SUPPLEMENT TO

Graphon games: A statistical framework for network games and intervention

Francesca Parise[1], Asuman Ozdaglar[2]

[1]*School of Electrical and Computer Engineering, Cornell University*

[2]*Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.*

Summary of Notation: We denote by $L^2([0,1])$ the space of square integrable functions defined on $[0,1]$ and by $L^2([0,1];\mathbb{R}^n)$ the space of square integrable vector valued functions defined on $[0,1]$. The norms in these spaces are denoted by $\|v\| := \sqrt{\sum_{h=1}^n [v]_h^2}$, $\|f\|_{L^2} := \sqrt{\int_0^1 f(x)^2 dx}$, $\|g\|_{L^2;\mathbb{R}^n} := \sqrt{\int_0^1 \|g(x)\|^2 dx}$, respectively. $[v]_h$ denotes the h-th component of the vector v. With the exception of $\mathbb{N}$ and $\mathbb{R}$ (that denote the sets of natural and real numbers, respectively), we use blackboard bold symbols (such as $\mathbb{O}$) to denote operators acting on $L^2([0,1])$ or on $L^2([0,1];\mathbb{R}^n)$. The induced operator norm are denoted by $|||\mathbb{O}||| := \sup_{\{f \mid \|f\|_{L^2}=1\}} \|\mathbb{O}f\|_{L^2}$ and $|||\mathbb{O}||| := \sup_{\{g \mid \|g\|_{L^2;\mathbb{R}^n}=1\}} \|\mathbb{O}g\|_{L^2;\mathbb{R}^n}$. We denote by $\lambda_{\max}(\mathbb{L})$ and by $r(\mathbb{L})$ the largest eigenvalue and the spectral radius of the linear integral operator $\mathbb{L}f := \int_0^1 L(x,y)f(y)dy$ with symmetric kernel $L(x,y) = L(y,x)$. We denote sets by using calligraphic symbols (such as $\mathcal{S}$) and the set of subsets of $\mathbb{R}^n$ by $2^{\mathbb{R}^n}$. The symbol 1_N denotes the vector of all ones in $\mathbb{R}^N$ and $1_{[0,1]}(x)$ the function constantly equal to one in $[0,1]$. $\mathbb{I}$ is the identity operator and I the identity matrix.

B. Incomplete information in sampled network games

In the main text we assumed that agents have perfect information about the network $A_s^{[N]}$.[21] In this appendix we generalize our analysis to sampled network games with *incomplete information*. As in the main text, we consider sampled network games with N agents whose types $\{t^i\}_{i=1}^N$ are drawn independently and uniformly at random from $[0,1]$ (recall that, e.g., in the community structure model of Example 3 an agent's type represents his community, while in the location model of Example 4 an agent's

[21] For simplicity we here focus on 0-1 adjacency matrices, similar results hold for the weighted case.

type is his location in the line segment $[0,1]$). Different from the main text, we here assume that agents do not have access to the exact structure of the sampled network $A_s^{[N]}$, but instead each agent i knows the stochastic network formation process (i.e., the graphon W in our framework) and his own type $t^i \in [0,1]$, which determines the probability $W(t^i, t^j)$ that he will connect to agent j of (random) type t^j. We next define a symmetric Bayesian Nash equilibrium for this incomplete information game and show that it is well approximated by the equilibrium of a graphon game with graphon W.

Note that the strategy $b(x)$ of each agent in an incomplete information sampled network game specifies the action that the agent will take as a function of his type x. Assuming that all other agents use the strategy b, the expected payoff of an agent i of type $t^i = x$ playing strategy $s(x) \in \mathtt{S}(x)$ is given by[22]

$$U_{\exp}(s(x) \mid b) = \mathbb{E}_{N,t^{-i},\text{links}}\left[U\left(s(x), \frac{1}{N-1}\sum_{j\neq i}[A_s^{[N]}]_{ij} b(t^j), \theta(x)\right)\right] \tag{B.1}$$

where U is as in (1) and $\mathbb{E}_{N,t^{-i},\text{links}}$ denotes the expectation with respect to the number of agents, their types (each agent knows its type $t^i = x$ but has no information about the other agents types $t^{-i} := \{t^j\}_{j\neq i}$, which are independent from t^i) and the link realizations (which are generated according to Bernoulli random variables with probability $\{W(t^i, t^j)\}_{j\neq i}$). We define a symmetric Bayesian Nash equilibrium as follows.

Definition 6. An *incomplete information sampled network game* $\mathcal{G}^{in}(\mathtt{S}, U, \theta, W)$, is a network game with a random number N of agents, whose types $\{t^i\}_{i=1}^N$ are sampled independently and uniformly at random from $[0,1]$, that interact according to a network $A_s^{[N]}$ sampled from the graphon W according to Definition 3.[23] Each agent i has information about the graphon W, his own type t^i, the strategy sets $\mathtt{S}$, the

[22] In this section we define the local aggregate by dividing by $N-1$ instead of N to account for the fact that agent i does not consider itself in the local aggregate. This allows us to obtain exact equivalence of the graphon game equilibrium and symmetric Bayesian Nash equilibrium in incomplete information sampled network games with linear quadratic payoffs. With the normalization $\frac{1}{N}$ instead of $\frac{1}{N-1}$ the equivalence would hold asymptotically in N.

[23] The distribution of N does not matter for our results, with the exception of Theorem 5 in which we assume that the support of such distribution is bounded from below by $N_{\min}$.

function θ and the payoff function U, while is uninformed about $A_s^{[N]}$ and the other agents types t^{-i}.

Definition 7. Consider a incomplete information sampled network game $\mathcal{G}^{in}(\mathtt{S}, U, \theta, W)$. A function b such that $b(x) \in \mathtt{S}(x)$ for all $x \in [0,1]$ is a *symmetric ε-Bayesian Nash equilibrium* if for all $x \in [0,1]$

$$U_{\exp}(b(x) \mid b) \geq U_{\exp}(\tilde{s} \mid b) - \varepsilon \text{ for all } \tilde{s} \in \mathtt{S}(x).$$

The function b is an exact symmetric Bayesian Nash equilibrium if the previous inequality holds for $\varepsilon = 0$.

Remark 7. Note that a strategy profile in both the graphon game and the incomplete information sampled network game is a function that maps $x \in [0,1]$ into a strategy $s(x) \in \mathtt{S}(x)$. In the graphon game this function specifies the action of a continuum of agents $x \in [0,1]$ interacting according to the graphon W, in the incomplete information sampled network game it specifies the action an agent with type x takes if he doesn't know the type of the other sampled agents and the realized links.

We start by focusing on linear quadratic games with payoff function as in (3).

Theorem 4. *Consider a linear quadratic game with payoff as in* (3) *and assume that the peer effect parameter α is the same for each agent while $\theta(x)$ is agent specific. A function $\bar{s}$ is a Nash equilibrium of the graphon game $\mathcal{G}(\mathtt{S}, U, \theta, W)$ if and only if it is a symmetric Bayesian Nash equilibrium for the incomplete information sampled network game $\mathcal{G}^{in}(\mathtt{S}, U, \theta, W)$.*

Proof. Let $\bar{z}(x) = \int_0^1 W(x,y)\bar{s}(y)dy$. By definition $\bar{s}$ is a graphon equilibrium if and only if for all $x \in [0,1]$, $\bar{s}(x) \in \mathtt{S}(x)$ and

$$U(\bar{s}(x), \bar{z}(x), \theta(x)) \geq U(s(x), \bar{z}(x), \theta(x)) \text{ for all } s(x) \in \mathtt{S}(x). \tag{B.2}$$

Note that for linear quadratic sampled network games with partial information, the expected payoff of an agent with type $t^i = x$ is

$$
\begin{aligned}
U_{\exp}(s(x) \mid \bar{s}) &= \mathbb{E}_{N,t^{-i},\text{links}}\left[-\frac{1}{2}(s(x))^2 + \left(\alpha\frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j) + \theta(x)\right)s(x)\right] \\
&= -\frac{1}{2}(s(x))^2 + \left(\alpha\mathbb{E}_{N,t^{-i},\text{links}}\left[\frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j)\right] + \theta(x)\right)s(x), \\
&= U(s(x), z_{\exp}(x), \theta(x))
\end{aligned}
$$

where we defined $z_{\exp}(x) := \mathbb{E}_{N,t^{-i},\text{links}}\left[\frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j)\right]$.[24] Hence $\bar{s}$ is a symmetric Bayesian Nash equilibrium if and only if for all $x \in [0,1]$, $\bar{s}(x) \in \mathtt{S}(x)$ and

$$
U(\bar{s}(x), z_{\exp}(x), \theta(x)) \geq U(s(x), z_{\exp}(x), \theta(x)) \text{ for all } s(x) \in \mathtt{S}(x). \tag{B.3}
$$

We conclude the proof by showing that $z_{\exp}(x) = \bar{z}(x)$ for all $x \in [0,1]$, proving that conditions (B.2) and (B.3) are equivalent. To this end, note that

$$
\begin{aligned}
z_{\exp}(x) &= \mathbb{E}_{N,t^{-i},\text{links}}\left[\frac{1}{N-1}\sum_{j\neq i} [A_s^{[N]}]_{ij}\bar{s}(t^j)\right] \\
&= \mathbb{E}_{\mathrm{N}}\mathbb{E}_{t^{-i}|\mathrm{N}}\mathbb{E}_{\text{links}|t^{-i},N}\left[\frac{1}{N-1}\sum_{j\neq i} [A_s^{[N]}]_{ij}\bar{s}(t^j)\right] \qquad \text{(B.4)} \\
&= \mathbb{E}_{\mathrm{N}}\mathbb{E}_{t^{-i}|\mathrm{N}}\left[\frac{1}{N-1}\sum_{j\neq i} W(x,t^j)\bar{s}(t^j)\right]
\end{aligned}
$$

and for any fixed N

$$
\begin{aligned}
&\mathbb{E}_{t^{-i}|\mathrm{N}}\left[\frac{1}{N-1}\sum_{j\neq i} W(x,t^j)\bar{s}(t^j)\right] = \frac{1}{N-1}\sum_{j\neq i}\mathbb{E}_{t^{-i}|\mathrm{N}}\left[W(x,t^j)\bar{s}(t^j)\right] \\
&= \frac{1}{N-1}\sum_{j\neq i}\mathbb{E}_{t^j}\left[W(x,t^j)\bar{s}(t^j)\right] = \frac{1}{N-1}\sum_{j\neq i}\int_0^1 W(x,y)\bar{s}(y)dy = \frac{1}{N-1}\sum_{j\neq i}\bar{z}(x) = \bar{z}(x)
\end{aligned}
$$

[24] Note that $\mathbb{E}_{N,t^{-i},\text{links}}\left[\frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j)\right]$ is a function of the type $t^i = x$ of agent i since a link between agent i and j forms (i.e., $[A_s^{[N]}]_{ij} = 1$) with Bernoulli probability $W(t^i, t^j) = W(x, t^j)$.

where we used the fact that the $\{t^j\}_{j=1}^N$ are independent and uniformly distributed in $[0,1]$. Hence

$$z_{\text{exp}}(x) = \mathbb{E}_{\text{N}}\mathbb{E}_{t^{-i}|\text{N}}\left[\frac{1}{N-1}\sum_{j\neq i} W(x,t^j)\bar{s}(t^j)\right] = \mathbb{E}_{\text{N}}\bar{z}(x) = \bar{z}(x). \tag{B.5}$$

Note that $z_{\text{exp}}(x)$ does not depend on the distribution of N. □

B.1. ***Generalization to Lipschitz payoff functions***

In the previous subsection we focused on games with linear quadratic payoff functions and we showed that $\bar{s}$ is a graphon equilibrium if and only if it is a symmetric Bayesian Nash equilibrium for an incomplete information sampled network game with any number of agents. We next consider a more general class of payoff functions, satisfying the following assumption.

Assumption 5. The payoff function $U(s,z,\theta)$ is Lipschitz continuous in z uniformly over s and θ, with constant L_U.

The expected payoff for an agent of type $t^i = x$ in this case is

$$\begin{aligned} U_{\text{exp}}(s(x) \mid b) &= \mathbb{E}_{N,t^{-i},\text{links}}\left[U\left(s(x), \frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij} b(t^j), \theta(x)\right)\right] \\ &= \mathbb{E}_{\zeta_b(x)}\left[U\left(s(x), \zeta_b(x), \theta(x)\right)\right], \end{aligned}$$

where $\zeta_b(x)$ is a random variable that describes the possible realizations of local aggregate perceived by an agent of type x over different network realizations when all agents play according to b. Note that, when the strategy b equals a graphon game equilibrium $\bar{s}$, $\mathbb{E}_{\zeta_{\bar{s}}(x)}[\zeta_{\bar{s}}(x)] = \bar{z}(x) = \int_0^1 W(x,y)\bar{s}(y)dy$ as shown in (B.4) and (B.5). For the payoff functions considered here however

$$\begin{aligned} U_{\text{exp}}(s(x) \mid \bar{s}) &= \mathbb{E}_{\zeta_{\bar{s}}(x)}\left[U\left(s(x), \zeta_{\bar{s}}(x), \theta(x)\right)\right] \\ &\neq U\left(s(x), \mathbb{E}_{\zeta_{\bar{s}}(x)}\left[\zeta_{\bar{s}}(x)\right], \theta(x)\right) = U(s(x), \bar{z}(x), \theta(x)), \end{aligned}$$

since the aggregate enters nonlinearly in the payoff function. Therefore it is not possible to use the argument of Theorem 4 to conclude that $\bar{s}$ is a symmetric Bayesian

Nash equilibrium. Nonetheless, we show in Lemma 14 (in Appendix D.5) that $\zeta_{\bar{s}}(x)$ concentrates around $\bar{z}(x)$ for large population sizes. Hence, for large populations, $U(s(x), \bar{z}(x), \theta(x))$ is indeed a good approximation of $U_{\exp}(s(x) \mid \bar{s})$. By exploiting this observation we show in the next theorem that, under the additional assumption that each agent has access to a lower bound ($N_{\min}$) on the population size in any realized sampled network game, the graphon equilibrium $\bar{s}$ is a symmetric ε-Bayesian Nash equilibrium with $\varepsilon \to 0$ as the lower bound on the population size $N_{\min} \to \infty$.

Theorem 5. *Consider an incomplete information sampled network game $\mathcal{G}^{in}(\mathtt{S}, U, \theta, W)$ where $\mathtt{S}(x) = \mathcal{S}$ for all $x \in [0,1]$. Suppose that all the agents know that the population size N is sampled from a distribution whose support is strictly lower bounded by $N_{\min}$ and suppose that Assumptions 1, 2, 3, 4 (with $\Omega = 0$) and 5 hold. Let $\bar{s}$ be the unique equilibrium of the corresponding graphon game $\mathcal{G}(\mathtt{S}, U, \theta, W)$. Then $\bar{s}$ is a symmetric ε-Bayesian Nash equilibrium with*

$$\varepsilon = \mathcal{O}\left(\sqrt{\frac{\log(N_{\min})}{N_{\min}}}\right).$$

Proof. It follows from the definition of graphon equilibrium that for all $x \in [0,1]$

$$U(\bar{s}(x), \bar{z}(x), \theta(x)) \geq U(s(x), \bar{z}(x), \theta(x)) \quad \forall s(x) \in \mathcal{S},$$

where $\bar{z}(x) = \int_0^1 W(x,y)\bar{s}(y)dy$. Consider an agent of type t^i. By the previous inequality specialized for $x = t^i$, it follows that for all $s(t^i) \in \mathcal{S}$

$$\begin{aligned}
&U_{\exp}(\bar{s}(t^i) \mid \bar{s}) = \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[U(\bar{s}(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i))] \\
&= \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[U(\bar{s}(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i)) - U(\bar{s}(t^i), \bar{z}(t^i), \theta(t^i))] + U(\bar{s}(t^i), \bar{z}(t^i), \theta(t^i)) \\
&\geq -L_U \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] + U(s(t^i), \bar{z}(t^i), \theta(t^i)) \\
&= -L_U \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] + \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[U(s(t^i), \bar{z}(t^i), \theta(t^i)) - U(s(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i))] \\
&\qquad + \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[U(s(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i))] \\
&\geq -2L_U \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] + \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[U(s(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i))] \\
&=: -\varepsilon + U_{\exp}(s(t^i) \mid \bar{s}).
\end{aligned}$$

The proof is concluded upon showing $\varepsilon := 2L_U \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] = \mathcal{O}\left(\sqrt{\frac{\log(N_{\min})}{N_{\min}}}\right)$. Since $\bar{z}(t^i) = \mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\zeta_{\bar{s}}(t^i)]$, we need to show that $\zeta_{\bar{s}}(t^i)$ concentrates around its mean when $N_{\min} \to \infty$. We show in Lemma 14 (given in Appendix D.5) that for any fixed population of size N and any fixed t^i with probability at least $1 - \frac{2n+1}{(N-1)^2}$ it holds $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \leq \varepsilon'$, with $\varepsilon' := \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$. It follows that

$$\begin{aligned}\mathbb{E}_{\zeta_{\bar{s}}(t^i)|N}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] &\leq \left(1 - \tfrac{2n+1}{(N-1)^2}\right)\varepsilon' + \tfrac{2n+1}{(N-1)^2} 2 s_{\max} \\ &\leq \varepsilon' + \tfrac{2(2n+1)s_{\max}}{(N-1)^2} = \mathcal{O}\left(\sqrt{\tfrac{\log(N-1)}{N-1}}\right),\end{aligned}$$

where we used that $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \leq 2s_{\max}$ for all realizations by Assumption 2. Consequently, if $N > N_{\min}$ $\mathbb{E}_{\zeta_{\bar{s}}(t^i)}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] = \mathbb{E}_N \mathbb{E}_{\zeta_{\bar{s}}(t^i)|N}[\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|] = \mathcal{O}\left(\sqrt{\frac{\log(N_{\min})}{N_{\min}}}\right)$. $\square$

C. Identification of unknown parameters

Consider a setting in which agents have payoffs as given in (2) which additionally depend on a common parameter $\bar{\eta}$ (for simplicity assume $n = 1$ so that agents strategies are scalars). This could be the case for example in a linear quadratic game with payoff

$$U_{\bar{\eta}}(s^i, z^i(s), \theta^i) = -\frac{1}{2}(s^i)^2 + (\theta^i + \bar{\eta} z^i(s))s^i, \tag{C.1}$$

where the common parameter $\bar{\eta} > 0$ represents the strength of peer effects. In this section we consider the problem of identifying the parameter $\bar{\eta}$ from the observation of a sampled equilibrium $\bar{s}^{[N]} \in \mathbb{R}^N$. We here assume that the realized network $A^{[N]}$ is unknown (so that results such as Bramoullé et al. (2009) cannot be applied). Instead we assume that the network $A^{[N]}$ is a realization from an underlying known graphon W (e.g., $A^{[N]}$ may be a realization from a stochastic block model).

Let us denote by Ξ the set of parameters η for which the corresponding graphon game $\mathcal{G}(\mathtt{S}, U_\eta, \theta, W)$ satisfies Assumptions 1 and 3. Moreover denote by $\bar{s}_\eta \in L^2([0,1])$ the unique equilibrium of $\mathcal{G}(\mathtt{S}, U_\eta, \theta, W)$. To identify the parameter $\bar{\eta}$ from an observation of $\bar{s}^{[N]}$, one could solve the following optimization problem

$$\hat{\eta} := \arg\min_{\eta \in \Xi} \|\bar{s}^{[N]} - \bar{s}_\eta\|_{L^2} \tag{C.2}$$

where $\bar{s}^{[N]}$ denotes the step function equilibrium. Intuitively, one can select as estimate the parameter η that minimizes the distance between the observed sampled equilibrium ($\bar{s}^{[N]}$) and the equilibrium ($\bar{s}_\eta$) of a graphon game with parameter η. We next show that if the parameter $\bar{\eta}$ is identifiable, as defined next, then $\|\hat{\eta} - \bar{\eta}\| \to 0$ as $N \to \infty$.

Definition 8. A parameter $\bar{\eta} \in \Xi$ is *identifiable* if there exists $L_{\bar{\eta}} > 0$ such that for any $\eta \in \Xi$ it holds

$$\|\bar{\eta} - \eta\| \leq L_{\bar{\eta}} \|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2}.$$

Intuitively, a parameter $\bar{\eta}$ is identifiable if equilibria that are close to $\bar{s}_{\bar{\eta}}$ are generated by parameters that are close to $\bar{\eta}$. Under this condition we can prove the following corollary of our main convergence theorem.

Corollary 1. Suppose that $\mathcal{G}(\mathtt{S}, U_{\bar{\eta}}, \theta, W)$ satisfies Assumptions 1, 2, 3 and 4 and that the parameter $\bar{\eta}$ is identifiable. Then for any $0 < \delta \leq e^{-1}$, with probability at least $1 - \frac{2\delta}{N}$, it holds

$$\|\bar{\eta} - \hat{\eta}\| = \mathcal{O}\left(\left(\frac{\log\left(\frac{N}{\delta}\right)}{N}\right)^{\frac{1}{4}}\right).$$

Proof. Set $\rho_s(N) := \|\bar{s}^{[N]} - \bar{s}_{\bar{\eta}}\|_{L^2}$. Since $\bar{\eta}$ is a feasible point of the optimization problem in (C.2) and $\hat{\eta}$ is the optimizer it must be

$$\|\bar{s}^{[N]} - \bar{s}_{\hat{\eta}}\|_{L^2} \leq \rho_s(N).$$

Combining these two inequalities yields

$$\|\bar{s}_{\hat{\eta}} - \bar{s}_{\bar{\eta}}\|_{L^2} \leq \|\bar{s}^{[N]} - \bar{s}_{\hat{\eta}}\|_{L^2} + \|\bar{s}^{[N]} - \bar{s}_{\bar{\eta}}\|_{L^2} \leq 2\rho_s(N).$$

The identifiability condition yields

$$\|\bar{\eta} - \hat{\eta}\| \leq L_{\bar{\eta}} \|\bar{s}_{\bar{\eta}} - \bar{s}_{\hat{\eta}}\|_{L^2} \leq 2L_{\bar{\eta}}\rho_s(N).$$

The conclusion follows by Theorem 2.

□

Assessing for which parameters and games the identifiability condition in Definition 8 holds is an interesting open problem. We here briefly comment on linear quadratic games with payoff as in (C.1). In this case, we recall from Example 1, that for any $\eta \in \Xi$, $\eta > 0$

$$\bar{s}_\eta = (\mathbb{I} - \eta \mathbb{W})^{-1}\theta \quad \Leftrightarrow \quad \bar{s}_\eta - \theta = \eta \bar{z}_\eta, \tag{C.3}$$

where $\bar{z}_\eta := \mathbb{W}\bar{s}_\eta$. It follows from $(\bar{\eta} - \eta)\bar{z}_{\bar{\eta}} = (\bar{\eta}\bar{z}_{\bar{\eta}} - \eta\bar{z}_\eta) - \eta(\bar{z}_{\bar{\eta}} - \bar{z}_\eta)$ that

$$\begin{aligned}|\bar{\eta} - \eta|\|\bar{z}_{\bar{\eta}}\|_{L^2} &\leq \|\bar{\eta}\bar{z}_{\bar{\eta}} - \eta\bar{z}_\eta\|_{L^2} + \eta\|\bar{z}_{\bar{\eta}} - \bar{z}_\eta\|_{L^2} = \|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2} + \eta\|\mathbb{W}(\bar{s}_{\bar{\eta}} - \bar{s}_\eta)\|_{L^2} \\ &\leq \|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2} + \eta\lambda_{\max}(\mathbb{W})\|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2} \leq 2\|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2},\end{aligned}$$

where we used that $\eta \in \Xi$ implies $\eta < \frac{1}{\lambda_{\max}(\mathbb{W})}$. Hence

$$|\bar{\eta} - \eta| \leq \frac{2}{\|\bar{z}_{\bar{\eta}}\|_{L^2}}\|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2} =: L_{\bar{\eta}}\|\bar{s}_{\bar{\eta}} - \bar{s}_\eta\|_{L^2},$$

proving that any $\bar{\eta} \in \Xi$ is identifiable in linear quadratic network games.

D. Auxiliary results

D.1. ***Properties of the game operator***

Lemma 5 (Properties of $\mathbb{W}_n$)**.** $\mathbb{W}_n$ is a linear, continuous, bounded and compact operator. The eigenvalues of $\mathbb{W}_n$ coincide (besides multiplicity) with those of $\mathbb{W}$ and are real. Finally, $|||\mathbb{W}_n||| = \lambda_{\max}(\mathbb{W})$.

Proof. This result is well-known for $n = 1$ since $\mathbb{W}$ is a self-adjoint Hilbert-Schmidt integral operator. The extension to $n > 1$ is immediate since $\mathbb{W}_n$ acts independently on each component. □

Lemma 6 (Properties of $\mathbb{B}_\theta$)**.** Under Assumption 1, the following holds:

1. *$\mathbb{B}_\theta$ is a Lipschitz continuous operator. That is, for any $f_1, f_2 \in L^2([0,1];\mathbb{R}^n)$ and $\theta_1, \theta_2 \in L^2([0,1];\mathbb{R}^m)$*

$$\|\mathbb{B}_{\theta_1}f_1 - \mathbb{B}_{\theta_2}f_2\|_{L^2;\mathbb{R}^n} \leq \frac{1}{\alpha_U}(\ell_U\|f_1 - f_2\|_{L^2;\mathbb{R}^n} + \ell_\theta\|\theta_1 - \theta_2\|_{L^2;\mathbb{R}^m});$$

2. *The codomain of $\mathbb{B}_\theta$ is $L^2([0,1];\mathbb{R}^n)$;*

3. *If Assumption 2 also holds, then the codomain of $\mathbb{B}_\theta$ is contained in*

$$L_{\mathcal{S}} := \{f \in L^2([0,1];\mathbb{R}^n) \mid \|f\|_{L^2;\mathbb{R}^n} \leq s_{\max}\}. \tag{D.1}$$

Proof. 1. Take any $f_1, f_2 \in L^2([0,1];\mathbb{R}^n)$ and $\theta_1, \theta_2 \in L^2([0,1];\mathbb{R}^m)$. For any $x \in [0,1]$

$$\begin{aligned}
&\|(\mathbb{B}_{\theta_1} f_1)(x) - (\mathbb{B}_{\theta_2} f_2)(x)\| = \\
&= \|\arg\max_{\tilde{s}\in \mathtt{S}(x)} U(\tilde{s}, f_1(x), \theta_1(x)) - \arg\max_{\tilde{s}\in \mathtt{S}(x)} U(\tilde{s}, f_2(x), \theta_2(x))\| \\
&\leq \frac{1}{\alpha_U}\| - \nabla_s U((\mathbb{B}_{\theta_2} f_2)(x), f_1(x), \theta_1(x)) + \nabla_s U((\mathbb{B}_{\theta_2} f_2)(x), f_2(x), \theta_2(x))\| \\
&\leq \frac{1}{\alpha_U}(\ell_U \|f_1(x) - f_2(x)\| + \ell_\theta \|\theta_1(x) - \theta_2(x)\|).
\end{aligned} \tag{D.2}$$

The first inequality in (D.2) can be proven by reformulating the optimization problem in (A.1) as the variational inequality $\mathrm{VI}(\mathtt{S}(x), -\nabla_s U(\cdot, z(x), \theta(x)))$. By Assumption 1, the operator $-\nabla_s U(\cdot, z, \theta)$ is strongly monotone with constant α_U for all $z \in \mathbb{R}^n, \theta \in \mathbb{R}^m$, (Scutari et al., 2010, Equation (12)) . The result then follows from a known bound on the distance of the solution of strongly monotone variational inequalities (Nagurney, 1993, Theorem 1.14). The second inequality in (D.2) comes from the assumption that $\nabla_s U(s, z, \theta)$ is uniformly Lipschitz in $[z, \theta]$ with constants ℓ_U, ℓ_θ for all $s \in \mathbb{R}^n$. Let us now compute $\|\mathbb{B}_{\theta_1} f_1 - \mathbb{B}_{\theta_2} f_2\|_{L^2;\mathbb{R}^n}$. For simplicity define $h(x) := \|(\mathbb{B}_{\theta_1} f_1)(x) - (\mathbb{B}_{\theta_2} f_2)(x)\|$, $h_f(x) := \frac{\ell_U}{\alpha_U}\|f_1(x) - f_2(x)\|$ and $h_\theta(x) := \frac{\ell_\theta}{\alpha_U}\|\theta_1(x) - \theta_2(x)\|$ for all $x \in [0,1]$. By (D.2), $0 \leq h(x) \leq h_f(x) + h_\theta(x)$ for all $x \in [0,1]$. Hence $\|h(x)\|_{L^2} \leq \|h_f(x) + h_\theta(x)\|_{L^2} \leq \|h_f(x)\|_{L^2} + \|h_\theta(x)\|_{L^2}$.

The conclusion follows from $\|h(x)\|_{L^2} = \|\mathbb{B}_{\theta_1} f_1 - \mathbb{B}_{\theta_2} f_2\|_{L^2;\mathbb{R}^n}$, $\|h_f(x)\|_{L^2} = \frac{\ell_U}{\alpha_U}\|f_1 - f_2\|_{L^2;\mathbb{R}^n}$ and $\|h_\theta(x)\|_{L^2} = \frac{\ell_\theta}{\alpha_U}\|\theta_1 - \theta_2\|_{L^2;\mathbb{R}^m}$.

2. We need to show that for any $z \in L^2([0,1];\mathbb{R}^n)$, $\|\mathbb{B}_\theta z\|_{L^2;\mathbb{R}^n} < \infty$. Consider the function $\hat{z}(x) := \hat{z}$ for all $x \in [0,1]$, where $\hat{z}$ is as in Assumption 1. Note that $\hat{z} \in L^2([0,1];\mathbb{R}^n)$ and

$$\|\mathbb{B}_\theta \hat{z}\|^2_{L^2;\mathbb{R}^n} = \int_0^1 \|(\mathbb{B}_\theta \hat{z})(x)\|^2 dx = \int_0^1 \|\arg\max_{\tilde{s}\in \mathtt{S}(x)} U(\tilde{s}, \hat{z}, \theta(x))\|^2 dx \leq M^2.$$

Consider now any $z \in L^2([0,1];\mathbb{R}^n)$. We have

$$\begin{aligned}\|\mathbb{B}_\theta z\|_{L^2;\mathbb{R}^n} &= \|\mathbb{B}_\theta z - \mathbb{B}_\theta \hat{z} + \mathbb{B}_\theta \hat{z}\|_{L^2;\mathbb{R}^n} \le \|\mathbb{B}_\theta z - \mathbb{B}_\theta \hat{z}\|_{L^2;\mathbb{R}^n} + \|\mathbb{B}_\theta \hat{z}\|_{L^2;\mathbb{R}^n} \\ &\le \left(\frac{\ell_U}{\alpha_U}\right) \|\hat{z} - z\|_{L^2;\mathbb{R}^n} + M \le \left(\frac{\ell_U}{\alpha_U}\right) (\|\hat{z}\|_{L^2;\mathbb{R}^n} + \|z\|_{L^2;\mathbb{R}^n}) + M < \infty,\end{aligned}$$

where the second inequality follows from statement 1).

3. Under Assumption 2 for any $x \in [0,1]$, $(\mathbb{B}_\theta z)(x) \in \mathtt{S}(x) \subseteq \mathcal{S}$ hence

$$\|\mathbb{B}_\theta z\|^2_{L^2;\mathbb{R}^n} = \int_0^1 \|(\mathbb{B}_\theta z)(x)\|^2 dx \le \int_0^1 s^2_{\max} dx = s^2_{\max}.$$

Consequently for any $z \in L^2([0,1];\mathbb{R}^n)$, $\mathbb{B}_\theta z \in L_{\mathcal{S}}$.

□

D.2. ***Statements in support of Section 5.2: Average instead of aggregate***

Lemma 7. If $\int_0^1 W(x,y)dy \ge d_{\min} > 0$ a.e. then $\mathbb{W}_d$ is a linear Hilbert-Schmidt integral operator and $\|\!|\mathbb{W}_d|\!\| \le \frac{\lambda_{\max}(\mathbb{W})}{d_{\min}}$.

Proof. Note that $\mathbb{W}_d$ is a linear integral operator with kernel $W_d(x,y) := \frac{W(x,y)}{\int_0^1 W(x,y)dy}$. Since

$$\int_0^1 \int_0^1 \left(\frac{W(x,y)}{\int_0^1 W(x,y)dy}\right)^2 dydx = \int_0^1 \frac{\int_0^1 W(x,y)^2 dy}{(\int_0^1 W(x,y)dy)^2} dx \le \int_0^1 \frac{\int_0^1 1dy}{(d_{\min})^2} dx = \left(\frac{1}{d_{\min}}\right)^2 < \infty,$$

$\mathbb{W}_d$ is a Hilbert-Schmidt integral operator. Moreover, by definition

$$\begin{aligned}\|\!|\mathbb{W}_d|\!\|^2 &= \sup_{f\in L^2([0,1])|\|f\|_{L^2}\le 1} \|\mathbb{W}_d f\|^2_{L^2} = \sup_{f\in L^2([0,1])|\|f\|_{L^2}\le 1} \int_0^1 (\mathbb{W}_d f)^2(x)dx \\ &= \sup_{f\in L^2([0,1])|\|f\|_{L^2}\le 1} \int_0^1 \left(\frac{\int_0^1 W(x,y)f(y)dy}{\int_0^1 W(x,y)dy}\right)^2 dx \\ &\le \sup_{f\in L^2([0,1])|\|f\|_{L^2}\le 1} \int_0^1 \left(\frac{\int_0^1 W(x,y)f(y)dy}{d_{\min}}\right)^2 dx \\ &= \frac{1}{(d_{\min})^2} \sup_{f\in L^2([0,1])|\|f\|_{L^2}\le 1} \|\mathbb{W} f\|^2_{L^2} = \left(\frac{\|\!|\mathbb{W}|\!\|}{d_{\min}}\right)^2 = \left(\frac{\lambda_{\max}(\mathbb{W})}{d_{\min}}\right)^2.\end{aligned}$$

□

Lemma 8. Consider a Lipschitz continuous graphon W and suppose that $\int_0^1 W(x,y)dy \ge d_{\min} > 0$ a.e. Let $\mathbb{W}_d$ be the normalized graphon operator as introduced in Section 5.2.

Morever, let $\mathbb{W}_d^{[N]}$ be the normalized graphon operators corresponding to the normalized version of the matrices $A_{w/s}^{[N]}$, as defined in Section 3.1. Fix any sequence $\{\delta_N\}_{N=1}^{\infty}$ such that $\delta_N \leq e^{-1}$ and $\frac{\log(N/\delta_N)}{N} \to 0$. Then, for N large enough, with probability at least $1-4\delta_N$,

$$\left|\left|\left|\mathbb{W}_d^{[N]} - \mathbb{W}_d\right|\right|\right| = \mathcal{O}\left(\sqrt{\frac{\log(N/\delta_N)}{N}}\right).$$

Proof. Note that for N large enough the condition $\delta_N \in (Ne^{-N/5}, e^{-1})$ is satisfied under the assumptions of this lemma. Hence Lemma 11 in Appendix D.5 applies. We distinguish the proof for simple and weighted sampled networks. To this end, let $W_{wd/sd}^{[N]}$ be the normalized graphon corresponding to the matrices $A_{w/s}^{[N]}$ and $\mathbb{W}_{wd/sd}^{[N]}$ be the corresponding normalized graphon operator.

1. For weighted networks, define

$$d_w^{[N]}(x) := \int_0^1 W_w^{[N]}(x,y)dy \quad \text{and} \quad d(x) := \int_0^1 W(x,y)dy.$$

Then with probability $1-\delta_N$ if $x \in \mathcal{U}_i^{[N]}$

$$\begin{aligned} &|d_w^{[N]}(x) - d(x)| \leq \int_0^1 |W_w^{[N]}(x,y) - W(x,y)|dy \\ &= \sum_j \int_{\mathcal{U}_j^{[N]}} |W(t^i,t^j) - W(x,y)|dy \leq \sum_j \int_{\mathcal{U}_j^{[N]}} 2Ld_N dy = 2Ld_N =: \epsilon_N, \end{aligned}$$

where we used the Lipschitz property and (D.5) from Lemma 11 in the last inequality. Hence $d_w^{[N]}(x) \geq d(x) - \epsilon_N \geq d_{\min} - \epsilon_N$. Similarly for $x \in \mathcal{U}_i^{[N]}$ and $y \in \mathcal{U}_j^{[N]}$ we obtain $|W_w^{[N]}(x,y) - W(x,y)| \leq \epsilon_N$. Let $D(x,y) := \frac{W_w^{[N]}(x,y)}{d_w^{[N]}(x)} - \frac{W(x,y)}{d(x)}$. Then

$$\begin{aligned} |D(x,y)| &= \left|\frac{W_w^{[N]}(x,y)}{d_w^{[N]}(x)} - \frac{W(x,y)}{d(x)}\right| = \frac{|W_w^{[N]}(x,y)d(x) - W(x,y)d_w^{[N]}(x)|}{d_w^{[N]}(x)d(x)} \\ &\leq \frac{|W_w^{[N]}(x,y)d(x) - W(x,y)d(x)| + |W(x,y)d(x) - W(x,y)d_w^{[N]}(x)|}{d_{\min}(d_{\min}-\epsilon_N)} \\ &\leq \frac{|W_w^{[N]}(x,y) - W(x,y)| + |d(x) - d_w^{[N]}(x)|}{d_{\min}(d_{\min}-\epsilon_N)} = \frac{2\epsilon_N}{d_{\min}(d_{\min}-\epsilon_N)} =: \gamma_N \to 0. \end{aligned}$$

Consider any $f \in L^2([0,1])$ such that $\|f\|_{L^2} = 1$. Using the inequalities above

$$\|\mathbb{W}_{wd}^{[N]} f - \mathbb{W}_d f\|_{L^2}^2 = \int_0^1 (\mathbb{W}_{wd}^{[N]} f - \mathbb{W}_d f)(x)^2 \mathrm{d}x = \int_0^1 \left(\int_0^1 D(x,y) f(y) \mathrm{d}y\right)^2 \mathrm{d}x$$
$$\leq \int_0^1 \left(\int_0^1 D(x,y)^2 \mathrm{d}y\right) \left(\int_0^1 f(y)^2 \mathrm{d}y\right) \mathrm{d}x = \int_0^1 \int_0^1 D(x,y)^2 \mathrm{d}y \mathrm{d}x \leq \gamma_N^2.$$

Hence with probability $1 - \delta_N$, (D.5) holds and

$$\left|\left|\left|\mathbb{W}_{wd}^{[N]} - \mathbb{W}_d\right|\right|\right| = \sup_{f \in L^2([0,1]) s.t. \|f\|_{L_2} = 1} \|\mathbb{W}_{wd}^{[N]} f - \mathbb{W}_d f\|_{L^2} \leq \gamma_N = \mathcal{O}(d_N).$$

2. For simple networks, note that $\left|\left|\left|\mathbb{W}_{sd}^{[N]} - \mathbb{W}_{wd}^{[N]}\right|\right|\right| \leq \|A_{sd}^{[N]} - A_{wd}^{[N]}\|$, where $A_{sd/wd}^{[N]}$ are the degree normalized versions of $A_{s/w}^{[N]}$. Recall that $A_w^{[N]} = \mathbb{E}[A_s^{[N]}]$, hence we can bound the term on the right hand side by employing matrix concentration inequalities. Define $d_{s/w}^i = \sum_{j=1}^N [A_{s/w}^{[N]}]_{ij}$.

   - By definition and by the previous point

   $$d_w^i = N d_w^{[N]}(t^i) \geq N(d_{\min} - \epsilon_N)$$

   $$\|A_w^{[N]}\| = |||\mathbb{W}_w^{[N]}||| N \leq N$$

   - By Hoeffding inequality for any fixed i and $t > 0$ $\Pr[|d_s^i - d_w^i| > t] < 2\exp\left(-\frac{2t^2}{N}\right)$. Setting $t = \sqrt{\frac{N}{2}\log\left(\frac{2N}{\delta_N}\right)}$ yields

   $$\Pr\left[|d_s^i - d_w^i| > \sqrt{\frac{N}{2}\log\left(\frac{2N}{\delta_N}\right)}\right] < 2\frac{\delta_N}{2N} = \frac{\delta_N}{N}$$

   and by the union bound with probability at least $1 - \delta_N$

   $$|d_s^i - d_w^i| \leq \sqrt{\frac{N}{2}\log\left(\frac{2N}{\delta_N}\right)} = t \quad \text{for all} \quad i \in \{1, \ldots, N\}.$$

   Let $D_{s/w} := \mathrm{diag}([d_{s/w}^i]_{i=1}^N)$. With the same probability

   $$\|D_s^{-1}\| = \max_i \frac{1}{d_s^i} \leq \max_i \frac{1}{d_w^i - t} \leq \frac{1}{(d_{\min} - \epsilon_N)N - t}$$

and

$$\|D_s^{-1} - D_w^{-1}\| = \max_i \left| \frac{1}{d_s^i} - \frac{1}{d_w^i} \right|$$
$$= \max_i \frac{|d_s^i - d_w^i|}{d_w^i d_s^i} \leq \frac{t}{(d_{\min} - \epsilon_N)N} \frac{1}{(d_{\min} - \epsilon_N)N - t}.$$

- The maximum expected degree $C_N^d := \max_i (\sum_{j=1}^N [A_w^{[N]}]_{ij})$ grows as order N. Hence for N large enough, it is greater than $\frac{4}{9} \log(\frac{2N}{\delta_N})$ since $\frac{\log(N/\delta_N)}{N} \to 0$ by assumption. Consequently, all the conditions of (Chung and Radcliffe, 2011, Theorem 1) are met and with probability $1 - \delta_N$

$$\|A_s^{[N]} - A_w^{[N]}\| \leq \sqrt{4C_N^d \log(2N/\delta_N)} \leq \sqrt{4N \log(2N/\delta_N)},$$

  where we used that $C_N^d \leq N$ since each element in $A_w^{[N]}$ belongs to $[0, 1]$.

- Combining the previous results yields that with probability $1 - 3\delta_N$

$$\begin{aligned}
\|A_{sd}^{[N]} - A_{wd}^{[N]}\| &= \|D_s^{-1} A_s^{[N]} - D_w^{-1} A_w^{[N]}\| \\
&\leq \|D_s^{-1} A_s^{[N]} - D_s^{-1} A_w^{[N]}\| + \|D_s^{-1} A_w^{[N]} - D_w^{-1} A_w^{[N]}\| \\
&\leq \|D_s^{-1}\| \|A_s^{[N]} - A_w^{[N]}\| + \|D_s^{-1} - D_w^{-1}\| \|A_w^{[N]}\| \\
&\leq \frac{\sqrt{4N \log(2N/\delta_N)}}{(d_{\min} - \epsilon_N)N - t} + \frac{t}{(d_{\min} - \epsilon_N)N} \cdot \frac{N}{(d_{\min} - \epsilon_N)N - t} \\
&= \frac{\sqrt{8} t/N}{(d_{\min} - \epsilon_N) - t/N} + \frac{t/N}{(d_{\min} - \epsilon_N)} \cdot \frac{1}{(d_{\min} - \epsilon_N) - t/N}
\end{aligned}$$

  Since $t/N = \sqrt{\frac{\log(2N/\delta_N)}{2N}} \to 0$, we obtain $\left\|\left\| \mathbb{W}_{sd}^{[N]} - \mathbb{W}_{wd}^{[N]} \right\|\right\| \leq \|A_{sd}^{[N]} - A_{wd}^{[N]}\| = \mathcal{O}(t/N) = \mathcal{O}\left(\sqrt{\frac{\log(N/\delta_N)}{N}}\right)$.

Using the fact that $\left\|\left\| \mathbb{W}_s^{[N]} - \mathbb{W} \right\|\right\| \leq \left\|\left\| \mathbb{W}_s^{[N]} - \mathbb{W}_w^{[N]} \right\|\right\| + \left\|\left\| \mathbb{W}_w^{[N]} - \mathbb{W} \right\|\right\|$ and the first statement concludes the proof.

□

D.3. ***Statements in support of Section 5.3: Directed networks***

Lemma 9. Consider a matrix $A_w^{[N]} \in [0,1]^{N\times N}$ with $\|A_w^{[N]}\|_\infty$ of order N and a random matrix $A_s^{[N]} \in \{0,1\}^{N\times N}$ such that

$$[A_s^{[N]}]_{ij} = Ber([A_w^{[N]}]_{ij}).$$

With probability $1-\delta_N$ for N large enough

$$\frac{1}{N}\|A_s^{[N]} - A_w^{[N]}\| \leq \sqrt{4\frac{\log(4N/\delta_N)}{N}}.$$

Proof. Construct the symmetric matrix

$$A_{s/w}^{[2N]} = \begin{bmatrix} 0 & A_{s/w}^{[N]} \\ (A_{s/w}^{[N]})^T & 0 \end{bmatrix} \in \mathbb{R}^{2N\times 2N}$$

and note that

1. $\|A_s^{[N]} - A_w^{[N]}\| = \|A_s^{[2N]} - A_w^{[2N]}\|$;
2. $\mathbb{E}[A_s^{[N]}] = A_w^{[N]}$ implies $\mathbb{E}[A_s^{[2N]}] = A_w^{[2N]}$;
3. the maximum degree Δ_A of $A_w^{[2N]}$ is of order N and is therefore greater than $\frac{4}{9}\log(4N/\delta_N)$ for N large enough.

Then by (Chung and Radcliffe, 2011, Theorem 1) with probability $1-\delta_N$ for N large enough

$$\frac{1}{N}\|A_s^{[N]} - A_w^{[N]}\| = \frac{1}{N}\|A_s^{[2N]} - A_w^{[2N]}\| \leq \frac{1}{N}\sqrt{4\Delta_A\log(4N/\delta_N)} \leq \sqrt{4\frac{\log(4N/\delta_N)}{N}}.$$

□

D.4. ***Section 7: Omitted details and proofs***

To recover the stochastic network formation model and the network effect parameter from aggregated relation data, the central planner can estimate:

1. The exact proportion of agents in each community (from the census data) as

$$\pi_h := \frac{N_h}{N} := \frac{\text{number of agents in community } h}{\text{total number of agents in the census}}.$$

Let Π be a diagonal matrix with π_h in position (h, h).

2. The maximum likelihood estimator of the interaction probability of agents of community h and h' (from the subset of agents interviewed in the aggregated survey) as

$$\hat{q}^{ARD}_{h,h'} := \frac{S_{h,h'} + S_{h',h}}{S_h N_{h'} + S_{h'} N_h},$$

where S_h is the total number of agents surveyed from community h and $S_{h,h'}$ is the total number of neighbors that they reported having in community h'. The superscript ARD denotes the use of aggregated data. Let $\hat{Q}^{ARD}_\kappa := [\hat{q}^{ARD}_{h,h'}]$ be the estimated interaction matrix (see Example 4 in Section 6.2) and $\hat{E}^{ARD}_\kappa := \hat{Q}^{ARD}_\kappa \Pi$.[25]

3. The average strategy of agents in community h before the intervention as

$$\hat{s}^{ARD}_h := \frac{\text{sum of effort of agents surveyed from community } h}{N_h}.$$

For $N \to \infty$, $\hat{s}^{ARD}_h$ converges almost surely to the strategy $\bar{s}^{com}_h$ played by agents of community h in the graphon game as shown in the following Corollary 2.

4. The parameter α_κ by

$$\hat{\alpha}^{ARD}_\kappa = (\hat{X}^\top \hat{X})^{-1} \hat{X}^\top \hat{Y},$$

with $\hat{X} := \hat{E}^{ARD}_\kappa \hat{s}^{ARD}$ and $\hat{Y} := \hat{s}^{ARD} - \theta^{com}$, where θ^{com} is the vector of marginal return per community (which can be recovered exactly from the census data). Corollary 3 below shows that $\hat{\alpha}^{ARD}_\kappa \to \alpha_\kappa$ almost surely for $N \to \infty$.

Corollary 2. For all $k = 1, \dots K$, $\hat{s}^{ARD}_k \to \bar{s}^{com}_k$ almost surely as $N \to \infty$.

Proof. Consider for simplicity the case with just one community and suppose that aggregated relational data is collected from all agents. In this case, in the graphon game each agent has the same equilibrium strategy, that is, $\bar{s}(x) = \bar{s}^{com}$ for all

[25]Technically, since we assume sublinear network growth these are the matrices Q and E as described in Example 4 *multiplied by* κ_N, this is not a problem because we can only estimate α *divided by* κ_N, hence the (unknown) κ_N factor cancels out, that is $\alpha_\kappa E_\kappa = \alpha E$.

$x \in [0,1]$ hence

$$\|\bar{s}^{[N]} - \bar{s}\|_{L^2}^2 = \int_0^1 (\bar{s}^{[N]}(x) - \bar{s}(x))^2 dx = \sum_{i=1}^N \int_{\mathcal{U}_i^{[N]}} (\bar{s}_i^{[N]} - \bar{s}^{com})^2 dx$$
$$= \frac{1}{N} \sum_{i=1}^N (\bar{s}_i^{[N]} - \bar{s}^{com})^2 = \frac{1}{N} \|\bar{s}^{[N]} - \bar{s}^{com} \mathbb{1}_N\|_2^2.$$

This yields

$$|\hat{s}^{ARD} - \bar{s}^{com}| = \left| \left(\frac{1}{N} \sum_{i=1}^N \bar{s}_i^{[N]} \right) - \bar{s}^{com} \right| = \left| \frac{1}{N} \sum_{i=1}^N (\bar{s}_i^{[N]} - \bar{s}^{com}) \right|$$
$$\leq \frac{1}{N} \sum_{i=1}^N |\bar{s}_i^{[N]} - \bar{s}^{com}| = \frac{1}{N} \|\bar{s}^{[N]} - \bar{s}^{com} \mathbb{1}_N\|_1$$
$$\leq \frac{\sqrt{N}}{N} \|\bar{s}^{[N]} - \bar{s}^{com} \mathbb{1}_N\|_2 = \frac{\sqrt{N}}{N} \sqrt{N} \|\bar{s}^{[N]} - \bar{s}\|_{L^2} = \|\bar{s}^{[N]} - \bar{s}\|_{L^2}.$$

Since by Theorem 2 $\|\bar{s}^{[N]} - \bar{s}\|_{L^2} \to 0$ almost surely,[26] we finally obtain that $|\hat{s}^{ARD} - \bar{s}^{com}| \to 0$ almost surely. A similar proof shows that in the case of K communities $|\hat{s}_k^{ARD} - \bar{s}_k^{com}| \to 0$ almost surely for all $k = 1, \dots K$. □

Corollary 3. $\hat{\alpha}_\kappa^{ARD} \to \alpha_\kappa$ almost surely as $N \to \infty$.

Proof. Recall from Corollary 2 that $\hat{s}^{ARD} \to \bar{s}^{com}$ and that $\hat{E}_\kappa^{ARD} \to E_\kappa$. Hence $\hat{X} \to \bar{X} := E_\kappa \bar{s}^{com}$, $\hat{Y} \to \bar{Y} := \bar{s}^{com} - \theta^{com}$ and

$$\hat{\alpha}_\kappa^{ARD} \to (\bar{X}^\top \bar{X})^{-1} \bar{X}^\top \bar{Y}. \tag{D.3}$$

Moreover by (15), $\bar{s}^{com} = (I - \alpha E)^{-1} \theta^{com} = (I - \alpha_\kappa E_\kappa)^{-1} \theta^{com}$ or equivalently $\bar{s}^{com} - \theta^{com} = \alpha_\kappa E_\kappa \bar{s}^{com}$ implying $\bar{Y} = \alpha_\kappa \bar{X}$. Substituting in (D.3) yields $\hat{\alpha}_\kappa^{ARD} \to (\bar{X}^\top \bar{X})^{-1} \bar{X}^\top \bar{Y} = \alpha_\kappa (\bar{X}^\top \bar{X})^{-1} \bar{X}^\top \bar{X} = \alpha_\kappa$, as desired. □

D.5. ***Auxiliary results***

We report here some auxiliary lemmas. Specifically,

- Lemma 10, 11 and 12 are immediate extensions of results in Avella-Medina et al. (2018);

[26]Theorem 2 requires Assumption 2 which is not met when $\mathcal{S} = \mathbb{R}_{\geq 0}$. Assumption 2 is only used within Theorem 2 to bound $\|\bar{s}_s^{[N]}\|$. We proved in Lemma 3 that for linear quadratic games, $\|\bar{s}_s^{[N]}\|$ can be bounded, with high probability, even without Assumption 2. Hence the conclusion of Theorem 2 holds.

- Lemma 13 derives sufficient conditions for the equilibrium of a graphon game to be Lipschitz continuous;

- Lemma 14 provides a concentration result for the local aggregate in incomplete information sampled network games;

- Lemma 15 proves that the graphon equilibrium is an ϵ-Nash equilibrium under additional regularity assumptions.

Lemma 10 (Avella-Medina et al. (2018))**.** Consider a stochastic block model graphon $\mathbb{W}_{\text{SBM}}$ which is piecewise constant over the partition $\{\mathcal{C}_k\}_{k=1}^K$. If (λ, ψ) is an eigenpair of $\mathbb{W}_{\text{SBM}}$, then there exists $v \in \mathbb{R}^K$ such that (λ, v) is an eigenpair of the matrix $E \in \mathbb{R}^{K\times K}$ defined in (11) and

$$\psi(x) = \gamma v_k, \text{ for all } x \in \mathcal{C}_k \tag{D.4}$$

where $\gamma > 0$ is a normalization parameter. Conversely, if (λ, v) is an eigenpair of the matrix $E \in \mathbb{R}^{K\times K}$ then (λ, ψ) is an eigenpair of $\mathbb{W}_{\text{SBM}}$ with ψ constructed from v as in (D.4).

Lemma 11 (Avella-Medina et al. (2018))**.** Let $\{t^i\}_{i=1}^N$ be the ordered statistics of N random samples from $\mathcal{U}[0,1]$. For any $\delta_N \in (Ne^{-N/5}, e^{-1})$ and N large, with probability at least $1-\delta_N$ it holds

$$|t^i - x| \le d_N \text{ for any } i \in \{1,\dots,N\} \text{ and any } x \in \mathcal{U}_i^{[N]} = [\tfrac{i-1}{N}, \tfrac{i}{N}), \tag{D.5}$$

where $d_N := \frac{1}{N} + \sqrt{\frac{8\log(N/\delta_N)}{N}} \to 0$.

Lemma 12 (Avella-Medina et al. (2018))**.** Consider a graphon W satisfying Assumption 4. Let $W_{w/s}^{[N]}$ be the step function graphons corresponding to the matrices $A_w^{[N]}$ and $\frac{A_s^{[N]}}{\kappa_N}$, as defined in Sections 3.1 and 5.1. Let $\theta^{[N]}$ be the step-function corresponding to $[\theta(t^i)]_{i=1}^N$. Fix any sequence $\{\delta_N, \kappa_N\}_{N=1}^\infty$ such that $\delta_N \le e^{-1}$ and $\frac{\log(N/\delta_N)}{N\kappa_N} \to 0$. Then, for N large enough, with probability at least $1-2\delta_N$ (D.5) holds and

1. $\|\theta^{[N]} - \theta\|_{L^2;\mathbb{R}^m} \le \rho_\theta(N) := \sqrt{(Ld_N)^2 + 8\Omega d_N \theta_{\max}^2} = \mathcal{O}\left(\left(\frac{\log(N/\delta_N)}{N}\right)^{1/4}\right)$

2. $|\lambda_{\max}(\mathbb{W}_w^{[N]}) - \lambda_{\max}(\mathbb{W})| \le \left|\left|\left|\mathbb{W}_w^{[N]} - \mathbb{W}\right|\right|\right| \le \tilde{\rho}(N) = \mathcal{O}\left(\left(\frac{\log(N/\delta_N)}{N}\right)^{1/4}\right)$

3. $|\lambda_{\max}(\mathbb{W}_s^{[N]}) - \lambda_{\max}(\mathbb{W})| \leq \left|\left|\left| \mathbb{W}_s^{[N]} - \mathbb{W} \right|\right|\right| \leq \rho_W(N) = \mathcal{O}\left(\left(\frac{\log(N/\delta_N)}{N}\right)^{1/4} + \left(\frac{\log(N/\delta_N)}{N\kappa_N}\right)^{\frac{1}{2}}\right)$

with $\tilde{\rho}(N) := 2\sqrt{(L^2 - \Omega^2)d_N^2 + \Omega d_N}$ and $\rho_W(N) := \tilde{\rho}(N) + \sqrt{\frac{4\log(2N/\delta_N)}{N\kappa_N}}$.

Proof. Note that for N large enough the condition $\delta_N \in (Ne^{-N/5}, e^{-1})$ is satisfied under the assumptions of this lemma. In fact, if $\delta_N \leq Ne^{-N/5}$ infinitely often then $\frac{\log(N/\delta_N)}{N\kappa_N} \geq \frac{\log(N/N \cdot e^{N/5})}{N} = \frac{1}{5}$ infinitely often and the assumption $\frac{\log(N/\delta_N)}{N\kappa_N} \to 0$ would be violated. Hence Lemma 11 applies and the result for piecewise Lipschitz graphons follows from (Avella-Medina et al., 2018, Theorem 1). We here report a simplified proof for Lipschitz continuous graphons (i.e. for the case $\Omega = 0$ in Footnote 11).

1. By Assumption 4 and Lemma 11, with probability $1 - \delta_N$

$$\begin{aligned}\|\theta^{[N]} - \theta\|_{L^2;\mathbb{R}^m}^2 &= \int_0^1 \|\theta^{[N]}(x) - \theta(x)\|^2 dx = \sum_i \int_{\mathcal{U}_i^{[N]}} \|\theta(t^i) - \theta(x)\|^2 dx \\ &\leq \sum_i \int_{\mathcal{U}_i^{[N]}} L^2 |t^i - x|^2 dx \leq \sum_i \int_{\mathcal{U}_i^{[N]}} (Ld_N)^2 dx = (Ld_N)^2.\end{aligned}$$

2. Consider any $f \in L^2([0,1])$ such that $\|f\|_{L^2} = 1$. Let $D(x,y) := W_w^{[N]}(x,y) - W(x,y)$. Then with probability $1 - \delta_N$ (independent of f)

$$\begin{aligned}\|\mathbb{W}_w^{[N]} f - \mathbb{W} f\|_{L^2}^2 &= \int_0^1 (\mathbb{W}_w^{[N]} f - \mathbb{W} f)(x)^2 \mathrm{d}x = \int_0^1 \left(\int_0^1 D(x,y) f(y) \mathrm{d}y\right)^2 \mathrm{d}x \\ &\leq \int_0^1 \left(\int_0^1 D(x,y)^2 \mathrm{d}y\right)\left(\int_0^1 f(y)^2 \mathrm{d}y\right) \mathrm{d}x \\ &= \int_0^1 \left(\int_0^1 D(x,y)^2 \mathrm{d}y\right) \|f\|_{L^2}^2 \mathrm{d}x = \int_0^1 \int_0^1 D(x,y)^2 \mathrm{d}y \mathrm{d}x \\ &= \sum_i \sum_j \int_{\mathcal{U}_i^{[N]}} \int_{\mathcal{U}_j^{[N]}} (W_w^{[N]}(x,y) - W(x,y))^2 \mathrm{d}y \mathrm{d}x \\ &= \sum_i \sum_j \int_{\mathcal{U}_i^{[N]}} \int_{\mathcal{U}_j^{[N]}} (W(t^i,t^j) - W(x,y))^2 \mathrm{d}y \mathrm{d}x \\ &\leq L^2 \sum_i \sum_j \int_{\mathcal{U}_i^{[N]}} \int_{\mathcal{U}_j^{[N]}} (|t^i - x| + |t^j - y|)^2 \mathrm{d}y \mathrm{d}x \\ &\leq L^2 \sum_i \sum_j \int_{\mathcal{U}_i^{[N]}} \int_{\mathcal{U}_j^{[N]}} (2d_N)^2 \mathrm{d}y \mathrm{d}x = (2Ld_N)^2\end{aligned}$$

where we used (D.5) from Lemma 11 in the last inequality. Hence with probability $1 - \delta_N$, (D.5) holds and

$$\left|\left|\left|\mathbb{W}_w^{[N]} - \mathbb{W}\right|\right|\right| = \sup_{f \in L^2([0,1]) s.t. \|f\|_{L_2}=1} \|\mathbb{W}_w^{[N]} f - \mathbb{W} f\|_{L^2} \leq 2Ld_N.$$

The fact that $|\lambda_{\max}(\mathbb{W}_w^{[N]}) - \lambda_{\max}(\mathbb{W})| \leq \left|\left|\left|\mathbb{W}_w^{[N]} - \mathbb{W}\right|\right|\right|$ can be proven by inverse triangular inequality upon noting that $\lambda_{\max}(\mathbb{W}_w^{[N]}) = \left|\left|\left|\mathbb{W}_w^{[N]}\right|\right|\right|$ and $\lambda_{\max}(\mathbb{W}) = |||\mathbb{W}|||$.

3. The operator $\mathbb{W}_s^{[N]} - \mathbb{W}_w^{[N]}$ can be seen as the graphon operator of a stochastic block model graphon with matrix $\frac{A_s^{[N]}}{\kappa_N} - A_w^{[N]}$ over the uniform partion $\{\mathcal{U}_i^{[N]}\}_{i=1}^N$. Note that for any graphon operator $\mathbb{A}$ over such partition (i.e. $\mathbb{A}(x,y) = A_{ij}$ for $x \in \mathcal{U}_i^{[N]}, y \in \mathcal{U}_i^{[N]}$) it holds $|||\mathbb{A}||| \leq \frac{1}{N}\|A\|$. Consequently,

$$\left|\left|\left|\mathbb{W}_s^{[N]} - \mathbb{W}_w^{[N]}\right|\right|\right| \leq \frac{1}{N}\left\|\frac{A_s^{[N]}}{\kappa_N} - A_w^{[N]}\right\| = \frac{1}{N\kappa_N}\|A_s^{[N]} - \kappa_N A_w^{[N]}\|.$$

Recall that $\kappa_N A_w^{[N]} = \mathbb{E}[A_s^{[N]}]$, hence we can bound the term on the right hand side by employing matrix concentration inequalities.

The maximum expected degree $C_N^d := \max_i(\sum_{j=1}^N \kappa_N [A_w^{[N]}]_{ij})$ grows as order $\kappa_N N$. Hence for N large enough, it is greater than $\frac{4}{9}\log(\frac{2N}{\delta_N})$ since $\frac{\log(N/\delta_N)}{N\kappa_N} \to 0$ by assumption. Consequently, all the conditions of (Chung and Radcliffe, 2011, Theorem 1) are met and with probability $1 - \delta_N$

$$\frac{1}{N\kappa_N}\|A_s^{[N]} - \kappa_N A_w^{[N]}\| \leq \frac{1}{N\kappa_N}\sqrt{4C_N^d \log(2N/\delta_N)} \leq \sqrt{\frac{4\log(2N/\delta_N)}{N\kappa_N}},$$

where we used that $C_N^d \leq \kappa_N N$ since each element in $A_w^{[N]}$ belongs to $[0,1]$. Using the fact that

$$\left|\left|\left|\mathbb{W}_s^{[N]} - \mathbb{W}\right|\right|\right| \leq \left|\left|\left|\mathbb{W}_s^{[N]} - \mathbb{W}_w^{[N]}\right|\right|\right| + \left|\left|\left|\mathbb{W}_w^{[N]} - \mathbb{W}\right|\right|\right|$$

and the previous statement concludes the proof. The fact that $|\lambda_{\max}(\mathbb{W}_s^{[N]}) - \lambda_{\max}(\mathbb{W})| \leq \left|\left|\left|\mathbb{W}_s^{[N]} - \mathbb{W}\right|\right|\right|$ can be proven as in the previous point.

□

Lemma 13. Consider a graphon game satisfying Assumptions 1, 2, 3 and 4 with

$\Omega = 0$ and suppose that $\mathtt{S}(x) = \mathcal{S}$ for all x. Then the unique graphon equilibrium is Lipschitz continuous with constant $L_s = \frac{\max\{\ell_U,\ell_\theta\}L(s_{\max}+1)}{\alpha_U}$.

Proof. Let $\bar{s}$ be the unique graphon equilibrium and $\bar{z} = \int_0^1 W(x,y)\bar{s}(y)dy$. For any $x_1, x_2 \in [0,1]$ it holds

$$\begin{aligned}
\|\bar{s}(x_1) - \bar{s}(x_2)\| &= \|\arg\max_{s\in\mathcal{S}} U(s,\bar{z}(x_1),\theta(x_1)) - \arg\max_{s\in\mathcal{S}} U(s,\bar{z}(x_2),\theta(x_2))\| \\
&\leq \frac{1}{\alpha_U}\|\nabla_s U(\bar{s}(x_1),\bar{z}(x_1),\theta(x_1)) - \nabla_s U(\bar{s}(x_1),\bar{z}(x_2),\theta(x_2))\| \\
&\leq \frac{\max\{\ell_U,\ell_\theta\}}{\alpha_U}\left(\|\bar{z}(x_1) - \bar{z}(x_2)\| + \|\theta(x_1) - \theta(x_2)\|\right).
\end{aligned} \tag{D.6}$$

Moreover,

$$\begin{aligned}
&\|\bar{z}(x_1) - \bar{z}(x_2)\| = \|\int_0^1 W(x_1,y)\bar{s}(y)dy - \int_0^1 W(x_2,y)\bar{s}(y)dy\| \\
&\leq \int_0^1 |W(x_1,y) - W(x_2,y)|\|\bar{s}(y)\|dy \leq \int_0^1 L|x_1 - x_2|s_{\max}dy = L|x_1 - x_2|s_{\max}, \\
&\|\theta(x_1) - \theta(x_2)\| \leq L|x_1 - x_2|.
\end{aligned} \tag{D.7}$$

Combining (D.6) and (D.7) yields

$$\|\bar{s}(x_1) - \bar{s}(x_2)\| \leq \frac{\max\{\ell_U,\ell_\theta\}L(s_{\max}+1)}{\alpha_U}|x_1 - x_2|.$$

□

Lemma 14. Suppose that the assumptions of Theorem 5 hold. Consider a fixed population size N, a fixed $t^i \in [0,1]$ and let $\zeta_{\bar{s}}(t^i)$ be a realization of $\frac{1}{N-1}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j)$, where $[A_s^{[N]}]$ is sampled from the graphon W according to Definition 3. Then with probability at least $1 - \frac{2n+1}{(N-1)^2}$ it holds $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \leq \varepsilon'$, with $\varepsilon' := \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$. By the union bound with probability at least $1 - \frac{(2n+1)N}{(N-1)^2}$ it holds $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \leq \varepsilon'$ for all $i \in \{1,\ldots,N\}$.

Proof. Let t^{-i} be the types of all the agents except for agent i. For each realization of t^{-i} we have

$$\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| = \|\frac{1}{N-1}\sum_{j\neq i}[A_s^{[N]}]_{ij}\bar{s}(t^j) - \bar{z}(t^i)\|$$

$$= \|\frac{1}{N-1}\sum_{j\neq i}\left([A_s^{[N]}]_{ij}\bar{s}(t^j) - W(t^i,t^j)\bar{s}(t^j) + W(t^i,t^j)\bar{s}(t^j)\right) - \bar{z}(t^i)\|$$

$$\leq \underbrace{\|\frac{1}{N-1}\sum_{j\neq i}([A_s^{[N]}]_{ij} - W(t^i,t^j))\bar{s}(t^j)\|}_{\text{Term 1}} + \underbrace{\|\frac{1}{N-1}\sum_{j\neq i}W(t^i,t^j)\bar{s}(t^j) - \bar{z}(t^i)\|}_{\text{Term 2}}.$$

We can bound the two terms separately.

- Term 1: Note that $\sum_{j\neq i}([A_s^{[N]}]_{ij} - W(t^i,t^j))\bar{s}(t^j) \in \mathbb{R}^n$. For each $h \in \{1,\ldots,n\}$, we denote by $S_h := \sum_{j\neq i}([A_s^{[N]}]_{ij} - W(t^i,t^j))\bar{s}_h(t^j)$ the h-th component of the previous vector and analyze each component separately.

  Let $X_j^h = ([A_s^{[N]}]_{ij} - W(t^i,t^j))\bar{s}_h(t^j)$ and note that for a fixed h the random variables $\{X_j^h\}_{j\neq i}$ are independent, zero mean and $-s_{\max} \leq X_j^h \leq s_{\max}$ for all $j \neq i$. Moreover, by definition $S_h = \sum_{j\neq i} X_j^h$. Note that $\mathbb{E}[S_h] = 0$. The Hoeffding's inequality then yields

$$\Pr\left[\frac{|S_h|}{N-1} > s_{\max}\sqrt{\frac{4\log(N-1)}{N-1}}\right] = \Pr\left[|S_h| > s_{\max}\sqrt{4\log(N-1)(N-1)}\right]$$
$$< 2\exp\left(-\frac{2s_{\max}^2 4\log(N-1)(N-1)}{(N-1)(2s_{\max})^2}\right) = 2\exp\left(-2\log(N-1)\right) = \frac{2}{(N-1)^2}.$$

  Hence for any $h \in \{1,\ldots,n\}$, with probability at least $1 - \frac{2}{(N-1)^2}$, it holds $\frac{|S_h|}{N-1} = \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$. By the union bound, with probability at least $1-\frac{2n}{(N-1)^2}$, it holds $\frac{|S_h|}{N-1} = \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$ for all $h \in \{1,\ldots,n\}$. With the same probability

$$[\text{term 1}] = \sqrt{\sum_{h=1}^n\left(\frac{S_h}{N-1}\right)^2} = \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right).$$

- Term 2:
  Let $\{t_{(k)}^{-i}\}_{k=1}^{N-1}$ be the ordered statistics of $\{t^j\}_{j\neq i}$ so that $t_{(1)}^{-i} \leq \ldots \leq t_{(N-1)}^{-i}$. By (Avella-Medina et al., 2018, Proposition 3) (see also Lemma 11) the set of

realizations of $\{t^j\}_{j\neq i}$ such that $|t^{-i}_{(k)}-y| \leq d_{N-1}$ for all $y \in \mathcal{U}_k^{[N-1]} := [\frac{k-1}{N-1}, \frac{k}{N-1})$ and for all $k \in \{1, \dots, N-1\}$ has measure at least $1-\delta_{N-1}$. Consequently, with this probability it holds

$$
\begin{aligned}
&[\text{term 2}] = \|\tfrac{1}{N-1}\textstyle\sum_{j\neq i} W(t^i,t^j)\bar{s}(t^j) - \bar{z}(t^i)\| \\
&= \|\frac{1}{N-1}\sum_{k=1}^{N-1} W(t^i,t^{-i}_{(k)})\bar{s}(t^{-i}_{(k)}) - \bar{z}(t^i)\| \\
&= \|\tfrac{1}{N-1}\textstyle\sum_{k=1}^{N-1} W(t^i,t^{-i}_{(k)})\bar{s}(t^{-i}_{(k)}) - \sum_{k=1}^{N-1}\int_{\mathcal{U}_k^{[N-1]}} W(t^i,y)\bar{s}(y)dy\| \\
&= \|\textstyle\sum_{k=1}^{N-1}\int_{\mathcal{U}_k^{[N-1]}} [W(t^i,t^{-i}_{(k)})\bar{s}(t^{-i}_{(k)}) - W(t^i,y)\bar{s}(y)]dy\| \\
&\leq \textstyle\sum_{k=1}^{N-1}\int_{\mathcal{U}_k^{[N-1]}} \|W(t^i,t^{-i}_{(k)})\bar{s}(t^{-i}_{(k)}) - W(t^i,t^{-i}_{(k)})\bar{s}(y)\| + \|W(t^i,t^{-i}_{(k)})\bar{s}(y) - W(t^i,y)\bar{s}(y)\|dy \\
&\leq \textstyle\sum_{k=1}^{N-1}\int_{\mathcal{U}_k^{[N-1]}} (L_s + Ls_{\max})|t^{-i}_{(k)} - y|dy \leq (L_s + Ls_{\max})d_{N-1},
\end{aligned}
$$

where the second to last inequality follows from the fact that, under the given assumptions, $\bar{s}$ is Lipschitz continuous with constant L_s (see Lemma 13), $\|\bar{s}\| \leq s_{\max}$ and W is Lipschitz continuous with constant L. By selecting $\delta_{N-1} = \frac{1}{(N-1)^2}$ with probability at least $1-\frac{1}{(N-1)^2}$, [term 2]$= \mathcal{O}(d_{N-1}) = \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$.

By the union bound with probability at least $1-\frac{2n+1}{(N-1)^2}$ it holds $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| = \mathcal{O}\left(\sqrt{\frac{\log(N-1)}{N-1}}\right)$.

□

Lemma 15. Consider a graphon game $\mathcal{G}(\mathtt{S}, U, \theta, W)$ in which $\mathtt{S}(x) = \mathcal{S}$ for all $x \in [0,1]$. Suppose that Assumptions 1, 2, 3, 4 (with $\Omega = 0$) and 5 hold. Let $\bar{s}$ be the unique equilibrium of the graphon game. Then with probability $1-\frac{(2n+1)}{N}$, the set $\{\tilde{s}^i := \bar{s}(t^i)\}_{i=1}^N$ is an ε Nash equilibrium of the sampled network game $\mathcal{G}^{[N]}(\{\mathcal{S}\}_{i=1}^N, U, \{\theta(t^i)\}_{i=1}^N, A_s^{[N]})$ with

$$\varepsilon = \mathcal{O}\left(\sqrt{\frac{\log(N)}{N}}\right).$$

Proof. For any agent i, let $\tilde{z}^i = \frac{1}{N}\sum_j [A_s^{[N]}]_{ij}\tilde{s}^j = \frac{1}{N}\sum_j [A_s^{[N]}]_{ij}\bar{s}(t^j) = \zeta_{\bar{s}}(t^i)$ then for any $s^i \in \mathcal{S}$

$$\begin{aligned} U(\tilde{s}^i, \tilde{z}^i, \theta^i) &= U(\bar{s}(t^i), \zeta_{\bar{s}}(t^i), \theta(t^i)) \geq U(\bar{s}(t^i), \bar{z}(t^i), \theta(t^i)) - L_U \|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \\ &\geq U(s^i, \bar{z}(t^i), \theta(t^i)) - L_U \|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| \\ &\geq U(s^i, \zeta_{\bar{s}}(t^i), \theta(t^i)) - 2L_U \|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| = U(s^i, \tilde{z}^i, \theta^i) - 2L_U \|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\|. \end{aligned}$$

The proof is concluded by noting that by Lemma 14, with probability $1 - \frac{(2n+1)}{N}$, $\|\zeta_{\bar{s}}(t^i) - \bar{z}(t^i)\| = \mathcal{O}\left(\sqrt{\frac{\log(N)}{N}}\right)$ for all agents $i = 1, \ldots, N$ (note that Lemma 14 is proven for normalization $\frac{1}{N-1}$ but similar arguments apply to $\frac{1}{N}$). □