

# GRAPHON GAMES: A STATISTICAL FRAMEWORK FOR NETWORK GAMES AND INTERVENTIONS

FRANCESCA PARISE

School of Electrical and Computer Engineering, Cornell University

ASUMAN OZDAGLAR

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

In this paper, we present a unifying framework for analyzing equilibria and designing interventions for large network games sampled from a stochastic network formation process represented by a graphon. To this end, we introduce a new class of infinite population games, termed *graphon games*, in which a continuum of heterogeneous agents interact according to a graphon, and we show that equilibria of graphon games can be used to approximate equilibria of large network games sampled from the graphon. This suggests a new approach for design of interventions and parameter inference based on the limiting infinite population graphon game. We show that, under some regularity assumptions, such approach enables the design of asymptotically optimal interventions via the solution of an optimization problem with much lower dimension than the one based on the entire network structure. We illustrate our framework on a synthetic data set and show that the graphon intervention can be computed efficiently and based solely on aggregated relational data.

**KEYWORDS:** Network games, aggregative games, large population games, Nash equilibrium, targeted interventions.

## 1. INTRODUCTION

RECENT DECADES have witnessed tremendous progress in the theory of network games, which have been used widely to model, understand, and predict behavior in a range of settings involving strategic interactions of agents embedded in networked environments. Despite this progress, several issues remain when considering interventions or regulation of economic behavior over *large scale* networks. First, in this case, the optimization problem that the central planner needs to solve for determining the optimal intervention is very high dimensional, often scaling with the size of the network. Second, assuming that the central planner has access to detailed information about the network structure is not a good approximation of reality since collection of exact network data is either extremely costly or, in many settings, not at all possible due to proprietary and privacy concerns.<sup>1</sup>

To overcome these issues, in this paper we define a new class of infinite population games, termed graphon games, which can approximate a wide variety of complex strategic interactions in large network environments. Graphon games involve a continuum of agents whose payoff depends on their own action as well as a weighted average of other

---

Francesca Parise: [fp264@cornell.edu](mailto:fp264@cornell.edu)

Asuman Ozdaglar: [asuman@mit.edu](mailto:asuman@mit.edu)

Research supported in part by SNSF Grant P300P2\_177805, by ARO MURI W911NF-12-1-0509, and by ARO W911NF-18-1-0407. The authors thank the anonymous referees, D. Acemoglu, M. Avella-Medina, M.T. Schaub, S. Segarra, R. Carmona, A. Jadbabaie, as well as the participants to the Fourth and Fifth Annual Conference on Network Science and Economics, for excellent comments and helpful conversations.

<sup>1</sup>Breza, Chandrasekhar, McCormick, and Pan (2020) estimated that conducting network surveys in 120 Indian villages would cost  $\approx$ \$190,000 and take over eight months. Proprietary and privacy concerns may arise, for example, when measuring high-risk populations or transactions between networks of financial intermediaries.

agents' actions, with heterogeneous weights specified by a graphon model. We stress that graphons have two interpretations in relation to networks. First, as shown in Lovász and Szegedy (2006), Lovász (2012), Borgs, Chayes, Lovász, Sós, and Vesztegombi (2008), they can be seen as the limit of a graph when the number of agents tends to infinity and can thus be used to capture heterogeneous interaction among a continuum of agents, as discussed above. Second, they can be used as a flexible representation of stochastic network formation models (that generalizes, e.g., Erdős–Rényi and stochastic block models).<sup>2</sup> To connect our infinite population analysis to finite network games, we consider this second interpretation of graphons and study large network games in which agents interact according to a finite network sampled from such a stochastic model. Our first key result shows that the equilibrium in such large sampled network games can be well approximated by the equilibrium of the corresponding graphon game. We provide bounds on the distance between sampled and graphon equilibria as a function of the network size and prove that it vanishes as the number of agents grows. This convergence result enables the study of large network games by considering their limit: the graphon game.

To illustrate why this is important, we turn to intervention design and show that the graphon approximation significantly reduces both the amount of data and the computational burden required to design interventions. In particular, we show that, under a finite-rank assumption on the graphon—leading to a low dimensional stochastic network formation model—the optimization problem faced by the central planner can be approximated by a low dimensional problem (with size corresponding to the rank of the graphon instead of the number of agents). Furthermore, under the same assumptions, graphon interventions can be designed with much less information than the entire network structure. To illustrate this second point, we show through a synthetic case study that easily collectable *aggregated relational data* are sufficient to estimate the parameters of network games sampled from a stochastic block model (a widely used type of graphon in which agents are partitioned into a finite number of communities).<sup>3</sup> For this case study, graphon interventions can be planned by solving an optimization problem with dimension equal to the number of blocks (communities) instead of number of agents, leading to an asymptotically optimal yet computationally tractable procedure for intervention design.

We discuss further applications of our framework to parameter estimation and incomplete information games in the Appendix, suggesting how our main convergence result can be exploited to derive an alternative approach to a number of classic network games problems beyond targeted intervention, by focusing on the limiting graphon game.

### 1.1. Detailed Contributions

We define a *graphon game* in terms of a continuum of agents indexed in  $[0, 1]$  and a graphon, represented by a symmetric measurable function  $W : [0, 1]^2 \rightarrow [0, 1]$  with  $W(x, y)$  denoting the influence of agent  $y$ 's strategy on agent  $x$ 's payoff function. We

<sup>2</sup>By the Aldous–Hoover theorem, any exchangeable infinite random graph is obtained as a mixture of graphons; Diaconis and Janson (2007).

<sup>3</sup>Aggregated relational data, as introduced in Breza et al. (2020), are collected through questions such as *how many of the agents you interact with have trait  $k$ ?*, instead of questions of the form *what is the identity of all the agents you interact with?* For villages in Karnataka, India, Breza et al. (2020) showed using J-PAL South Asia cost estimates that collecting aggregated relational data leads to a 70–80% cost reduction with respect to the cost of the exact network data collected for the study in Banerjee, Chandrasekhar, Duffo, and Jackson (2013).

assume that agent  $x$ 's payoff function depends on his strategy  $s(x)$  as well as a weighted average of other agents' strategies computed according to the graphon  $W$ .

Our first main contribution is to show that, beyond the natural interpretation of graphon games as a model for strategic heterogeneous interactions among a *continuum* of nonatomic agents, equilibria of graphon games can also be used to approximate strategic behavior in *large but finite* populations of agents. To this end, we build on the interpretation of graphons as stochastic network formation models and show that equilibria of finite network games, in which agents interact according to a network sampled from a graphon, converge almost surely to the equilibrium of the corresponding graphon game. For simplicity of exposition, we first present our convergence result for the case of *dense undirected networks*, in which the number of neighbors grows linearly with the population size and the local aggregate is defined as the sum of neighbors' actions (normalized by the network size). We then show in Section 5 that our results can be generalized from undirected to directed networks and from games in which the sum of neighbors' actions is normalized by the population size to games in which it is normalized by each agent's degree. Additionally, we show that our results can be extended to classes of networks in which the number of neighbors grows sublinearly in the population size so that the edge density converges to zero. We stress that for our convergence result to hold, the number of neighbors still needs to increase at least logarithmically. This condition is used in our technical results to ensure that concentration inequality bounds apply (see also [Jackson and Storms \(2019\)](#)). Nonetheless, we show within our case study that, in practice, the graphon approach leads to useful insights even when the average degree in a network of 1000 agents is around 20, illustrating the applicability of our framework to realistic networks.

As a second main contribution, we show how our main convergence result can be exploited to suggest an alternative approach for design of targeted interventions. To this end, we formulate a novel optimization problem in the graphon space which, through sampling, provides interventions for finite sampled network games. We show that such graphon-based interventions are asymptotically optimal and that, for finite rank graphons, the graphon optimization problem is a tractable finite dimensional problem. We apply our results to both linear quadratic network games, as studied in [Galeotti, Golub, and Goyal \(2020\)](#), and to a class of nonlinear and non-monotone network games considered in [Parise and Ozdaglar \(2019\)](#).

To illustrate the computational and informational gains obtained with the graphon approach, we perform a case study on a simulated data set of 80 different networks, drawn as independent realizations of a stochastic block model with four communities. For this case study, the graphon approach results in a near optimal solution and reduces to a four-dimensional optimization problem (whereas designing the optimal intervention or a network heuristic suggested in [Galeotti, Golub, and Goyal \(2020\)](#) would necessitate solving a problem of dimension equal to the size  $N$  of the network which we set to  $N = 300, 600, 1200$  in our simulations). We also stress that our procedure can be applied based on aggregated data (the graphon model) instead of requiring exact network data. Finally, within this case study, we suggest how to estimate peer effects under partial network data. This is a topic of recent interest, as discussed, for example, in [Chandrasekhar and Lewis \(2016\)](#), [De Paula, Rasul, and Souza \(2018\)](#), [Boucher and Houndetoungan \(2019\)](#), [Lewbel, Qu, and Tang \(2019\)](#). Estimating peer effects is not the subject of our work, hence we do not develop this aspect beyond the intuition given in the case study and a preliminary analysis given in Appendix C of the Supplemental Material ([Parise and Ozdaglar \(2023\)](#)).

The results discussed so far are derived under the assumption that agents have perfect information about the sampled network. In Appendix B of the Supplemental Material, we analyze an incomplete information version of sampled network games and develop a close relation between the corresponding Bayesian Nash equilibrium and the graphon equilibrium discussed above. In particular, we show that, under suitable regularity conditions and under the assumption that the agents know the graphon generating the sampled network (but not the realization), the graphon equilibrium is an  $\varepsilon$ -Bayesian Nash equilibrium for the incomplete information game.

## 1.2. *Related Literature*

The idea of dimensionality reduction via stochastic network modeling is a key concept of statistical network analysis and has been successfully applied in the previous literature to specific instances of network games and specific instances of network models. For example, [Golub and Jackson \(2012a,b\)](#) used stochastic block models to achieve dimensionality reduction in opinion dynamic models. In more recent work, contagion processes have been considered in [Akbarpour, Malladi, and Saberi \(2018\)](#) (which focuses on random seeding in a generalized Erdős–Rényi model), [Jackson and Storms \(2019\)](#) (which focuses on the effect of behavioral communities in stochastic block models), and [Sadler \(2020\)](#) (which studies the Bayesian equilibrium of a game-theoretic model of diffusion). The present paper adds to this literature by (i) bringing the idea of dimensionality reduction via stochastic network modeling to a general class of strategic interactions (moving beyond opinion dynamics and contagion processes),<sup>4</sup> (ii) focusing on how to exploit such dimensionality reduction to derive computationally tractable methods, based on the limiting graphon game, to design targeted interventions and to perform parameter inference, and (iii) using graphons as a broad description of the generating process to unify several different stochastic models.

Our results on incomplete information network games, reported in Appendix B of the Supplemental Material, are related to two previous works: [Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv \(2010\)](#) and [Kalai \(2004\)](#). [Galeotti et al. \(2010\)](#) focused on finite network games over a specific random network model in which agents only know their degree and did not discuss dimensionality reduction. [Kalai \(2004\)](#) proved that the Bayesian Nash equilibrium of a game with anonymous payoffs is an almost Nash equilibrium of the complete information game when the number of agents tends to infinity. Two points are noteworthy in relating this paper to the network game literature and our paper in particular. First, network games capture heterogeneous interactions, hence do not satisfy the anonymity assumption in [Kalai \(2004\)](#). Second, [Kalai \(2004\)](#) showed that the Bayesian Nash equilibrium is an  $\varepsilon$ -Nash equilibrium; instead, we prove that the Bayesian Nash equilibrium converges (in strategies) to the equilibrium of the corresponding graphon game, thus providing a characterization of the limiting behavior which is, under suitable assumptions, low dimensional.

While our goal is to use graphon games to approximate equilibria of sampled network games, we note that graphon games can also be of independent interest as a model of

---

<sup>4</sup>While contagion models have discrete (typically 0-1) strategies, we focus here on games with continuous strategies. The type of continuous games considered in our framework has been broadly used in the literature, in both theoretical and empirical works, for example to model applications in which agents need to decide on their level of effort or investment in a certain activity (see [Vives \(2005\)](#), [Ballester, Calvó-Armengol, and Zenou \(2006\)](#), [Acemoglu, Ozdaglar, and Tahbaz-Salehi \(2015\)](#), [Bramoullé and Kranton \(2007\)](#), [Bramoullé, Kranton, and D'Amours \(2014\)](#), [Allouch \(2015\)](#)).

nonatomic games. In this context, our work complements results derived for finite population network games (see, e.g., Ballester, Calvó-Armengol, and Zenou (2006), Bramoullé and Kranton (2007), Bramoullé, Kranton, and D’Amours (2014), Jackson and Zenou (2014), Bramoullé and Kranton (2016), Galeotti, Golub, and Goyal (2020)) as well as previous infinite population models by incorporating heterogeneous local effects in infinite population games (see, e.g., Lasry and Lions (2007), Huang, Caines, and Malhamé (2007), Sandholm (2010), Kukushkin (2004), Jensen (2010), Cornes and Hartley (2012), Dubey, Haimanko, and Zapechelnyuk (2006), Ma, Callaway, and Hiskens (2013), Altman, Boulogne, El-Azouzi, Jiménez, and Wynter (2006)).

We finally remark that the idea of using graphons as a support for large population analysis has been successfully applied recently in different areas such as community detection in Eldridge, Belkin, and Wang (2016), crowd-sourcing in Lee and Shah (2017), signal processing in Morency and Leus (2017), and optimal control of dynamical systems in Gao and Caines (2017). The concurrent work by Caines and Huang (2018) suggests the use of graphons to extend the setup of mean-field games (which, differently from network games, are dynamic and stochastic games) to heterogeneous settings. Moreover, the idea of interpreting observed graphs as random realizations from an underlying random graph model has recently been used in the study of centrality measures in Dasaratha (2020) for stochastic block models and in Avella-Medina, Parise, Schaub, and Segarra (2018) for graphon models. The authors of these papers studied, among others, Bonacich centrality, which is the equilibrium of a specific type of network games with scalar nonnegative strategies, quadratic payoff functions, and strategic complements.

## 2. GRAPHON GAMES

### 2.1. Recap on Network Games

Network games capture settings in which agents make decisions while interacting with others through a network; see, for example, Jackson and Zenou (2014). In the following, we describe such a network with an adjacency matrix  $A^{[N]} \in \mathbb{R}^{N \times N}$ , where  $N$  is the population size and  $A_{ij}^{[N]}$  denotes the level of interaction between agents  $i$  and  $j$ . For simplicity, we start by assuming that the network is undirected so that  $A^{[N]}$  is symmetric. Each agent  $i \in \{1, \dots, N\}$  selects a strategy  $s^i \in \mathbb{R}$  in its feasible set  $\mathcal{S}^i \subseteq \mathbb{R}$  to maximize a payoff function

$$U(s^i, z^i(s), \theta^i), \tag{1}$$

where  $s := [s^i]_{i=1}^N \in \mathbb{R}^N$ ,  $z^i(s) := \frac{1}{N} \sum_{j=1}^N [A^{[N]}]_{ij} s^j$  denotes the *local aggregate*<sup>5</sup> (defined as the weighted average of other agents’ strategies computed according to the heterogeneous weights of the network  $A^{[N]}$ ), and  $\theta^i \in \mathbb{R}$  is a parameter modeling heterogeneity in the payoff functions of different agents.<sup>6</sup> We denote compactly a network game with the notation  $\mathcal{G}^{[N]}(\{\mathcal{S}^i\}_{i=1}^N, U, \{\theta^i\}_{i=1}^N, A^{[N]})$  and we say a network game  $\mathcal{G}^{[N]}$  with network  $A^{[N]}$  if we need to stress the role of the network.

<sup>5</sup>In network games, typically there is no factor  $\frac{1}{N}$  in the definition of  $z^i(s)$ . Since we study the behavior when  $N$  changes, we find it useful to consider this factor explicitly. A different normalization in terms of agents’ degree instead of population size is discussed in Section 5.

<sup>6</sup>For simplicity of exposition, in the main text we consider games in which both  $s^i$  and  $\theta^i$  are scalars; the extension to the vector case is immediate (as presented in the Appendix).

### 2.2. Graphon Games: The Model

To extend network games to infinite populations, we consider a continuum of agents indexed by the variable  $x \in [0, 1]$  (instead of the finite index  $i \in \{1, \dots, N\}$ ). As in the finite population case, each agent selects a scalar strategy denoted by  $s(x) \in \mathbb{R}$  (instead of  $s^i \in \mathbb{R}$ ) under local constraints of the form  $s(x) \in S(x)$ . In finite network games, each agent computes its best response to the local aggregate  $z^i(s) := \frac{1}{N} \sum_{j=1}^N A_{ij}^{[N]} s^j$  according to the weights of the underlying graph  $A^{[N]}$ . In the infinite population case, the natural mathematical object to describe the network of interactions is a *graphon*. Graphons are bounded symmetric measurable functions  $W : [0, 1]^2 \mapsto [0, 1]$  and can be used to formally define the limit of a sequence of graphs when the number of nodes tends to infinity; see, for example, Lovász (2012). The value  $W(x, y)$  can thus be interpreted as measuring the level of interaction between two infinitesimal agents  $x$  and  $y$  belonging to the  $[0, 1]$  interval, exactly as  $A_{ij}^{[N]}$  denotes the level of interaction between agents  $i$  and  $j$  in  $\{1, \dots, N\}$ . For any graphon  $W$ , we can then define the network effects experienced by agent  $x$  as the local aggregate of the other agents' actions according to the graphon

$$z(x | s) := \int_0^1 W(x, y) s(y) dy.$$

REMARK 1: For graphon games, a strategy profile  $s : [0, 1] \rightarrow \mathbb{R}$  is a *function*. In the following, we require that any strategy profile is square integrable, that is,  $s(x) \in L^2([0, 1])$ , where  $L^2([0, 1])$  denotes the space of square integrable functions on  $[0, 1]$ .

Similarly to network games, we assume that the payoff function of agent  $x$ ,

$$U(s(x), z(x | s), \theta(x)), \tag{2}$$

depends on his strategy  $s(x)$ , on his local aggregate  $z(x | s)$ , and on a heterogeneity parameter  $\theta(x)$ .

DEFINITION 1: A *graphon game*  $\mathcal{G}$  is defined in terms of a continuum set of agents indexed by  $[0, 1]$ , a graphon  $W$ , a payoff function  $U$  as in (2), and, for each agent  $x \in [0, 1]$ , a parameter  $\theta(x)$  and a strategy set  $S(x)$ .

Note that the payoff function for graphon games has the same structural form as in network games. The difference in the two setups is the way in which the local aggregate ( $z^i(s)$  for network games and  $z(x | s)$  for graphon games) is evaluated. In the following, we say a graphon game  $\mathcal{G}$  with graphon  $W$  if we need to stress the role of the graphon, and we explicitly write  $\mathcal{G}(S, U, \theta, W)$  if we want to stress the role of all the game primitives.

### 2.3. Graphon Games: Equilibrium Properties

To derive properties of the Nash equilibrium in graphon games, we focus on smooth and convex payoff functions (see, e.g., Rosen (1965)).

ASSUMPTION 1—Smooth and Convex Game: *The function  $U(s, z, \theta)$  in (2) is continuously differentiable and strongly concave in  $s$  with uniform strong concavity constant  $\alpha_U$*

each value of  $z, \theta$ .<sup>7</sup> Moreover,  $\nabla_s U(s, z, \theta)$  is uniformly Lipschitz in  $[z, \theta]$  with constants  $\ell_U, \ell_\theta$  for all  $s$ . For each  $x \in [0, 1]$ , the set  $S(x)$  is convex, closed, and there exist  $\hat{z}$  and  $M > 0$  such that  $\| \arg \max_{s \in S(x)} U(\tilde{s}, \hat{z}, \theta(x)) \| \leq M$  for all  $x \in [0, 1]$ .

ASSUMPTION 2—Strategy Set: *There exists a compact set  $S$  such that  $S(x) \subseteq S$  for all  $x \in [0, 1]$  so that  $s_{\max} := \max_{s \in S} \|s\| < \infty$ .*

Lipschitz continuity of  $\nabla_s U(s, z, \theta)$  in Assumption 1 guarantees that the effect of the network aggregate  $z$  and the heterogeneity parameter  $\theta$  on the marginal payoff is continuous and bounded. Under Assumptions 1 and 2, it follows from standard fixed point theory that a Nash equilibrium exists. Uniqueness in games can typically be guaranteed if the best response mapping is a contraction. To specify such a contraction property in the context of graphon games, we introduce the *graphon operator*; see also Lovász (2012, Section 7.5).

DEFINITION 2: For a given graphon  $W$ , we define the associated *graphon operator*  $\mathbb{W}$  as the integral operator  $\mathbb{W} : L^2([0, 1]) \mapsto L^2([0, 1])$  given by  $f(x) \mapsto (\mathbb{W}f)(x) = \int_0^1 W(x, y)f(y) dy$ . A complex number  $\lambda$  is an *eigenvalue* of the operator  $\mathbb{W}$  if there exists a nonzero function  $\psi \in L^2([0, 1])$ , called the *eigenfunction*, such that  $(\mathbb{W}\psi)(x) = \lambda\psi(x)$ .

As summarized in Lemma 5 in the Supplemental Material, all the eigenvalues of the graphon operator  $\mathbb{W}$  are real and the operator norm, that is,  $\|\mathbb{W}\| := \sup_{f \in L^2([0, 1]), \|f\|_{L^2} = 1} \|\mathbb{W}f\|_{L^2}$ , coincides with the largest eigenvalue of  $\mathbb{W}$  which we denote by  $\lambda_{\max}(\mathbb{W})$ . We next show that if  $\lambda_{\max}(\mathbb{W})$  is not too large, as formalized in Assumption 3, then the best response operator is a contraction, guaranteeing uniqueness of the graphon equilibrium.

ASSUMPTION 3—Contraction: *Suppose that*

$$\frac{\ell_U}{\alpha_U} \cdot \lambda_{\max}(\mathbb{W}) < 1,$$

where  $\ell_U$  and  $\alpha_U$  are regularity constants related to the utility function, as defined in Assumption 1, while  $\lambda_{\max}(\mathbb{W})$  is the largest eigenvalue of the graphon operator  $\mathbb{W}$ .

REMARK 2: Assumption 3 is similar to assumptions used to obtain uniqueness in finite network games (see, e.g., Ballester, Calvó-Armengol, and Zenou (2006)) and guarantees that the effect of the neighbor’s aggregate on an agent’s marginal payoff, quantified by  $\ell_U \lambda_{\max}(\mathbb{W})$ , is not too large with respect to the effect of its own strategy, quantified by  $\alpha_U$ . The only difference is that while in the network game literature the effect of the network is captured by the maximum eigenvalue of the finite network  $\mathcal{A}^{[N]}$ , in the case of graphon games the corresponding role is played by the dominant eigenvalue of the graphon, that is,  $\lambda_{\max}(\mathbb{W})$ . In both cases, this quantity captures the maximum amount by which the network/graphon can amplify a unitary vector/function.

<sup>7</sup>That is,  $(\nabla_s U(s, z, \theta) - \nabla_s U(s', z, \theta))^\top (s - s') \leq -\alpha_U \|s - s'\|^2$  for all  $s, s', z, \theta$ .

**THEOREM 1—Existence and Uniqueness:** *If the graphon game  $\mathcal{G}(S, U, \theta, W)$  satisfies Assumptions 1 and 2, then it admits at least one Nash equilibrium. If it satisfies Assumptions 1 and 3, then the Nash equilibrium exists and is unique.*<sup>8</sup>

To illustrate this result, we consider the framework of linear quadratic games.

**EXAMPLE 1—Linear quadratic graphon games:** Consider a linear quadratic game in which the strategy of each agent is scalar and nonnegative, so that  $S(x) = \mathbb{R}_{\geq 0}$  for all  $x \in [0, 1]$  and the payoff function of an arbitrary agent playing strategy  $s$  and subject to the local aggregate  $z$  is

$$U(s, z, \theta) = -\frac{1}{2}s^2 + s[\alpha z + \theta]. \tag{3}$$

The parameter  $\alpha \in \mathbb{R}$  in (3) captures how much the local aggregate affects each agent’s marginal return. The parameter  $\theta > 0$  represents the *standalone marginal return* that does not depend on other’s actions. We refer to Jackson and Zenou (2014), Bramoullé and Kranton (2016) for a detailed review of this utility model and its applications. The best response for each agent  $x$  is

$$s_{\text{br}}(x \mid s) = \max\{0, [\alpha z(x \mid s) + \theta]\}. \tag{4}$$

Hence, the payoff function  $U$  satisfies Assumption 1 with  $\alpha_U = 1, \ell_U = |\alpha|$ . Consequently, by Theorem 1, a unique graphon equilibrium exists if  $|\alpha| < \frac{1}{\lambda_{\max}(\mathbb{W})}$ , which is a similar condition as the one derived in Ballester, Calvó-Armengol, and Zenou (2006) for finite network games. If additionally  $\alpha > 0$ , we can immediately see from (4) that the best response of each agent is an increasing function of the local aggregate  $z(x \mid s)$ , that is, this is a game of strategic complements (Ballester, Calvó-Armengol, and Zenou (2006)) and the unique Nash equilibrium  $\bar{s}$  is *internal* (i.e., it satisfies  $\bar{s}(x) > 0$  for all  $x \in [0, 1]$ ). For  $\theta(x) \equiv \theta$ , from (4) we then have  $\bar{s}(x) = \alpha z(x \mid \bar{s}) + \theta \Rightarrow \bar{s}(x) = \alpha(\mathbb{W}\bar{s})(x) + \theta \Rightarrow (\mathbb{I}\bar{s})(x) = \alpha(\mathbb{W}\bar{s})(x) + \theta \Rightarrow ((\mathbb{I} - \alpha\mathbb{W})\bar{s})(x) = \theta 1_{[0,1]}(x)$ . The condition  $|\alpha|\lambda_{\max}(\mathbb{W}) < 1$  implies invertibility of the operator  $(\mathbb{I} - \alpha\mathbb{W})$ . Hence,

$$\bar{s}(x) = \theta((\mathbb{I} - \alpha\mathbb{W})^{-1}1_{[0,1]})(x) = \theta \sum_{k=0}^{\infty} \alpha^k (\mathbb{W}^k 1_{[0,1]})(x), \tag{5}$$

which corresponds to the Bonacich centrality of agent  $x$  in the graphon  $W$ , as defined in Avella-Medina et al. (2018).

### 3. SAMPLED NETWORK GAMES: DEFINITION AND EXAMPLES

#### 3.1. Graphons as a Stochastic Network Formation Model

A graphon describes a probability distribution over the space of networks and can thus be used to construct sampled networks as discussed in Lovász (2012, Chapter 10) and illustrated in Figure 1.

---

<sup>8</sup>Note that when Assumption 3 holds, the strategy sets  $S(x)$  do not need to be bounded (i.e., Assumption 2 is not needed). This is because for contraction mappings, existence and uniqueness of the fixed point can be guaranteed under the sole assumption that the domain is closed and convex, without the need for compactness.



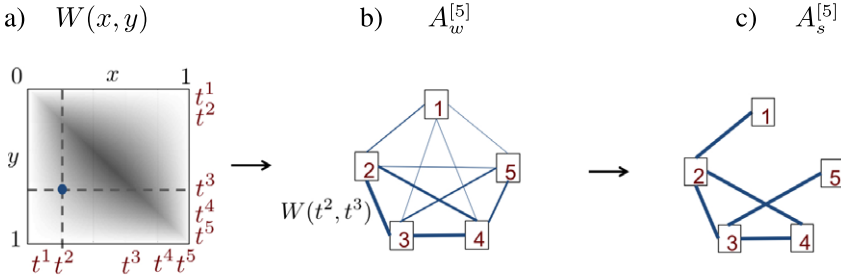


FIGURE 1.—Illustration of the sampling procedure described in Definition 3 for  $N = 5$ . (a) The graphon. (b) The *weighted adjacency matrix*  $A_w^{[5]}$  associated with the random sample  $[t^1, \dots, t^5] = [0.03, 0.31, 0.69, 0.82, 0.95]$ . (c) A realization of the 0-1 *adjacency matrix*  $A_s^{[5]}$ . For the graphon, a linear grayscale colormap is used with white associated to  $W = 0$  and black to  $W = 1$ . For  $A_w^{[5]}$ , the width of the line is proportional to the weight of the edge. In  $A_s^{[5]}$ , any edge has weight 1.

DEFINITION 3: Given any graphon  $W$  and any desired number  $N$  of nodes, a *weighted sampled network* can be obtained by uniformly and independently sampling  $N$  points  $\{t^i\}_{i=1}^N$  from  $[0, 1]$  and defining a *weighted adjacency matrix*  $A_w^{[N]}$  as follows:

$$[A_w^{[N]}]_{ij} = \begin{cases} W(t^i, t^j) & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

Starting from  $A_w^{[N]}$ , a simple sampled network can be obtained by defining the 0-1 *adjacency matrix*  $A_s^{[N]}$  as the adjacency matrix corresponding to a graph with  $N$  nodes obtained by randomly connecting nodes  $i, j \in [1, N]$  with Bernoulli probability  $[A_w^{[N]}]_{ij}$ . In the following, we use the symbol  $A_w^{[N]}$  for statements that hold for both weighted and simple sampled networks and simply refer to  $A_w^{[N]}$  as a *sampled network*.

REMARK 3: The random points  $\{t^i\}_{i=1}^N$  can be thought of as *agent types* (an agent’s type may, e.g., represent the community to which the agent belongs to or its geographical location, as discussed in the following Examples 2 and 3). The graphon value  $W(t^i, t^j)$  is thus encoding information about the level of interaction between two arbitrary agents of type  $t^i$  and  $t^j$ . From here on, we are going to assume that the  $\{t^i\}_{i=1}^N$  are ordered such that  $t^i \leq t^{i+1}$  for all  $i \in \{1, \dots, N - 1\}$ . This is without loss of generality, since it simply corresponds to a relabeling of the nodes. Figure 1 illustrates the sampling procedure described in Definition 3. Note that both  $A_w^{[N]}$  and  $A_s^{[N]}$  are stochastic matrices. The difference between the two is that  $A_w^{[N]} \in [0, 1]^{N \times N}$  while  $A_s^{[N]} \in \{0, 1\}^{N \times N}$ . Finally, note that an agent of type  $t^i$  has an expected number of neighbors that grows as  $N \int_0^1 W(t^i, t^j) dt^j$ . Hence, networks sampled according to Definition 3 are dense.

To develop more intuition on the framework of graphons and its connection to other well-known stochastic network formation models, we start by noting that for any  $p \in [0, 1]$ , the constant graphon  $W(x, y) \equiv p$  coincides with the Erdős–Rényi random graph model in which each pair of agents is connected with probability  $p$ . In the next example, we show how graphons can be used to represent stochastic block models, which are an extension of Erdős–Rényi models to a setting with finitely many communities.

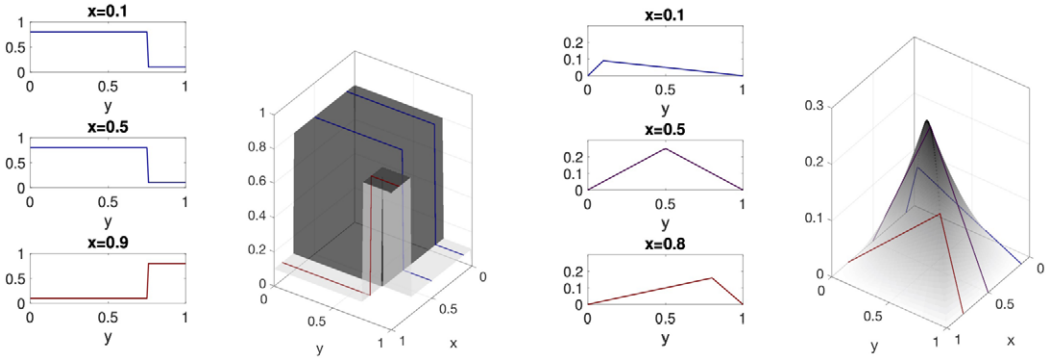


FIGURE 2.—Left: The stochastic block model graphon of Example 2. Note that the agents in position  $x = 0.1$  and  $x = 0.5$  belong to the same community and are thus affected in the same way by the rest of the agents. Right: The minmax graphon of Example 3. In both cases,  $W(x, \cdot)$  for three different values of  $x$  and  $W(x, y)$  as a function of both  $x$  and  $y$  are shown. A linear grayscale colormap is used with white associated to  $W = 0$  and black to  $W = 1$ .

EXAMPLE 2—Community Structure: Consider networks in which agents are divided into  $K$  communities and let  $\pi_k$  be the probability that a random agent belongs to community  $k$ , with  $\sum_{k=1}^K \pi_k = 1$ . Additionally, assume that agents belonging to the same community form a link with Bernoulli probability  $g_{in}$ , while agents from different communities form a link with probability  $g_{out}$  (typically smaller than  $g_{in}$ ).<sup>9</sup> To generate such a community structure from a graphon, one can partition  $[0, 1]$  into  $K$  disjoint intervals  $\{\mathcal{C}_k\}_{k=1}^K$ , with  $|\mathcal{C}_k| = \pi_k$ , and use the piecewise constant graphon

$$W_{SBM}(x, y) = \begin{cases} g_{in} & \text{if there exists } k \text{ s.t. } x \in \mathcal{C}_k, y \in \mathcal{C}_k, \\ g_{out} & \text{otherwise.} \end{cases}$$

We denote this graphon with the label SBM because of its relation to Stochastic Block Models. Figure 2 (left) illustrates a stochastic block model graphon of this type with  $K = 2$  communities (e.g., red and blue agents) of size  $[w_1, w_2] = [0.75, 0.25]$  and with  $g_{in} = 0.8$ ,  $g_{out} = 0.1$ . In this case, we selected  $\mathcal{C}_1 = [0, 0.75]$ ,  $\mathcal{C}_2 = (0.75, 1]$ .

Graphons can also be used to model situations in which agent types can take infinitely many values. The next example illustrates one such case in which an agent’s type is given by its location.

EXAMPLE 3—Location Model: Consider a model in which  $N$  agents are independently located uniformly at random along a line segment represented by the interval  $[0, 1]$  (e.g., homeowners along a street) and assume that the level of interaction between agents  $i$  and  $j$  is a decreasing function of their spatial distance, capturing the natural observation that the cost of forming links increases with geographical distance, as motivated in Johnson and Gilles (2003). This type of interaction can be represented, for example, by using

<sup>9</sup>The parameters  $\pi_k$  are exogenous and model the probability that agents are born with type  $k$ , for example, male or female. The exogenous parameters  $g_{in}$  and  $g_{out}$  are instead a result of the different costs borne by each agent when forming a link to someone from the same and from the other community (see, e.g., Jackson and Rogers (2005)).

the *minmax graphon*  $W_{MM}(x, y) = \min(x, y)(1 - \max(x, y))$ , where  $x \in [0, 1]$  denotes the agent's position along the line; see Figure 2 (right).

### 3.2. Sampled Network Games

We define a *sampled network game* as a network game in which agents interact over a *network sampled from a graphon*.

DEFINITION 4: Consider a graphon  $W$ , a payoff function  $U$ , a set valued function  $S$ , and a parameter function  $\theta$ . Let  $A_W^{[N]}$  be a network sampled from  $W$  with types  $\{t^i\}_{i=1}^N$ . We define a *sampled network game* as  $\mathcal{G}^{[N]}(\{S(t^i)\}_{i=1}^N, U, \{\theta(t^i)\}_{i=1}^N, A_W^{[N]})$ .

Figures 3 and 4 show the equilibria of three realizations of sampled network games with linear quadratic payoffs, when the networks are sampled from the graphons described in Examples 2 and 3, for different values of  $N$ . In both examples, one can observe similarities

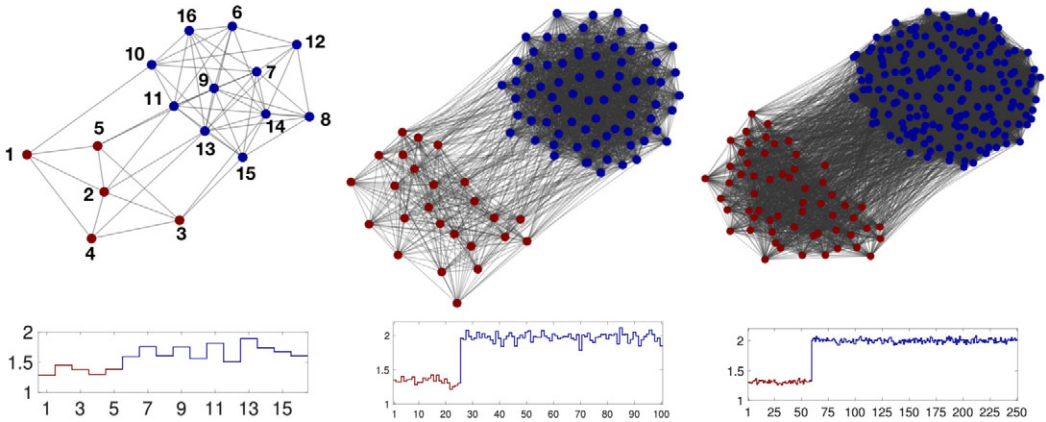


FIGURE 3.—Three realizations of networks formed according to the two-community model described in Example 2 (with  $\pi_{red} = 0.25$ ,  $\pi_{blue} = 0.75$ ,  $g_{in} = 0.8$ , and  $g_{out} = 0.1$ ) for  $N = 10, 100, 250$  and their corresponding equilibria (for payoff as in (3) with  $\alpha = 0.8$ ,  $\theta = 1$ ).

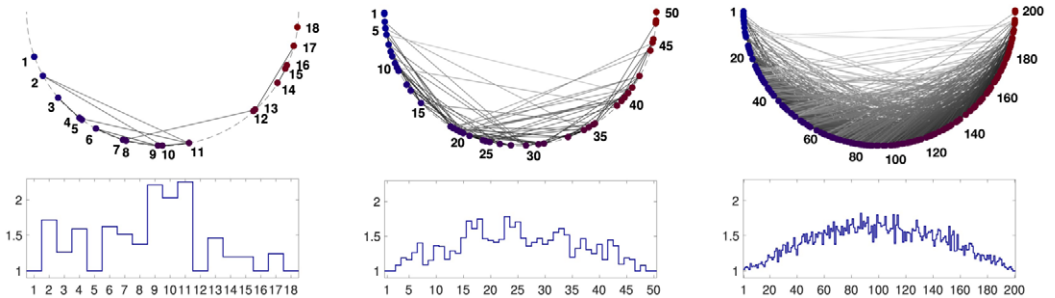


FIGURE 4.—Three realizations of networks formed according to the location model described in Example 3 and their corresponding equilibria (for payoff as in (3) with  $\alpha = 3$  and  $\theta = 1$ ). The line along which agents are located is represented as a semicircle for simplicity of visualization. The color of the nodes is associated to the agent's location along the line (blue being one extreme and red the other extreme). Edges between agents that are further apart are in lighter color.

between equilibria of different sampled network games. For instance, in Example 2, red agents tend to exert lower efforts at equilibrium than blue agents, while in Example 3, agents at more central locations exert higher efforts at equilibrium. This trend becomes sharper and more deterministic as the population size increases. It is then natural to ask whether equilibria of sampled network games converge with increasing population size and whether we can provide a characterization of their limit. Graphon games provide the answer.

#### 4. SAMPLED NETWORK GAMES: CONVERGENCE

##### 4.1. Network Games Are Graphon Games

In network games Nash equilibria are vectors of  $\mathbb{R}^N$ , while in graphon games they are functions of  $L^2([0, 1])$ . To compare these two objects, we define a one-to-one correspondence between vector and functions using a partition  $\mathcal{U}^{[N]}$  of  $[0, 1]$  in equal intervals  $\mathcal{U}_k^{[N]}$  for  $k \in \{1, \dots, N-1\}$ , each of length  $1/N$ . Using this partition, we pair each agent  $i$  in the finite network with the interval  $\mathcal{U}_i^{[N]}$ . We can then define: (i) the *step function equilibrium*  $\bar{s}_{[N]}(x) \in L^2([0, 1])$  corresponding to any equilibrium  $\bar{s}_{[N]} \in \mathbb{R}^N$  as

$$\bar{s}_{[N]}(x) := \bar{s}_{[N]}^i, \quad \forall x \in \mathcal{U}_i^{[N]}, \quad \forall i \in \{1, \dots, N\},$$

and (ii) the *step function graphon*  $W^{[N]}$  corresponding to any graph  $A^{[N]} \in \mathbb{R}^{N \times N}$  as

$$W^{[N]}(x, y) := A_{ij}^{[N]}, \quad \forall (x, y) \in \mathcal{U}_i^{[N]} \times \mathcal{U}_j^{[N]}, \quad \forall i, j \in \{1, \dots, N\}. \quad (6)$$

We show in Lemma 1 in Appendix A that a vector  $\bar{s}_{[N]} \in \mathbb{R}^N$  is a Nash equilibrium of  $\mathcal{G}^{[N]}(\{\mathcal{S}_{j=1}^N, U, \{\theta^i\}_{j=1}^N, A^{[N]}\})$  if and only if the corresponding step function equilibrium  $\bar{s}_{[N]}(x) \in L^2([0, 1])$  is a Nash equilibrium of the graphon game  $\mathcal{G}(S^{[N]}, U, \theta^{[N]}, W^{[N]})$  with payoff function as in (2), set valued function  $S^{[N]}(x) := S^i$  for all  $x \in \mathcal{U}_i^{[N]}$ , parameter function  $\theta^{[N]}(x) := \theta^i$  for all  $x \in \mathcal{U}_i^{[N]}$ , and step function graphon  $W^{[N]}$  corresponding to  $A^{[N]}$ . Using this equivalence, we can compare equilibria of sampled network games and graphon games in  $L^2([0, 1])$  by comparing equilibria of two graphon games: the original graphon game and the one constructed from the sampled network game as discussed above.

##### 4.2. Equilibria in Sampled Network Games

To derive a bound on the distance between equilibria of sampled network and graphon games, we impose the following additional regularity condition.<sup>10</sup>

**ASSUMPTION 4**—Lipschitz Continuity: *The graphon  $W$  and the parameter function  $\theta$  are piecewise Lipschitz over the same partition of  $[0, 1]$ .<sup>11</sup> Moreover, there exists  $\theta_{\max}$  such that  $\|\theta(x)\| \leq \theta_{\max}$  for all  $x \in [0, 1]$ .*

<sup>10</sup>A more general result that requires only Assumption 3 is given in Parise and Ozdaglar (2018). We here focus on Lipschitz graphons to obtain simpler bounds.

<sup>11</sup>Rigorously, we assume that there exist a constant  $L > 0$  and a sequence of non-overlapping intervals  $\mathcal{I}_k$ ,  $k \in \{1, \dots, \Omega + 1\}$ , such that for any pairs  $(x, y), (x', y') \in \mathcal{I}_{k_1} \times \mathcal{I}_{k_2}$ , we have that  $|W(x, y) - W(x', y')| \leq L(|x - x'| + |y - y'|)$  and  $|\theta(x) - \theta(x')| \leq L|x - x'|$  for any  $x, x' \in \mathcal{I}_k$ .

Assumption 4 is satisfied by all graphons typically used in practice. For example, both the minmax graphon and any stochastic block model graphon satisfy this assumption.

Since networks are sampled randomly from the graphon, our statement on convergence of equilibria of sampled network games to equilibria of the corresponding graphon game holds in probability.

**THEOREM 2—Convergence of Equilibria:** *Consider a graphon game  $\mathcal{G}$  in which each player has homogeneous strategy sets, that is,  $S(x) = S$  for all  $x \in [0, 1]$ . Suppose that  $\mathcal{G}$  satisfies Assumptions 1, 2, 3, 4, and let  $\bar{s}$  be its unique Nash equilibrium. Let  $\bar{s}_W^{[N]}$  be an arbitrary step function equilibrium of the sampled network game  $\mathcal{G}^{[N]}$ , as introduced in Section 3.2. Then*

1.  $\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2} \rightarrow 0$  almost surely when  $N \rightarrow \infty$ ;
2. for any  $0 < \delta \leq e^{-1}$ , with probability at least  $1 - \frac{2\delta}{N}$ , it holds that

$$\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2} = \mathcal{O}\left(\left(\frac{\log(N^2/\delta)}{N}\right)^{\frac{1}{4}}\right).$$

**REMARK 4:** Statement 1 in Theorem 2 proves almost sure convergence of equilibria in sampled network games to the equilibrium of the limiting graphon game. Statement 2 further specifies the convergence rate and can be equivalently reformulated as a bound on the probability of graphon and sampled network equilibria being more than  $\epsilon$  apart (i.e., for any  $\epsilon > 0$ ,  $\exists C > 0$  s.t. for  $N$  large enough  $\mathbb{P}[\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2} > \epsilon] \leq \frac{2N}{\exp(\frac{\epsilon^4 N}{C})}$ ).

In many practical contexts, it might also be of interest to quantify the distance between the equilibria of two network games sampled from the same graphon. Such a result can be used to judge the robustness of the equilibrium outcome to stochastic variations in the realized links or in the number of players. Theorem 2 can be used to obtain such a bound by triangular inequality. Finally, we note that Theorem 2 bounds the distance of the equilibria of the sampled network game to the graphon equilibrium in  $\|\cdot\|_{L^2}$ . This does not directly imply that playing the graphon equilibrium strategy in the sampled network game is an (approximate) Nash equilibrium; we show that this is the case under additional regularity assumptions in Lemma 15 in Appendix D of the Supplemental Material.

## 5. EXTENSIONS

### 5.1. Sublinear Network Growth

The sampling procedure given in Definition 3 generates dense networks, that is, networks in which the number of neighbors per agent grows as  $N$  (thus implying that the number of edges grows roughly as the square of the number of nodes). Our theory can be generalized to a class of sparser networks for which the number of neighbors per agent grows *sublinearly* with  $N$  so that  $\frac{\sqrt{\#\text{edges}}}{\#\text{nodes}} \rightarrow 0$ . To this end, we introduce a sparsity parameter  $\kappa_N$  and consider the following (generalized) procedure to sample simple networks from a graphon; see, for example, Borgs, Chayes, Cohn, and Zhao (2019).

**DEFINITION 5:** Given any graphon  $W$ , a sequence  $\{\kappa_N\}_{N=1}^\infty$  with  $0 < \kappa_N \leq 1$ , and any desired number  $N$  of nodes, a (generalized) sampled network can be constructed by uniformly and independently sampling  $N$  points  $\{t^i\}_{i=1}^N$  from  $[0, 1]$  and constructing the (generalized)

0-1 adjacency matrix  $A_s^{[N]}$  as the adjacency matrix corresponding to a simple network obtained by randomly connecting nodes  $i, j \in [1, N]$  with Bernoulli probability

$$\kappa_N W(t^i, t^j).$$

REMARK 5: Definition 3 is a special case of Definition 5 obtained by setting  $\kappa_N = 1$ . It is easy to see that the expected number of neighbors in  $A_s^{[N]}$  is of order  $\kappa_N N$ . Hence, for these sampled networks,  $\frac{\sqrt{\# \text{ edges}}}{\# \text{ nodes}} \approx \sqrt{\kappa_N}$  converges to zero if  $\kappa_N \rightarrow 0$ . In the following, we will require that  $\lim_{N \rightarrow \infty} \frac{\log(N)}{N \kappa_N} = 0$ . Hence, this generalized framework allows the number of neighbors to grow sublinearly in  $N$  but still requires a growth faster than  $\log(N)$ . This is a necessary condition for being able to use concentration inequalities guaranteeing accumulation in the neighbors aggregate.

The new Definition 5 affects only how a sampled network is generated from the graphon but has no repercussions on the limit for infinite number of agents. In other words, the infinite population game is exactly the same graphon game described in Section 2 and the same theorem on existence and uniqueness continues to hold. Instead, we need to modify the definition of local aggregate in a sampled network game to account for the fact that the number of neighbors may now be sublinear. In fact, if we were to use as aggregate the quantity  $z^i(s) = \frac{1}{N} \sum_{j=1}^N [A_s^{[N]}]_{ij} s^j$  as introduced in Section 3.2, then we may have that  $z^i(s) \rightarrow 0$  as  $N$  grows larger, leading to vanishing network effects. To overcome this issue, we need to scale the local aggregate  $\sum_{j=1}^N [A_s^{[N]}]_{ij} s^j$  by the expected order of neighbors which, according to Definition 5, is  $\kappa_N N$  instead of  $N$ . Overall, we can define a sampled network game exactly as in Section 3.2, but using as aggregate

$$z_k^i(s) = \frac{1}{\kappa_N N} \sum_{j=1}^N [A_s^{[N]}]_{ij} s^j. \tag{7}$$

Our main convergence result can be extended to this broader class of sampled networks; the formal statements and proofs can be found in Appendix A.2.

### 5.2. Average Instead of Aggregate

In the results derived so far, we defined the *local aggregate* as  $z^i(s) = \frac{1}{N} \sum_{j=1}^N [A_s^{[N]}]_{ij} s^j$ , that is, the sum of neighbors' strategies normalized by the population size. While this model is used widely in both theoretical and empirical works, for some applications, an alternative model using a *local average* obtained by normalizing the network effect by the agent's degree may be favored (see Patacchini and Zenou (2012) and Ushchev and Zenou (2020)). This corresponds to the choice

$$z_d^i(s) := \frac{\sum_{j=1}^N [A_W^{[N]}]_{ij} s^j}{\sum_{j=1}^N [A_W^{[N]}]_{ij}}.$$

Our results can be extended to this setting. The first step is to define the *local average* for a continuum of agents as  $z_d(x | s) := \frac{\int_0^1 W(x, y) s(y) dy}{\int_0^1 W(x, y) dy}$ . For this quantity to be well defined, we assume from here on that  $\int_0^1 W(x, y) dy \geq d_{\min} > 0$  for all  $x \in [0, 1]$ . This definition of

local average leads to a graphon game as defined in Section 2, played over the *normalized graphon*  $W_d(x, y) := \frac{W(x, y)}{\int_0^1 W(x, y) dy}$ . One can then define the associated *normalized graphon operator*  $\mathbb{W}_d$  as the operator  $\mathbb{W}_d : L^2([0, 1]) \mapsto L^2([0, 1])$  given by  $f(x) \mapsto (\mathbb{W}_d f)(x) = \frac{\int_0^1 W(x, y) f(y) dy}{\int_0^1 W(x, y) dy}$ . Under the assumption that  $\int_0^1 W(x, y) dy \geq d_{\min}$ , we show in Appendix D.2 of the Supplemental Material that all the results derived in Section 2 about existence and uniqueness of the graphon equilibrium continue to hold.<sup>12</sup> For example, uniqueness holds if  $\frac{\ell_U}{\alpha_U} \cdot \frac{\lambda_{\max}(\mathbb{W})}{d_{\min}} < 1$ . For the convergence result in Theorem 2, the key step is to show that the distance between the normalized operators  $\mathbb{W}_d$  and  $\mathbb{W}_d^{[N]}$  (corresponding to the graphon  $W$  and sampled network  $A_W^{[N]}$ , respectively) converges to zero with high probability. We provide a proof of this fact for Lipschitz continuous graphons in Appendix D.2 of the Supplemental Material.

### 5.3. Directed Networks

So far, we assumed that the graphon is a symmetric function and we thus generated undirected sampled networks. The results of Section 2 continue to hold even when the generating graphon is not symmetric, with the only caveat that the eigenvalues of the corresponding operator are not necessarily real, hence one needs to use  $\|\mathbb{W}\|$  instead of  $\lambda_{\max}(W)$ . The only place where symmetry is used in Theorem 2 is to prove that the matrix  $A_s^{[N]}$  accumulates around its expectation  $A_w^{[N]}$ . To prove this fact, we used a matrix concentration result from Chung and Radcliffe (2011) which holds for symmetric matrices. However, a similar result can be obtained for the directed case as well (see Lemma 9 in Appendix D.3 of the Supplemental Material). Using such a result, one can obtain convergence also for directed networks.

## 6. THEORY OF TARGETED INTERVENTIONS

The convergence results derived so far suggest a new approach for designing near optimal and computationally tractable interventions in sampled network games, by exploiting knowledge of the graphon game limit. To illustrate this point, we assume that a central planner aims at maximizing the *average social welfare* (defined as the average of the agents' payoffs at equilibrium) through targeted interventions that modify the payoff of each agent  $i$  by changing the parameter  $\theta^i$  to  $\theta^i + \hat{\theta}^i$  in (1), leading to the modified payoff function

$$U(s^i, z^i, \theta^i + \hat{\theta}^i). \tag{8}$$

The central planner is subject to the convex budget constraint  $\sum_{i=1}^N (\hat{\theta}^i)^2 \leq CN$ , capturing the fact that interventions are increasingly costly, leading to the question of how to optimally allocate the available budget among the agents to maximize welfare.<sup>13</sup> Mathe-

<sup>12</sup>Since  $W_d$  is not symmetric, results need to be stated in terms of  $\|\mathbb{W}_d\|$  instead of  $\lambda_{\max}(\mathbb{W})$ . Note, however, that the bound  $\|\mathbb{W}_d\| \leq \frac{\lambda_{\max}(\mathbb{W})}{d_{\min}}$  holds (see Appendix D.2 of the Supplemental Material).

<sup>13</sup>Note that we allow the budget to scale with the population size  $N$  to model the fact that networks with more agents are allocated a proportionally higher budget.

matically, the central planner aims at solving the following optimization problem:

$$\begin{aligned}
 T_{\text{opt}}^{[N]} &:= \max_{\hat{\theta}^{[N]} \in \mathbb{R}^N} T^{[N]}(\hat{\theta}^{[N]}) = \frac{1}{N} \sum_{i=1}^N U(\bar{s}_i^{[N]}, \bar{z}_i^{[N]}, \theta_i^{[N]} + \hat{\theta}_i^{[N]}) \\
 \text{s.t. } \bar{s}^{[N]} &= \text{Nash equilibrium of } \mathcal{G}^{[N]}(\mathcal{S}, U, \theta^{[N]} + \hat{\theta}^{[N]}, A^{[N]}), \quad \bar{z}^{[N]} = A^{[N]}\bar{s}^{[N]} \quad (9) \\
 \frac{1}{N} \|\hat{\theta}^{[N]}\|^2 &\leq C,
 \end{aligned}$$

where we added the apex  $^{[N]}$  to stress the dependence on the population size.

### 6.1. Graphon Intervention

Problem (9) scales with the size of the network and becomes computationally challenging for networks with more than a few hundreds of agents. We therefore suggest an alternative approach for sampled network games (i.e., for cases when  $A^{[N]} = A_W^{[N]}$  is a realization from an underlying graphon  $W$  and  $\theta_i^{[N]} = \theta(t^i)$ ) based on the corresponding optimization problem in the graphon space

$$\begin{aligned}
 \theta^* \in \arg \max_{\hat{\theta} \in L^2([0,1])} T(\hat{\theta}) &= \int_0^1 U(\bar{s}_{\hat{\theta}}(x), \bar{z}_{\hat{\theta}}(x), \theta(x) + \hat{\theta}(x)) dx \\
 \text{s.t. } \bar{s}_{\hat{\theta}} &= \text{Nash equilibrium of } \mathcal{G}(\mathcal{S}, U, \theta + \hat{\theta}, W), \quad \bar{z}_{\hat{\theta}}(x) = \mathbb{W}\bar{s}_{\hat{\theta}}(x) \quad (10) \\
 \|\hat{\theta}\|_{L^2}^2 &\leq C.
 \end{aligned}$$

The solution  $\theta^*$  to Problem (10) is a function specifying the optimal allocation for the infinite population graphon game. The proposed intervention for finite sampled network games allocates to any sampled agent  $i$  (of type  $t^i$ ) an intervention proportional to  $\theta^*(t^i)$ , that is,

$$[\hat{\theta}_{\text{graphon}}^{[N]}]_i = \frac{\theta^*(t^i)}{\eta^{[N]}},$$

where  $\eta^{[N]}$  is a normalization to guarantee that the budget constraint is met with equality (i.e.,  $\frac{1}{N} \|\hat{\theta}_{\text{graphon}}^{[N]}\|^2 = C$ ).

**THEOREM 3:** *Consider a sampled network game  $\mathcal{G}^{[N]}$  as described in Definition 3. Suppose that Assumptions 1, 2, 3, and 4 hold, that  $U(s, z, \theta)$  is jointly Lipschitz in  $[s, z, \theta]$ , and that  $\theta^*$  solution to (10) is piecewise Lipschitz and bounded. Then, for any  $0 < \delta \leq e^{-1}$ , with probability at least  $1 - \frac{2\delta}{N}$ ,*

$$T_{\text{opt}}^{[N]} - T^{[N]}(\hat{\theta}_{\text{graphon}}^{[N]}) \leq \mathcal{O}\left(\left(\frac{\log(N^2/\delta)}{N}\right)^{\frac{1}{4}}\right).$$

**REMARK 6:** Following Remark 4, Theorem 3 can be equivalently reformulated as a bound on the probability of  $\hat{\theta}_{\text{graphon}}^{[N]}$  being more than  $\epsilon$  suboptimal with respect to the optimal solution to Problem (10). In other words, Theorem 3 guarantees that, for any



$\epsilon > 0$ , there exists  $C > 0$  s.t. for  $N$  large enough,  $\mathbb{P}[T_{\text{opt}}^{[N]} - T^{[N]}(\hat{\theta}_{\text{graphon}}^{[N]}) > \epsilon] \leq \frac{2N}{\exp(\frac{\epsilon^4 N}{C})}$ . Such a bound vanishes with  $N$ , leading to asymptotically optimal performance.

The advantage of the graphon approach is that, to solve Problem (10), the central planner only needs information about the stochastic network formation model (the graphon  $W$ ) and not about the exact network realization ( $A_W^{[N]}$ ), thus overcoming the need for acquisition of exact network information. Regarding computational tractability, we remark that, while solving Problem (10) in general may be challenging, in many settings the stochastic network formation model is a low dimensional object. In this case, Problem (10) can be solved efficiently.

### 6.2. Tractability of Problem (10)

To obtain computational tractability, we restrict our attention to graphons in which only a finite number  $R$  of eigenvalues  $\{\lambda_r\}_{r=1}^R$  are different from zero (i.e., finite-rank graphons). For this class of graphons and for different structures of payoff functions, a solution to Problem (10) can be obtained by solving a finite dimensional optimization problem whose dimension is typically much smaller than the population size  $N$ . Before proving this fact, we provide some examples of finite-rank graphons, illustrating the fact that, while a refinement, finite-rank graphons are general enough to nest a large number of random graph models.

EXAMPLE 4—Community structure: Consider a generalization of the community model with  $K$  communities introduced in Example 2, in which we allow agents across different communities to interact with different probabilities. Specifically, let  $Q \in [0, 1]^{K \times K}$  be a symmetric matrix whose element in position  $(k, l)$  denotes the probability that agents of community  $k$  and  $l$  are interacting (the graphon in Example 2 corresponds to the special case  $Q = [g_{\text{in}}I_K + g_{\text{out}}(\mathbb{1}_K\mathbb{1}_K^\top - I_K)]$ ). Let  $\mathcal{C}_k$  be the subset of  $[0, 1]$  associated with community  $k$ , with  $|\mathcal{C}_k| = \pi_k$  and  $\sum_k \pi_k = 1$ , and construct the stochastic block model graphon

$$W_{\text{SBM}}(x, y) = Q_{ij} \quad \text{for all } x \in \mathcal{C}_i, y \in \mathcal{C}_j.$$

The stochastic block model graphon is finite rank with rank equal to the number of communities. In fact, as shown, for example, in Avella-Medina et al. (2018), eigenvalues and eigenfunctions of the corresponding graphon operator can be easily computed by considering the auxiliary matrix

$$E := Q\Pi \in \mathbb{R}^{K \times K}, \tag{11}$$

where  $\Pi$  is a diagonal matrix whose diagonal elements correspond to the community sizes, that is,  $\Pi_{kk} = \pi_k$  for all  $k$ . Lemma 10 provided in the Supplemental Material shows that  $\mathbb{W}_{\text{SBM}}$  and  $E$  have the same eigenvalues and the eigenfunctions of  $\mathbb{W}_{\text{SBM}}$  are piecewise constant over the partition  $\{\mathcal{C}_k\}_{k=1}^K$ , with constant value in each community  $\mathcal{C}_k$  given by the  $k$ th element of the corresponding eigenvector of  $E$ .

The example above considered a graphon with a finite number of communities. A graphon can, however, have finite rank even when there is a continuum of types. For example, randomly grown ranked attachment graph sequences as described in Borgs, Chayes, Lovász, Sós, and Vesztegombi (2011) converge to a graphon that has rank 2 (see Avella-Medina et al. (2018, Section 4.2)), and uniform attachment graph sequences

converge to a graphon which can be very well approximated with a rank 5 graphon. The next example provides a simple rank 1 graphon.

**EXAMPLE 5—Rank 1 Graphon:** Consider a graphon  $W_v(x, y) = v(x)v(y)$  for some nonnegative function  $v \in L^2([0, 1])$  s.t.  $W_v(x, y) \in [0, 1]$  for all  $x, y$ . This graphon has degree  $d(x) \propto v(x)$  and is rank 1 with eigenfunction  $v(x)$  and eigenvalue  $\lambda = \|v\|_{L^2}^2$ . For example, by choosing  $v = \frac{1}{\sqrt{2(x+0.5)^\gamma}}$  with  $\gamma \in (0, 1/2)$ , one obtains a rank 1 graphon  $W_\gamma$  with power law degree distribution (as often observed in social networks).

6.2.1. *Linear Quadratic Setting*

As a first application, we consider network games with scalar nonnegative strategies and linear quadratic payoff

$$U(s^i, z^i, \theta^i) = -\frac{1}{2}(s^i)^2 + (\alpha z^i + \theta^i)s^i. \tag{12}$$

For simplicity, we focus on games with strategic complementarities (i.e., with  $\alpha > 0$ ). Following Galeotti, Golub, and Goyal (2020), we assume that the central planner can directly modify the standalone marginal return for an arbitrary agent  $i$  from  $\theta^i$  to  $\theta^i + \hat{\theta}^i$ , leading to the modified payoff function

$$U(s^i, z^i, \theta^i + \hat{\theta}^i) = -\frac{1}{2}(s^i)^2 + s^i[\alpha z^i + \theta^i + \hat{\theta}^i]. \tag{13}$$

While Galeotti, Golub, and Goyal (2020) focused on design of interventions when the central planner has perfect knowledge of the network of interactions, we here consider a setting in which the central planner only knows the graphon model and show how Problem (10) can be solved efficiently leading to asymptotically optimal interventions without the need for exact network data.

**PROPOSITION 1:** *Suppose that  $0 < \alpha < \frac{1}{\lambda_{\max}(\mathbb{W})}$  and assume that  $\mathbb{W}$  has finite rank  $R < \infty$ , let  $\mathcal{K}$  be the kernel of  $\mathbb{W}$ , and  $\{\psi_r\}_{r=1}^R$  be an orthonormal basis of  $\mathcal{K}^\perp$  composed of eigenfunctions of  $\mathbb{W}$  corresponding to the eigenvalues  $\{\lambda_r\}_{r=1}^R$ . Set  $b_r = \langle \theta, \psi_r \rangle$  for all  $r = 1, \dots, R$  and let  $b_0\psi_0$  be the projection of  $\theta$  in  $\mathcal{K}$ , with  $\|\psi_0\|_{L^2} = 1$ . Set  $\lambda_0 = 0$ . A maximizer of (10) can be computed as  $\theta^* = \sum_{r=0}^R \hat{b}_r^* \psi_r$ , where  $\{\hat{b}_r^*\}_{r=0}^R$  solves<sup>14</sup>*

$$\begin{aligned} \max_{\{\hat{b}_0, \dots, \hat{b}_R\}} & \frac{1}{2} \sum_{r=0}^R \frac{(b_r + \hat{b}_r)^2}{(1 - \alpha\lambda_r)^2} \\ \text{s.t.} & \sum_{r=0}^R \hat{b}_r^2 \leq C. \end{aligned} \tag{14}$$

---

<sup>14</sup>We stress that under the assumption of this proposition, both Problem (10) and Problem (9) can be reformulated as semi definite programs with two variables and one inequality constraint; see Boyd and Vandenberghe (2004, Appendix B.1). The main computational difference is that while the inequality constraint associated with Problem (9) involves a matrix of dimension  $N + 1$ , the one associated with Problem (10) is limited to dimension  $R + 2$ .

EXAMPLE 6: Consider again the community structure discussed in Example 4, with payoff functions as in (13). First note that if we assume that all agents within the same community have the same standalone marginal return (i.e.,  $\theta(x) = \theta_k^{\text{com}}$  for all  $x \in \mathcal{C}_k$ ), then it is immediate to see that at the graphon equilibrium, each agent belonging to the same community has the same strategy  $\bar{s}_k^{\text{com}}$  and the vector of such strategies  $\bar{s}^{\text{com}} \in \mathbb{R}^K$  satisfies the relation

$$\bar{s}^{\text{com}} = (I_K - \alpha E)^{-1} \theta^{\text{com}}. \tag{15}$$

Turning to the optimal intervention problem (10), we note that since  $\theta(x) = \theta_k^{\text{com}}$  for all  $x \in \mathcal{C}_k$ ,  $\theta$  can be written as a linear combination of the eigenfunctions of  $\mathbb{W}$  (i.e.,  $b_0$  as defined in Proposition 1 is zero). One can then conclude that  $\hat{b}_0^* = 0$  and  $\theta^* = \sum_{r=1}^R \hat{b}_r^* \psi_r$ , hence the optimal intervention is constant within each community and can be computed by solving an optimization problem that has dimension equal to the number of communities instead of agents.

### 6.2.2. Beyond Linear Quadratic Setting

Games with linear quadratic payoffs (or more generally, affine best responses) as discussed in the previous section cover a large number of applications. Nonetheless, there may be settings that would be better modeled using a nonlinear and non-monotone best response. We illustrate how our framework can be applied to such a case by using a game in which agents have quadratic nonlinear payoffs

$$U(s^i, z^i, \theta^i + \hat{\theta}^i) = -\frac{1}{2}(s^i)^2 + s^i[\alpha z^i(1 - z^i) + \theta^i]. \tag{16}$$

As shown in Parise and Ozdaglar (2019), this payoff leads to a non-affine and non-monotone best response function, modeling, for instance, race and tournaments. We next show that, nonetheless, it is possible to reformulate Problem (10) as a lower dimensional optimization problem. The key insight on why this is possible is that, since the graphon is low rank, the aggregate  $z(x)$  in the graphon game is a low dimensional object that can be described as a linear combination of the eigenfunctions of the graphon. Finding the Nash equilibrium and the optimal intervention then reduces to finding the coefficients of such a linear combination.

For simplicity, in the next proposition we illustrate this argument for rank 1 graphons as described in Example 5 and for games with quadratic dependence on the local aggregate as given in (16), but similar arguments could be generalized to graphons with any finite rank and payoffs with higher order dependence on  $z$ .

PROPOSITION 2: Consider a graphon game  $\mathcal{G}([0, 1], U, \theta + \hat{\theta}, W_v)$  with payoff functions as in (16) and a rank 1 graphon  $\mathbb{W}_v$  as in Example 5, with corresponding eigenvalue  $\lambda$ . Suppose that  $\alpha \in [0, \min(1, 1/\lambda)]$ ,  $\theta(x) \in [0, 3/4]$  for all  $x \in [0, 1]$ , and  $\langle \theta, v \rangle > 0$ . For small budget, a solution to Problem (10) is given by  $\hat{\theta}^* = \hat{b}_1^* \psi_1 + \hat{b}_2^* \psi_2 + \hat{b}_0^* \psi_0$ , where the functions  $\psi_i$  form an orthonormal basis for the space generated by  $v, v^2$ , and  $\theta$  (computed as in Lemma 4 in

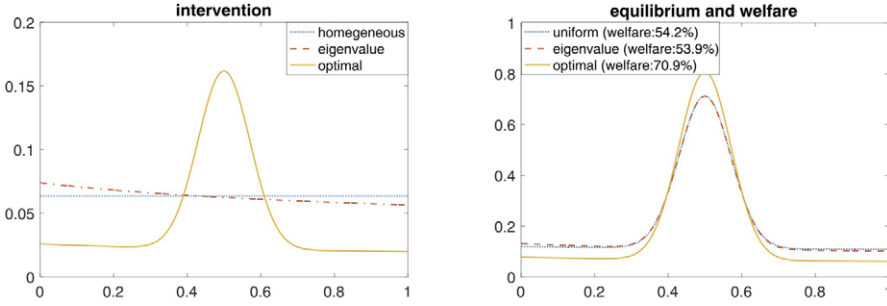


FIGURE 5.—Intervention design for a graphon game  $\mathcal{G}([0, 1], U, \theta + \hat{\theta}, W_\gamma)$  with payoff function as in (16) (for  $\alpha = 0.5$ ), power law rank 1 graphon  $W_\gamma$  as described in Example 5 (with  $\gamma = 0.25$ ), and  $\theta(x) = 0.6e^{-\frac{(x-0.5)^2}{0.1}}$ . Left: Three possible interventions  $\hat{\theta}$  satisfying the budget constraint  $\|\hat{\theta}\|_{L^2} \leq 4$ . Right: The corresponding equilibria with welfare improvement noted in the legend (expressed as percentage of improvement with respect to the no-intervention equilibrium given by  $\hat{\theta} = 0$ ).

the Appendix) and  $\hat{b}_0^*, \hat{b}_1^*, \hat{b}_2^*$  solve the following optimization problem:

$$\begin{aligned} \max_{\hat{b}_1, \hat{b}_2, \hat{b}_0} \quad & s_1^2 + (-\alpha s_1^2 \lambda p_2 + b_2 + \hat{b}_2)^2 + (b_0 + \hat{b}_0)^2 \\ \text{s.t.} \quad & s_1 = \frac{-(1 - \alpha\lambda) + \sqrt{(1 - \alpha\lambda)^2 + 4(\alpha\lambda p_1)(b_1 + \hat{b}_1)}}{2\alpha\lambda p_1}, \\ & \hat{b}_0^2 + \hat{b}_1^2 + \hat{b}_2^2 = C, \end{aligned} \quad (17)$$

where  $p_i = \langle v^2, \psi_i \rangle$  and  $b_i = \langle \theta, \psi_i \rangle$ .

Intuitively, Proposition 2 guarantees that the optimal intervention is a combination of the  $\psi_i$  functions with coefficients given by the optimizer of (17). Figure 5 shows a comparison of the graphon equilibrium obtained under the optimal intervention (computed as in Proposition 2), a uniform intervention, and an intervention proportional to the eigenfunction for a game with rank 1 graphon constructed as in Example 5. It is important to remark that in this example, the best response is nonlinear (not affine and not monotone) and the network generating model is non-standard (having a continuum of heterogeneous agents instead of a more typical community structure with finite types), yet the proposed framework still leads to a tractable analysis and an intervention that outperforms classic ones, illustrating the usefulness of the proposed approach beyond network games typically studied in the literature.

## 7. AN ILLUSTRATIVE CASE STUDY

To illustrate the differences (in terms of information, computation, and optimality) between the intervention procedure described in Section 6 and a more direct approach based on detailed network information, we construct a simulated data set of 80 network games (which, e.g., could model interactions among the inhabitants of 80 different rural villages deciding the level of investment in a microfinance program). Each agent  $i$  has payoff

$$U(s^i, z_\kappa^i(s), \theta^i) = -\frac{1}{2}(s^i)^2 + s^i[\alpha z_\kappa^i(s) + \theta^i],$$

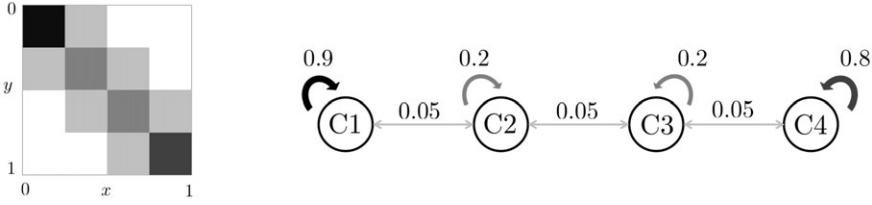


FIGURE 6.—The stochastic block model graphon used to generate sampled networks in the case study of Section 7. The figure on the right is a visualization of the interactions among the four communities (the width of the arrows is larger the stronger the interaction), the figure on the left shows the corresponding graphon (with white representing 0 and black representing 1). Note that communities 1 and 4 are tightly intra-connected, while communities 2 and 3 are less intra-connected. Communities are only slightly inter-connected with neighboring communities. For sampling, we used a sparsity parameter  $\kappa_N = 40/N^{0.8}$  (see Section 5.1).

where the parameter  $\alpha$  is the same for all agents,  $\theta^i$  is agent specific, and  $\kappa$  is a sparsity parameter as introduced in Section 5.1. We assume that agents in each network are equally likely to belong to one of four different communities (e.g., four different caste in the case of rural villages) and that the probability of agents interacting depends on community identity according to the community structure illustrated in Figure 6. Finally, agents belonging to the same community have the same parameter  $\theta^i$ , which we denote by  $\theta_h^{\text{com}}$  for community  $h = 1, \dots, 4$ .

Our main interest is to understand how a central planner can allocate a limited budget in each sampled network (village) to maximize agents’ welfare, by designing interventions as discussed in Section 6.

### 7.1. Data Acquisition

Our aim is to simulate the procedure that the central planner would have to follow in a field experiment. To this end, from here on we are going to assume that the central planner *does not* have access to the information detailed above, but instead needs to rely on surveys to reconstruct agents’ attributes and interactions. Regarding the latter, we are going to assume that the central planner can use two different types of relational surveys:

- Detailed Relational Data: The central planner is able to ask to *each agent* in each network (village) *the exact identity of each of his neighbors*.
- Aggregated Relational Data: The central planner is able to ask to a *subset of the agents* in some of the networks (villages) *how many of his neighbors belong to each community*.

As argued in Breza et al. (2020), aggregated relational data of the second type are much easier to obtain in the field than the information required by the detailed relational survey. Furthermore, the aggregated information required by the second type of survey can allow data acquisition in settings in which detailed information is not possible because of proprietary data or privacy concerns.

While relational data are typically hard to obtain, it is instead common in empirical studies to collect detailed agent-level information through an exhaustive census; see, for example, Banerjee et al. (2013). We here assume that the central planner can perform a census of all the agents asking about agent-level information such as: (i) agent type (e.g., the community to which the agent belongs), (ii) equilibrium strategy before the intervention (e.g., the current level of investment in the microfinance program), and (iii) the standalone marginal return  $\theta^i$ .

Importantly, the central planner has no information about the strength of peer effects,  $\alpha$ , or about the parameters of the network formation model.<sup>15</sup> These parameters would not be available in field work and therefore need to be estimated from the relational survey and census data described above.

### 7.2. Intervention Design and Estimation Based on Different Relational Data

If the central planner has access to the information contained in the census and in the *detailed relational data*, then he can reconstruct, for each village: (i) the exact network of interactions among the agents  $A_s^{[N]}$ , (ii) the vector of parameters  $\theta^{[N]} \in \mathbb{R}^N$ , and (iii) the Nash equilibrium  $\bar{s}^{[N]}$  before the intervention. He can then use this information to infer the unknown normalized network parameter  $\alpha_\kappa := \frac{\alpha}{\kappa_N}$  by performing least squares regression given the Nash equilibrium relation  $\bar{s}^{[N]} = (I_N - \frac{\alpha_\kappa}{N} A_s^{[N]})^{-1} \theta^{[N]}$ , leading to an estimator for  $\alpha_\kappa$  which we denote by  $\hat{\alpha}_\kappa^{\text{DRD}}$ .<sup>16</sup> Using  $A_s^{[N]}$ ,  $\theta^{[N]}$ , and  $\hat{\alpha}_\kappa^{\text{DRD}}$ , the central planner can either solve exactly the optimal intervention problem in (9) (if this is computationally feasible) or otherwise he can use the heuristic suggested in Galeotti, Golub, and Goyal (2020) and allocate the budget according to the dominant eigenvector of  $A_s^{[N]}$ . We refer to these two interventions as *network optimal* and *network heuristic*, respectively.<sup>17</sup>

If instead the central planner has only access to the census and *aggregated relational data*, the previous analysis cannot be performed. Still, we show in Appendix D that aggregated relational data are sufficient to infer the stochastic network model (via maximum likelihood estimation) and the graphon equilibrium before the intervention (by averaging the strategies of agents belonging to the same community).<sup>18</sup> The central planner can then obtain an estimator  $\hat{\alpha}^{\text{ARD}}$  by performing least squares regression on the graphon equilibrium relation  $\hat{s}^{\text{ARD}} = (I_K - \hat{\alpha}_\kappa^{\text{ARD}} \hat{E}_\kappa^{\text{ARD}})^{-1} \theta^{\text{com}}$ , where  $\theta^{\text{com}}$  is the vector of marginal return per community (which can be recovered exactly from the census data) and the superscript ARD denotes estimators computed from aggregated relational data. Based on  $\theta^{\text{com}}$ ,  $\hat{\alpha}_\kappa^{\text{ARD}}$ , and  $\hat{E}_\kappa^{\text{ARD}}$ , the central planner can solve Problem (10) (by equivalently solving Problem (14)) and obtain the optimal graphon intervention. Note that for this case study, Problem (14) is a problem of dimension 4 and outputs the intervention that the central planner should apply in each community. The central planner then knows which intervention to apply to each agent because he collected information about agent's type in the census. We refer to this intervention as *graphon optimal*.

### 7.3. Comparison

Figure 7 illustrates the *network optimal*, *network heuristic*, and *graphon optimal* interventions for two sampled networks of size  $N = 300$  and  $N = 600$ . A first observation is that while the first two interventions are tailored to the specific network realization (and thus prescribe a different intervention to each agent), the graphon intervention prescribes the

<sup>15</sup>We assume that the planner knows the sampled networks are generated from a stochastic block model with four communities, but our result could be extended to cases when the number of communities is unknown.

<sup>16</sup>In our simulated data we assumed no noise, hence this procedure allows the planner to recover  $\alpha_\kappa$  exactly.

<sup>17</sup>The central planner could also employ an in-between strategy by allocating the budget according to the  $r$  dominant eigenvectors for some  $r > 1$ . This strategy still requires the detailed relation data set and will have performances that are in between the *network optimal* and *network heuristic*.

<sup>18</sup>Recall that since the graphon in this case is a stochastic block model, each agent in community  $h$  has the same graphon equilibrium strategy; see (15).

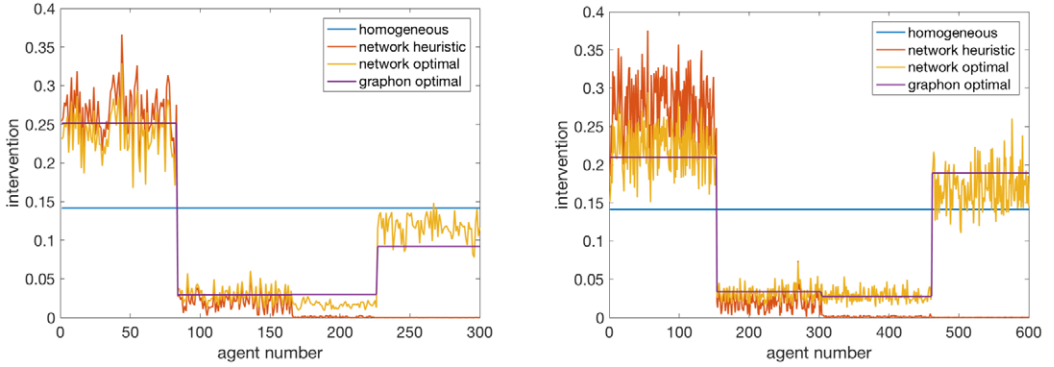


FIGURE 7.—Interventions for a sampled network of size  $N = 300$  (left) and  $N = 600$  (right).

same intervention to each agent belonging to the same community. We next compare the performances of the three interventions in terms of optimality, information, and computation.

1. *Information*: as discussed above, the *network optimal* and *network heuristic* interventions require detailed relational data, while the *graphon optimal* intervention can be computed based solely on aggregated relational data;
2. *Computation*: the *network optimal* intervention requires the solution of Problem (9) whose complexity is polynomial in  $N$ ; the *network heuristic* intervention requires the computation of the dominant eigenvector of  $A_s^{[N]} \in \mathbb{R}^{N \times N}$  which is again polynomial in  $N$ ; the *graphon optimal* intervention requires the solution of Problem (14) which is polynomial in  $K = 4$ .
3. *Optimality*: Table I illustrates the percentage of welfare improvement under the three different policies (averaged over the 80 networks) with respect to the homogeneous intervention that splits the budget equally for all the agents. Different columns represent repetitions of the case study for networks with increasing size.

## 8. CONCLUSION

In this work, we introduced the novel class of graphon games for modeling strategic behavior in infinite populations while accounting for local heterogeneity. We then showed that graphon games can be used to approximate strategic behavior in large but finite sampled network games by interpreting the graphon as a stochastic network formation process. This statistical interpretation of network games allows for the design of simple intervention policies that are computationally tractable and do not require detailed information about the network realization.

We believe that the initial investigation of graphons as a tool to model strategic behavior presented in this work can be extended in a number of different directions. First, in this paper, to guarantee uniqueness of the Nash equilibrium we used Assumption 3, which is formulated in terms of the maximum eigenvalue of the graphon. Previous works showed that alternative conditions for uniqueness can be formulated in finite network games by using conditions involving the maximum degree or the minimum eigenvalue (for games with strategic substitutes). Extending those results to graphon games is an interesting open direction. Similarly, we believe it should be possible to extend our convergence results beyond games with unique equilibria; such analysis requires tools related to set convergence and is left as future work. Second, as an application of our framework,

TABLE I  
COMPARISON OF NETWORK OPTIMAL (NO), GRAPHON OPTIMAL (GO), AND NETWORK HEURISTIC (NH)  
INTERVENTION FOR THE CASE STUDY IN SECTION 7.

	Case Study 1 ( $N = 300$ )	Case Study 2 ( $N = 600$ )	Case Study 3 ( $N = 1200$ )
Avg. Improvement Network Optimal	25.6% ( $\pm 6.6$ ) [15.4; 61.5]%	23.0% ( $\pm 3.6$ ) [19.0; 37.8]%	– [–]
Avg. Improvement Graphon Optimal	22.8% ( $\pm 7.1$ ) [5.6; 59.7]%	20.8% ( $\pm 3.9$ ) [14.8; 35.7]%	20.6% ( $\pm 2.3$ ) [17.6; 31.5]%
Avg. Improvement Network Heuristic	12.6% ( $\pm 15.7$ ) [–12.6; 60.73]%	4.5% ( $\pm 12.3$ ) [–12.2; 33.9]%	2.8% ( $\pm 10.5$ ) [–12.4; 27.3]%
Average Degree	18.5	21.6	24.9

Note: The average improvement is computed as

$$\text{Avg. Improvement} = \frac{1}{80} \sum_{g=1}^{80} \left( \frac{\text{welfare according to NO/NH/GO intervention in network } g}{\text{welfare according to homogeneous intervention in network } g} - 1 \right).$$

One standard deviation is reported in round brackets; minimum and maximum are reported in square brackets. The average degree is shown to illustrate that the graphon optimal intervention is a good approximation in a range of degrees that is realistic (and does not increase too quickly in  $N$  thanks to the sparsity parameter  $\kappa_N$ ). In all the case studies, we used  $C = 0.02N$ ,  $\alpha = 2.65$ ,  $\hat{\theta}^{\text{com}} = [0.1, 0.1, 0.1, 0.25]$ , and aggregated relational surveys completed by 10% of the agents in each network.

we showed how the graphon approach allows the computation of almost optimal targeted interventions, overcoming the computational intractability of approaches based on full network information. We believe that our results can be generalized to other types of interventions, such as selecting the key player as introduced in [Ballester, Calvó-Armengol, and Zenou \(2006\)](#). Third, we here defined graphon games as nonatomic games. It might be interesting to extend this framework to allow for a small number of atomic (major) agents that influence a mass of nonatomic (minor) agents interacting over a graphon, similarly to previous results derived for mean field games in [Nourian and Caines \(2013\)](#). Finally, in our case study and in Appendix C of the Supplemental Material, we hinted at how the graphon game framework could be used to estimate peer effects when information about the realized network is not available. We believe that extending these results would be of practical interest.

## APPENDIX A: OMITTED PROOFS

Throughout the appendix, we consider the general case in which strategies are vectors in  $\mathbb{R}^n$  instead of scalars and the parameter  $\theta$  is a vector of  $\mathbb{R}^m$  instead of a scalar. Consequently, a strategy profile  $s : [0, 1] \rightarrow \mathbb{R}^n$  is a *vector valued function*. In other words,  $s(x) = [s_1(x), \dots, s_n(x)]^\top$  for all  $x \in [0, 1]$ . In the following, we require that any strategy profile is square integrable, that is,  $s(x) \in L^2([0, 1]; \mathbb{R}^n)$ .<sup>19</sup> For any strategy function

<sup>19</sup>This implies that each component is square integrable, that is,  $s_k(x) \in L^2([0, 1])$  for all  $k \in \{1, \dots, n\}$ . For any  $k \in \{1, \dots, n\}$ , it holds  $\|s_k\|_{L^2} = \sqrt{\int_0^1 s_k(x)^2 dx} \leq \sqrt{\int_0^1 \|s(x)\|^2 dx} = \|s\|_{L^2; \mathbb{R}^n}$ .



$s \in L^2([0, 1]; \mathbb{R}^n)$ , the corresponding local aggregate is

$$z(x | s) := \int_0^1 W(x, y)s(y) dy := \begin{bmatrix} \int_0^1 W(x, y)s_1(y) dy \\ \vdots \\ \int_0^1 W(x, y)s_n(y) dy \end{bmatrix} = \begin{bmatrix} (\mathbb{W}s_1)(x) \\ \vdots \\ (\mathbb{W}s_n)(x) \end{bmatrix} =: (\mathbb{W}_n s)(x),$$

where  $\mathbb{W}_n : L^2([0, 1]; \mathbb{R}^n) \rightarrow L^2([0, 1]; \mathbb{R}^n)$  is defined by applying  $\mathbb{W}$  component-wise.<sup>20</sup> We additionally define the operator  $\mathbb{B}_\theta : L^2([0, 1]; \mathbb{R}^n) \rightarrow L^2([0, 1]; \mathbb{R}^n)$  point-wise as follows:

$$(\mathbb{B}_\theta z)(x) := \arg \max_{\bar{s} \in \mathcal{S}(x)} U(\bar{s}, z(x), \theta(x)), \tag{18}$$

where  $z(x)$  is any function of  $L^2([0, 1]; \mathbb{R}^n)$  (i.e., not necessarily  $z(x | s)$ ). In words,  $(\mathbb{B}_\theta z)(x)$  is the best response of agent  $x$  to the *fixed* local aggregate  $z(x)$ . Note that, under Assumption 1, such best response operator is well defined since the maximization problem in (18) has a unique solution. The fact that, under the given assumptions, the codomain of the best response operator  $\mathbb{B}_\theta$  is  $L^2([0, 1]; \mathbb{R}^n)$  will be proven in Appendix D.1 of the Supplemental Material. Overall, a strategy profile  $\bar{s} \in L^2([0, 1]; \mathbb{R}^n)$  is a Nash equilibrium if and only if

$$\bar{s} = \mathbb{B}_\theta \mathbb{W}_n \bar{s}, \tag{19}$$

that is, the function  $\bar{s}$  is a fixed point of the composite operator  $\mathbb{B}_\theta \mathbb{W}_n$ , which we term the *game operator*.

### A.1. Section 2: Omitted Proofs

**PROOF OF THEOREM 1: Existence:** We aim at applying Schauder fixed point theorem Smart (1974, Theorem 4.1.1) to  $\mathbb{B}_\theta \mathbb{W}_n : L_S \rightarrow K := (\mathbb{B}_\theta \mathbb{W}_n(L_S))^{cl}$ , where  $L_S := \{f \in L^2([0, 1]; \mathbb{R}^n) \mid \|f\|_{L^2; \mathbb{R}^n} \leq s_{\max}\}$ . We proceed in four steps. (i) The set  $L_S$  is non-empty, convex, closed and bounded. (ii) In Lemmas 5 and 6 in Appendix D.1 of the Supplemental Material, we prove that both  $\mathbb{W}_n$  and  $\mathbb{B}_\theta$  are continuous operators, hence  $\mathbb{B}_\theta \mathbb{W}_n$  is continuous. (iii) We show  $K \subseteq L_S$ . In fact, since  $\mathbb{W}_n : L_S \rightarrow L^2([0, 1]; \mathbb{R}^n)$  and  $\mathbb{B}_\theta : L^2([0, 1]; \mathbb{R}^n) \rightarrow L_S$ , it holds  $\mathbb{B}_\theta \mathbb{W}_n(L_S) \subseteq L_S$ . Since  $L_S$  is closed,  $K = (\mathbb{B}_\theta \mathbb{W}_n(L_S))^{cl} \subseteq (L_S)^{cl} = L_S$ . (iv) Finally, we show that  $K$  is compact. To this end, note that  $\mathbb{W}_n$  is a compact operator by Lemma 5 in Appendix D.1 of the Supplemental Material and  $L_S$  is bounded, hence  $(\mathbb{W}_n(L_S))^{cl}$  is compact, see Hutson, Pym, and Cloud (2005, Definition 7.2.1). We prove in Lemma 6 in Appendix D.1 that  $\mathbb{B}_\theta$  is Lipschitz (and thus continuous); consequently,  $\mathbb{B}_\theta((\mathbb{W}_n(L_S))^{cl})$  is compact, see Aliprantis and Border (2006, Theorem 2.34). Clearly  $\mathbb{B}_\theta(\mathbb{W}_n(L_S)) \subseteq \mathbb{B}_\theta((\mathbb{W}_n(L_S))^{cl})$  and thus  $K := (\mathbb{B}_\theta(\mathbb{W}_n(L_S)))^{cl} \subseteq (\mathbb{B}_\theta((\mathbb{W}_n(L_S))^{cl}))^{cl} = \mathbb{B}_\theta((\mathbb{W}_n(L_S))^{cl})$ .  $K$  is thus a closed subset of a compact set, which implies that  $K$  is compact, see Aliprantis and Border (2006, p. 40). Schauder fixed point theorem thus guarantees existence of a fixed point.

**Uniqueness:** We show that the game operator is a contraction in the Hilbert space  $L^2([0, 1]; \mathbb{R}^n)$ . The conclusion then follows from Banach fixed point theorem, see Smart

<sup>20</sup>Note that  $s(x) \in L^2([0, 1]; \mathbb{R}^n) \Rightarrow z(x | s) \in L^2([0, 1]; \mathbb{R}^n)$ .

(1974, Theorem 4.3.4). For any  $f, g \in L^2([0, 1]; \mathbb{R}^n)$ ,

$$\begin{aligned} \|\mathbb{B}_\theta \mathbb{W}_n f - \mathbb{B}_\theta \mathbb{W}_n g\|_{L^2; \mathbb{R}^n} &\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n f - \mathbb{W}_n g\|_{L^2; \mathbb{R}^n} = \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n(f - g)\|_{L^2; \mathbb{R}^n} \\ &\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n\| \|f - g\|_{L^2; \mathbb{R}^n} = \frac{\ell_U}{\alpha_U} \lambda_{\max}(\mathbb{W}) \|f - g\|_{L^2; \mathbb{R}^n}, \end{aligned}$$

where we used Lemma 6, given in Appendix D.1, for the first inequality, the fact that  $\mathbb{W}_n$  is linear in the first equality, and the fact that  $\|\mathbb{W}_n\| = \lambda_{\max}(\mathbb{W})$ , as proven in Lemma 5, in the last line. The conclusion follows from Assumption 3. *Q.E.D.*

**PROPOSITION 3—Continuity:** *Suppose that the graphon game  $\mathcal{G}(S, U, \theta, W)$  satisfies Assumptions 1, 2, 3 and let  $\bar{s}$  be its unique Nash equilibrium. Consider a perturbed graphon  $\tilde{W}$ , a perturbed function  $\tilde{\theta}$ , and let  $\tilde{s}$  be any Nash equilibrium of the graphon game  $\mathcal{G}(S, U, \tilde{\theta}, \tilde{W})$ . Then it holds*

$$\|\bar{s} - \tilde{s}\|_{L^2; \mathbb{R}^n} \leq \frac{1/\alpha_U}{1 - \ell_U/\alpha_U \lambda_{\max}(\mathbb{W})} (\ell_U \|\mathbb{W} - \tilde{\mathbb{W}}\| s_{\max} + \ell_\theta \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m}). \quad (20)$$

**PROOF:** To prove that (20) holds, note that  $\bar{s} = \mathbb{B}_\theta \mathbb{W}_n \bar{s}$  and  $\tilde{s} = \mathbb{B}_{\tilde{\theta}} \tilde{\mathbb{W}}_n \tilde{s}$ ; hence,

$$\begin{aligned} \|\bar{s} - \tilde{s}\|_{L^2; \mathbb{R}^n} &= \|\mathbb{B}_\theta \mathbb{W}_n \bar{s} - \mathbb{B}_{\tilde{\theta}} \tilde{\mathbb{W}}_n \tilde{s}\|_{L^2; \mathbb{R}^n} \leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \bar{s} - \tilde{\mathbb{W}}_n \tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m} \\ &\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \bar{s} - \mathbb{W}_n \tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \tilde{s} - \tilde{\mathbb{W}}_n \tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m} \\ &\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n\| \|\bar{s} - \tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n - \tilde{\mathbb{W}}_n\| \|\tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m} \\ &= \frac{\ell_U}{\alpha_U} \lambda_{\max}(\mathbb{W}) \|\bar{s} - \tilde{s}\|_{L^2; \mathbb{R}^n} + \left( \frac{\ell_U}{\alpha_U} \|\mathbb{W} - \tilde{\mathbb{W}}\| \|\tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m} \right), \end{aligned}$$

where we used that  $\mathbb{B}_\theta$  is Lipschitz, as proven in Lemma 6, the fact that  $\|\mathbb{W}_n\| = \lambda_{\max}(\mathbb{W})$ , and the fact that  $\|\mathbb{W}_n - \tilde{\mathbb{W}}_n\| = \|\mathbb{W} - \tilde{\mathbb{W}}\|$ . The conclusion follows from the fact that  $1 - \ell_U/\alpha_U \lambda_{\max}(\mathbb{W}) > 0$  by Assumption 3; hence,

$$\|\bar{s} - \tilde{s}\|_{L^2; \mathbb{R}^n} \leq \frac{1}{1 - \frac{\ell_U}{\alpha_U} \lambda_{\max}(\mathbb{W})} \left( \frac{\ell_U}{\alpha_U} \|\mathbb{W} - \tilde{\mathbb{W}}\| \|\tilde{s}\|_{L^2; \mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|\theta - \tilde{\theta}\|_{L^2; \mathbb{R}^m} \right), \quad (21)$$

and the fact that, under Assumption 2,  $\|\tilde{s}\|_{L^2; \mathbb{R}^n} \leq s_{\max}$ , as proven in Lemma 6. *Q.E.D.*

### A.2. Section 4 and 5.1: Omitted Proofs

We here report a formal lemma for the equivalence of network games to graphon games discussed in Section 4.1 and a more general statement for Theorem 2 that includes the sparsity parameter  $\kappa_N$ , as discussed in Section 5.1. The statements in Section 4.1 and Theorem 2 are obtained as special cases by setting  $\delta_N = \delta/N$  and  $\kappa_N = 1$ .

LEMMA 1: A vector  $\bar{s}_{[N]} \in \mathbb{R}^{Nn}$  is a Nash equilibrium of  $\mathcal{G}_{\kappa}^{[N]}(\{\mathcal{S}^i\}_{i=1}^N, U, \{\theta^i\}_{i=1}^N, A^{[N]})$  with  $N$  players, local aggregate as in (7) for some sparsity parameter  $\kappa_N$ , strategy sets  $\mathcal{S}^i$ , parameters  $\theta^i$ , and graph  $A^{[N]}$  if and only if the corresponding step function equilibrium  $\bar{s}_{[N]}(x) \in L^2([0, 1]; \mathbb{R}^n)$  is a Nash equilibrium of the graphon game  $\mathcal{G}(\mathcal{S}^{[N]}, U, \theta^{[N]}, W_{\kappa}^{[N]})$  with payoff function as in (2), set valued function  $\mathcal{S}^{[N]}(x) := \mathcal{S}^i$  for all  $x \in \mathcal{U}_i^{[N]}$ , parameter function  $\theta^{[N]}(x) := \theta^i$  for all  $x \in \mathcal{U}_i^{[N]}$ , and step function graphon  $W_{\kappa}^{[N]}$  corresponding to  $\frac{A^{[N]}}{\kappa_N}$ .

PROOF: Suppose that  $\bar{s}_{[N]}$  is a Nash equilibrium of  $\mathcal{G}(\mathcal{S}^{[N]}, U, \theta^{[N]}, W_{\kappa}^{[N]})$ . Since  $W_{\kappa}^{[N]}$  is a step function over the partition  $\mathcal{U}^{[N]}$ , the aggregate  $\bar{z}_{\kappa}(x) = \int_0^1 W_{\kappa}^{[N]}(x, y)\bar{s}_{[N]}(y) dy$  is a step function with respect to the same partition. Let  $\bar{z}_{\kappa}^i$  be the value of  $\bar{z}_{\kappa}(x)$  in  $\mathcal{U}_i^{[N]}$  and recall that  $\theta^{[N]}(x) = \theta^i$  in  $\mathcal{U}_i^{[N]}$ . From the definition of Nash equilibrium for the graphon game,

$$\bar{s}_{[N]}(x) = \arg \max_{s \in \mathcal{S}^{[N]}(x)} U(s, \bar{z}_{\kappa}(x), \theta^{[N]}(x)) = \arg \max_{s \in \mathcal{S}^i} U(s, \bar{z}_{\kappa}^i, \theta^i) \quad \text{for all } x \in \mathcal{U}_i^{[N]}.$$

Consequently, also  $\bar{s}_{[N]}(x)$  is a step function with respect to  $\mathcal{U}^{[N]}$ . Let  $\bar{s}_{[N]}^i$  be the value of  $\bar{s}_{[N]}(x)$  in  $\mathcal{U}_i^{[N]}$ . Then  $\bar{z}_{\kappa}^i = \int_0^1 W_{\kappa}^{[N]}(x, y)\bar{s}_{[N]}(y) dy = \frac{1}{N} \sum_{j=1}^N \frac{A_{ij}^{[N]}}{\kappa_N} \bar{s}_{[N]}^j$  and  $\bar{s}_{[N]}(x)$  is a graphon Nash equilibrium if and only if, for each  $i \in \{1, \dots, N\}$ , it holds

$$\bar{s}_{[N]}^i = \arg \max_{s \in \mathcal{S}^i} U(s, \bar{z}_{\kappa}^i, \theta^i), \quad \bar{z}_{\kappa}^i = \frac{1}{\kappa_N N} \sum_{j=1}^N A_{ij}^{[N]} \bar{s}_{[N]}^j.$$

The latter is the definition of Nash equilibrium in the sampled network game with network  $A^{[N]}$ , thus concluding the proof. Q.E.D.

THEOREM 2—Generalized: Consider a graphon game  $\mathcal{G}(\mathcal{S}, U, \theta, W)$  in which each player has homogeneous strategy set, that is,  $\mathcal{S}(x) = \mathcal{S}$  for all  $x \in [0, 1]$ . Suppose that  $\mathcal{G}$  satisfies Assumptions 1, 2, 3, and 4. Let  $\bar{s}$  be its unique Nash equilibrium and fix any sequences  $\{\delta_N, \kappa_N\}_{N=1}^{\infty}$  such that  $\delta_N \leq e^{-1}$  and  $\frac{\log(N/\delta_N)}{N\kappa_N} \rightarrow 0$ . Let  $\bar{s}_W^{[N]}$  be an arbitrary step function equilibrium of the sampled network game corresponding to either the sampled network  $A_w^{[N]}$  (as defined in Section 3.2) or  $A_s^{[N]}$  (as defined in Section 5.1). Then

1.  $\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} \rightarrow 0$  almost surely when  $N \rightarrow \infty$ ;
2. with probability at least  $1 - 2\delta_N$ ,

$$\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} = \mathcal{O}\left(\left(\frac{\log(N/\delta_N)}{N}\right)^{1/4} + \left(\frac{\log(N/\delta_N)}{\kappa_N N}\right)^{\frac{1}{2}}\right);$$

3. if  $\kappa_N = 1$  for all  $N$  and any  $\epsilon > 0$ , there exists  $C > 0$  s.t. for  $N$  large enough,

$$\mathbb{P}[\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2} > \epsilon] \leq \frac{2N}{\exp(\epsilon^4 N/C)}.$$

PROOF: • We start by proving statement 2. Let  $\theta^{[N]}$  be the step function corresponding to the vector  $[\theta(t^i)]_{i=1}^N$  and  $W_w^{[N]}, W_s^{[N]}$  be the step function graphons corresponding to  $A_w^{[N]}$  and  $\frac{A_s^{[N]}}{\kappa_N}$ , respectively, so that  $\bar{s}_w^{[N]}$  and  $\bar{s}_s^{[N]}$  are the equilibria of the graphon games

played over the graphon  $W_w^{[N]}$  and  $W_s^{[N]}$  with parameter function  $\theta^{[N]}$ . By Proposition 3, it follows

$$\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} \leq K(\|\mathbb{W}_W^{[N]} - \mathbb{W}\|_{s_{\max}} + \|\theta^{[N]} - \theta\|_{L^2; \mathbb{R}^m}). \quad (22)$$

The bound on  $\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n}$  follows from (22) and the fact that for  $N$  large enough with probability at least  $1 - 2\delta_N$ ,

$$\|\theta^{[N]} - \theta\|_{L^2; \mathbb{R}^m} \leq \rho_\theta(N) \quad \text{and} \quad \|\mathbb{W}_W^{[N]} - \mathbb{W}\| \leq \rho_W(N), \quad (23)$$

for  $\rho_\theta(N)$  and  $\rho_W(N)$  as defined in Lemma 12 in Appendix D.5 of the Supplemental Material. The bounds in (23) are proven in Avella-Medina et al. (2018, Theorem 1) (reported in Lemma 12 in Appendix D.5) and follow from the fact that the  $\{t^i\}_{i=1}^N$  are the ordered statistic of  $N$  uniform samples from  $[0, 1]$  combined with the fact that  $W$  is piecewise Lipschitz by Assumption 4.

Overall, there exists  $M > 0$  such that for  $N$  sufficiently large with probability at least  $1 - 2\delta_N$ , it holds

$$\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} \leq M \left( \left( \frac{\log(N/\delta_N)}{N} \right)^{\frac{1}{4}} + \left( \frac{\log(N/\delta_N)}{\kappa_N N} \right)^{\frac{1}{2}} \right) =: \Theta_N.$$

• To prove statement 1, let us define the sequence of events  $\mathcal{E}_N := \{\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} > \Theta_N\}$ . It follows by the previous point that  $\Pr[\mathcal{E}_N] < 2\delta_N$ . Note that  $\frac{\log(N/\delta_N)}{N\kappa_N} \rightarrow 0$  implies  $\frac{\log(N)}{N\kappa_N} \rightarrow 0$  (since  $\delta_N \leq 1$ ). Hence  $\delta_N = \frac{1}{N^2}$  is an admissible choice and leads to  $\sum_{N=1}^{\infty} \Pr[\mathcal{E}_N] < \sum_{N=1}^{\infty} \frac{2}{N^2} < \infty$ . By the Borel–Cantelli lemma, there exists a positive integer  $\bar{N}$  such that for all  $N \geq \bar{N}$ , the complement of  $\mathcal{E}_N$ , that is,  $\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} \leq \Theta_N$ , holds a.s. Since,  $\Theta_N \rightarrow 0$ , we obtain  $\|\bar{s}_W^{[N]} - \bar{s}\|_{L^2; \mathbb{R}^n} \rightarrow 0$  a.s.

• Finally, to prove statement 3, fix  $\delta_N = \delta/N$  for any  $0 \leq \delta \leq e^{-1}$ . Statement 2 then implies that there exists  $\tilde{C} > 0$  s.t. for  $N$  large enough,

$$\mathbb{P} \left[ \|\bar{s}_W^{[N]} - \bar{s}\|_{L^2} > \tilde{C} \left( \frac{\log(N^2/\delta)}{N} \right)^{\frac{1}{4}} \right] \leq \frac{2\delta}{N}.$$

Statement 3 follows by defining  $\epsilon = \tilde{C} \left( \frac{\log(N^2/\delta)}{N} \right)^{\frac{1}{4}}$ , explicitly deriving  $\delta$  from such expression, and then substituting it into the probability bound  $\frac{2\delta}{N}$  with  $C = \tilde{C}^4$ . *Q.E.D.*

### A.3. Section 6: Omitted Proofs

PROOF OF THEOREM 3: Problem (9) can be equivalently reformulated as a problem in the space of functions instead of vectors by using the reformulation given in Section 4.1, so that

$$\begin{aligned} T_{\text{opt}}^{[N]} &:= \max_{\hat{\theta}^{[N]} \in L^{[N]}} T^{[N]}(\hat{\theta}^{[N]}) = \int_0^1 U(\bar{s}_{\hat{\theta}^{[N]}}^{[N]}(x), \bar{z}_{\hat{\theta}^{[N]}}^{[N]}(x), \theta^{[N]}(x) + \hat{\theta}^{[N]}(x)) \\ \text{s.t. } &\bar{s}_{\hat{\theta}^{[N]}}^{[N]} = \text{Nash equilibrium of } \mathcal{G}(\mathcal{S}, U, \theta^{[N]} + \hat{\theta}^{[N]}, W^{[N]}), \\ &\bar{z}_{\hat{\theta}^{[N]}}^{[N]} = \mathbb{W}^{[N]} \bar{s}_{\hat{\theta}^{[N]}}^{[N]}, \\ &\|\hat{\theta}^{[N]}\|_{L^2}^2 \leq C, \end{aligned} \quad (24)$$

where we used  $L^{[N]}$  to denote the subspace of  $L^2([0, 1])$  composed by functions that are piecewise constant w.r.t. the partition  $\{\mathcal{U}_i^{[N]}\}_{i=1}^N$  and we used the same symbol  $\theta^{[N]}$  to denote the vector  $[\theta(t^i)]_{i=1}^N$  and its corresponding piecewise constant function.

Let  $\hat{\theta}_{\text{opt}}^{[N]}$  be an optimizer of Problem (24) and note that

$$\begin{aligned} T^{[N]}(\hat{\theta}_{\text{graphon}}^{[N]}) &\geq T(\hat{\theta}^*) - \underbrace{|T^{[N]}(\hat{\theta}_{\text{graphon}}^{[N]}) - T(\hat{\theta}^*)|}_{T_1} \geq T(\hat{\theta}_{\text{opt}}^{[N]}) - T_1 \\ &\geq T^{[N]}(\hat{\theta}_{\text{opt}}^{[N]}) - T_1 - \underbrace{|T^{[N]}(\hat{\theta}_{\text{opt}}^{[N]}) - T(\hat{\theta}_{\text{opt}}^{[N]})|}_{T_2} = T_{\text{opt}}^{[N]} - T_1 - T_2, \end{aligned}$$

where we used  $T(\hat{\theta}^*) \geq T(\hat{\theta}_{\text{opt}}^{[N]})$  since  $\theta^*$  is a maximizer. Our next objective is to upper bound the terms  $T_1$  and  $T_2$ . Using Lemma 2 given after this proof, we obtain

$$T_1 \leq L \sqrt{\|\bar{s}_{\hat{\theta}_{\text{graphon}}^{[N]}}^{[N]} - \bar{s}_{\hat{\theta}^*}\|_{L^2}^2 + \|\bar{z}_{\hat{\theta}_{\text{graphon}}^{[N]}}^{[N]} - \bar{z}_{\hat{\theta}^*}\|_{L^2}^2 + \|\hat{\theta}_{\text{graphon}}^{[N]} - \hat{\theta}^*\|_{L^2}^2 + \|\theta^{[N]} - \theta\|_{L^2}^2},$$

and similarly

$$T_2 \leq L \sqrt{\|\bar{s}_{\hat{\theta}_{\text{opt}}^{[N]}}^{[N]} - \bar{s}_{\hat{\theta}_{\text{opt}}^{[N]}}\|_{L^2}^2 + \|\bar{z}_{\hat{\theta}_{\text{opt}}^{[N]}}^{[N]} - \bar{z}_{\hat{\theta}_{\text{opt}}^{[N]}}\|_{L^2}^2 + \|\theta^{[N]} - \theta\|_{L^2}^2}.$$

Note that for any function  $\hat{\theta}_1, \hat{\theta}_2 \in L^2([0, 1])$  by the proof of Proposition 3 (Equation (21)) and Lemma 12, with probability  $1 - 2\delta_N$ , there exists  $K > 0$  s.t.

$$\begin{aligned} \|\bar{s}_{\hat{\theta}_1}^{[N]} - \bar{s}_{\hat{\theta}_2}\|_{L^2} &\leq K(\|\mathbb{W}\bar{s}_{\hat{\theta}_1}^{[N]} - \mathbb{W}\bar{s}_{\hat{\theta}_2}\|_{L^2} + \|(\theta^{[N]} + \hat{\theta}_1) - (\theta + \hat{\theta}_2)\|_{L^2}) \\ &\leq K(\rho_W(N)s_{\max} + \|\theta^{[N]} - \theta\|_{L^2} + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}) \\ &\leq K(\rho_W(N)s_{\max} + \rho_\theta(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}) =: K(\rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}), \end{aligned}$$

where  $\rho_W, \rho_\theta$  are as defined in Lemma 12.

Similarly,

$$\begin{aligned} \|\bar{z}_{\hat{\theta}_1}^{[N]} - \bar{z}_{\hat{\theta}_2}\|_{L^2} &= \|\mathbb{W}\bar{s}_{\hat{\theta}_1}^{[N]} - \mathbb{W}\bar{s}_{\hat{\theta}_2}\|_{L^2} \leq \|\mathbb{W}\bar{s}_{\hat{\theta}_1}^{[N]} - \mathbb{W}\bar{s}_{\hat{\theta}_1}\|_{L^2} + \|\mathbb{W}\bar{s}_{\hat{\theta}_1}^{[N]} - \mathbb{W}\bar{s}_{\hat{\theta}_2}\|_{L^2} \\ &\leq \|\mathbb{W}^{[N]} - \mathbb{W}\| \|\bar{s}_{\hat{\theta}_1}^{[N]}\|_{L^2} + \|\mathbb{W}\| \|\bar{s}_{\hat{\theta}_1}^{[N]} - \bar{s}_{\hat{\theta}_2}\|_{L^2} \\ &\leq \rho_W(N)s_{\max} + \lambda_{\max}(W)K(\rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}) \\ &\leq K'(\rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}), \end{aligned}$$

with  $K' = (1 + \lambda_{\max}(W)K)$ .

We finally bound  $\|\hat{\theta}_{\text{graphon}}^{[N]} - \hat{\theta}^*\|_{L^2}$ . To this end, let  $\hat{\theta}^{[N]*}$  be the piecewise function corresponding to  $[\theta^*(t^i)]_{i=1}^N$ , so that (i) by Lemma 12 with the same probability  $1 - 2\delta_N$ ,  $\|\hat{\theta}^{[N]*} - \hat{\theta}^*\|_{L^2} \leq \rho_\theta^*(N) := \sqrt{(L^*d_N)^2 + 4\Omega^*d_N\theta_{\max}^*}$  (where  $L^*$  is the Lipschitz constant of  $\hat{\theta}^*$ ,  $\Omega^*$  is the number of points in which  $\hat{\theta}^*$  is not Lipschitz continuous and  $\hat{\theta}^*(x) \leq \theta_{\max}^*$  for

all  $x$ ), and (ii) by definition,  $\hat{\theta}_{\text{graphon}}^{[N]} = \hat{\theta}^{[N]*} \frac{\sqrt{C}}{\|\hat{\theta}^{[N]*}\|_{L^2}} = \hat{\theta}^{[N]*} \frac{\|\hat{\theta}^*\|_{L^2}}{\|\hat{\theta}^{[N]*}\|_{L^2}}$ . Overall,

$$\begin{aligned} \|\hat{\theta}_{\text{graphon}}^{[N]} - \hat{\theta}^*\|_{L^2} &\leq \|\hat{\theta}_{\text{graphon}}^{[N]} - \hat{\theta}^{[N]*}\|_{L^2} + \|\hat{\theta}^{[N]*} - \hat{\theta}^*\|_{L^2} \leq \left| \frac{\|\hat{\theta}^*\|_{L^2}}{\|\hat{\theta}^{[N]*}\|_{L^2}} - 1 \right| \|\hat{\theta}^{[N]*}\|_{L^2} + \rho_{\theta}^*(N) \\ &= \left| \|\hat{\theta}^*\|_{L^2} - \|\hat{\theta}^{[N]*}\|_{L^2} \right| + \rho_{\theta}^*(N) \leq \|\hat{\theta}^* - \hat{\theta}^{[N]*}\|_{L^2} + \rho_{\theta}^*(N) = 2\rho_{\theta}^*(N). \end{aligned}$$

Hence  $T_1 + T_2 \leq 2L(\sqrt{K'^2 + K^2}(\rho_M(N) + 2\rho_{\theta}^*(N))^2 + (2\rho_{\theta}^*(N))^2 + \rho_{\theta}(N)^2) =: \rho_T(N) = \mathcal{O}((\frac{\log(N/\delta_N)}{N})^{\frac{1}{4}})$ . The result follows setting  $\delta_N = \frac{\delta}{N}$ . Q.E.D.

LEMMA 2: Suppose that  $U(s, z, \theta)$  is jointly Lipschitz continuous in  $[s, z, \theta]$  with constant  $L$ . For any function  $s, z, \theta, \tilde{s}, \tilde{z}, \tilde{\theta} \in L^2([0, 1])$ ,

$$\begin{aligned} &\left| \int_0^1 U(s(x), z(x), \theta(x)) dx - \int_0^1 U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x)) dx \right| \\ &\leq L\sqrt{\|s - \tilde{s}\|_{L^2}^2 + \|z - \tilde{z}\|_{L^2}^2 + \|\theta - \tilde{\theta}\|_{L^2}^2}. \end{aligned}$$

PROOF:

$$\begin{aligned} &\left| \int_0^1 U(s(x), z(x), \theta(x)) dx - \int_0^1 U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x)) dx \right| \\ &\leq \int_0^1 |U(s(x), z(x), \theta(x)) - U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x))| dx \\ &\leq L \int_0^1 (\sqrt{\|s(x) - \tilde{s}(x)\|^2 + \|z(x) - \tilde{z}(x)\|^2 + \|\theta(x) - \tilde{\theta}(x)\|^2}) dx \\ &\leq L\sqrt{\int_0^1 (\|s(x) - \tilde{s}(x)\|^2 + \|z(x) - \tilde{z}(x)\|^2 + \|\theta(x) - \tilde{\theta}(x)\|^2) dx} \\ &= L\sqrt{\|s - \tilde{s}\|_{L^2}^2 + \|z - \tilde{z}\|_{L^2}^2 + \|\theta - \tilde{\theta}\|_{L^2}^2}. \end{aligned} \quad \text{Q.E.D.}$$

PROOF OF PROPOSITION 1: First, note that for linear quadratic games with  $\alpha > 0$ , the payoff at equilibrium is

$$U(\bar{s}_{\hat{\theta}}(x), \bar{z}_{\hat{\theta}}(x), \theta(x) + \hat{\theta}(x)) = \frac{1}{2}\bar{s}_{\hat{\theta}}(x)^2.$$

Hence  $T(\hat{\theta}) = \frac{1}{2}\|\bar{s}_{\hat{\theta}}\|_{L^2}^2$ . Similarly to the proof of Lemma 3, given next, it can be shown that with probability  $1 - 2\delta_N$ , both  $\|\bar{s}_{\hat{\theta}^{[N]}}\|_{L^2}$  and  $\|\bar{s}_{\hat{\theta}}\|_{L^2}$  (as defined in (24) and in (10), respectively) can be bounded by some constant  $M_s$  for any feasible  $\hat{\theta}^{[N]}, \hat{\theta}$ . Let  $\mathcal{K}$  be the kernel of  $\mathbb{W}$ , so that  $\mathbb{W}\psi = 0$  for any function  $\psi \in \mathcal{K}$  and let  $\mathcal{K}^{\perp}$  be its orthogonal complement. By the spectral theorem, it is possible to construct an orthonormal basis for  $\mathcal{K}^{\perp}$  made of eigenfunctions of  $\mathbb{W}$ . We denote such basis by  $\{\psi_r\}_{r=1}^R$ . Let  $\hat{\theta}$  be a generic function of  $L^2([0, 1])$ . Let  $b_0\psi_0, \hat{b}_0\hat{\psi}_0$  be the projection of  $\theta, \hat{\theta}$  in  $\mathcal{K}$  (with  $\|\psi_0\|_{L^2} = \|\hat{\psi}_0\|_{L^2} = 1$ ) and let

$b_r = \langle \theta, \psi_r \rangle$ ,  $\hat{b}_r = \langle \hat{\theta}, \psi_r \rangle$ . Since  $\mathcal{K}$  is a closed linear subspace, it holds  $\theta = b_0\psi_0 + \sum_{r=1}^R b_r\psi_r$  and  $\hat{\theta} = \hat{b}_0\hat{\psi}_0 + \sum_{r=1}^R \hat{b}_r\psi_r$ . Using the fact that  $\{\psi_r\}_{r=1}^R$  are orthogonal to each other and orthogonal to any function in  $\mathcal{K}$  yields  $\|\hat{\theta}\|_{L^2}^2 = \|\hat{b}_0\hat{\psi}_0 + \sum_{r=1}^R \hat{b}_r\psi_r\|_{L^2}^2 = \|\hat{b}_0\hat{\psi}_0\|_{L^2}^2 + \sum_{r=1}^R \hat{b}_r^2 = \sum_{r=0}^R \hat{b}_r^2$ . Since we are considering linear quadratic games, from Example 1, the equilibrium induced by  $\theta + \hat{\theta}$  can be rewritten as

$$\begin{aligned} \bar{s}_\delta &= (\mathbb{I} - \alpha\mathbb{W})^{-1}(\theta + \hat{\theta}) = \sum_{h=0}^{\infty} \alpha^h \mathbb{W}^h \left( b_0\psi_0 + \hat{b}_0\hat{\psi}_0 + \sum_{r=1}^R (b_r + \hat{b}_r)\psi_r \right) \\ &= \sum_{h=0}^{\infty} \alpha^h \mathbb{W}^h (b_0\psi_0 + \hat{b}_0\hat{\psi}_0) + \sum_{r=1}^R (b_r + \hat{b}_r) \sum_{h=0}^{\infty} \alpha^h \mathbb{W}^h (\psi_r) \\ &= (b_0\psi_0 + \hat{b}_0\hat{\psi}_0) + \sum_{r=1}^R (b_r + \hat{b}_r) \sum_{h=0}^{\infty} \alpha^h \lambda_r^h \psi_r = (b_0\psi_0 + \hat{b}_0\hat{\psi}_0) + \sum_{r=1}^R \frac{(b_r + \hat{b}_r)}{1 - \alpha\lambda_r} \psi_r. \end{aligned}$$

Hence  $2T(\hat{\theta}) = \|\bar{s}_\delta\|_{L^2}^2 = \|b_0\psi_0 + \hat{b}_0\hat{\psi}_0\|_{L^2}^2 + \sum_{r=1}^R \frac{(b_r + \hat{b}_r)^2}{(1 - \alpha\lambda_r)^2}$ . Note that for any fixed value of  $b_0, \psi_0, \hat{b}_0$ , the quantity  $\|b_0\psi_0 + \hat{b}_0\hat{\psi}_0\|_{L^2}^2$  is maximized when  $\hat{\psi}_0 = \psi_0$ , in which case  $\|b_0\psi_0 + \hat{b}_0\hat{\psi}_0\|_{L^2}^2 = (b_0 + \hat{b}_0)^2$ . Hence Problem (10) can be reformulated as (14). *Q.E.D.*

LEMMA 3: Let  $\bar{s}_W^{[N]}$  be an arbitrary step function equilibrium of the sampled network game  $\mathcal{G}^{[N]}(\{\mathbb{R}_{\geq 0}\}_{i=1}^N, U, \{\theta(t^i)\}_{i=1}^N, A_W^{[N]})$ , as introduced in Section 3.2, with linear quadratic payoff  $U$  as in (12) and set  $0 < \alpha < \frac{1}{\lambda_{\max}(\mathbb{W})}$ . Then there exists  $M_s$  such that for any admissible confidence sequence  $\{\delta_N\}_{N=1}^{\infty}$  and  $N$  large, with probability  $1 - 2\delta_N$ ,  $\|\bar{s}_W^{[N]}\|_{L^2} \leq M_s$ .

PROOF: Fix  $\bar{k} \in (1, \frac{1}{\alpha\lambda_{\max}(\mathbb{W})})$  (this interval has non-empty interior by assumption). Since by Lemma 12 (given in Appendix D.5 of the Supplemental Material) with probability  $1 - 2\delta_N$  for  $N$  large  $|\lambda_{\max}(\mathbb{W}_W^{[N]}) - \lambda_{\max}(\mathbb{W})| \leq \rho_W(N)$ , it follows that for  $N$  large enough,  $\lambda_{\max}(\mathbb{W}_W^{[N]}) \leq \bar{k}\lambda_{\max}(\mathbb{W})$ , and by the formula for the equilibrium in linear quadratic graphon games derived in Example 1, we obtain  $\|\bar{s}_W^{[N]}\|_{L^2} \leq \|(\mathbb{I} - \alpha\mathbb{W}_W^{[N]})^{-1}\theta^{[N]}\|_{L^2} \leq \frac{1}{1 - \alpha\lambda_{\max}(\mathbb{W}_W^{[N]})} \|\theta^{[N]}\|_{L^2} \leq \frac{1}{1 - \alpha\bar{k}\lambda_{\max}(\mathbb{W})} 2\|\theta\|_{L^2} =: M_s$ , where we used that, for large  $N$ ,  $\|\theta^{[N]}\|_{L^2} \leq 2\|\theta\|_{L^2}$  since  $\|\theta^{[N]} - \theta\|_{L^2} \rightarrow 0$ , as shown in Lemma 12 in Appendix D.5. *Q.E.D.*

PROOF OF PROPOSITION 2: By Lemma 4, given next, at the unique internal equilibrium,

$$\bar{s}_\delta(x) = \alpha\bar{z}_\delta(x)(1 - \bar{z}_\delta(x)) + \theta(x) + \hat{\theta}(x).$$

Hence  $U(\bar{s}_\delta(x), \bar{z}_\delta(x), \theta(x) + \hat{\theta}(x)) = \frac{1}{2}\bar{s}_\delta(x)^2$  and  $T(\hat{\theta}) = \|\bar{s}_\delta\|_{L^2}^2$ . Moreover, by Lemma 4, the unique graphon equilibrium and the intervention  $\hat{\theta}$  can be rewritten as

$$\begin{aligned} \bar{s}_\delta &= s_1\psi_1 + s_2\psi_2 + s_0\psi_0 + \tilde{b}\tilde{\psi}, \\ \hat{\theta} &= \hat{b}_1\psi_1 + \hat{b}_2\psi_2 + \hat{b}_0\psi_0 + \tilde{b}\tilde{\psi}, \end{aligned}$$

where  $\psi_1, \psi_2, \psi_0, \tilde{\psi}$  are orthonormal functions defined in Lemma 4. We then immediately obtain  $\|\bar{s}_\delta\|_{L^2}^2 = s_0^2 + s_1^2 + s_2^2 + \tilde{b}^2$ . Using the expression for  $s_0, s_1, s_2$  given in Lemma 4,

Problem (10) can be equivalently reformulated as

$$\begin{aligned} & \max_{\hat{b}_1, \hat{b}_2, \hat{b}_3} s_1^2 + (-\alpha s_1^2 \lambda p_2 + b_2 + \hat{b}_2)^2 + (b_0 + \hat{b}_0)^2 \\ \text{s.t. } & s_1 = \frac{-(1 - \alpha\lambda) + \sqrt{(1 - \alpha\lambda)^2 + 4(\alpha\lambda p_1)(b_1 + \hat{b}_1)}}{2\alpha\lambda p_1}, \\ & \hat{b}_1^2 + \hat{b}_2^2 + \hat{b}_3^2 = C, \end{aligned}$$

where we used that, at the optimum,  $\tilde{b} = 0$ . Note that to use Lemma 4, we should verify that  $b_1 + \hat{b}_1 > 0$  and  $\theta(x) + \hat{\theta}(x) \in [0, 3/4]$  for all  $x \in [0, 1]$ . However, this is true in the small budget regime under the given assumption that  $b_1 > 0$  and  $\theta(x) \in [0, 3/4]$  for all  $x \in [0, 1]$ . Q.E.D.

LEMMA 4: Consider a graphon game  $\mathcal{G}([0, 1], U, \theta + \hat{\theta}, W_v)$  with payoff function  $U$  as in (16) and graphon  $\mathbb{W}_v$  as in Example 5. Suppose that  $\alpha \in [0, \min(1, 1/\lambda)]$  and  $\theta(x) + \hat{\theta}(x) \in [0, 3/4]$  for all  $x \in [0, 1]$ . Then there is a unique equilibrium and is internal. To compute an explicit formula, let  $\psi_1(x) = v(x)/\sqrt{\lambda} \in L^2([0, 1])$  be the unique eigenfunction of  $W_v$  normalized such that  $\|\psi_1\|_{L^2} = 1$ . Moreover, define (i)  $\psi_2$  such that  $v^2(x) = p_1\psi_1(x) + p_2\psi_2(x)$  with  $p_1, p_2 > 0$ ,  $\|\psi_2\|_{L^2} = 1$ ,  $\langle \psi_1, \psi_2 \rangle = 0$ , (ii)  $\psi_0$  such that  $\theta = b_1\psi_1 + b_2\psi_2 + b_0\psi_0$  with  $\|\psi_0\|_{L^2} = 1$ ,  $\langle \psi_0, \psi_i \rangle = 0$  for  $i = 1, 2$ , and (iii)  $\tilde{\psi}$  such that  $\hat{\theta} = \hat{b}_1\psi_1 + \hat{b}_2\psi_2 + \hat{b}_0\psi_0 + \tilde{b}\tilde{\psi}$  with  $\|\tilde{\psi}\|_{L^2} = 1$ ,  $\langle \tilde{\psi}, \psi_i \rangle = 0$ , for all  $i = 0, 1, 2$ .<sup>21</sup> Suppose  $b_1 + \hat{b}_1 > 0$ . Then

$$\bar{s}_\theta = s_1\psi_1 + s_2\psi_2 + s_0\psi_0 + \tilde{b}\tilde{\psi},$$

with  $s_1 = \frac{-(1-\alpha\lambda) + \sqrt{(1-\alpha\lambda)^2 + 4(\alpha\lambda p_1)(b_1 + \hat{b}_1)}}{2\alpha\lambda p_1}$ ,  $s_2 = -\alpha s_1^2 \lambda p_2 + b_2 + \hat{b}_2$ , and  $s_0 = b_0 + \hat{b}_0$ .

PROOF: First note that all assumptions of Theorem 1 are met, hence there is a unique equilibrium. The best response is given by

$$\bar{s}_\theta(x) = \Pi_{[0,1]}[\alpha \bar{z}_\theta(x)(1 - \bar{z}_\theta(x)) + \theta(x) + \hat{\theta}(x)]. \tag{25}$$

However,  $\bar{s}_\theta(x) \in [0, 1]$  implies  $\bar{z}_\theta(x) \in [0, 1]$ , which implies  $\bar{z}_\theta(x)(1 - \bar{z}_\theta(x)) \in [0, \frac{1}{4}]$ . Since  $\alpha \in [0, 1]$  and  $\theta(x) + \hat{\theta}(x) \in [0, \frac{3}{4}]$ , we obtain that the argument inside of the projection in (25) is always in  $[0, 1]$ . Consequently, (25) can be rewritten as

$$\bar{s}_\theta(x) = \alpha \bar{z}_\theta(x)(1 - \bar{z}_\theta(x)) + \theta(x) + \hat{\theta}(x). \tag{26}$$

Note that

$$\bar{z}_\theta(x) = \int_0^1 v(x)v(y)\bar{s}_\theta(y) dy = \lambda \int_0^1 \psi_1(x)\psi_1(y)\bar{s}_\theta(y) dy = \lambda\psi_1(x)\langle \psi_1, \bar{s}_\theta \rangle.$$

Set  $\xi := \langle \psi_1, \bar{s}_\theta \rangle$ , then  $\bar{z}_\theta(x) = \xi\lambda\psi_1(x) = \xi\sqrt{\lambda}v(x)$ . Consequently,  $\bar{z}_\theta^2(x) = \xi^2\lambda v^2(x) = \xi^2\lambda[p_1\psi_1(x) + p_2\psi_2(x)]$ .

<sup>21</sup>Such decomposition can be obtained by using the Gram–Schmidt process to construct an orthonormal basis.



Substituting this notation in (26), we obtain

$$\begin{aligned}
 \bar{s}_\theta &= \alpha \bar{z}_\theta - \alpha \bar{z}_\theta^2 + \theta + \hat{\theta} \\
 &= \alpha \xi \lambda \psi_1 - \alpha \xi^2 \lambda [p_1 \psi_1 + p_2 \psi_2] + (b_0 + \hat{b}_0) \psi_0 + (b_1 + \hat{b}_1) \psi_1 \\
 &\quad + (b_2 + \hat{b}_2) \psi_2 + \tilde{b} \tilde{\psi} \\
 &= \underbrace{(\alpha \xi \lambda - \alpha \xi^2 \lambda p_1 + b_1 + \hat{b}_1)}_{s_1} \psi_1 + \underbrace{(-\alpha \xi^2 \lambda p_2 + b_2 + \hat{b}_2)}_{s_2} \psi_2 \\
 &\quad + \underbrace{(b_0 + \hat{b}_0)}_{s_0} \psi_0 + \tilde{b} \tilde{\psi}.
 \end{aligned} \tag{27}$$

Plugging this expression in  $\xi := \langle \psi_1, \bar{s}_\theta \rangle$ , we obtain

$$\xi = (\alpha \xi \lambda - \alpha \xi^2 \lambda p_1 + b_1 + \hat{b}_1) \Leftrightarrow \alpha \lambda p_1 \xi^2 + (1 - \alpha \lambda) \xi - (b_1 + \hat{b}_1) = 0,$$

which has two solutions  $\xi_\pm = \frac{- (1 - \alpha \lambda) \pm \sqrt{(1 - \alpha \lambda)^2 + 4(\alpha \lambda p_1)(b_1 + \hat{b}_1)}}{2\alpha \lambda p_1}$ . Since  $\alpha, p_1, \lambda, (b_1 + \hat{b}_1) > 0$ , it holds  $\xi_+ > 0$  and  $\xi_- < 0$ . Since  $\bar{z}_\theta(x) = \sqrt{\lambda} \xi v(x) \in [0, \frac{1}{4}]$  and  $v(x) > 0$ , it must be  $\xi > 0$ , hence  $\xi = \xi_+ = s_1$ . *Q.E.D.*

REFERENCES

ACEMOGLU, DARON, ASUMAN OZDAGLAR, AND ALIREZA TAHBAZ-SALEHI (2015): “Networks, Shocks, and Systemic Risk,” in *The Oxford Handbook of the Economics of Networks*, ed. by Yann Bramoullé, Andrea Galeotti, and Brian Rogers. Oxford University Press. [194]

AKBARPOUR, MOHAMMAD, SURAJ MALLADI, AND AMIN SABERI (2018): “Just a few Seeds More: Value of Network Information for Diffusion,” Available at SSRN 3062830. [194]

ALIPRANTIS, CHARALAMBOS D., AND KIM C. BORDER (2006): *Infinite Dimensional Analysis a Hitchhiker’s Guide*. Springer. [215]

ALLOUCH, NIZAR (2015): “On the Private Provision of Public Goods on Networks,” *Journal of Economic Theory*, 157, 527–552. [194]

ALTMAN, EITAN, THOMAS BOULOGNE, RACHID EL-AZOUZI, TANIA JIMÉNEZ, AND LAURA WYNTER (2006): “A Survey on Networking Games in Telecommunications,” *Computers & Operations Research*, 33 (2), 286–311. [195]

AVELLA-MEDINA, MARCO, FRANCESCA PARISE, MICHAEL T. SCHAUB, AND SANTIAGO SEGARRA (2018): “Centrality Measures for Graphons: Accounting for Uncertainty in Networks,” *IEEE Transactions on Network Science and Engineering*, 7 (1), 520–537. [195,198,207,218]

BALLESTER, CORALIO, ANTONI CALVÓ-ARMENGOL, AND YVES ZENOU (2006): “Who’s Who in Networks. Wanted: The Key Player,” *Econometrica*, 74 (5), 1403–1417. [194,195,197,198,214]

BANERJEE, ABHIJIT, ARUN G. CHANDRASEKHAR, ESTHER DUFLU, AND MATTHEW O. JACKSON (2013): “The Diffusion of Microfinance,” *Science*, 341 (6144), 1236498. [192,211]

BORGS, CHRISTIAN, JENNIFER CHAYES, HENRY COHN, AND YUFEI ZHAO (2019): “An  $L^p$  Theory of Sparse Graph Convergence I: Limits, Sparse Random Graph Models, and Power Law Distributions,” *Transactions of the American Mathematical Society*, 372 (5), 3019–3062. [203]

BORGS, CHRISTIAN, JENNIFER CHAYES, LÁSZLÓ LOVÁSZ, VERA SÓS, AND KATALIN VESZTERGOMBI (2011): “Limits of Randomly Grown Graph Sequences,” *European Journal of Combinatorics*, 32 (7), 985–999. [207]

BORGS, CHRISTIAN, JENNIFER T. CHAYES, LÁSZLÓ LOVÁSZ, VERA T. SÓS, AND KATALIN VESZTERGOMBI (2008): “Convergent Sequences of Dense Graphs I: Subgraph Frequencies, Metric Properties and Testing,” *Advances in Mathematics*, 219 (6), 1801–1851. [192]

BOUCHER, VINCENT, AND ARISTIDE HOUNDETOUNGAN (2019): “Estimating Peer Effects Using Partial Network Data,” Working paper. [193]

- BOYD, STEPHEN, AND LIEVEN VANDENBERGHE (2004): *Convex Optimization*. Cambridge University Press. [208]
- BRAMOULLÉ, YANN, AND RACHEL KRANTON (2007): “Public Goods in Networks,” *Journal of Economic Theory*, 135 (1), 478–494. [194,195]
- (2016): “Games Played on Networks,” in *The Oxford Handbook of the Economics of Networks*, ed. by Yann Bramoullé, Andrea Galeotti, and Brian Rogers. Oxford University Press. [195,198]
- BRAMOULLÉ, YANN, RACHEL KRANTON, AND MARTIN D’AMOURS (2014): “Strategic Interaction and Networks,” *The American Economic Review*, 104 (3), 898–930. [194,195]
- BREZA, EMILY, ARUN G. CHANDRASEKHAR, TYLER H. MCCORMICK, AND MENGJIE PAN (2020): “Using Aggregated Relational Data to Feasibly Identify Network Structure Without Network Data,” *American Economic Review*, 110 (8), 2454–2484. [191,192,211]
- CAINES, PETER E., AND MINYI HUANG (2018): “Graphon Mean Field Games and the GMFG Equations,” in *2018 IEEE Conference on Decision and Control (CDC)*, 4129–4134, IEEE. [195]
- CHANDRASEKHAR, ARUN, AND RANDALL LEWIS (2016): “Econometrics of Sampled Networks,”. [193]
- CHUNG, FAN, AND MARY RADCLIFFE (2011): “On the Spectra of General Random Graphs,” *The Electronic Journal of Combinatorics*, 18 (1), 215. [205]
- CORNES, RICHARD, AND ROGER HARTLEY (2012): “Fully Aggregative Games,” *Economics Letters*, 116 (3), 631–633. [195]
- DASARATHA, KRISHNA (2020): “Distributions of Centrality on Networks,” *Games and Economic Behavior*, 122, 1–27. [195]
- DE PAULA, AUREO, IMRAN RASUL, AND PEDRO SOUZA (2018): “Recovering Social Networks From Panel Data: Identification, Simulations and an Application,” CEPR Discussion Paper No. DP12792. [193]
- DIACONIS, PERSI, AND SVANTE JANSON (2007): “Graph Limits and Exchangeable Random Graphs,” arXiv preprint arXiv:0712.2749. [192]
- DUBEY, PRADEEP, ORI HAIMANKO, AND ANDRIY ZAPECHELNYUK (2006): “Strategic Complements and Substitutes, and Potential Games,” *Games and Economic Behavior*, 54 (1), 77–94. [195]
- ELDRIDGE, JUSTIN, MIKHAIL BELKIN, AND YUSU WANG (2016): “Graphons, Mergeons, and so on!” in *Advances in Neural Information Processing Systems*, 2307–2315. [195]
- GALEOTTI, ANDREA, BENJAMIN GOLUB, AND SANJEEV GOYAL (2020): “Targeting Interventions in Networks,” *Econometrica*, 88 (6), 2445–2471. [193,195,208,212]
- GALEOTTI, ANDREA, SANJEEV GOYAL, MATTHEW O. JACKSON, FERNANDO VEGA-REDONDO, AND LEEAT YARIV (2010): “Network Games,” *The Review of Economic Studies*, 77 (1), 218–244. [194]
- GAO, SHUANG, AND PETER E. CAINES (2017): “The Control of Arbitrary Size Networks of Linear Systems via Graphon Limits: An Initial Investigation,” in *Proceedings of the IEEE Conference on Decision and Control*. [195]
- GOLUB, BENJAMIN, AND MATTHEW O. JACKSON (2012a): “Does Homophily Predict Consensus Times? Testing a Model of Network Structure via a Dynamic Process,” *Review of Network Economics*, 11 (3). [194]
- (2012b): “How Homophily Affects the Speed of Learning and Best-Response Dynamics,” *The Quarterly Journal of Economics*, 127 (3), 1287–1338. [194]
- HUANG, MINYI, PETER E. CAINES, AND ROLAND P. MALHAMÉ (2007): “Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized  $\epsilon$ -Nash Equilibria,” *IEEE Transactions on Automatic Control*, 52 (9), 1560–1571. [195]
- HUTSON, VIVIAN, JOHN PYM, AND MICHAEL CLOUD (2005): *Applications of Functional Analysis and Operator Theory*, Vol. 200. Elsevier. [215]
- JACKSON, MATTHEW O., AND BRIAN W. ROGERS (2005): “The Economics of Small Worlds,” *Journal of the European Economic Association*, 3 (2–3), 617–627. [200]
- JACKSON, MATTHEW O., AND EVAN STORMS (2019): “Behavioral Communities and the Atomic Structure of Networks,” Available at SSRN 3049748. [193,194]
- JACKSON, MATTHEW O., AND YVES ZENOU (2014): “Games on Networks,” in *Handbook of Game Theory*, Vol. 4, ed. by Peyton Young and Shmuel Zamir. Elsevier. [195,198]
- JENSEN, MARTIN KAAE (2010): “Aggregative Games and Best-Reply Potentials,” *Economic Theory*, 43 (1), 45–66. [195]
- JOHNSON, CATHLEEN, AND ROBERT P. GILLES (2003): “Spatial Social Networks,” in *Networks and Groups*. Springer, 51–77. [200]
- KALAI, EHUD (2004): “Large Robust Games,” *Econometrica*, 72 (6), 1631–1665. [194]
- KUKUSHKIN, NIKOLAI S. (2004): “Best Response Dynamics in Finite Games With Additive Aggregation,” *Games and Economic Behavior*, 48 (1), 94–110. [195]
- LASRY, JEAN-MICHEL, AND PIERRE-LOUIS LIONS (2007): “Mean Field Games,” *Japanese Journal of Mathematics*, 2, 229–260. [195]

- LEE, CHRISTINA E., AND DEVAVRAT SHAH (2017): “Unifying Framework for Crowd-Sourcing via Graphon Estimation,” arXiv preprint arXiv:1703.08085. [195]
- LEWBEL, ARTHUR, XI QU, AND XUN TANG (2019): “Social Networks With Misclassified or Unobserved Links,” Manuscript. [193]
- LOVÁSZ, LÁSZLÓ (2012): *Large Networks and Graph Limits*, Vol. 60. Providence: American Mathematical Society. [192,196-198]
- LOVÁSZ, LÁSZLÓ, AND BALÁZS SZEGEDY (2006): “Limits of Dense Graph Sequences,” *Journal of Combinatorial Theory, Series B*, 96 (6), 933–957. [192]
- MA, ZHONGJING, DUNCAN S. CALLAWAY, AND IAN A. HISKENS (2013): “Decentralized Charging Control of Large Populations of Plug-in Electric Vehicles,” *IEEE Transactions on Control Systems Technology*, 21 (1), 67–78. [195]
- MORENCY, MATTHEW W., AND GEERT LEUS (2017): “Signal Processing on Kernel-Based Random Graphs,” in *Proceedings of the European Signal Processing Conference (EUSIPCO)*. [195]
- NOURIAN, MOJTABA, AND PETER E. CAINES (2013): “ $\epsilon$ -Nash Mean Field Game Theory for Nonlinear Stochastic Dynamical Systems With Major and Minor Agents,” *SIAM Journal on Control and Optimization*, 51 (4), 3302–3331. [214]
- PARISE, FRANCESCA, AND ASUMAN OZDAGLAR (2018): “Graphon Games,” arXiv preprint arXiv:1802.00080 (v4). [202]
- (2019): “A Variational Inequality Framework for Network Games: Existence, Uniqueness, Convergence and Sensitivity Analysis,” *Games and Economic Behavior*, 114, 47–82. [193,209]
- (2023): “Supplement to ‘Graphon Games: A Statistical Framework for Network Games and Interventions’,” *Econometrica Supplemental Material*, 91, <https://doi.org/10.3982/ECTA17564>. [193]
- PATACCHINI, ELEONORA, AND YVES ZENOU (2012): “Juvenile Delinquency and Conformism,” *The Journal of Law, Economics, & Organization*, 28 (1), 1–31. [204]
- ROSEN, J. BEN (1965): “Existence and Uniqueness of Equilibrium Points for Concave n-Person Games,” *Econometrica: Journal of the Econometric Society*, 33 (3), 520–534. [196]
- SADLER, EVAN (2020): “Diffusion Games,” *American Economic Review*, 110 (1), 225–270. [194]
- SANDHOLM, WILLIAM H. (2010): *Population Games and Evolutionary Dynamics*. The MIT press. [195]
- SMART, D. R. (1974): *Fixed Point Theorems*. Cambridge University Press Archive. [215,216]
- USHCHEV, PHILIP, AND YVES ZENOU (2020): “Social Norms in Networks,” *Journal of Economic Theory*, 185, 104969. [204]
- VIVES, XAVIER (2005): “Complementarities and Games: New Developments,” *Journal of Economic Literature*, 43 (2), 437–479. [194]

---

*Co-editor Barton L. Lipman handled this manuscript.*

*Manuscript received 9 August, 2019; final version accepted 5 August, 2022; available online 29 August, 2022.*