Graphon games: A statistical framework for network games and interventions

Francesca Parise\textsuperscript{1}, Asuman Ozdaglar\textsuperscript{2}

\textsuperscript{1}School of Electrical and Computer Engineering, Cornell University
\textsuperscript{2}Dep. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Abstract

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 dataset and show that the graphon intervention can be computed efficiently and based solely on aggregated relational data.

Keywords: Network games, graphons, 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.\footnote{Breza et al. (2020) estimated that conducting network surveys in 120 Indian villages would cost \$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.}

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 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 et al. (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 generalize e.g. Erdős-Rényi and stochastic block models).\footnote{Graphons are a flexible class of nonparametric random graph models, including Erdős-Rényi and stochastic block models as special cases. By the Aldous-Hoover theorem any exchangeable infinite random graph is obtained as a mixture of graphons, Diaconis and Janson (2007).}

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). Moreover, 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). 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 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 heterogenous interactions among a continuum

---

3 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 the villages in Karnataka, India, Breza et al. (2020) shows using J-PAL South Asia cost estimated that collecting aggregated relational data leads to a 70-80% cost reduction with respect to the cost of data collected in Banerjee et al. (2013).
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 et al. (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 dataset of 80 different networks, drawn as independent realizations of a stochastic block model with 4 communities. For this case study, the graphon approach results in a near optimal solution and reduces to a 4 dimensional optimization problem (whereas designing the optimal in-
tervention or a network heuristic suggested in Galeotti et al. (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 et al. (2018); Boucher and Hounde\-toungan (2019); Lewbel et al. (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 online Appendix C.

The results discussed so far are derived under the assumption that agents have perfect information about the sampled network. In online Appendix B, 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) and Golub and Jackson (2012b) used stochastic block models to achieve dimensionality reduction in opinion dynamic models. In more recent work, contagion processes have been considered in Akbarpour et al. (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 conta-
tion processes), ii) by 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) by 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 online Appendix B, are related to two previous works: Galeotti et al. (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) shows 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 nonatomic games. In this context, our work complements results derived for finite population network games (see e.g. Ballester et al. (2006); Bramoullé and Kranton (2007); Bramoullé et al. (2014); Jackson and Zenou (2014); Bramoullé and Kranton (2016); Galeotti et al. (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 et al. (2007); Sandholm (2010); Kukushkin (2004); Jensen (2010); Cornes and Hartley (2012); Dubey et al. (2006); Ma et al. (2013); Altman et al. (2006)).

\[^{4}\text{While contagion models have discrete (typically 0 – 1) strategies, we focus here on games with continuous strategies. The type of continuous games considered in our framework has been broadly used in the literature, both in theoretical and empirical works, for example to model applications in which agents need to decide on their level of effort or investment in a certain activity (see Vives (2005); Ballester et al. (2006); Acemoglu et al. (2015); Bramoullé and Kranton (2007); Bramoullé et al. (2014); Allouch (2015)).}\]
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 et al. (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 et al. (2018) for graphon models. The authors of these papers study 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 e.g. 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 \( A^{[N]}_{ij} \) 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 \( i \in \{1, \ldots, N\} \) selects a strategy \( s^i \in \mathbb{R} \) in its feasible set \( S^i \subseteq \mathbb{R} \) to maximize a payoff function

\[
U(s^i, z^i(s), \theta^i)
\]

where \( s := [s^i]_i=1 \in \mathbb{R}^N \), \( z^i(s) := \frac{1}{N} \sum_{j=1}^N [A^{[N]}_{ij}]s^j \) denotes the local aggregate\(^5\) (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.\(^6\) We denote compactly a

\(^5\)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.

\(^6\)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).
network game with the notation \( G^{[N]}(\{S^i\}^N_{i=1}, U, \{\theta^i\}^N_{i=1}, A^{[N]}) \) and we say a network game \( G^{[N]} \) with network \( A^{[N]} \) if we need to stress the role of the network.

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 \( i \in \{1, \ldots, N\} \)). 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 \( z^i(s) := \frac{1}{N} \sum_{j=1}^{N} P^{[N]}_{ij} s^j \) 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 \( W : [0, 1]^2 \rightarrow [0, 1] \) and can be used to formally define the limit of a sequence of graphs when the number of nodes tends to infinity, see e.g. 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 \( A^{[N]}_{ij} \) denotes the level of interaction between agents \( i \) and \( j \) in \( \{1, \ldots, N\} \). 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:

\[
z(x \mid s) := \int_0^1 W(x, y)s(y)dy.
\]

**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 \([0, 1]\).

Similar to network games, we assume that the payoff function of agent \( x \)

\[
U(s(x), z(x \mid s), \theta(x)). \tag{2}
\]

depends on his strategy \( s(x) \), on his local aggregate \( z(x \mid s) \) 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 \(z(x \mid s)\) 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 \(\mathcal{G}(\mathcal{S}, U, \theta, W)\) is 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 \(\alpha_\theta\) for each value of \(z, \theta\).\(^7\) Moreover, \(\nabla_s U(s, z, \theta)\) is uniformly Lipschitz in \([z, \theta]\) with constants \(l_\theta, l_\theta\) for all \(s\). For each \(x \in [0, 1]\) the set \(\mathcal{S}(x)\) is convex, closed and there exists \(\hat{z}\) and \(M > 0\) such that \(\|\arg \max_{\tilde{s} \in \mathcal{S}(x)} U(\tilde{s}, \hat{z}, \theta(x))\| \leq M\) for all \(x \in [0, 1]\).

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

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 can typically be guaranteed if the best response mapping is a contraction. To specify such a contraction property for graphon games we introduce the **graphon operator**, (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 online Appendix, all the eigenvalues of the graphon operator \(\mathbb{W}\) are real and the operator norm, i.e. \(\|\mathbb{W}\| := \sup_{f \in L^2([0, 1]): \|f\|_{L^2}=1} \|\mathbb{W}f\|_{L^2},\)

\(^7\) That is, \((\nabla_s U(s, z, \theta) - \nabla_s U(s', z, \theta))^\top (s - s') \leq -\alpha_\theta \|s - s'\|^2\) for all \(s, s', z, \theta\).
coincides with the largest eigenvalue of $\mathbb{W}$ which we denote by $\lambda_{\text{max}}(\mathbb{W})$. We next show that if $\lambda_{\text{max}}(\mathbb{W})$ 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

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

where $\ell_U$ and $\alpha_U$ are regularity constants related to the utility function, as defined in Assumption 1, while $\lambda_{\text{max}}(\mathbb{W})$ 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 for example Ballester et al. (2006), and guarantees that the effect of the neighbors aggregate on an agent’s marginal payoff, quantified by $\ell_U \lambda_{\text{max}}(\mathbb{W})$ is not too large with respect to 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 $A^N$, in the case of graphon games the corresponding role is played by the dominant eigenvalue of the graphon, that is, $\lambda_{\text{max}}(\mathbb{W})$. In both cases this quantity captures the maximum amount by which the network/graphon can amplify a unitary vector/function.

**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.\(^8\)

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

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

\(^8\)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.
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

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

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 \( |\alpha| < \frac{1}{\lambda_{\text{max}}(W)} \), which is a similar condition as the one derived in Ballester et al. (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 \( z(x \mid s) \), i.e., this is a game of strategic complements [Ballester et al. (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

\[
\bar{s}(x) = \alpha z(x \mid \bar{s}) + \theta \Rightarrow \bar{s}(x) = \alpha(W\bar{s})(x) + \theta \Rightarrow (I\bar{s})(x) = \alpha(W\bar{s})(x) + \theta \Rightarrow ((I - \alpha W)\bar{s})(x) = \theta_1_{[0,1]}(x).
\]

The condition \( |\alpha|\lambda_{\text{max}}(W) < 1 \) implies invertibility of the operator \((I - \alpha W)\). Hence

\[
\bar{s}(x) = \theta((I - \alpha W)^{-1}1_{[0,1]})(x) = \theta \sum_{k=0}^{\infty} \alpha^k(W^k1_{[0,1]})(x) \tag{5}
\]

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, (Lovász, 2012, Chapter 10).

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

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

Starting from $A^{[N]}_w$, a simple sampled network can be obtained by defining the 0-1 adjacency matrix $A^{[N]}_s$ as the adjacency matrix corresponding to a graph with $N$ nodes obtained by randomly connecting nodes $i, j \in [1, N]$ with Bernoulli probability $[A^{[N]}_w]_{ij}$. In the following, we use the symbol $A^{[N]}_W$ for statements that hold for both weighted and simple sampled networks and simply refer to $A^{[N]}_W$ 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 for example 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 $i \in \{1, \ldots, N-1\}$. 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 $A^{[N]}_w$ and $A^{[N]}_s$ are stochastic matrices. The difference between the two is that $A^{[N]}_w \in [0, 1]^{N \times N}$ while $A^{[N]}_s \in \{0, 1\}^{N \times N}$. 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.

**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 $g_{in}$).\(^9\) To generate such a community structure from a graphon, one can partition $[0, 1]$ into $K$ disjoint intervals $\{C_k\}_{k=1}^{K}$, with $|C_k| = \pi_k$, and use the piecewise constant graphon

\(^9\) The parameters $\pi_k$ are exogenous and model the probability that agents are born with type $k$, e.g. 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 for example Jackson and Rogers (2005)).
\[ W_{SBM}(x, y) = \begin{cases} 
  g_{in} & \text{if there exists } k \text{ s.t. } x \in C_k, y \in C_k, \\
  g_{out} & \text{otherwise}. 
\end{cases} \]

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 \( C_1 = [0, 0.75], C_2 = (0.75, 1]\).

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 agent \( i \) 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 the *minmax graphon* \( W_{MM}(x, y) = \min(x,y)(1 - \max(x,y)) \), where \( x \in [0,1] \) denotes the agents 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 \( G^{[N]}(\{S(t^i)\}_{i=1}^N, U, \{\theta(t^i)\}_{i=1}^N, A_W^{[N]}). \)

Figure 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 Example 2 and 3, for different values of \( N \). In both examples, one can observe similarities 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
Figure 3: Three realizations of networks formed according to the two community model described in Example 2 (with $\pi_{\text{red}} = 0.25$, $\pi_{\text{blue}} = 0.75$, $g_{\text{in}} = 0.8$ and $g_{\text{out}} = 0.1$) 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 agents location along the line (blue being one extreme and red the other extreme). Edges between agents that are further apart are in lighter color.
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 \( U_i^{[N]} \) for \( k \in \{1, \ldots, N-1\} \), each of length \( 1/N \). Using this partition, we pair each agent \( i \) in the finite network with the interval \( U_i^{[N]} \). 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

\[
\bar{s}^{[N]}(x) := \bar{s}^{i}, \quad \forall x \in U_i^{[N]}, \forall i \in \{1, \ldots, N\},
\]

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

\[
W^{[N]}(x, y) := A_{ij}^{[N]}, \quad \forall (x, y) \in U_i^{[N]} \times U_j^{[N]}, \forall i, j \in \{1, \ldots, N\}.
\] (6)

We show in Lemma 1 in Appendix A that a vector \( \bar{s}^{[N]} \in \mathbb{R}^N \) is a Nash equilibrium of \( G^{[N]}(\{S^i\}_{i=1}^{N}, U, \{\theta^i\}_{i=1}^{N}, A^{[N]} \) 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 \( G(S^{[N]}, U, \theta^{[N]}, W^{[N]}) \) with payoff function as in (2), set valued function \( S^{[N]}(x) := S^i \) for all \( x \in U_i^{[N]} \), parameter function \( \theta^{[N]}(x) := \theta^i \) for all \( x \in U_i^{[N]} \) 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.\(^{10}\)

\(^{10}\)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.
Assumption 4 (Lipschitz continuity). The graphon $W$ and the parameter function $\theta$ are piece-wise Lipschitz over the same partition of $[0, 1]$. Moreover, there exists $\theta_{\text{max}}$ such that $\|\theta(x)\| \leq \theta_{\text{max}}$ for all $x \in [0, 1]$.

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 statements on convergence of equilibria of sampled network games to equilibria of the corresponding graphon game hold in probability.

Theorem 2 (Convergence of equilibria). Consider a graphon game $G$ in which each player has homogeneous strategy sets, that is, $S(x) = S$ for all $x \in [0, 1]$. Suppose that $G$ 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 $G_{[N]}$, as introduced in Section 3.2. Then

1. $\|\bar{s}_{W}^{[N]} - \bar{s}\|_{L^2} \to 0$ almost surely when $N \to \infty$;

2. for any $0 < \delta \leq e^{-1}$, with probability at least $1 - \frac{2\delta}{N}$ it holds

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

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 $\mathbb{P}\left[\|\bar{s}_{W}^{[N]} - \bar{s}\|_{L^2} > \epsilon\right] \leq \frac{2N}{\exp\left(\frac{N^2}{2}\right)}$).

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

---

11Rigorously we assume that there exists a constant $L > 0$ and a sequence of non-overlapping intervals $\mathcal{I}_k$, $k \in \{1, \ldots, \Omega + 1\}$, 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 $x, x' \in \mathcal{I}_k$. 

17
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 online Appendix D.

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 $\sqrt{\frac{\#\text{ edges}}{\#\text{ nodes}}} \to 0$. To this end, we introduce a sparsity parameter $\kappa_N$ and consider the following (generalized) procedure to sample simple networks from a graphon, see e.g. Borgs et al. (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 $A^{[N]}_s$ 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 $A^{[N]}_s$ is of order $\kappa_N N$. Hence for these sampled networks $\sqrt{\frac{\#\text{ edges}}{\#\text{ nodes}}} \approx \sqrt{\kappa_N}$ converges to zero if $\kappa_N \to 0$. In the following, we will require that $\lim_{N \to \infty} \frac{\log(N)}{N\kappa_N} = 0$. 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 
\[ z^i(s) = \frac{1}{N} \sum_{j=1}^{N} [A^{[N]}_s]_{ij} s^j \]
as introduced in Section 3.2 then we may have that \( z^i(s) \to 0 \) as \( N \) grows larger, leading to vanishing network effects. To overcome this issue, we need to scale the local aggregate \( \sum_{j=1}^{N} [A^{[N]}_s]_{ij} s^j \) by the expected order of neighbors which according to Definition 5 is \( \kappa N \) instead of \( N \). Overall, we can define a sampled network game exactly as in Section 3.2, but using as aggregate
\[ z^i_{\kappa}(s) = \frac{1}{\kappa N} \sum_{j=1}^{N} [A^{[N]}_s]_{ij} s^j. \] (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 
\[ z^i(s) = \frac{1}{N} \sum_{j=1}^{N} [A^{[N]}_s]_{ij} s^j, \]
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
\[ z^i_d(s) := \frac{\sum_{j=1}^{N} [A^{[N]}_d]_{ij} s^j}{\sum_{j=1}^{N} [A^{[N]}_d]_{ij}}. \]

Our results can be extended to this setting. The first step is to define the local average for a continuum of agents as 
\[ z_d(x \mid s) := \frac{\int_0^1 W(x,y)s(y)dy}{\int_0^1 W(x,y)dy}. \]
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 \( \mathbb{W}_d : L^2([0,1]) \to L^2([0,1]) \) 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 online Appendix D.2 that all the results derived in Section 2 about
existence and uniqueness of the graphon equilibrium continue to hold.\textsuperscript{12} For example, uniqueness holds if $\frac{\ell_W}{\alpha_U} \cdot \frac{\lambda_{\text{max}}(W)}{\delta_{\text{min}}}$ $< 1$. For the convergence result in Theorem 2, the key step is to show that the distance between the normalized operators $W_d$ and $W_d^{[N]}$ (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 online Appendix D.2.

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 need to use $\|W_d\|$ instead of $\lambda_{\text{max}}(W)$. 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 the online Appendix D.3). 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

\textsuperscript{12}Since $W_d$ is not symmetric, results need to be stated in terms of $\|W_d\|$ instead of $\lambda_{\text{max}}(W)$. Note however that the bound $\|W_d\| \leq \frac{\lambda_{\text{max}}(W)}{\delta_{\text{min}}}$ holds (see online Appendix D.2).
of how to optimally allocate the available budget among the agents to maximize welfare. Mathematically, the central planner aims at solving the following optimization problem

\[
T_{\text{opt}}^{[N]} := \max_{\hat{\theta}^{[N]} \in \mathbb{R}^N} T^{[N]}(\hat{\theta}^{[N]}) = \frac{1}{N} \sum_{i=1}^{N} U(s_i^{[N]}, \bar{z}_i^{[N]}, \theta_i^{[N]} + \hat{\theta}_i^{[N]})
\]

\[
\text{s.t. } \bar{s}^{[N]} = \text{Nash equilibrium of } G^{[N]}(S, U, \theta^{[N]} + \hat{\theta}^{[N]}, A^{[N]}), \bar{z}^{[N]} = A^{[N]} \bar{s}^{[N]}, \frac{1}{N} \|\hat{\theta}^{[N]}\|_2^2 \leq C,
\]

(9)

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 \(A^{[N]} = A^{[N]}_W\) 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

\[
\theta^* \in \arg \max_{\hat{\theta} \in L^2([0,1])} T(\hat{\theta}) = \int_0^1 U(\bar{s}_\hat{\theta}(x), \bar{z}_\hat{\theta}(x), \theta(x) + \hat{\theta}(x))dx
\]

\[
\text{s.t. } \bar{s}_\hat{\theta} = \text{Nash equilibrium of } G(S, U, \theta + \hat{\theta}, W), \bar{z}_\hat{\theta}(x) = W \bar{s}_\hat{\theta}(x) \|\hat{\theta}\|_{L^2}^2 \leq C.
\]

(10)

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,

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

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

---

\(^{13}\)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.
Theorem 3. Consider a sampled network game $G^{[N]}$ 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 $1 - \frac{2\delta}{N}$,

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

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 $\mathbb{P}\left[T_{opt}^{[N]} - T^{[N]}(\hat{\theta}_{\text{graphon}}^{[N]}) > \epsilon\right] \leq \frac{2N}{\exp(2\epsilon^2 N)}$. 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 ($A^{[N]}_W$), 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 $Q = [g_{in}I_K + g_{out}(1_K1_K^T - I_K)]$). Let $C_k$ be the subset of $[0, 1]$ associated with community $k$, with $|C_k| = \pi_k$ and $\sum_k \pi_k = 1$, and construct the stochastic block model graphon

$$W_{SBM}(x, y) = Q_{ij}$$

for all $x \in C_i, y \in C_j$.

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

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

(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 online Appendix shows that $W_{SBM}$ and $E$ have the same eigenvalues and the eigenfunctions of $W_{SBM}$ are piecewise constant over the partition $\{C_k\}_{k=1}^K$, with constant value in each community $C_k$ 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 et al. (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 one graphon.

Example 5 (Rank one graphon). Consider a graphon $W_v(x, y) = v(x)v(y)$ for some non-negative 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 one with eigenfunction $v(x)$ and eigenvalue $\lambda = \|v\|_{L^2}^2$. For example, by choosing $v = \frac{1}{\sqrt{2(x+0.5)^7}}$ with $\gamma \in (0, 1/2)$ one obtains a rank one
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

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

For simplicity we focus on games with strategic complementarities (i.e. with $\alpha > 0$). Following Galeotti et al. (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

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

While Galeotti et al. (2020) focus 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 $0 < \alpha < \frac{1}{\lambda_{\text{max}}(W)}$ and assume that $W$ has finite rank $R < \infty$, let $\mathcal{K}$ be the kernel of $W$ and $\{\psi_r\}_{r=1}^R$ be an orthonormal basis of $\mathcal{K}^\perp$ composed of eigenfunctions of $W$ corresponding to the eigenvalues $\{\lambda_r\}_{r=1}^R$. Set $b_r = \langle \theta, \psi_r \rangle$ for all $r = 1, \ldots, R$ 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\(^{14}\)

$$\max_{[b_0, \ldots, b_R]} \frac{1}{2} \sum_{r=0}^R \frac{(b_r + \hat{b}_r)^2}{(1 - \alpha \lambda_r)^2},$$  \hspace{1cm} (14)

s.t. \quad $\sum_{r=0}^R \hat{b}_r^2 \leq C$.

\(^{14}\)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 Problem (10) is limited to dimension $R + 2$. 

24
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., $\theta(x) = \theta^\text{com}_k$ for all $x \in C_k$) then it is immediate to see that at the graphon equilibrium each agent belonging to the same community has the same strategy $\bar{s}^\text{com}_k$ and the vector of such strategies $\bar{s}^\text{com} \in \mathbb{R}^K$ satisfies the relation

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

Turning to the optimal intervention problem (10), we note that since $\theta(x) = \theta^\text{com}_k$ for all $x \in C_k$, $\theta$ 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 that $\theta^* = \sum_{r=1}^\mathcal{R} \hat{b}^*_r \psi_r$, 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 number 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 non-linear payoffs

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

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

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

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 one graphon constructed as in Example 5. It is important to remark that in this example the best response is non-linear (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 dataset of 80 network games (which for example 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

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

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 4 different communities (e.g., 4 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 $\theta^i_{com}$ for community $h = 1, \ldots, 4$.

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
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 4 communities (the color 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 community 1 and 4 are tightly intra-connected, while 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).

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 is 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 is 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, α, or about the parameters of the network formation model. 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 square regression given the Nash equilibrium relation $\bar{s}^{[N]} = \left( I_N - \frac{\alpha_\kappa}{N} A_s^{[N]} \right)^{-1} \theta^{[N]}$ leading to an estimator for $\alpha_\kappa$ which we denote by $\hat{\alpha}_\kappa^{DRD}$.  

Using $A_s^{[N]}$, $\theta^{[N]}$ and $\hat{\alpha}_\kappa^{DRD}$, 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 et al. (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.

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 is 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). The central planner can then obtain an estimator $\hat{\alpha}_\kappa^{ARD}$ by performing least square regression on the graphon equilibrium relation $\hat{s}^{ARD} = (I_K - \hat{\alpha}_\kappa^{ARD} \hat{E}_\kappa^{ARD})^{-1} \theta^{\text{com}}$ 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 aggre-

---

15 We here assume that the central planner knows the sampled networks are generated from a stochastic block model with 4 communities, but our result could be extended to cases when the number of communities are unknown.

16 In our simulated data we assumed no noise hence this procedure allows the central planner to recover $\alpha_\kappa$ exactly.

17 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 dataset and will have performances that are in between the network optimal and network heuristic.

18 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).
gated relational data. Based on $\theta_{com}, \hat{\alpha}_{\kappa}^{ARD}$ and $\hat{E}_{\kappa}^{ARD}$ 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 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 $K = 4$. 

Figure 7: Interventions for a sampled network of size $N = 300$ (left) and $N = 600$ (right).
3. Optimality: the following table 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.

<table>
<thead>
<tr>
<th>Case study 1 ( (N = 300) )</th>
<th>Case study 2 ( (N = 600) )</th>
<th>Case study 3 ( (N = 1200) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avg. Impr. Network Optimal</strong></td>
<td>25.6% (±6.6) ( [15.4;61.5]% )</td>
<td>23.0% (±3.6) ( [19.0;37.8]% )</td>
</tr>
<tr>
<td><strong>Avg. Impr. Graphon Optimal</strong></td>
<td>22.8% (±7.1) ( [5.6;59.7]% )</td>
<td>20.8% (±3.9) ( [14.8;35.7]% )</td>
</tr>
<tr>
<td><strong>Avg. Impr. Network Heuristic</strong></td>
<td>12.6% (±15.7) ( [-12.6;60.73]% )</td>
<td>4.5% (±12.3) ( [-12.2;33.9]% )</td>
</tr>
</tbody>
</table>

Table I: Comparison of network optimal (NO), graphon optimal (GO) and network heuristic (NH) intervention for the case study described in Section 7. The average improvement is computed as \( \text{Avg. Impr} = \frac{1}{80} \sum_{g=1}^{80} \left( \frac{\text{welfare according to NO/NH/GO intervention in network } g}{\text{welfare according to homogeneous intervention in network } g} - 1 \right) \), one standard deviation is reported in round brackets, minimum and maximum are reported in square brackets. We also show the average degree to illustrate that the graphon optimal (GO) 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 case studies, we used \( C = 0.02N, \alpha = 2.65, \theta^{com} = [0.1, 0.1, 0.1, 0.25] \) and aggregated relational surveys completed by 10% of the agents in each network.

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 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 type of interventions, such as selecting the key player as introduced in Ballester et al. (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 online Appendix C, 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.

Appendices

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 $s : [0, 1] \to \mathbb{R}^n$ is a *vector valued function*. In other words, $s(x) = [s_1(x), \ldots, s_n(x)]^\top$ for all $x \in [0, 1]$. In the following, we require that any strategy profile is square integrable, that is $s(x) \in L^2([0, 1]; \mathbb{R}^n)$.\footnote{This implies that each component is square integrable, that is, $s_k(x) \in L^2([0, 1])$ for all $k \in \{1, \ldots, n\}$. For any $k \in \{1, \ldots, n\}$ it holds $\|s_k\|_{L^2} = \sqrt{\int_0^1 s_k(x)^2 dx} \leq \sqrt{\int_0^1 \|s(x)\|^2 dx} = \|s\|_{L^2}\mathbb{R}^n$.} For any
strategy function \( s \in L^2([0,1];\mathbb{R}^n) \), the corresponding local aggregate is

\[
z(x \mid s) := \int_0^1 W(x,y)s(y)dy := \begin{bmatrix}
\int_0^1 W(x,y)s_1(y)dy \\
\vdots \\
\int_0^1 W(x,y)s_n(y)dy
\end{bmatrix} = \begin{bmatrix}
(Ws_1)(x) \\
\vdots \\
(Ws_n)(x)
\end{bmatrix} =: (W_n s)(x),
\]

where \( W_n : L^2([0,1];\mathbb{R}^n) \to L^2([0,1];\mathbb{R}^n) \) is defined by applying \( W \) component-wise. Note that \( s(x) \in L^2([0,1];\mathbb{R}^n) \Rightarrow z(x \mid s) \in L^2([0,1];\mathbb{R}^n) \). We additionally define the operator \( B_\theta : L^2([0,1];\mathbb{R}^n) \to L^2([0,1];\mathbb{R}^n) \) point-wise as

\[
(B_\theta z)(x) := \arg \max_{s \in \mathcal{U}(x)} U(\bar{s},z(x),\theta(x)), \quad (A.1)
\]

where \( z(x) \) is any function of \( L^2([0,1];\mathbb{R}^n) \) (i.e., not necessarily \( z(x \mid s) \)). In words, \( (B_\theta z)(x) \) 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 (A.1) has a unique solution. The fact that, under the given assumptions, the codomain of the best response operator \( B_\theta \) is \( L^2([0,1];\mathbb{R}^n) \) will be proven in the online Appendix D.1. Overall, a strategy profile \( \bar{s} \in L^2([0,1];\mathbb{R}^n) \) is a Nash equilibrium if and only if

\[
\bar{s} = B_\theta W_n \bar{s}, \quad (A.2)
\]

that is, the function \( \bar{s} \) is a fixed point of the composite operator \( B_\theta W_n \), 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 \( B_\theta W_n : L_S \to K := (B_\theta W_n(L_S))^{cl} \), where \( L_S := \{ f \in L^2([0,1];\mathbb{R}^n) \mid \|f\|_{L^2;\mathbb{R}^n} \leq s_{\text{max}} \} \). We proceed in 4 steps: i) the set \( L_S \) is non-empty, convex, closed and bounded. ii) In Lemma 5 and 6 in the online Appendix D.1 we prove that both \( W_n \) and \( B_\theta \) are continuous operators, hence \( B_\theta W_n \) is continuous. iii) We show \( K \subseteq L_S \). In fact, since \( W_n : L_S \to L^2([0,1];\mathbb{R}^n) \) and \( B_\theta : L^2([0,1];\mathbb{R}^n) \to L_S \) it holds \( B_\theta W_n(L_S) \subseteq L_S \). Since \( L_S \) is closed, \( K = (B_\theta W_n(L_S))^{cl} \subseteq (L_S)^{cl} = L_S \). iv) Finally, we show that \( K \) is compact. To this end note that \( W_n \) is a compact operator by Lemma 5 in the online Appendix D.1 and \( L_S \) is bounded, hence \( (W_n(L_S))^{cl} \) is compact (Hutson et al.,
2005, Definition 7.2.1). We prove in Lemma 6 in the online Appendix D.1 that $\mathbb{B}_\theta$ is Lipschitz (and thus continuous), consequently $\mathbb{B}_\theta(\mathbb{W}_n(L_S))$ is compact (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 $K := (\mathbb{B}_\theta(\mathbb{W}_n(L_S)))^{cl} \subseteq (\mathbb{B}_\theta(\mathbb{W}_n((L_S))^{cl}) = \mathbb{B}_\theta(\mathbb{W}_n((L_S))^{cl}))$. $K$ is thus a closed subset of a compact set, which implies that $K$ is compact (Aliprantis and Border, 2006, pg.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 (Smart, 1974, Theorem 4.3.4). For any $f, g \in L^2([0,1];\mathbb{R}^n)$,

$$
\|\mathbb{B}_\theta \mathbb{W}_n f - \mathbb{B}_\theta \mathbb{W}_n g\|_{L^2;\mathbb{R}^n} \leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n f - \mathbb{W}_n g\|_{L^2;\mathbb{R}^n} = \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n (f - g)\|_{L^2;\mathbb{R}^n}
$$

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

where we used Lemma 6, given in the online 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.

**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

$$
\|\tilde{s} - \bar{s}\|_{L^2;\mathbb{R}^n} \leq \frac{1/\alpha_U}{1 - \ell_U/\alpha_U \lambda_{\max}(\mathbb{W})} \left(\ell_U \|\mathbb{W} - \tilde{W}\| s_{\max} + \ell_\theta \|\theta - \tilde{\theta}\|_{L^2;\mathbb{R}^m}\right). \quad (A.3)
$$

**Proof.** To prove that (A.3) holds, note that $\bar{s} = \mathbb{B}_\theta \mathbb{W}_n \bar{s}$ and $\tilde{s} = \mathbb{B}_{\tilde{\theta}} \tilde{W}_n \tilde{s}$ hence

$$
\|\tilde{s} - \bar{s}\|_{L^2;\mathbb{R}^n} = \|\mathbb{B}_{\tilde{\theta}} \tilde{W}_n \tilde{s} - \mathbb{B}_\theta \mathbb{W}_n \bar{s}\|_{L^2;\mathbb{R}^n} \leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \tilde{s} - \tilde{W}_n \bar{s}\|_{L^2;\mathbb{R}^n} + \ell_\theta \|\theta - \tilde{\theta}\|_{L^2;\mathbb{R}^m}
$$

$$
\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \tilde{s} - \mathbb{W}_n \bar{s}\|_{L^2;\mathbb{R}^n} + \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n \bar{s} - \tilde{W}_n \bar{s}\|_{L^2;\mathbb{R}^n} + \ell_\theta \|\theta - \tilde{\theta}\|_{L^2;\mathbb{R}^m}
$$

$$
\leq \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n\| \|\tilde{s} - \bar{s}\|_{L^2;\mathbb{R}^n} + \frac{\ell_U}{\alpha_U} \|\mathbb{W}_n - \tilde{W}_n\| \|\bar{s}\|_{L^2;\mathbb{R}^n} + \ell_\theta \|\theta - \tilde{\theta}\|_{L^2;\mathbb{R}^m}
$$

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


34
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 $\|\mathbb{W}_n - \mathbb{W}\| = \|\mathbb{W} - \mathbb{W}\|$. The conclusion follows from the fact that $1 - \ell_U/\alpha_U \lambda_{\max}(\mathbb{W}) > 0$ by Assumption 3 hence

$$
\|\bar{s} - \bar{s}\|_{L^2;\mathbb{R}^n} \leq \frac{1}{1 - \frac{\ell_U}{\alpha_U} \lambda_{\max}(\mathbb{W})} \left( \frac{\ell_U}{\alpha_U} \|\mathbb{W} - \mathbb{W}\| \|\bar{s}\|_{L^2;\mathbb{R}^n} + \frac{\ell_\theta}{\alpha_U} \|ar{\theta} - \tilde{\theta}\|_{L^2;\mathbb{R}^m} \right) \quad (A.4)
$$

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

\medskip

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 a special cases for $\delta_N = \delta/N$ 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 $\tilde{s}_{\{N\}}(x) := \mathcal{S}^i$ for all $x \in \mathcal{U}^\{N\}_i$, parameter function $\theta_{\{N\}}(x) := \theta^i$ for all $x \in \mathcal{U}^\{N\}_i$, and step function graphon $W_{\kappa}^\{N\}$ corresponding to $A_{\kappa^\{N\}}$.

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 $\bar{z}^i_\kappa$ be the value of $\bar{z}_\kappa(x)$ in $\mathcal{U}^\{N\}_i$ and recall that $\theta_{\{N\}}(x) = \theta^i$ in $\mathcal{U}^\{N\}_i$. From the definition of Nash equilibrium for the graphon game

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

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

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

35
The latter is the definition of Nash equilibrium in the sampled network game with network \( A^{[N]} \), thus concluding the proof.

**Theorem 2 (generalized).** Consider a graphon game \( G(S, U, \theta, W) \) in which each player has homogeneous strategy set, i.e., \( S(x) = S \) for all \( x \in [0, 1] \). Suppose that \( G \) 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 \( \frac{\log(N/\delta_N)}{N\kappa_N} \to 0 \). Let \( \bar{s}^{[N]}_W \) be an arbitrary step function equilibrium of the sampled network game corresponding to either the sampled network \( A^{[N]}_w \) (as in Section 3.2) or \( A^{[N]}_s \) (as in Section 5.1). Then

1. \( \| \bar{s}^{[N]}_W - \bar{s} \|_{L^2; \mathbb{R}^n} \to 0 \) almost surely when \( N \to \infty \);

2. with probability at least \( 1 - 2\delta_N \)

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

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

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

**Proof.** We start by proving statement 2. Let \( \theta^{[N]} \) be the step function corresponding to the vector \( \left[ \theta(t_i) \right]_{i=1}^{N} \) and \( W^{[N]}_w, W^{[N]}_s \) be the step function graphons corresponding to \( A^{[N]}_w \) and \( A^{[N]}_s \), respectively, so that \( \bar{s}^{[N]}_W \) and \( \bar{s}^{[N]}_s \) are the equilibria of the graphon games played over the graphon \( W^{[N]}_w \) and \( W^{[N]}_s \) with parameter function \( \theta^{[N]} \). By Proposition 3 it follows

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

The bound on \( \| \bar{s}^{[N]}_W - \bar{s} \|_{L^2; \mathbb{R}^n} \) follows from (A.5) and the fact that for \( N \) large enough with probability at least \( 1 - 2\delta_N \)

\[
\| \theta^{[N]} - \theta \|_{L^2; \mathbb{R}^m} \leq \rho_{\theta}(N) \quad \text{and} \quad \| W^{[N]}_W - W \| \leq \rho_{W}(N), \quad (A.6)
\]

for \( \rho_{\theta}(N) \) and \( \rho_{W}(N) \) as defined in Lemma 12 in online Appendix D.5. The bounds in (A.6) are proven in (Avella-Medina et al., 2018, Theorem 1) (reported
in Lemma 12 in online 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

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

- To prove statement 1, let us define the infinite sequence of events \( \mathcal{E}_N := \left\{ \| \bar{s}^{[N]}_W - \bar{s} \|_{L^2;\mathbb{R}^n} > \Theta_N \right\} \). It follows by the previous point that \( \Pr[\mathcal{E}_N] < 2\delta_N \). Note that \( \frac{\log(N/\delta_N)}{N} \to 0 \) implies \( \frac{\log(N)}{\kappa N N} \to 0 \) (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 Borel-Cantelli lemma there exists a positive integer \( \bar{N} \) such that for all \( N \geq \bar{N} \), the complement of \( \mathcal{E}_N \), i.e., \( \| \bar{s}^{[N]}_W - \bar{s} \|_{L^2;\mathbb{R}^n} \leq \Theta_N \), holds a.s. Since, \( \Theta_N \to 0 \) we obtain \( \| \bar{s}^{[N]}_W - \bar{s} \|_{L^2;\mathbb{R}^n} \to 0 \) 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

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

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

\( \square \)

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 equivalent reformulation given
in Section 4.1. So that

\[
T_{\text{opt}}^{[N]} := \max_{\hat{\theta}^{[N]} \in \mathcal{L}^{[N]}} T^{[N]}(\hat{\theta}^{[N]}) = \int_0^1 U(\bar{s}^{[N]}_{\hat{\theta}^{[N]}}(x), \bar{z}^{[N]}_{\hat{\theta}^{[N]}}(x), \theta^{[N]}(x) + \hat{\theta}^{[N]}(x))
\]

\[
\text{s.t. } \bar{s}^{[N]}_{\hat{\theta}^{[N]}} = \text{Nash equilibrium of } \mathcal{G}(\mathcal{S}, U, \theta^{[N]} + \hat{\theta}^{[N]}, W^{[N]}), \quad (A.7)
\]

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 \(\{U_i^{[N]}\}_{i=1}^{N}\) 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 \(\hat{\theta}^{[N]}_{\text{opt}}\) be an optimizer of Problem (A.7) and note that

\[
T^{[N]}(\hat{\theta}^{[N]}_{\text{graphon}}) \geq T(\theta^*) - |T^{[N]}(\hat{\theta}^{[N]}_{\text{graphon}}) - T(\theta^*)| \geq T(\hat{\theta}^{[N]}_{\text{opt}}) - T_1
\]

\[
\geq T^{[N]}(\hat{\theta}^{[N]}_{\text{opt}}) - T_1 - |T^{[N]}(\hat{\theta}^{[N]}_{\text{opt}}) - T(\hat{\theta}^{[N]}_{\text{opt}})| = T^{[N]}_{\text{opt}} - T_1 - T_2,
\]

where we used \(T(\theta^*) \geq T(\hat{\theta}^{[N]}_{\text{opt}})\) 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

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

and similarly

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

Note that for any function \(\hat{\theta}_1, \hat{\theta}_2 \in L^2([0, 1])\) by the proof of Proposition 3 (Equation (A.4)) and Lemma 12, with probability \(1 - 2 \delta_N\), there exists \(K > 0\) s.t.

\[
\|\bar{s}^{[N]}_{\hat{\theta}_1} - \bar{s}^{[N]}_{\hat{\theta}_2}\|_{L^2} \leq K \left(\left\|\hat{W}^{[N]}_{\hat{\theta}_1} - \hat{W}^{[N]}_{\hat{\theta}_2}\right\|_{L^2} + \|\theta^{[N]}_{\hat{\theta}_1} + \hat{\theta}_1 - (\theta + \hat{\theta}_2)\|_{L^2}\right)
\]

\[
\leq K \left(\rho_W(N) s_{\text{max}} + \|\theta^{[N]} - \theta\|_{L^2} + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}\right)
\]

\[
\leq K \left(\rho_W(N) s_{\text{max}} + \rho_0(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}\right) =: K \left(\rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2}\right).
\]
where $\rho_W, \rho_0$ are as defined in Lemma 12. Similarly
\[
\|z^{[N]}_{\hat{\theta}_1} - \bar{z}_{\hat{\theta}_2}\|_{L^2} = \|W[N] s_{\hat{\theta}_1}^{[N]} - W s_{\hat{\theta}_2}\|_{L^2} \leq \|W[N] s_{\hat{\theta}_1}^{[N]} - W s_{\hat{\theta}_1}\|_{L^2} + \|W s_{\hat{\theta}_1} - W s_{\hat{\theta}_2}\|_{L^2} \\
\leq \|W[N] - W\| \|s_{\hat{\theta}_1}^{[N]}\|_{L^2} + \|W\| \|s_{\hat{\theta}_1}^{[N]} - \bar{s}_{\hat{\theta}_2}\|_{L^2} \\
\leq \rho_W(N) s_{\max} + \lambda_{\max}(W) K \left( \rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2} \right) \leq K' \left( \rho_M(N) + \|\hat{\theta}_1 - \hat{\theta}_2\|_{L^2} \right),
\]
with $K' = (1 + \lambda_{\max}(W) K)$.

We finally bound $\|\hat{\theta}^{[N]}_{\text{graphon}} - \hat{\theta}^*\|_{L^2}$. To this end, let $\hat{\theta}^{[N]}$ be the piecewise function corresponding to $[\theta^*(t')]_{i=1}^N$, so that i) by Lemma 12 with the same probability $1 - 2\delta_N$, $\|\hat{\theta}^{[N]} - \theta^*\|_{L^2} \leq \rho^*_0(N) := \sqrt{(L^* d_N^*)^2 + 4\Omega^* d_N^* \theta_{\max}^*}$ (where $L^*$ is the Lipschitz constant of $\theta^*$, $\Omega^*$ is the number of points in which $\theta^*$ is not Lipschitz continuous and $\theta^*(x) \leq \theta_{\max}^*$ for all $x$) and ii) by definition $\hat{\theta}^{[N]}_{\text{graphon}} = \hat{\theta}^{[N]} - \frac{\sqrt{C}}{\|\hat{\theta}^{[N]}\|_{L^2}} = \hat{\theta}^{[N]} - \|\theta^*\|_{L^2} \frac{\|\hat{\theta}^{[N]}\|_{L^2}}{\|\theta^*\|_{L^2}}$. Overall
\[
\|\hat{\theta}^{[N]}_{\text{graphon}} - \theta^*\|_{L^2} \leq \|\hat{\theta}^{[N]}_{\text{graphon}} - \hat{\theta}^{[N]}\|_{L^2} + \|\hat{\theta}^{[N]} - \theta^*\|_{L^2} \\
\leq \left( \frac{\|\theta^*\|_{L^2}}{\|\hat{\theta}^{[N]}\|_{L^2}} - 1 \right) \|\hat{\theta}^{[N]}\|_{L^2} + \rho^*_0(N) \\
= \|\hat{\theta}^*\|_{L^2} - \|\hat{\theta}^{[N]}\|_{L^2} + \rho^*_0(N) \leq \|\theta^* - \hat{\theta}^{[N]}\|_{L^2} + \rho^*_0(N) = 2\rho^*_0(N).
\]
Hence $T_1 + T_2 \leq 2L \left( \sqrt{(K'^2 + K^2)(\rho_M(N) + 2\rho^*_0(N))^2 + (2\rho^*_0(N))^2 + \rho_0(N)^2} \right) =: \rho_T(N) = O \left( \left( \frac{\log(N/\delta_N)}{N} \right)^{\frac{1}{2}} \right)$. The result follows setting $\delta_N = \frac{\delta}{N}$.

**Lemma 2.** Suppose that $U(s, z, \theta)$ is jointly Lipschitz continuous in $[s, z, \theta]$ with constant $L$. For any function $s, z, \theta, \tilde{s}, \tilde{z}, \tilde{\theta} \in L^2([0, 1])$
\[
| \int_0^1 U(s(x), z(x), \theta(x))dx - \int_0^1 U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x))dx | \leq L \sqrt{\|s - \tilde{s}\|_{L^2}^2 + \|z - \tilde{z}\|_{L^2}^2 + \|\theta - \tilde{\theta}\|_{L^2}^2}
\]

**Proof.** $| \int_0^1 U(s(x), z(x), \theta(x))dx - \int_0^1 U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x))dx | \leq \int_0^1 |U(s(x), z(x), \theta(x)) - U(\tilde{s}(x), \tilde{z}(x), \tilde{\theta}(x))|dx \leq L \int_0^1 (\|s(x) - \tilde{s}(x)\|_{L^2}^2 + \|z(x) - \tilde{z}(x)\|_{L^2}^2 + \|\theta(x) - \tilde{\theta}(x)\|_{L^2}^2)dx \leq L \sqrt{\int_0^1 (\|s(x) - \tilde{s}(x)\|_{L^2}^2 + \|z(x) - \tilde{z}(x)\|_{L^2}^2 + \|\theta(x) - \tilde{\theta}(x)\|_{L^2}^2)dx} = L \sqrt{\|s - \tilde{s}\|_{L^2}^2 + \|z - \tilde{z}\|_{L^2}^2 + \|\theta - \tilde{\theta}\|_{L^2}^2}.$
**Proof of Proposition 1:** First note that for linear quadratic games with \( \alpha > 0 \) the payoff at equilibrium is

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

Hence \( T(\hat{\theta}) = \frac{1}{2} \| \hat{s}_\theta \|_{L^2}^2 \). Similarly to the proof of Lemma 3, given next, it can be shown that with probability \( 1 - 2\delta_N \) both \( \| \hat{s}^{[N]}_\theta \|_{L^2} \) and \( \| \hat{s}_\theta \|_{L^2} \) (as defined in (A.7) and in (10), respectively) can be bounded by some constant \( M_* \) for any feasible \( \hat{\theta}^{[N]}, \hat{\theta} \).

Let \( \mathcal{K} \) be the kernel of \( \mathcal{W} \), so that \( \mathcal{W}\psi = 0 \) 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 \( \mathcal{W} \). We denote such basis by \( \{ \psi_r \}_{r=1}^\infty \). Let \( \hat{\theta} \) be a generic function of \( L^2([0,1]) \). Let \( b_0\hat{\psi}_0, \hat{b}_0 \hat{\psi}_0 \) 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 \( \theta = b_0\hat{\psi}_0 + \sum_{r=1}^R b_r \psi_r \) and \( \hat{\theta} = \hat{b}_0 \hat{\psi}_0 + \sum_{r=1}^R \hat{b}_r \psi_r \). Using the fact that \( \{ \psi_r \}_{r=1}^\infty \) are orthogonal to each other and orthogonal to any function in \( \mathcal{K} \) yields \( \| \hat{\theta} \|_{L^2}^2 = \| \hat{b}_0 \hat{\psi}_0 + \sum_{r=1}^R \hat{b}_r \psi_r \|_{L^2}^2 = \| \hat{b}_0 \hat{\psi}_0 \|_{L^2}^2 + \sum_{r=1}^R \hat{b}_r^2 = \sum_{r=1}^R \hat{b}_r^2 \). Since we are considering linear quadratic games, from Example 1, the equilibrium induced by \( \theta + \hat{\theta} \) can be rewritten as

\[
\hat{s}_\theta = (I - \alpha \mathcal{W})^{-1}(\theta + \hat{\theta}) = \sum_{r=1}^\infty \alpha^r \mathcal{W}^r (b_0\psi_0 + \hat{b}_0 \hat{\psi}_0 + \sum_{r=1}^R (b_r + \hat{b}_r) \psi_r)
\]

\[
= \sum_{r=1}^\infty \alpha^r \mathcal{W}^r (b_0\psi_0 + \hat{b}_0 \hat{\psi}_0) + \sum_{r=1}^R (b_r + \hat{b}_r) \sum_{s=0}^\infty \alpha^s \mathcal{W}^s \psi_r
\]

\[
= (b_0\psi_0 + \hat{b}_0 \hat{\psi}_0) + \sum_{r=1}^R (b_r + \hat{b}_r) \sum_{s=0}^\infty \alpha^s \mathcal{W}^s \psi_r
\]

\[
= (b_0\psi_0 + \hat{b}_0 \hat{\psi}_0) + \sum_{r=1}^R (b_r + \hat{b}_r) \sum_{s=0}^\infty \alpha^s \mathcal{W}^s \psi_r
\]

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

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

**Proof.** Fix \( \bar{k} \in (1, \frac{1}{\alpha \max(\mathcal{W})}) \) (this interval has non empty interior by assumption). Since by Lemma 12 (given in online Appendix D.5) with probability \( 1 - 2\delta_N \) for
Lemma 4. Consider a graphon game

\[ G([0,1], U, \theta + \hat{\theta}, W_v) \]

with payoff function \( U \) as in (16) and graphon \( W_v \) 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 $W_\varepsilon$ 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 $\hat{\theta} = \tilde{b}_1 \psi_1 + \tilde{b}_2 \psi_2 + \tilde{b}_0 \psi_0 + \tilde{b} \tilde{\psi}$ with $\|\tilde{\psi}\|_{L^2} = 1$, $\langle \tilde{\psi}, \psi_i \rangle = 0$, for all $i = 0, 1, 2$.\footnote{Such decomposition can be obtained by using the Gram-Schmidt process to construct an orthonormal basis.}

Suppose $b_1 + \tilde{b}_1 > 0$. Then

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

with $s_1 = \frac{-\sqrt{1-\alpha^2+4\alpha \lambda \psi}}{2\alpha \lambda \psi}$, $s_2 = -\alpha s_1 \lambda p_2 + b_2 + \tilde{b}_2$ and $s_0 = b_0 + \tilde{b}_0$.

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

$$
\bar{s}_\theta(x) = \Pi_{[0,1]}[\alpha \bar{z}_\theta(x)(1 - \bar{z}_\theta(x)) + \theta(x) + \hat{\theta}(x)]. \tag{A.8}
$$

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 (A.8) is always in $[0, 1]$. Consequently, (A.8) can be rewritten as

$$
\bar{s}_\theta(x) = \alpha \bar{z}_\theta(x)(1 - \bar{z}_\theta(x)) + \theta(x) + \hat{\theta}(x). \tag{A.9}
$$

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, $\bar{z}_\theta^2(x) = \xi^2 \lambda v^2(x) = \xi^2 \lambda [p_1 \psi_1(x) + p_2 \psi_2(x)]$. Substituting this notation in (A.9) we obtain

$$
\bar{s}_\theta = \alpha \bar{z}_\theta - \alpha \bar{z}_\theta^2 + \theta + \hat{\theta} = \alpha \xi \lambda \psi_1 - \alpha \xi^2 \lambda [p_1 \psi_1 + p_2 \psi_2] + (b_0 + \tilde{b}_0) \psi_0 + (b_1 + \tilde{b}_1) \psi_1 + (b_2 + \tilde{b}_2) \psi_2 + \tilde{b} \tilde{\psi} = \left(\alpha \xi \lambda - \alpha \xi^2 \lambda p_1 + b_1 + \tilde{b}_1\right) \psi_1 + \left(-\alpha \xi^2 \lambda p_2 + b_2 + \tilde{b}_2\right) \psi_2 + \left(b_0 + \tilde{b}_0\right) \psi_0 + \tilde{b} \tilde{\psi} \tag{A.10}
$$
Plugging this expression in \( \xi := \langle \psi_1, \bar{s}_\theta \rangle \) we obtain

\[
\xi = \left( \alpha \xi \lambda - \alpha^2 \lambda p_1 + b_1 + \hat{b}_1 \right) \Leftrightarrow \alpha \lambda p_1 \xi^2 + (1 - \alpha \lambda) \xi - (b_1 + \hat{b}_1) = 0
\]

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

\( \square \)

REFERENCES


Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave n-person