q}_i||^2 \\
&= -\sum_{i=1}^{n}\left(||\dot{q}_i||^2 + \lambda^2||r_i - \mathcal{A}_i(q_{-i})||^2\right) \;+\; \frac{1}{2}\sum_{i=1}^{n}\left(\lambda^2||r_i - \mathcal{A}_i(q_{-i})||^2 + ||\dot{q}_i||^2\right) \\
&= -\frac{1}{2}\sum_{i=1}^{n}\left(\lambda^2||r_i - \mathcal{A}_i(q_{-i})||^2 + ||\dot{q}_i||^2\right) \\
&\leq 0,
\end{aligned}$$

which implies that $\dot{\mathcal{V}}$ is locally negative semi-definite. The local positive definiteness of $\mathcal{V}$ may be argued similarly to the proof of Theorem AA3 and we are done. □