The Best of Both Worlds in Network Population Games: Reaching Consensus & Convergence to Equilibrium

Anonymous Author(s)
Affiliation
Address
`email`

Contents

1 Experiments on Network Competition

We have shown in Theorem 5 (of the main paper) that SFP converges to a unique QRE in any weighted zero-sum network population game even if there are multiple Nash equilibria underlying that game. In the following, we corroborate this by providing empirical evidence in agent-based simulations with different belief initialization.

Game Description. Consider a five-population asymmetric matching pennies game [3], where the network structure is a line (depicted in Figure 1). Each agent has two actions $\{H, T\}$. Agents in populations 1 and 5 do not learn; they always play strategies H and T, respectively. For agents in populations 2 to 4, they receive $+1$ if they match the strategy of the opponent in the next population, and receive -1 if they mismatch. On the contrary, they receive $+1$ if they mismatch the strategy of the opponent in the previous population, and receive -1 if they match. Hence, this game has infinitely many Nash equilibria of the form: agents in populations 2 and 4 play strategy T, whereas agents in population 3 are indifferent between strategies H and T.

Experimental Setups. In this game, agents in each population form two beliefs (one for the previous population and one for the next population). We are mainly interested in the strategies of population 3, as the Nash equilibria differ in the strategies in population 3. Thus, we let the initial beliefs about populations 1, 3 and 5 remain unchanged across different cases, and vary population 3's initial beliefs about populations 2 and 4. The initial beliefs about populations 1, 3 and 5, denoted by μ_{1H}, μ_{3H} and μ_{5H}, are distributed according to the distributions Beta$(20, 10)$, Beta$(6, 4)$, and Beta$(10, 5)$, respectively. The initial beliefs about populations are given in the legends of Figure 2. In all cases, the initial sum of weights $\lambda = 10$ and the temperature $\beta = 10$. Note that $\mu_{iT} = 1 - \mu_{iH}$ for all populations $i = 1, 2, 3, 4, 5$. We run 100 simulation runs for each initialization, and each simulation run consists of $1,000$ agents in each population.

Results. As shown in Figure 2, given differential initialization of beliefs, agents in population 3 converge to the same equilibrium where they all take strategy H with probability 0.5. Therefore, even when the underlying zero-sum game has many Nash equilibria, SFP with different initial belief heterogeneity selects a unique equilibrium, addressing the problem of equilibrium selection.

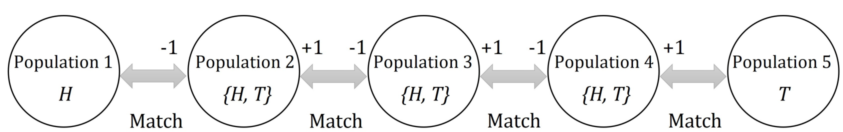

Figure 1: Asymmetric Matching Pennies.

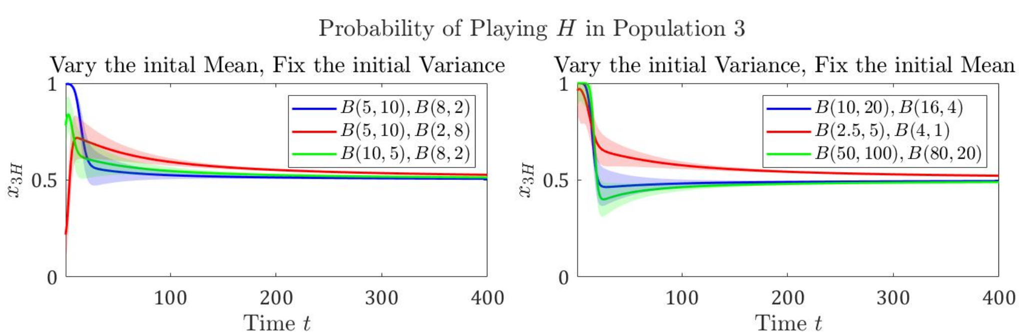

Figure 2: With different belief initialization, SFP selects a unique equilibrium where all agents in population 3 play strategy H with probability 0.5. The thin lines represent the mean mixed strategy (the choice probability of H) and the shaded areas represent the variance of the mixed strategies in the population. In the legends, B denotes Beta distribution; the two Beta distributions correspond to the initial beliefs about the neighbor populations 2 and 4, respectively.

2 Proof of Proposition 1

It follows from Equation 2 and Equation 3 of the main paper that the change in $\mu_j^i(k,t)$ between two discrete time steps is

$$\mu_j^i(k,t+1) = \mu_j^i(k,t) + \frac{\bar{\mathbf{x}}_j(t) - \mu_j^i(k,t)}{\lambda + t + 1}. \tag{1}$$

Lemma 1. *Under Assumption 1 of the main paper, for an arbitrary agent k in population i, its belief $\mu_j^i(k,t)$ about a neighbor population j will never reach the extreme belief, i.e., the probability density for the boundary of the simplex Δ_i will remain zero.*

Proof. Assumption 1 ensures that $\bar{\mathbf{x}}_j(0)$ is in the interior of the simplex Δ_j. Moreover, the logit choice function (Equation 5 in the main paper) also ensures that $\bar{\mathbf{x}}_j(t)$ stays in the interior of Δ_j afterwards for a finite temperature β. Hence, from Equation 1, one can see that $\mu_j^i(k,t)$ for every time step t will stay in the interior of Δ_j. $\square$

In the following, for notation convenience, we sometimes drop the agent index k and the time index t depending on the context. Consider a population i. We rewrite the change in the beliefs about this population as follows.

$$\mu_i(t+1) = \mu_i(t) + \frac{\bar{\mathbf{x}}_i(t) - \mu_i(t)}{\lambda + t + 1}. \tag{2}$$

Suppose that the amount of time that passes between two successive time steps is $\delta \in (0,1]$. We rewrite the above equation as

$$\mu_i(t+\delta) = \mu_i(t) + \delta\frac{\bar{\mathbf{x}}_i(t) - \mu_i(t)}{\lambda + t + 1}. \tag{3}$$

Next, we consider a test function $\theta(\mu_i)$. Define

$$Y = \frac{\mathbb{E}[\theta(\mu_i(t+\delta))] - \mathbb{E}[\theta(\mu_i(t))]}{\delta}. \tag{4}$$

Applying Taylor series for $\theta(\mu_i(t+\delta))$ at $\mu_i(t)$, we obtain

$$\begin{aligned}\theta(\mu_i(t+\delta)) = \theta(\mu_i(t)) &+ \frac{\delta}{\lambda+t+1}\partial_{\mu_i}\theta(\mu_i)\left[\bar{\mathbf{x}}_i(t) - \mu_i(t)\right] \\ &+ \frac{\delta^2}{2(\lambda+t+1)^2}\left[\bar{\mathbf{x}}_i(t) - \mu_i(t)\right]^\top \mathbf{H}\theta(\mu_i)\left[\bar{\mathbf{x}}_i(t) - \mu_i(t)\right] \\ &+ o\left(\left[\delta\frac{\bar{\mathbf{x}}_i(t) - \mu_i(t)}{\lambda+t+1}\right]^2\right)\end{aligned} \tag{5}$$

where $\mathbf{H}$ denotes the Hessian matrix. Hence, the expectation $\mathbb{E}[\theta(\mu_i(t+\delta))]$ is

$$\begin{aligned}\mathbb{E}[\theta(\mu_i(t+\delta))] = \mathbb{E}[\theta(\mu_i(t))] &+ \frac{\delta}{\lambda+t+1}\mathbb{E}[\partial_{\mu_i}\theta(\mu_i(t))(\bar{\mathbf{x}}_i(t) - \mu_i(t))] \\ &+ \frac{\delta^2}{2(\lambda+t+1)^2}\mathbb{E}\left[\left[\bar{\mathbf{x}}_i(t) - \mu_i(t)\right]^\top \mathbf{H}\theta(\mu_i)\left[\bar{\mathbf{x}}_i(t) - \mu_i(t)\right]\right] \\ &+ \frac{\delta^2}{2(\lambda+t+1)^2}\mathbb{E}[o([\bar{\mathbf{x}}_i(t) - \mu_i(t)]^2)]\end{aligned} \tag{6}$$

Moving the term $\mathbb{E}[\theta(\mu_i(t))]$ to the left hand side and dividing both sides by δ, we recover the quantity Y, i.e.,

$$\begin{aligned}Y = &\frac{1}{\lambda+t+1}\mathbb{E}[\partial_{\mu_i}\theta(\mu_i(t))(\bar{\mathbf{x}}_i(t) - \mu_i(t))] \\ &+ \frac{\delta}{2(\lambda+t+1)^2}\mathbb{E}[[\bar{\mathbf{x}}_i(t) - \mu_i(t)]^\top \mathbf{H}\theta(\mu_i(t))[\bar{\mathbf{x}}_i(t) - \mu_i(t)] + o\left((\bar{\mathbf{x}}_i(t) - \mu_i(t))^2\right)]\end{aligned} \tag{7}$$

Taking the limit of Y with $\delta \to 0$, the contribution of the second term on the right hand side vanishes, yielding

$$\lim_{\delta\to 0} Y = \frac{1}{\lambda+t+1}\mathbb{E}[\partial_{\boldsymbol{\mu}_i}\theta(\boldsymbol{\mu}_i(t))(\bar{\mathbf{x}}_i(t)-\boldsymbol{\mu}_i(t))] \tag{8}$$

$$= \frac{1}{\lambda+t+1}\int p(\boldsymbol{\mu}_i(t),t)\left[\partial_{\boldsymbol{\mu}_i}\theta(\boldsymbol{\mu}_i(t))(\bar{\mathbf{x}}_i(t)-\boldsymbol{\mu}_i(t))\right]d\boldsymbol{\mu}_i(t). \tag{9}$$

Apply integration by parts. We obtain

$$\lim_{\delta\to 0} Y = 0 - \frac{1}{\lambda+t+1}\int \theta(\boldsymbol{\mu}_i(t))\nabla\cdot\left[p(\boldsymbol{\mu}_i(t),t)(\bar{\mathbf{x}}_i(t)-\boldsymbol{\mu}_i(t))\right]d\boldsymbol{\mu}_i(t) \tag{10}$$

where we have leveraged that the probability mass $p(\boldsymbol{\mu}_i,t)$ at the boundary $\partial\Delta_i$ remains zero as a result of Lemma 1. On the other hand, according to the definition of Y,

$$\lim_{\delta\to 0} Y = \lim_{\delta\to 0}\int \theta(\boldsymbol{\mu}_i(t))\frac{p(\boldsymbol{\mu}_i,t+\delta)-p(\boldsymbol{\mu}_i,t)}{\delta}d\boldsymbol{\mu}_i = \int \theta(\boldsymbol{\mu}_i(t))\partial_t p(\boldsymbol{\mu}_i,t)d\boldsymbol{\mu}_i. \tag{11}$$

Therefore, we have the equality

$$\int \theta(\boldsymbol{\mu}_i(t))\partial_t p(\boldsymbol{\mu}_i,t)d\boldsymbol{\mu}_i = -\frac{1}{\lambda+t+1}\int \theta(\boldsymbol{\mu}_i(t))\nabla\cdot\left[p(\boldsymbol{\mu}_i(t),t)(\bar{\mathbf{x}}_i(t)-\boldsymbol{\mu}_i(t))\right]d\boldsymbol{\mu}_i(t). \tag{12}$$

As θ is a test function, this leads to

$$\partial_t p(\boldsymbol{\mu}_i,t) = -\frac{1}{\lambda+t+1}\nabla\cdot\left[p(\boldsymbol{\mu}_i(t),t)(\bar{\mathbf{x}}_i(t)-\boldsymbol{\mu}_i(t))\right]. \tag{13}$$

Rearranging the terms, we obtain Equation 7 of the main paper. By the definition of expectation given a probability distribution, it is straightforward to obtain Equation 8 of the main paper. Q.E.D.

3 Proof of Theorem 1

Without loss of generality, we consider the variance of the belief μ_{is_i} about strategy s_i of population i. Note that

$$\mathrm{Var}(\mu_{is_i}) = \mathbb{E}[(\mu_{is_i})^2] - (\bar{\mu}_{is_i})^2. \tag{14}$$

Hence, we have

$$\frac{d\mathrm{Var}(\mu_{is_i})}{dt} = \frac{d\mathbb{E}[(\mu_{is_i})^2]}{dt} - 2\bar{\mu}_{is_i}\frac{d\bar{\mu}_{is_i}}{dt}. \tag{15}$$

Consider the first term on the right hand side. We apply the Leibniz rule to interchange differentiation and integration, and then substitute $\frac{\partial p(\boldsymbol{\mu}_i,t)}{\partial t}$ with Equation 8 in the main paper.

$$\frac{d\mathbb{E}[(\mu_{is_i})^2]}{dt}$$

$$= \int (\mu_{is_i})^2\frac{\partial p(\boldsymbol{\mu}_i,t)}{\partial t}d\boldsymbol{\mu}_i \tag{16}$$

$$= -\int (\mu_{is_i})^2\nabla\cdot\left(p(\boldsymbol{\mu}_i,t)\frac{\bar{\mathbf{x}}_i-\boldsymbol{\mu}_i}{\lambda+t+1}\right)d\boldsymbol{\mu}_i \tag{17}$$

$$= -\int (\mu_{is_i})^2\sum_{s_i\in S_i}\partial_{\mu_{is_i}}\left(p(\boldsymbol{\mu}_i,t)\frac{\bar{x}_{is_i}-\mu_{is_i}}{\lambda+t+1}\right)d\boldsymbol{\mu}_i \tag{18}$$

$$= \gamma\int (\mu_{is_i})^2\sum_{s_i\in S_i}\partial_{\mu_{is_i}}p(\boldsymbol{\mu}_i,t)\left(\bar{x}_{is_i}-\mu_{is_i}\right)d\boldsymbol{\mu}_i + \gamma\int (\mu_{is_i})^2 p(\boldsymbol{\mu}_i,t)\sum_{s_i\in S_i}\partial_{\mu_{is_i}}\left(\bar{x}_{is_i}-\mu_{is_i}\right)d\boldsymbol{\mu}_i \tag{19}$$

where $\gamma := -\frac{1}{\lambda+t+1}$. Applying integration by parts to the first term in Equation 19 yields

$$\int (\mu_{is_i})^2\sum_{s_i\in S_i}\partial_{\mu_{is_i}}p(\boldsymbol{\mu}_i,t)\left(\bar{x}_{is_i}-\mu_{is_i}\right)d\boldsymbol{\mu}_i$$

$$= -\int (\mu_{is_i})^2 p(\boldsymbol{\mu}_i,t)\left[\sum_{s'_i\in S_i}\partial_{\mu_{is'_i}}\left(\bar{x}_{is'_i}-\mu_{is'_i}\right)\right] + p(\boldsymbol{\mu}_i,t)\partial_{\mu_{is_i}}\left[(\mu_{is_i})^2(\bar{x}_{is_i}-\mu_{is_i})\right]d\boldsymbol{\mu}_i \tag{20}$$

where we have leveraged that the probability mass at the boundary remains zero (Lemma 1). Combining the above two equations, we obtain

$$\begin{aligned}
&\frac{d\mathbb{E}[(\mu_{is_i})^2]}{dt} \\
&= -\gamma \int (\mu_{is_i})^2 p(\boldsymbol{\mu}_i, t) \left[\sum_{s_i' \in S_i} \partial_{\mu_{is_i'}} (\bar{x}_{is_i'} - \mu_{is_i'}) \right] + p(\boldsymbol{\mu}_i, t) \partial_{\mu_{is_i}} \left[(\mu_{is_i})^2 (\bar{x}_{is_i} - \mu_{is_i}) \right] d\boldsymbol{\mu}_i \\
&\qquad + \gamma \int (\mu_{is_i})^2 p(\boldsymbol{\mu}_i, t) \sum_{s_i \in S_i} \partial_{\mu_{is_i}} (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i && (21) \\
&= \gamma \int \left[-p(\boldsymbol{\mu}_i, t) \partial_{\mu_{is_i}} \left[(\mu_{is_i})^2 (\bar{x}_{is_i} - \mu_{is_i}) \right] \right] + (\mu_{is_i})^2 p(\boldsymbol{\mu}_i, t) \partial_{\mu_{is_i}} (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i && (22) \\
&= \gamma \int 2(\mu_{is_i})^2 p(\boldsymbol{\mu}_i, t) d\boldsymbol{\mu}_i - \gamma \int 2\bar{x}_{is_i} \mu_{is_i} p(\boldsymbol{\mu}_i, t) d\boldsymbol{\mu}_i && (23) \\
&= -\frac{2\mathbb{E}[(\mu_{is_i})^2] - 2\bar{x}_{is_i} \bar{\mu}_{is_i}}{\lambda + t + 1}. && (24)
\end{aligned}$$

Next, we consider the second term in Equation 15. By Lemma 2, we have

$$2\bar{\mu}_{is_i} \frac{d\bar{\mu}_{is_i}}{dt} = \frac{2\bar{\mu}_{is_i}(\bar{x}_{is_i} - \bar{\mu}_{is_i})}{\lambda + t + 1}. \tag{25}$$

Combining Equations 24 and 25, the dynamics of the variance is

$$\begin{aligned}
\frac{d\mathrm{Var}(\mu_{is_i})}{dt} &= -\frac{2\mathbb{E}[(\mu_{is_i})^2] - 2\bar{x}_{is_i} \bar{\mu}_{is_i}}{\lambda + t + 1} - \frac{2\bar{\mu}_{is_i}(\bar{x}_{is_i} - \bar{\mu}_{is_i})}{\lambda + t + 1} && (26) \\
&= \frac{2(\bar{\mu}_{is_i})^2 - 2\mathbb{E}[(\mu_{is_i})^2]}{\lambda + t + 1} && (27) \\
&= -\frac{2\mathrm{Var}(\mu_{is_i})}{\lambda + t + 1}. && (28)
\end{aligned}$$

Q.E.D.

4 Proof of Proposition 2

Lemma 2. *The dynamics of the mean belief $\bar{\boldsymbol{\mu}}_i$ about each population $i \in V$ is governed by a differential equation*

$$\frac{d\bar{\mu}_{is_i}}{dt} = \frac{\bar{x}_{is_i} - \bar{\mu}_{is_i}}{\lambda + t + 1}, \qquad \forall s_i \in S_i. \tag{29}$$

Proof. The time derivative of the mean belief about strategy s_i is

$$\frac{d\bar{\mu}_{is_i}}{dt} = \frac{d}{dt} \int \mu_{is_i} p(\boldsymbol{\mu}_i, t) d\boldsymbol{\mu}_i. \tag{30}$$

We apply the Leibniz rule to interchange differentiation and integration, and then substitute $\frac{\partial p(\boldsymbol{\mu}_i,t)}{\partial t}$ with Equation 8 in the main paper.

$$\frac{d}{dt}\int \mu_{is_i} p(\boldsymbol{\mu}_i,t) d\boldsymbol{\mu}_i \tag{31}$$
$$= \int \mu_{is_i} \frac{\partial p(\boldsymbol{\mu}_i,t)}{\partial t} d\boldsymbol{\mu}_i \tag{32}$$
$$= -\int \mu_{is_i} \nabla \cdot \left(p(\boldsymbol{\mu}_i,t) \frac{\bar{\mathbf{x}}_i - \boldsymbol{\mu}_i}{\lambda+t+1} \right) d\boldsymbol{\mu}_i \tag{33}$$
$$= -\int \mu_{is_i} \sum_{s_i\in S_i} \partial_{\mu_{is_i}} \left(p(\boldsymbol{\mu}_i,t) \frac{\bar{x}_{is_i} - \mu_{is_i}}{\lambda+t+1} \right) d\boldsymbol{\mu}_i \tag{34}$$
$$= \gamma \left[\int \mu_{is_i} \sum_{s_i\in S_i} \left(\partial_{\mu_{is_i}} p(\boldsymbol{\mu}_i,t) \right) (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i + \int \mu_{is_i} p(\boldsymbol{\mu}_i,t) \sum_{s_i\in S_i} \partial_{\mu_{is_i}} (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i \right] \tag{35}$$

where $\gamma := -\frac{1}{\lambda+t+1}$. Apply integration by parts to the first term in Equation 35.

$$\int \mu_{is_i} \sum_{s_i\in S_i} \left(\partial_{\mu_{is_i}} p(\boldsymbol{\mu}_i,t) \right) (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i$$
$$= -\int \mu_{is_i} p(\boldsymbol{\mu}_i,t) \left[\sum_{s'_i\in S_i} \partial_{\mu_{is'_i}} (\bar{x}_{is'_i} - \mu_{is'_i}) \right] + p(\boldsymbol{\mu}_i,t) \partial_{\mu_{is_i}} \left[\mu_{is_i} (\bar{x}_{is_i} - \mu_{is_i}) \right] d\boldsymbol{\mu}_i \tag{36}$$

where we have leveraged that the probability mass at the boundary remains zero. Hence, it follows from Equation 35 that

$$\frac{d}{dt}\int \mu_{is_i} p(\boldsymbol{\mu}_i,t) d\boldsymbol{\mu}_i \tag{37}$$
$$= -\gamma \int \mu_{is_i} p(\boldsymbol{\mu}_i,t) \sum_{s'_i\in S_i} \partial_{\mu_{is'_i}} (\bar{x}_{is'_i} - \mu_{is'_i}) d\boldsymbol{\mu}_i - \gamma \int p(\boldsymbol{\mu}_i,t) \partial_{\mu_{is_i}} \left[\mu_{is_i} (\bar{x}_{is_i} - \mu_{is_i}) \right] d\boldsymbol{\mu}_i$$
$$+ \gamma \int \mu_{is_i} p(\boldsymbol{\mu}_i,t) \sum_{s_i\in S_i} \partial_{\mu_{is_i}} (\bar{x}_{is_i} - \mu_{is_i}) \, d\boldsymbol{\mu}_i \tag{38}$$
$$= \gamma \int p(\boldsymbol{\mu}_i,t) \left[\mu_{is_i} \partial_{\mu_{is_i}} (\bar{x}_{is_i} - \mu_{is_i}) - \partial_{\mu_{is_i}} \left[\mu_{is_i} (\bar{x}_{is_i} - \mu_{is_i}) \right] \right] d\boldsymbol{\mu}_i \tag{39}$$
$$= \gamma \int p(\boldsymbol{\mu}_i,t) \mu_{is_i} d\boldsymbol{\mu}_i - \int p(\boldsymbol{\mu}_i,t) \bar{x}_{is_i} d\boldsymbol{\mu}_i \tag{40}$$
$$= \frac{\bar{x}_{is_i} - \bar{\mu}_{is_i}}{\lambda+t+1} \tag{41}$$

$\square$

We repeat the mean probability $\bar{x}_{is_i}$, which has been given in Equation 8 in the main paper, as follows:

$$\bar{x}_{is_i} = \int \frac{\exp(\beta u_{is_i})}{\sum_{s'_i\in S_i} \exp(\beta u_{is'_i})} \prod_{j\in V_i} p(\boldsymbol{\mu}_j,t) \left(\prod_{j\in V_i} d\boldsymbol{\mu}_j \right) \tag{42}$$

where $u_{is_i} = \sum_{j\in V_i} \mathbf{e}_{s_i}^\top \mathbf{A}_{ij} \boldsymbol{\mu}_j$. Define $\bar{\boldsymbol{\mu}} := \{\bar{\boldsymbol{\mu}}_j\}_{j\in V_i}$ and

$$f_{s_i}(\{\boldsymbol{\mu}_j\}_{j\in V_i}) := \frac{\exp(\beta \sum_{j\in V_i} \mathbf{e}_{s_i}^\top \mathbf{A}_{ij} \boldsymbol{\mu}_j)}{\sum_{s'_i\in S_i} \exp(\beta \sum_{j\in V_i} \mathbf{e}_{s'_i}^\top \mathbf{A}_{ij} \boldsymbol{\mu}_j)}. \tag{43}$$

Applying the Taylor expansion to approximate this function at the mean belief $\bar{\boldsymbol{\mu}}$, we have

$$f_{s_i}(\{\boldsymbol{\mu}_j\}_{j\in V_i}) = f_{s_i}(\bar{\boldsymbol{\mu}}) + \nabla f_{s_i}(\bar{\boldsymbol{\mu}}) \cdot (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) + \frac{1}{2!} (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^\top \mathbf{H} f_{s_i}(\bar{\boldsymbol{\mu}}) (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) + \frac{1}{3!} O(||\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}||^3) \tag{44}$$

where $\mathbf{H}$ denotes the Hessian matrix. Hence, we can rewrite Equation 42 as

$$\bar{x}_{is_i} = \int f_{s_i}(\{\boldsymbol{\mu}_j\}_{j\in V_i}) \prod_{j\in V_i} p(\boldsymbol{\mu}_j, t) \left(\prod_{j\in V_i} d\boldsymbol{\mu}_j\right) \tag{45}$$

$$\begin{aligned} &\approx f_{s_i}(\bar{\boldsymbol{\mu}}) + \int \nabla f_{s_i}(\bar{\boldsymbol{\mu}}) \cdot \boldsymbol{\mu} \prod_{j\in V_i} p(\boldsymbol{\mu}_j, t) \left(\prod_{j\in V_i} d\boldsymbol{\mu}_j\right) - \nabla f_{s_i}(\bar{\boldsymbol{\mu}}) \cdot \bar{\boldsymbol{\mu}} \\ &+ \int \frac{1}{2}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^\top \mathbf{H} f_{s_i}(\bar{\boldsymbol{\mu}})(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) \prod_{j\in V_i} p(\boldsymbol{\mu}_j, t) \left(\prod_{j\in V_i} d\boldsymbol{\mu}_j\right) \\ &+ \int \frac{1}{3!} O(||\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}||)^3 \prod_{j\in V_i} p(\boldsymbol{\mu}_j, t) \left(\prod_{j\in V_i} d\boldsymbol{\mu}_j\right) \end{aligned} \tag{46}$$

Observe that in Equation 46, the second and the third term can be canceled out. Moreover, for any two neighbor populations $j, k \in V_i$, the beliefs $\boldsymbol{\mu}_j, \boldsymbol{\mu}_k$ about these two populations are updated separately and independently. Hence, the covariance of these beliefs are zero. We apply the moment closure approximation [4, 2] with the second order and obtain

$$\bar{x}_{is_i} \approx f_{s_i}(\bar{\boldsymbol{\mu}}) + \frac{1}{2} \sum_{j\in V_i} \sum_{s_j \in S_j} \frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial \mu_{js_j})^2} \mathrm{Var}(\mu_{js_j}). \tag{47}$$

Hence, substituting $\bar{x}_{is_i}$ in Lemma 2 with the above approximation, we have the mean belief dynamics

$$\frac{d\bar{\mu}_{is_i}}{dt} \approx \frac{f_{s_i}(\bar{\boldsymbol{\mu}}) - \bar{\mu}_{is_i}}{\lambda + t + 1} + \frac{\sum_{j\in V_i} \sum_{s_j\in S_j} \frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial \mu_{js_j})^2} \mathrm{Var}(\mu_{js_j})}{2(\lambda + t + 1)}. \tag{48}$$

Q.E.D.

5 Proof of Proposition 3

It follows from Equation 2 and Equation 3 of the main paper that the change in beliefs between two successive time steps is as follows.

$$\boldsymbol{\mu}_i(t+1) = \boldsymbol{\mu}_i(t) + \frac{\mathbf{x}_i(t) - \boldsymbol{\mu}_i(t)}{\lambda + t + 1}. \tag{49}$$

Suppose that the amount of time that passes between two successive time steps is $\delta \in (0, 1]$. We rewrite the above equation as

$$\boldsymbol{\mu}_i(t+\delta) = \boldsymbol{\mu}_i(t) + \delta \frac{\mathbf{x}_i(t) - \boldsymbol{\mu}_i(t)}{\lambda + t + 1}. \tag{50}$$

Move the term $\boldsymbol{\mu}_i(t)$ to the right hand side and divide both sides by δ,

$$\frac{\boldsymbol{\mu}_i(t+\delta) - \boldsymbol{\mu}_i(t)}{\delta} = \frac{\mathbf{x}_i(t) - \boldsymbol{\mu}_i(t)}{\lambda + t + 1}. \tag{51}$$

Assume that the amount of time δ between two successive time steps goes to zero. we have

$$\frac{d\boldsymbol{\mu}_i}{dt} = \lim_{\delta\to 0} \frac{\boldsymbol{\mu}_i(t+\delta) - \boldsymbol{\mu}_i(t)}{\delta} = \frac{\mathbf{x}_i(t) - \boldsymbol{\mu}_i(t)}{\lambda + t + 1}. \tag{52}$$

Q.E.D.

6 Proof of Proposition 4

It is straightforward to see that

$$\frac{d\boldsymbol{\mu}_i}{dt} = \frac{\mathbf{x}_i - \boldsymbol{\mu}_i}{\lambda + t + 1} = 0 \implies \mathbf{x}_i = \boldsymbol{\mu}_i. \tag{53}$$

Denote the equilibrium points of the system dynamics, which satisfies the above equation, by $(\mathbf{x}_i^*, \boldsymbol{\mu}_i^*)$ for each population i. By the logit choice rule, we have

$$x_{is_i}^* = \frac{\exp\left(\beta u_{is_i}\right)}{\sum_{s_i' \in S_i} \exp\left(\beta u_{is_i'}\right)} = \frac{\exp\left(\beta \sum_{j \in V_i} \mathbf{e}_{s_i}^\top \mathbf{A}_{ij} \boldsymbol{\mu}_j^*\right)}{\sum_{s_i' \in S_i} \exp\left(\beta \sum_{j \in V_i} \mathbf{e}_{s_i'}^\top \mathbf{A}_{ij} \boldsymbol{\mu}_j^*\right)}. \tag{54}$$

Leveraging that $\mathbf{x}_i^* = \boldsymbol{\mu}_i^*, \forall i \in V$ at equilibrium, we can replace $\boldsymbol{\mu}_j^*$ with $\mathbf{x}_j^*$. Q.E.D.

7 Proof of Theorem 2

Consider an agent i in a classic network game. The set of neighbors is V_i, the set of beliefs about the neighbors is $\{\boldsymbol{\mu}_j\}_{j \in V_i}$, and the choice distribution is $\mathbf{x}_i$. Given a classic network game, the expected payoff is given by $\mathbf{x}_i^\top \sum_{(i,j) \in E} A_{ij} \boldsymbol{\mu}_j$. Define a perturbed payoff function

$$\pi_i\left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i}\right) := \mathbf{x}_i^\top \sum_{j \in V_i} A_{ij} \boldsymbol{\mu}_j + v(\mathbf{x}_i) \tag{55}$$

where $v(\mathbf{x}_i) = -\frac{1}{\beta} \sum_{s_i \in S_i} x_{is_i} \ln(x_{is_i})$. Under this form of $v(\mathbf{x}_i)$, the maximization of π_i yields the choice distribution $\mathbf{x}_i$ from the logit choice function [1]. Based on this, we establish the following lemma.

Lemma 3. *For a choice distribution $\mathbf{x}_i$ of SFP in a network game,*

$$\partial_{\mathbf{x}_i} \pi_i\left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i}\right) = \mathbf{0} \quad \textit{and} \quad \sum_{j \in V_i} \left(A_{ij} \boldsymbol{\mu}_j\right)^\top = -\partial_{\mathbf{x}_i} v(\mathbf{x}_i). \tag{56}$$

Proof. This lemma immediately follows from the fact that the maximization of π_i will yield the choice distribution $\mathbf{x}_i$ from the logit choice function [1]. $\square$

The belief dynamics of an agent can be simplified after time-reparameterization.

Lemma 4. *Given $\tau = \ln \frac{\lambda+t+1}{\lambda+1}$, the belief dynamics of homogeneous systems (given in Equation 11 in the main paper) is equivalent to*

$$\frac{d\boldsymbol{\mu}_i}{d\tau} = \mathbf{x}_i - \boldsymbol{\mu}_i. \tag{57}$$

Proof. From $\tau = \ln \frac{\lambda+t+1}{\lambda+1}$, we have

$$t = (\lambda+1)(\exp(\tau) - 1). \tag{58}$$

By the chain rule, for each dimension s_i,

$$\begin{aligned}
\frac{d\mu_{is_i}}{d\tau} &= \frac{d\mu_{is_i}}{dt} \frac{dt}{d\tau} && (59)\\
&= \frac{x_{is_i} - \mu_{is_i}}{\lambda + t + 1} \frac{d\left((\lambda+1)(\exp(\tau) - 1)\right)}{d\tau} && (60)\\
&= \frac{x_{is_i} - \mu_{is_i}}{\lambda + (\lambda+1)(\exp(\tau) - 1) + 1} (\lambda+1) \exp(\tau) && (61)\\
&= x_{is_i} - \mu_{is_i}. && (62)
\end{aligned}$$

$\square$

Next, we define the Lyapunov function L as

$$L := \sum_{i \in V} \omega_i L_i \quad \text{s.t.} \quad L_i := \pi_i\left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i}\right) - \pi_i\left(\boldsymbol{\mu}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i}\right). \tag{63}$$

where $\{\omega_i\}_{i \in V}$ is the set of positive weights defined in the weighted zero-sum Γ. The function L is non-negative because for every $i \in V$, $\mathbf{x}_i$ maximizes the function π_i. When for every $i \in V$, $\mathbf{x}_i = \boldsymbol{\mu}_i$, the function L reaches the minimum value 0.

Rewrite L as

$$L = \sum_{i \in V} \left[\omega_i \pi_i \left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) - \omega_i \boldsymbol{\mu}_i^\top \sum_{j \in V_i} A_{ij} \boldsymbol{\mu}_j - \omega_i v(\boldsymbol{\mu}_i) \right]. \tag{64}$$

We observe that $\pi_i \left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right)$ is convex in $\boldsymbol{\mu}_j, j \in V_i$ by Danskin's theorem, and $-v(\boldsymbol{\mu}_i)$ is strictly convex in $\boldsymbol{\mu}_i$. Moreover, by the weighted zero-sum property given in Equation 2 in the main paper, we have

$$\sum_{i \in V} \left(\omega_i \boldsymbol{\mu}_i^\top \sum_{j \in V_i} A_{ij} \boldsymbol{\mu}_j \right) = 0 \tag{65}$$

since $\mu_i \in \Delta_i, \mu_j \in \Delta_j$ for every $i, j \in V$. Therefore, the function L is a strictly convex function and attains its minimum value 0 at a unique point $\mathbf{x}_i = \boldsymbol{\mu}_i, \forall i \in V$.

Consider the function L_i. Its time derivative is

$$\begin{aligned} \dot{L}_i = {} & \partial_{\mathbf{x}_i} \pi_i \left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \dot{\mathbf{x}}_i + \sum_{j \in V_i} \left[\partial_{\boldsymbol{\mu}_j} \pi_i \left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \dot{\boldsymbol{\mu}}_j \right] \\ & - \partial_{\boldsymbol{\mu}_i} \pi_i \left(\boldsymbol{\mu}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \dot{\boldsymbol{\mu}}_i - \sum_{j \in V_i} \left[\partial_{\boldsymbol{\mu}_j} \pi_i \left(\boldsymbol{\mu}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \dot{\boldsymbol{\mu}}_j \right]. \end{aligned} \tag{66}$$

Note that the partial derivative $\partial_{\mathbf{x}_i} \pi_i$ equals $\mathbf{0}$ by Lemma 3. Thus, we can rewrite this as

$$\dot{L}_i = \partial_{\boldsymbol{\mu}_i} \pi_i \left(\boldsymbol{\mu}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \dot{\boldsymbol{\mu}}_i + \sum_{j \in V_i} \left[\partial_{\boldsymbol{\mu}_j} \pi_i \left(\mathbf{x}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) - \partial_{\boldsymbol{\mu}_j} \pi_i \left(\boldsymbol{\mu}_i, \{\boldsymbol{\mu}_j\}_{j \in V_i} \right) \right] \dot{\boldsymbol{\mu}}_j \tag{67}$$

$$= - \left[\sum_{j \in V_i} \left(A_{ij} \boldsymbol{\mu}_j \right)^\top + \partial_{\boldsymbol{\mu}_i} v(\boldsymbol{\mu}_i) \right] (\mathbf{x}_i - \boldsymbol{\mu}_i) + \sum_{j \in V_i} \left(\mathbf{x}_i^\top A_{ij} - \boldsymbol{\mu}_i^\top A_{ij} \right) (\mathbf{x}_j - \boldsymbol{\mu}_j) \tag{68}$$

$$= \left[\partial_{\mathbf{x}_i} v(\mathbf{x}_i) - \partial_{\boldsymbol{\mu}_i} v(\boldsymbol{\mu}_i) \right] (\mathbf{x}_i - \boldsymbol{\mu}_i) + \sum_{j \in V_i} \left(\mathbf{x}_i^\top A_{ij} \mathbf{x}_j - \boldsymbol{\mu}_i^\top A_{ij} \mathbf{x}_j - \mathbf{x}_i^\top A_{ij} \boldsymbol{\mu}_j + \boldsymbol{\mu}_i^\top A_{ij} \boldsymbol{\mu}_j \right). \tag{69}$$

where from Equation 68 to 69, we apply Lemma 3 to substitute $\sum_{j \in V_i} \left(A_{ij} \boldsymbol{\mu}_j \right)^\top$ with $-\partial_{\mathbf{x}_i} v(\mathbf{x}_i)$. Hence, summing over all the populations, the time derivative of L is

$$\begin{aligned} \dot{L} = {} & \sum_{i \in V} \omega_i \left[\partial_{\mathbf{x}_i} v(\mathbf{x}_i) - \partial_{\boldsymbol{\mu}_i} v(\boldsymbol{\mu}_i) \right] (\mathbf{x}_i - \boldsymbol{\mu}_i) \\ & + \sum_{i \in V} \sum_{j \in V_i} \omega_i \left(\mathbf{x}_i^\top A_{ij} \mathbf{x}_j - \boldsymbol{\mu}_i^\top A_{ij} \mathbf{x}_j - \mathbf{x}_i^\top A_{ij} \boldsymbol{\mu}_j + \boldsymbol{\mu}_i^\top A_{ij} \boldsymbol{\mu}_j \right). \end{aligned} \tag{70}$$

The summation in the second line is equivalent to

$$\sum_{(i,j) \in E} (\omega_i \mathbf{x}_i^\top A_{ij} \mathbf{x}_j + \omega_j \mathbf{x}_j^\top A_{ji} \mathbf{x}_i) - (\omega_i \boldsymbol{\mu}_i^\top A_{ij} \mathbf{x}_j + \omega_j \mathbf{x}_j^\top A_{ji} \boldsymbol{\mu}_i) \tag{71}$$

$$- (\omega_i \mathbf{x}_i^\top A_{ij} \boldsymbol{\mu}_j + \omega_j \boldsymbol{\mu}_j^\top A_{ji} \mathbf{x}_i) + (\omega_i \boldsymbol{\mu}_i^\top A_{ij} \boldsymbol{\mu}_j + \omega_j \boldsymbol{\mu}_j^\top A_{ji} \boldsymbol{\mu}_i). \tag{72}$$

By the weighted zero-sum property given in Equation 2 in the main paper, this summation equals 0, yielding

$$\dot{L} = \sum_{i \in V} \omega_i \left[\partial_{\mathbf{x}_i} v(\mathbf{x}_i) - \partial_{\boldsymbol{\mu}_i} v(\boldsymbol{\mu}_i) \right] (\mathbf{x}_i - \boldsymbol{\mu}_i). \tag{73}$$

Note that the function v is strictly concave such that its second derivative is negative definite. By this property, $\dot{L} \leqslant 0$ with equality only if $\mathbf{x}_i = \boldsymbol{\mu}_i, \forall i \in V$, which corresponds to the QRE. Therefore, L is a strict Lyapunov function, and the global asymptotic stability of the QRE follows. Q.E.D.

8 Proof of Theorem 3

Consider a root agent j of a star structure. Its set of leaf (neighbor) agents is V_j, the set of beliefs about the leaf agents is $\{\boldsymbol{\mu}_i\}_{i\in V_j}$, and the choice distribution is $\mathbf{x}_j$. Given the game Γ, the expected payoff is $\mathbf{x}_j^\top \sum_{i\in V_j} A_{ji}\boldsymbol{\mu}_i$. Define a perturbed payoff function

$$\pi_j\left(\mathbf{x}_j, \{\boldsymbol{\mu}_i\}_{i\in V_j}\right) := \mathbf{x}_j^\top \sum_{i\in V_j} A_{ji}\boldsymbol{\mu}_i + v(\mathbf{x}_j) \tag{74}$$

where $v(\mathbf{x}_j) = -\frac{1}{\beta}\sum_{s_j\in S_j} x_{js_j}\ln(x_{js_j})$. Under this form of $v(\mathbf{x}_j)$, the maximization of π_j yields the choice distribution $\mathbf{x}_j$ from the logit choice function [1].

Consider a leaf agent i of the root agent j. It has only one neighbor, which is population j. Thus, given the game Γ, the expected payoff is $\mathbf{x}_i^\top A_{ij}\boldsymbol{\mu}_j$. Define a perturbed payoff function

$$\pi_i\left(\mathbf{x}_i, \boldsymbol{\mu}_j\right) := \mathbf{x}_i^\top A_{ij}\boldsymbol{\mu}_j + v(\mathbf{x}_i) \tag{75}$$

where $v(\mathbf{x}_i) = -\frac{1}{\beta}\sum_{s_i\in S_i} x_{is_i}\ln(x_{is_i})$. Similarly, the maximization of π_i yields the choice distribution $\mathbf{x}_i$ from the logit choice function [1]. Based on this, we establish the following lemma.

Lemma 5. *For choice distributions of SFP in a network game with a star structure,*

$$\partial_{\mathbf{x}_j}\pi_j\left(\mathbf{x}_j, \{\boldsymbol{\mu}_i\}_{i\in V_j}\right) = \mathbf{0} \quad \textit{and} \quad \sum_{i\in V_j}\left(A_{ji}\boldsymbol{\mu}_i\right)^\top = -\partial_{\mathbf{x}_j}v(\mathbf{x}_j) \qquad \textit{if } j \textit{ is a root agent,} \tag{76}$$

$$\partial_{\mathbf{x}_i}\pi_i\left(\mathbf{x}_i, \boldsymbol{\mu}_j\right) = \mathbf{0} \quad \textit{and} \quad \left(A_{ij}\boldsymbol{\mu}_j\right)^\top = -\partial_{\mathbf{x}_i}v(\mathbf{x}_i) \qquad \textit{if } i \textit{ is a leaf agent.} \tag{77}$$

Proof. This lemma immediately follows from the fact that the maximization of π_j and π_i , respectively, yield the choice distributions $\mathbf{x}_j$ and $\mathbf{x}_i$ from the logit choice function [1]. □

Let $\mathcal{R} \subset V$ be the set of all root agents. We define

$$L := \sum_{j\in\mathcal{R}} L_j \quad \text{s.t.} \quad L_j := \boldsymbol{\mu}_j^\top \sum_{i\in V_j} A_{ji}\boldsymbol{\mu}_i + v(\boldsymbol{\mu}_j) + \sum_{i\in V_j} v(\boldsymbol{\mu}_i). \tag{78}$$

Consider the function L_j. Its time derivative $\dot{L}_j$ is

$$\dot{L}_j = \left[\partial_{\boldsymbol{\mu}_j}(\boldsymbol{\mu}_j^\top \sum_{i\in V_j} A_{ji}\boldsymbol{\mu}_i)\dot{\boldsymbol{\mu}}_j + \sum_{i\in V_j}\partial_{\boldsymbol{\mu}_i}(\boldsymbol{\mu}_j^\top \sum_{i\in V_j} A_{ji}\boldsymbol{\mu}_i)\dot{\boldsymbol{\mu}}_i\right] + \partial_{\boldsymbol{\mu}_j}v(\boldsymbol{\mu}_j)\dot{\boldsymbol{\mu}}_j + \sum_{i\in V_j}\partial_{\boldsymbol{\mu}_i}v(\boldsymbol{\mu}_i)\dot{\boldsymbol{\mu}}_i \tag{79}$$

$$= \sum_{i\in V_j}(A_{ji}\boldsymbol{\mu}_i)^\top(\mathbf{x}_j - \boldsymbol{\mu}_j) + \left[\sum_{i\in V_j}\boldsymbol{\mu}_j^\top A_{ji}(\mathbf{x}_i - \boldsymbol{\mu}_i)\right] + \partial_{\boldsymbol{\mu}_j}v(\boldsymbol{\mu}_j)(\mathbf{x}_j - \boldsymbol{\mu}_j) + \sum_{i\in V_j}\partial_{\boldsymbol{\mu}_i}v(\boldsymbol{\mu}_i)(\mathbf{x}_i - \boldsymbol{\mu}_i). \tag{80}$$

Since we have $\left(A_{ij}\boldsymbol{\mu}_j\right)^\top = \boldsymbol{\mu}_j^\top A_{ij}^\top = \boldsymbol{\mu}_j^\top A_{ji}$, applying Lemma 5, we can substitute $\sum_{i\in V_j}(A_{ji}\boldsymbol{\mu}_i)^\top$ with $-\partial_{\mathbf{x}_j}v(\mathbf{x}_j)$, and $\boldsymbol{\mu}_j^\top A_{ji}$ with $-\partial_{\mathbf{x}_i}v(\mathbf{x}_i)$, yielding

$$\begin{aligned}\dot{L}_j &= -\partial_{\mathbf{x}_j}v(\mathbf{x}_j)(\mathbf{x}_j - \boldsymbol{\mu}_j) + \left[\sum_{i\in V_j}(-\partial_{\mathbf{x}_i}v(\mathbf{x}_i))(\mathbf{x}_i - \boldsymbol{\mu}_i)\right] + \partial_{\boldsymbol{\mu}_j}v(\boldsymbol{\mu}_j)(\mathbf{x}_j - \boldsymbol{\mu}_j) \\ &\quad + \sum_{i\in V_j}\partial_{\boldsymbol{\mu}_i}v(\boldsymbol{\mu}_i)(\mathbf{x}_i - \boldsymbol{\mu}_i) \qquad (81) \\ &= (\partial_{\boldsymbol{\mu}_j}v(\boldsymbol{\mu}_j) - \partial_{\mathbf{x}_j}v(\mathbf{x}_j))(\mathbf{x}_j - \boldsymbol{\mu}_j) + \sum_{i\in V_j}(\partial_{\boldsymbol{\mu}_i}v(\boldsymbol{\mu}_i) - \partial_{\mathbf{x}_i}v(\mathbf{x}_i))(\mathbf{x}_i - \boldsymbol{\mu}_i) \qquad (82)\end{aligned}$$

Note that the function v is strictly concave such that its second derivative is negative definite. By this property, $\dot{L}_j \geqslant 0$ with equality only if $\mathbf{x}_i = \boldsymbol{\mu}_i, \forall i \in V_j$ and $\mathbf{x}_j = \boldsymbol{\mu}_j$. Thus, the time derivative of the function L, i.e., $\dot{L} = \sum_{j\in\mathcal{R}}\dot{L}_j \geqslant 0$ with equality only if $\mathbf{x}_i = \boldsymbol{\mu}_i, \forall i \in V_j, \mathbf{x}_j = \boldsymbol{\mu}_j, \forall j \in \mathcal{R}$. Q.E.D.

9 Proof of Lemma 1

Definition 1. *A nonautonomous system of differential equations in R^n*

$$x' = f(t,x) \tag{83}$$

is said to be asymptotically autonomous with limit equation

$$y' = g(y), \tag{84}$$

if $f(t,x) \to g(x), t \to \infty$, where the convergence is uniform on each compact subset of R^n. Conventionally, the solution flow of Eq. 83 is called the asymptotically autonomous semiflow (denoted by ϕ) and the solution flow of Eq. 84 is called the limit semiflow (denoted by Θ).

We first time-reparameterize the mean belief dynamics of heterogeneous systems. Assume $\tau = \ln \frac{\lambda+t+1}{\lambda+1}$. By the chain rule and Equation 48, for each dimension s_i,

$$\frac{d\bar{\mu}_{is_i}}{d\tau} = \frac{d\bar{\mu}_{is_i}}{dt}\frac{dt}{d\tau} \tag{85}$$

$$= \left[\frac{f_{s_i}(\bar{\boldsymbol{\mu}}) - \bar{\mu}_{is_i}}{\lambda+t+1} + \frac{\sum_{j\in V_i}\sum_{s_j\in S_j}\frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial\mu_{js_j})^2}\mathrm{Var}(\mu_{js_j})}{2(\lambda+t+1)} \right] \frac{d\left((\lambda+1)(\exp(\tau)-1)\right)}{d\tau} \tag{86}$$

$$= \frac{f_{s_i}(\bar{\boldsymbol{\mu}}) - \bar{\mu}_{is_i} + \frac{1}{2}\sum_{j\in V_i}\sum_{s_j\in S_j}\frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial\mu_{js_j})^2}\left(\frac{\lambda+1}{\lambda+t+1}\right)^2\sigma^2(\mu_{js_j})}{\lambda+(\lambda+1)(\exp(\tau)-1)+1}(\lambda+1)\exp(\tau) \tag{87}$$

$$= f_{s_i}(\bar{\boldsymbol{\mu}}) - \bar{\mu}_{is_i} + \frac{1}{2}\sum_{j\in V_i}\sum_{s_j\in S_j}\frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial\mu_{js_j})^2}\sigma^2(\mu_{js_j})\exp(-2\tau). \tag{88}$$

Observe that $\exp(-2\tau)$ decays to zero exponentially fast and that both $\sigma^2(\mu_{js_j})$ and $\frac{\partial^2 f_{s_i}(\bar{\boldsymbol{\mu}})}{(\partial\mu_{js_j})^2}$ are bounded for every $\boldsymbol{\mu}$ in the simplex $\prod_{j\in V_i}\Delta_j$. Hence, Equation 88 converges locally and uniformly to the following equation:

$$\frac{d\bar{\mu}_{is_i}}{d\tau} = f_{s_i}(\bar{\boldsymbol{\mu}}) - \bar{\mu}_{is_i}. \tag{89}$$

Note that $x_{is_i} = f_{s_i}(\bar{\boldsymbol{\mu}})$ for a single representative agent, and thus the above equation is algebraically equivalent to the limit equation in Lemma 1 of the main paper. Q.E.D.

10 Numerical Methods, Source Code, and Computing Resource

Numerical Method for the PDE model. Only limited types of PDEs allow analytic solutions. Hence, we numerically solve the PDE using the finite difference method [5]. The theoretical predictions in Figure 1 of the main paper are generated using the finite difference method given a specific initial setting (the initial sum of weights is $\lambda = 10$, the temperature is $\beta = 10$, the initial belief distribution is specified in the caption of the figures).

Source Code and Computing Resource. We have attached the source code for reproducing our main experiments. The Matlab script *finitedifference.m* numerically solves our PDE model presented in Proposition 1 in the main paper. The Matlab script *regionofattraction.m* visualizes the region of attraction of different equilibria in stag hunt games, which are depicted in Figure 2. The Python scripts *simulation(staghunt).py* and *simulation(matchingpennies).py* correspond to the agent-based simulations in two-population stag hunt games and five-population asymmetric matching pennies games, respectively. We use a laptop (CPU: AMD Ryzen 7 5800H) to run all the experiments.

References

[1] Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[2] Colin S Gillespie. Moment-closure approximations for mass-action models. *IET systems biology*, 3(1):52–58, 2009.

[3] Stefanos Leonardos and Georgios Piliouras. Exploration-exploitation in multi-agent learning: Catastrophe theory meets game theory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11263–11271, 2021.

[4] Timothy I Matis and Ivan G Guardiola. Achieving moment closure through cumulant neglect. *The Mathematica Journal*, 12:12–2, 2010.

[5] Gordon Dennis Smith. *Numerical solution of partial differential equations: finite difference methods*. Clarendon Press, 1985.