Les Cahiers du GERAD ISSN: 0711–2440

Stochastic graphon field tracking games with finite and infinite horizons

A. Dunyak, P. E. Caines

G–2024–68

Octobre 2024

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2024
– Bibliothèque et Archives Canada, 2024

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2024
– Library and Archives Canada, 2024

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Stochastic graphon field tracking games with finite and infinite horizons

Alex Dunyak [a, b]

Peter E. Caines [a, b]

[a] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[b] *McGill University, Montréal (Qc), Canada, H3A 0B9*

alexander.dunyak@mail.mcgill.ca
peterc@cim.mcgill.ca

Octobre 2024
Les Cahiers du GERAD
G–2024–68

Abstract : Linear quadratic games on very large dense networks are modelled with discrete time linear quadratic graphon field games with Q-noise. In such a game, the agents are interconnected via an undirected network with one agent per node. Brownian motion which is correlated over nodes affects each agent. The limit of the finite-sized linear quadratic network tracking game in discrete time is formulated, and it is shown that under the proper assumptions, the game has a graphon limit system with Q-noise. Then, the optimal control of the discrete time system is found in closed-form and the Nash equilibrium behavior of the game is demonstrated numerically. The infinite time horizon discounted case is also analyzed, and a closed form feedback solution is presented in the special case where the underlying graphon is finite rank.

Keywords: Graphon, Mean Field Games, Infinite dimensional noise

Résumé : Les jeux quadratiques linéaires sur de très grands réseaux denses sont modélisés par des jeux de champs de graphes quadratiques linéaires à temps discret avec un bruit de qualité. Dans un tel jeu, les agents sont interconnectés via un réseau non dirigé avec un agent par nœud. Le mouvement brownien corrélé aux nœuds affecte chaque agent. La limite du jeu de suivi de réseau linéaire quadratique de taille finie en temps discret est formulée, et il est montré que sous les hypothèses appropriées, le jeu a un système limite de graphon avec Q-bruit. Ensuite, le contrôle optimal du système en temps discret est trouvé en forme fermée et le comportement d'équilibre de Nash du jeu est démontré numériquement. Le cas de l'horizon temporel infini actualisé est également analysé, et une solution de rétroaction en forme fermée est présentée dans le cas spécial où le graphon sous-jacent est de rang fini.

Mots clés : Graphon, jeux de champ moyen, bruit de dimension infinie

Acknowledgements: Work supported by *A. Dunyak NSERC 2019–05336, **P.E. Caines: NSERC 2019–05336, ARL W911NF1910110, Air Force OSR Grant FA9550–23–1–0015

1 Introduction

Large systems composed of interacting non-cooperative agents arise in many applications such as cellular networks, financial markets, and electrical grids. The modelling and control of such systems is intractable due to the size and complexity of their respective networks.

In the case of a game with a large number of identical agents, one can use Mean Field Game theory to find the approximate system behavior by simulating a system with an infinite number of agents and finding the distribution of agents' states under different control methods [1,7,12]. The mean field game approach simplifies the game, as rather than optimizing with respect to the actions of every individual player, one can optimize with respect to a single representative player.

Standard mean field games can be extended to games on networks where each node has an infinite population of players by the use of graphons [3–5]. Graphon theory [14] allows the adjacency matrix of an infinitely large graph to be represented as a bounded, symmetric operator, providing a limiting object for large, dense graphs. Under the Graphon Mean Field Games (GMFG) model each node in a network contains a separate, infinite population interacting with the agents local to their node uniformly and the agents in the network through the graphon. As there are an infinite number of agents, the actions of a single agent do not change the mean field.

Expanding on previous work [9], a linear quadratic game with correlated Gaussian disturbances on an infinitely large dense graph is investigated where each node represents a single agent. A continuous time, deterministic model with stochastic initial conditions for this type of linear-quadratic game was investigated by Gao, Foguen-Tchuendom, and Caines [11]. To distinguish this from the infinite-agent-per-node GMFG model, this approach was termed the Graphon Field Game model. As in the GMFG model, the actions of any individual agent do not directly affect the field of the system. The Nash equilibria of such a system requires the optimality of each agent's chosen actions with respect to the field generated by the ensemble of agents' optimal strategies.

This work extends the work of Gao, Foguen-Tchuendom, and Caines [11] by applying the Q-noise foundations of Dunyak and Caines [8] to discrete time systems. This model is analogous to the limit behavior of a finite dimensional graph system with a correlated Gaussian disturbance impacting each node at each time step. It is demonstrated that the discrete time linear quadratic Q-noise tracking game has an adapted Nash equilibrium solution, and the behavior of the equilibrium solution is demonstrated numerically.

Section 3 presents the finite network, finite time field-tracking game as motivation for taking the infinite node limit, called the graphon field-tracking game. The general approach used to solve the graphon field tracking game in (Section IV) is:

1. Use the mean field game principle: a given agent optimizes against an exogenous signal that is indifferent to their own control and state input,
2. Use dynamic programming to find the solution of an arbitrary stochastic tracking problem (part 4.1), and
3. Use the graphon field to recursively generate the tracked signal for each time step (part 4.2).

The second step generates the tracking signal solution in reference to a stochastic adjoint process, which is defined recursively backwards. The third step uses the graphon field to generate the adjoint process for each agent by finding a deterministic backwards operator equation. This is called the Consistency Condition, as it ensures that the graphon field is consistent with each agent's optimizing control input. This behavior is demonstrated numerically in two scenarios: where the maximal eigenvalue of the graphon $\boldsymbol{M}$ is less than one, and where the maximal eigenvalue is equal to one. This causes the state of all players of the game to either stabilize about zero or stabilize about an eigenfunction of the game respectively.

Part V extends the analysis to infinite time horizon games with multiplicative discounting. Unlike in the finite time horizon case, the operator equation solutions for the discounted game have multiple solutions. For the case where the graphon $\boldsymbol{M}$ is low rank, a finite number of closed form solutions are found. It is demonstrated that it is nontrivial to determine which solution is rational for all agents numerically.

For simplicity, the initial formulation is presented where each agent has a scalar state and a scalar control. The extension to games where each agent has multiple states and controls is straightforward, and the notation is presented in Appendix A.1.

2 Preliminaries

2.1 Notation

- The set of vectors of real numbers of dimension m is denoted $\mathcal{R}^m$.
- Graphons (i.e. bounded symmetric $[0,1]^2$ functions used as the kernels of linear integral operators) are denoted in italicized bold capital letters, and in this article is typically written as $\boldsymbol{M}$.
- $L_2[0,1]$ denotes the Hilbert space of real square-integrable functions on the unit interval. In addition, $L_2[0,1]$ is equipped with the standard inner product, denoted $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$. For any function $\boldsymbol{v}$, $\boldsymbol{v}^*$ denotes the adjoint of $\boldsymbol{v}$. As such, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ is sometimes written as $\boldsymbol{v}^*\boldsymbol{u}$.
- The identity operator in both $L_2[0,1]$ and finite dimensional spaces is denoted $\mathbb{I}$.
- A linear integral operator with the kernel $\mathbf{Q} : [0,1]^2 \to \mathcal{R}$ acting on a function $\boldsymbol{f} \in L_2[0,1]$ is defined by

$$(\mathbf{Q}\boldsymbol{f})(x) = \int_0^1 \mathbf{Q}(x,y)\boldsymbol{f}(y)dy, \quad \forall\, x \in [0,1]. \tag{1}$$

- The operators $\mathbf{Q}$ are equipped with the standard $L_2[0,1]$ operator norm $||\mathbf{Q}||_{\text{op}}$.
- A symmetric function $\mathbf{Q} : [0,1]^2 \to \mathcal{R}$ is non-negative if the following inequality is satisfied for every function $\boldsymbol{f} \in L_2[0,1]$,

$$\begin{aligned} 0 &\leq \int_0^1 \int_0^1 \mathbf{Q}(x,y)\boldsymbol{f}^*(x)\boldsymbol{f}(y)dxdy \\ &:= \langle \mathbf{Q}\boldsymbol{f}, \boldsymbol{f} \rangle < \infty. \end{aligned} \tag{2}$$

  Additionally, denote $\mathcal{Q}$ to be the set of bounded symmetric non-negative functions. All valid Q-noise covariance functions are members of $\mathcal{Q}$.
- Discrete time Q-noise processes (stochastic processes over the time interval $(0,1,...,T)$) will be denoted by the bold font $\boldsymbol{g}_k$. For each $k \in (0,1,...,T)$, $\boldsymbol{g}_k$ is an $L_2[0,1]$ function. The precise definition of a Q-noise process is given in Section 2.2.
- The expectation of a random variable at time k with respect to a sigma algebra $\mathcal{F}_k$ is denoted:

$$\mathbb{E}[\cdot|\mathcal{F}_k] := \mathbb{E}_k[\cdot]. \tag{3}$$

This article focuses on systems on games where each agent possesses a scalar state. The extension to vectors of states is straightforward, and is presented in Appendix A.1.

2.2 Discrete time Q-noise processes

Discrete time Q-noise processes are $L_2([0,1])$ valued random processes satisfying the following axioms (modified from [8] for discrete time processes):

1. Let $\mathbf{Q} \in \mathcal{Q}$, and let $([0,1] \times \{0,1,...,T\} \times \Omega, \mathcal{B}([0,1] \times \{0,1,...,T\} \times \Omega), \mathbb{P})$ be a probability space with the measurable random variable $\boldsymbol{g}_k(\alpha,\omega) : [0,1] \times \{0,1,...,T\} \times \Omega \to \mathbb{R}$ for all $k \in \{0,1,...,T\}, \alpha \in [0,1]$, and $\omega \in \Omega$. For notation, ω is suppressed when the meaning is clear.
2. For all $\alpha \in [0,1]$, $\boldsymbol{g}_k(\alpha) \sim \mathcal{N}(0, \mathbf{Q}(\alpha,\alpha))$.
3. For all α and β, $\mathbb{E}[\boldsymbol{g}_k(\alpha)\boldsymbol{g}_k(\beta)] = \mathbf{Q}(\alpha,\beta)$.

An orthonormal basis example: Let $\{W_k^1, W_k^2, \cdots\}$ be a sequence of independent standard normal random variables for each $k \in \{0,1,...,T\}$. Let $\mathbf{Q} \in \mathcal{Q}$ have a diagonalizing orthonormal basis $\{\phi_r\}_{r=1}^{\infty}$ with eigenvalues $\{\lambda_r\}_{r=1}^{\infty}$. Then

$$\boldsymbol{g}_k(\alpha,\omega) = \sum_{r=1}^{\infty} \sqrt{\lambda_r}\phi_r(\alpha) W_k^r(\omega) \tag{4}$$

is a discrete time Q-noise process.

3 Problem statement

3.1 Discrete-time network system games

Consider a discrete time game on a graph $G^N = (V^N, E^N)$ where each node i represents an agent. The state of agent i at time k (denoted x_k^i with control u_k^i) evolves with the following stochastic difference equation:

$$x_{k+1}^i = (ax_k^i + bu_k^i + c\frac{1}{N}\sum_{j=1}^{N} M_{ij}^N x_k^j) + W_k^i \tag{5}$$

where $a, b, c \in \mathcal{R}$, M^N is the weighted adjacency matrix of G^N, and $\{W_k^i\}$ is a collection of Gaussian disturbances with covariance matrix Q^N for each k. Subject to the actions of all other agents, each agent i minimizes the expected quadratic cost function with respect to their information set $\mathcal{F}_k^i$, which is the sigma algebra generated by the set $\{x_k^i, z_k^i\}_{i=0}^N$,

$$J^i(u^i, u^{-i} | \{z_k^i\}_{i=1}^N) \tag{6}$$

$$= \mathbb{E}\left[\sum_{k=0}^{T-1} ||x_k^i - z_k^i||_S^2 + ||u_k^i||_R^2 + ||x_T^i - z_T^i||_S^2 \middle| \mathcal{F}_0^i\right], \tag{7}$$

where $z_k^i = \frac{1}{N}\sum_{j=1}^N M_{ij}^N x_k^j$, $||v||_S^2 = Sv^2$ for some $S \in \mathcal{R}$, $R \in \mathcal{R}$, $S \geq 0$ and $R > 0$.

A Nash Equilibrium of the game exists when no agent can benefit by deviating from its current strategy. If the optimal strategy tuple is $\{u^{i*}\}_{i=1}^N$, this implies

$$J^i(u^{i*}, u^{-i*}) \leq J^i(u^i, u^{-i*}) \quad \forall i \in \{1,...,N\}. \tag{8}$$

As the network size grows, the networked system adjacency matrix M^N approaches its associated *graphon* which is a bounded measurable function mapping $[0,1] \times [0,1] \to [0,1]$, denoted $\boldsymbol{M}$ (see [10, 14]). When the underlying graph is undirected, its graphon is also symmetric. An example of this network convergence is shown with the two finite networks in Fig. 1 converging to the graphon limit (Fig. 2).

3.2 Graphon field tracking games

In graphon analysis, as the size of the network tends to infinity, each agent in the system is associated with a point α on the unit interval. Define the discrete time Q-noise $\boldsymbol{g}_k^{\alpha}$, $\alpha \in [0,1]$, and the resulting discrete time system evolves according to

$$\boldsymbol{x}_{k+1}^{\alpha} = (a\boldsymbol{x}_k^{\alpha} + b\boldsymbol{u}_k^{\alpha} + c\boldsymbol{z}_k^{\alpha}) + \boldsymbol{g}_k^{\alpha}, \tag{9}$$

$$\boldsymbol{z}_k^{\alpha} = \int_0^1 \boldsymbol{M}(\alpha,\beta)\boldsymbol{x}_k^{\beta} d\beta \quad \forall\ \alpha \in [0,1]. \tag{10}$$

The local field for an agent designated by α refers to the value $\boldsymbol{z}_k^{\alpha}$ found using the above integral and expectation.

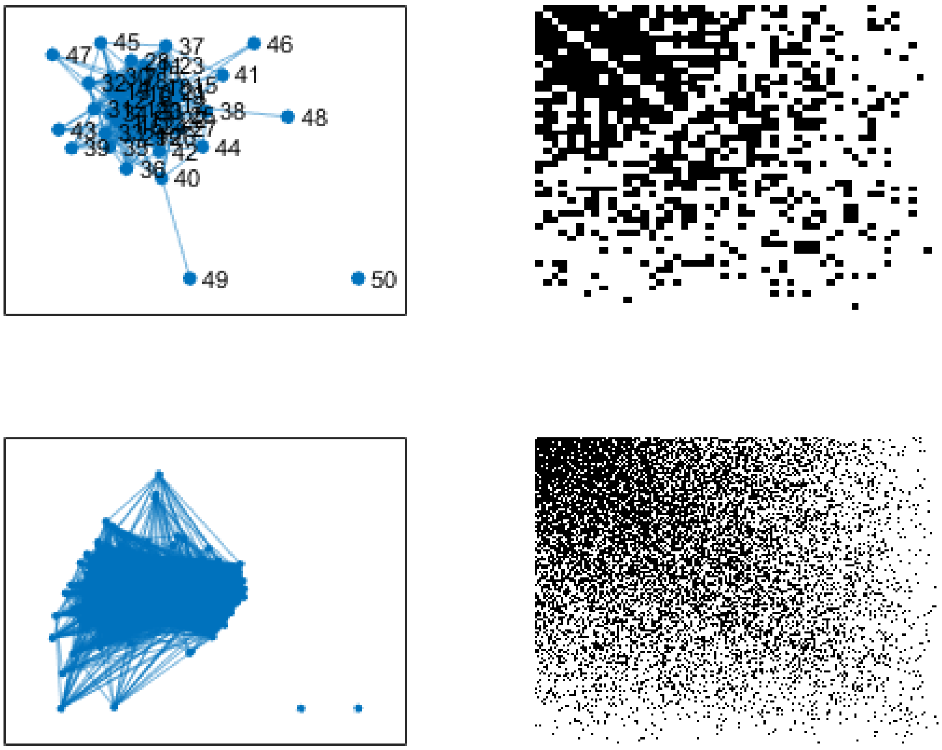

Figure 1: Graphs of graphs with 50 and 500 nodes, respectively, where their associated adjacency matrices converge to the graphon in Fig. 2 when mapped to the unit square. In this example, nodes with an index closer to zero are more likely to have many connections than nodes with an index closer to one.

Figure 2: The graph sequence shown in Fig. 1 converges to the Uniform Attachment Graphon $\boldsymbol{W}(\alpha,\beta) = 1 - \max(\alpha,\beta), \quad \alpha,\beta \in [0,1]$.

The objective function for the single agent at node α has the limit

$$J^{\alpha}(\boldsymbol{u}^{\alpha},\boldsymbol{x}_0) = \mathbb{E}\left[\sum_{k=0}^{T-1} ||\boldsymbol{x}_k^{\alpha} - \boldsymbol{z}_k^{\alpha}||_S^2 + ||\boldsymbol{u}_k^{\alpha}||_R^2 \middle| \mathcal{F}_0^{\alpha}\right], \tag{11}$$

and its goal is to minimize the objective function for the control strategy $\boldsymbol{u}^\alpha$ adapted to the information pattern $\mathcal{F}_k^\alpha$. For the purposes of this article, the full-state information pattern is used, where $\mathcal{F}_k^\alpha = \mathcal{F}_k := \{\boldsymbol{x}_k^\beta,\ \beta \in [0,1]\}$ for each agent α. This is sufficient for each agent to calculate the entire graphon field $\boldsymbol{z}_k$ at time k. This has the value function form:

$$\begin{aligned} V_k^\alpha(\mathcal{F}_k^\alpha) =& \mathbb{E}\big[||\boldsymbol{x}_k^\alpha - \boldsymbol{z}_k^\alpha||_S^2 + ||\boldsymbol{u}_k^\alpha||_R^2 \\ &+ V_{k+1}(\mathcal{F}_{k+1}^\alpha)|\mathcal{F}_k^\alpha\big], \quad k = (0,1,...,T-1), \end{aligned} \tag{12}$$

$$V_T^\alpha(\mathcal{F}_T^\alpha) = \mathbb{E}_T[||\boldsymbol{x}_T^\alpha - \boldsymbol{z}_T^\alpha||_S^2] = ||\boldsymbol{x}_T^\alpha - \boldsymbol{d}_T^\alpha||_S^2. \tag{13}$$

The agents are in a Nash equilibrium when the following inequality holds,

$$J^\alpha(\boldsymbol{u}^{\alpha *}, \boldsymbol{u}^{-\alpha *}) \leq J^\alpha(\boldsymbol{u}^\alpha, \boldsymbol{u}^{-\alpha *}),\ \forall\ \alpha \in [0,1]. \tag{14}$$

Unlike in the case of finite agent games, taking the limit as the number of agents tends to infinity yields an indifference to the costs of a particular agent α with respect to the strategies of any other specific agent β. Only the strategies of the agents as a mass (taken as a function over the unit interval) affect the cost of a given agent.

As with mean field games and graphon mean field games, the graphon field term $\boldsymbol{z}_k^\alpha$ is not dependent on both the state $\boldsymbol{x}_k^\alpha$ and the action $\boldsymbol{u}_k^\alpha$ of any single agent, in the sense that altering $\{\boldsymbol{x}_r^\alpha\}_{0\leq r\leq k}$ or $\{\boldsymbol{u}_r^\alpha\}_{0\leq r\leq k}$ for a particular α does not change $\boldsymbol{z}_k^\alpha$. This is evident from the integral operator definition of $\boldsymbol{z}_k^\alpha$. As with many mean field game problems, this changes the limit problem from a game to a tracking control problem where each node in the network is penalized for deviating from its associated graphon field.

4 Solution to the Q-noise graphon field tracking game

The game is solved in two steps, first by formulating the response of an individual agent $\alpha \in [0,1]$ as a stochastic tracking problem, then by showing that the individual actions of each agent generate a Nash equilibrium.

4.1 Solution to the stochastic control tracking problem

Assume an agent α is tracking an exogenous square-integrable drift process $\boldsymbol{d}_k(\alpha),\ \alpha \in [0,1]$. Define the state transition equation as

$$\boldsymbol{x}_{k+1}^\alpha = (a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha) + \boldsymbol{g}_k^\alpha, \tag{15}$$

The value function is found using dynamic programming, and, as above, has the form

$$\begin{aligned} V_k^\alpha(\mathcal{F}_k^\alpha) =& \mathbb{E}\big[||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||\boldsymbol{u}_k^\alpha||_R^2 \\ &+ V_{k+1}(\mathcal{F}_{k+1}^\alpha)|\mathcal{F}_k^\alpha\big], \quad k = (0,1,...,T-1), \end{aligned} \tag{16}$$

$$V_T^\alpha(\mathcal{F}_T^\alpha) = \mathbb{E}_T[||\boldsymbol{x}_T^\alpha - \boldsymbol{d}_T^\alpha||_S^2] = ||\boldsymbol{x}_T^\alpha - \boldsymbol{d}_T^\alpha||_S^2. \tag{17}$$

As stated above, this work considers the case where all agents have the full-information set $\mathcal{F}_k$ consisting of $\boldsymbol{x}_k^\eta$ and $\boldsymbol{d}_k^\eta$ for all $\eta \in [0,1]$.

Lemma 1. *The value function of agent α at time k, V_k^α, is given by*

$$\begin{aligned} V_k^\alpha(\mathcal{F}_k^\alpha) =& \mathbb{E}_k[P_k(\boldsymbol{x}_k^\alpha)^2 + 2\boldsymbol{x}_k^\alpha \boldsymbol{s}_k^\alpha + m_k^\alpha], \\ & k = \{0,...,T\}, \end{aligned} \tag{18}$$

where $\mathbb{E}[\cdot|\mathcal{F}_k] := \mathbb{E}_k[\cdot]$. Here, P_k is a positive scalar, and $\boldsymbol{s}_k$ and m_k^α are $L_2([0,1])$ valued functions for all $k = \{0,...,T\}$ derived from the following backwards recurrence relations:

$$F_k = (R + b^2 P_{k+1})^{-1} ab P_{k+1}, \tag{19}$$

$$
\begin{aligned}
G_k =&(R + b^2 P_{k+1})^{-1} bc P_{k+1}, && (20)\\
H_k =&(R + b^2 P_{k+1})^{-1} b, && (21)\\
P_k =&S + RF_k^2 + P_{k+1}(a - bF_k)^2, && (22)\\
\boldsymbol{s}_k^\alpha =& - S\boldsymbol{d}_k^\alpha + F_k R(G_k \boldsymbol{d}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]) && (23)\\
&+ (a - bF_k)P_{k+1}\\
&\quad \times \left[(c - bG_k)\boldsymbol{d}_k^\alpha - bH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\
&+ (a - bF_k)\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha],\\
m_k^\alpha =&S\boldsymbol{d}_k^\alpha \boldsymbol{d}_k^\alpha + (G_k \boldsymbol{d}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])R && (24)\\
&\quad \times (G_k \boldsymbol{d}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])\\
&+ \left[(c - G_k)\boldsymbol{d}_k^\alpha - bH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] P_{k+1}\\
&\quad \times \left[(c - G_k)\boldsymbol{d}_k^\alpha - bH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\
&+ 2\left[(c - G_k)\boldsymbol{d}_k^\alpha - bH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\\
&+ \mathbf{Q}(\alpha, \alpha) + \mathbb{E}_k[m_{k+1}^\alpha],
\end{aligned}
$$

with the terminal conditions

$$
\begin{aligned}
P_T =&S, && (25)\\
\boldsymbol{s}_T^\alpha =& - S\boldsymbol{d}_T^\alpha, && (26)\\
m_T^\alpha =&S||\boldsymbol{d}_T^\alpha||_2^2. && (27)
\end{aligned}
$$

Further, the optimal control is given by

$$
\begin{aligned}
u_k^{o,\alpha} =& - (R + b^2 P_{k+1})^{-1} b[P_{k+1}(a\boldsymbol{x}_k^\alpha + c\boldsymbol{d}_k^\alpha) && (28)\\
&+ \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]]\\
=:& - F_k \boldsymbol{x}_k^\alpha - G_k \boldsymbol{d}_k^\alpha - H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]. && (29)
\end{aligned}
$$

The proof follows from the ansatz (18). See A.2.

The cost m_k^α does not affect the control input $\boldsymbol{u}_k^\alpha$ and does not have a simple closed form solution, so it is not calculated here. The structure of the tracking control solution (with the feedforward term in the costate $\boldsymbol{s}_{k+1}$) is common in discrete time tracking problems [13].

The value function above solves the general discrete-time stochastic optimal control problem where an agent α tracks an exogenous signal $\boldsymbol{d}_k^\alpha$. The problem is ill-defined in general since it requires the computation of the expectation of the offset, $\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]$ in terms of the expected terminal value of $\boldsymbol{d}_T^\alpha$, which requires additional assumptions on the process $\boldsymbol{d}_k$.

However, in the graphon field game setting, at each time step k the chosen strategy must generate the local field term $\boldsymbol{z}$, i.e. the optimal input $\{\boldsymbol{u}_k^o,\ k = 0, ..., T-1\}$ must generate a trajectory satisfying $\boldsymbol{z}_k^\alpha = [\boldsymbol{M}\boldsymbol{x}_k](\alpha) = \boldsymbol{d}_k^\alpha$ for all α. This provides additional structure to the tracked stochastic process, and is known as the Consistency Condition for the Nash equilibrium in the limit game [11]. In the full state feedback case, the Consistency Condition allows the expectation of the offset $\boldsymbol{s}_k$ to be explicitly calculated as a linear state process in terms of $\boldsymbol{z}_k$.

4.2 Nash equilibrium Consistency Condition with full state information

By the Consistency Condition, for each k, the local field $\boldsymbol{z}_k$ is given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$. As $\boldsymbol{x}_k$ is square-integrable for each k when generated by the optimal strategy $\boldsymbol{u}_k$ and $\boldsymbol{M}$ is an $L_2[0,1]$ to $L_2[0,1]$ operator, the graphon field $\boldsymbol{z}_k$ is square-integrable as well. For the game to yield a Nash equilibrium, it is necessary for all agents to apply their respective control $\boldsymbol{u}_k^\alpha$ and generate the local field process $\boldsymbol{z}_k^\alpha$. To denote the function over the whole index set the superscript α is omitted.

To do this, define two new, time varying operators: $\Gamma_k : L_2[0,1] \to L_2[0,1]$ which calculates the expected graphon field $\boldsymbol{z}_{k+1}$ at the next time-step given the state of all agents at the current time-step, and $\Psi_k : L_2[0,1] \to L_2[0,1]$ which calculates the tracking adjoint process $\boldsymbol{s}_k$ as a linear function of the current graphon field $\boldsymbol{z}_k$.

Lemma 2. *Let the signal to be tracked be given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$ for time k. Let Γ_k and Ψ_k be $L_2([0,1])$ operators which are defined by the backwards recursion equations*

$$\begin{aligned}\Psi_k =& - S\mathbb{I} + F_k R(G_k\mathbb{I} + H_k\Psi_{k+1}\Gamma_k) \\ &+ (a - bF_k)P_{k+1}\big[(c - bG_k)\mathbb{I} \\ &- b\Psi_{k+1}H_k\Gamma_k\big] \\ &+ (a - bF_k)\Psi_{k+1}\Gamma_k,\end{aligned} \tag{30}$$

$$\begin{aligned}\Gamma_k =&(\mathbb{I} + bH_k\boldsymbol{M}\Psi_{k+1})^{-1}[(a - bF_k)\mathbb{I} \\ &+ (c - bG_k)\boldsymbol{M}]\end{aligned} \tag{31}$$

with the terminal condition

$$\Psi_T = - S\mathbb{I}. \tag{32}$$

Assume that for all $k = \{0, ..., T-1\}$, the inverse $(\mathbb{I} + bH_k\boldsymbol{M}\Psi_{k+1})^{-1}$ exists. Then,

$$\mathbb{E}_k[\boldsymbol{z}_{k+1}] =\Gamma_k\boldsymbol{z}_k, \tag{33}$$

$$\boldsymbol{s}_k =\Psi_k\boldsymbol{z}_k, \tag{34}$$

and the trajectory generated by

$$\boldsymbol{u}_k = -F_k\boldsymbol{x}_k - (G_k\mathbb{I} + H_k\Psi_{k+1}\Gamma_k)\boldsymbol{z}_k \tag{35}$$

gives the optimal tracking trajectory for each α.

See proof A.3.

Combining Lemma 1 and 2 yields the Nash equilibrium of the game.

Theorem 1. *Given the limit graphon tracking game of the type (11) for the family of systems (9), where each agent α indexed by $[0,1]$ has the total information pattern (that is, each agent knows the states of all other agents at the current time step), $\mathcal{F}_k^\alpha = \{\boldsymbol{x}_k, \boldsymbol{z}_k\}$ for all $\alpha \in [0,1]$, the control strategy given in Equations (A17), (A18), and (35) yields a Nash equilibrium.*

4.3 Numerical simulation

To demonstrate the behavior of the field tracking game, There are two general phenomena depending on the maximal eigenvalue of the graphon $\boldsymbol{M}$. Namely, as the state $\boldsymbol{x}_k^\alpha$ attempts to track the field average $\boldsymbol{z}_k^\alpha$, the optimal trajectory (without noise) would satisfy

$$\boldsymbol{x}_k = \boldsymbol{z}_k := \boldsymbol{M}\boldsymbol{x}_k. \tag{36}$$

This is an eigenvalue and eigenfunction relation, satisfied by either the trivial function $\boldsymbol{x}_k^\alpha = 0$ for all agents $\alpha \in [0,1]$, or when $\boldsymbol{x}_k$ is an eigenfunction of the operator $\boldsymbol{M}$ with associated eigenvalue $\lambda = 1$. The only $[0,1]$-graphon $\boldsymbol{M}$ (where $0 \le \boldsymbol{M}(\alpha,\beta) \le 1$ for all $\alpha,\beta \in [0,1]$) is the Erdos-Renyi graphon with $p = 1$, $\boldsymbol{M}(\alpha,\beta) = 1$. When $\max(\lambda(\boldsymbol{M})) < 1$, tracking the graphon field contracts the state of all agents to zero to zero. Two contracting examples, with varying graphons and Q-noise disturbances are presented first, and Two normalized examples are presented second. In order to create examples where the maximal eigenvalue is one, the graphon $\boldsymbol{M}$ is normalized by its maximal eigenvalue.

In all examples, the states are scalar, with the system parameters $a = 1$, $b = 1$, and $c = 0.1$. To illustrate that the system effectively tracks the perturbed graphon fields with noise, the disturbance covariance will be scaled to be small relative to the effect of the graphon field.

To verify the definitions of Γ_k and Ψ_k, for all examples we compute $\Delta(\Gamma_k)$, the L_2 norm between the expected value of the graphon field $\boldsymbol{z}_{k+1}$ given $\boldsymbol{z}_k$ and the operator Γ_k acting on the current graphon field $\boldsymbol{z}_k$, i.e.,

$$\Delta(\Gamma) := \max_{0\leq k\leq T} ||\boldsymbol{M}(a\boldsymbol{x}_k + b\boldsymbol{u}_k + c\boldsymbol{z}_k) - \Gamma_k \boldsymbol{z}_k||_2. \tag{37}$$

In all four sample paths below, $\Delta(\Gamma)$ was calculated to be below the order of 10^{-9}, often on the order of 10^{-14}. This confirms numerically that the operator Γ_k is equivalent to the expected value of the graphon field $\boldsymbol{z}_k$ at the next time step.

4.4 Contracting graphon, $\max(\lambda(\boldsymbol{M})) < 1$

4.4.1 Erdos-Renyi: $\boldsymbol{M}(\alpha,\beta) = 0.9$, $\mathbf{Q}(\alpha,\beta) = (1 - \max(\alpha,\beta))/20$

Here, set the initial state for each agent to $\boldsymbol{x}_0^\alpha = 3\sin(\pi\alpha)$. This initial condition is arbitrary, it only needs to be set to be non-constant to demonstrate the graphon field behavior. The state $\boldsymbol{x}_k$ and field $\boldsymbol{z}_k$ are shown in Fig. 3, as well as the trajectory of the error between the two, $\boldsymbol{x}_k - \boldsymbol{z}_k$. The error is low, but due to the addition of noise at each time step, is never zero.

For the contracting Erdos-Renyi graphon, $\Delta(\Gamma)$ is zero on the order of 10^{-14}.

4.5 $\boldsymbol{M}(\alpha,\beta) = \sqrt{|x-y|}$, $\mathbf{Q}(\alpha,\beta) = \cos(\alpha-\beta)$

The trajectory of the state in this example is shown in Fig. 4. Unlike in the Erdos-Renyi case, the graphon field $\boldsymbol{z}_k$ is non-constant at each k, but the controlled game is still stable about the origin $x_k^\alpha = 0,\ \alpha \in [0,1]$. For the contracting square root graphon, $\Delta(\Gamma)$ is zero on the order of 10^{-15}.

4.6 Normalized graphon, $\max(\lambda(\boldsymbol{M})) = 1$

When the graphon $\boldsymbol{M}$ is normalized by its maximal eigenvalue, instead of sending the state of each agent to zero, the calculated optimal control $\boldsymbol{u}_k$ instead moves the state $\boldsymbol{x}_k$ towards the associated eigenfunction of $\boldsymbol{M}$. As $c = 0.1$ and $a = 1$, the system is unstable, and the system tracks a scaling of an eigenfunction of $\boldsymbol{M}$.

4.6.1 Erdos-Renyi: $\boldsymbol{M}(\alpha,\beta) = 1$, $\mathbf{Q}(\alpha,\beta) = (1 - \max(\alpha,\beta))/20$

This trajectory is shown in Fig. 5. Unlike the case where $\boldsymbol{M}(\alpha,\beta) = 0.9$, the controlled game is stable near the eigenfunction initial condition. For the normalized Erdos-Renyi graphon, $\Delta(\Gamma)$ is zero on the order of 10^{-9}.

4.7 $\boldsymbol{M}(\alpha,\beta) = \sqrt{|x-y|}$ (Normalized), $\mathbf{Q}(\alpha,\beta) = \cos(\alpha-\beta)$

The trajectory shown in Fig. 6 shows that the controlled state is attracted to a scaled eigenfunction of the system. For the normalized square root graphon, $\Delta(\Gamma)$ is zero on the order of 10^{-10}.

5 Infinite horizon discounted cost

Due to the addition of disturbance of bounded variation, the standard horizon field tracking problem does not have a well-defined value function in the infinite time horizon case. This can be addressed in the standard manner by using a multiplicative, stage-wise discount factor ρ.

Even then, this poses some conceptual questions. The graphon field process to be tracked is non-deterministic and non-constant, and the backwards equation method of deriving the operators Ψ and

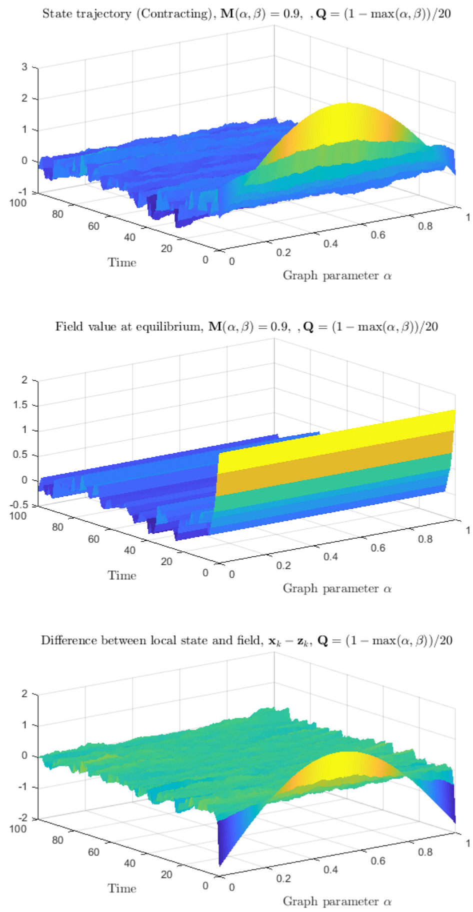

Figure 3: Top: The state trajectory of the system x_k when the graphon field is given by the Erdos-Renyi graphon $M = 0.9$. As the graphon is contracting, the controlled state trajectory is near zero for all agents. Middle: The associated graphon field. As this is a rank one graphon equivalent for all agents, the field is flat at each time step. Bottom: The difference between the graphon field and state, which determines the cost to particular agents.

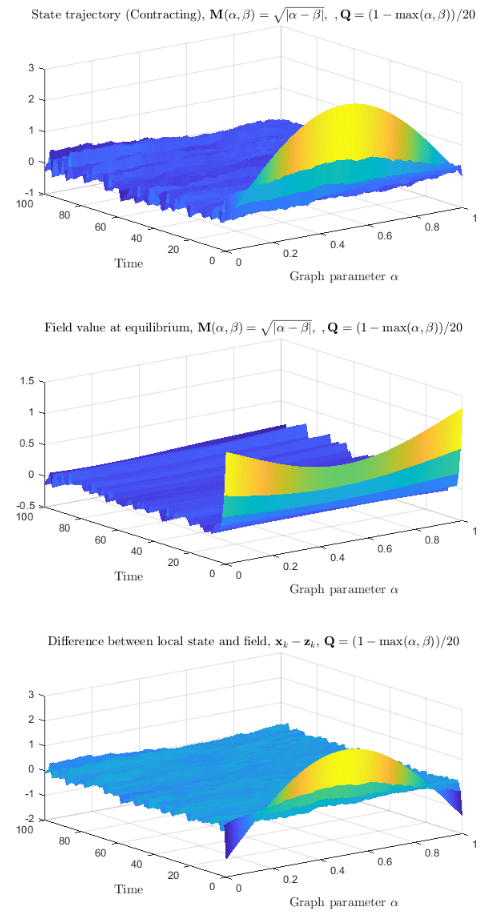

Figure 4: Top: The state trajectory of the system x_k when the graphon field is given by the graphon $M(\alpha,\beta) = \sqrt{|\alpha-\beta|}$. This graphon is also contracting, and the controlled state trajectory approaches zero. Middle: The associated graphon field. Bottom: The difference between the agents' states and graphon fields.

Γ cannot be used directly. As a starting point, consider the finite time horizon discounted case; for $0<\rho<1$, define the finite horizon discounted tracking problem for a system of the form (15),

$$J_\rho^\alpha(\boldsymbol{u},\boldsymbol{x}_0)=\sum_{k=0}^{T-1}\rho^k\mathbb{E}[||\boldsymbol{x}_k^\alpha-\boldsymbol{z}_k^\alpha||_S^2+||\boldsymbol{u}_k^\alpha||_R^2 \mid \mathcal{F}_0^\alpha] \tag{38}$$

$$+\rho^T\mathbb{E}[||\boldsymbol{x}_T^\alpha-\boldsymbol{z}_T^\alpha||_S^2|\mathcal{F}_T]. \tag{39}$$

By the same proof approach to the finite time non-discounted game, this is associated with the sequence of value functions

$$V_k^\alpha(\mathcal{F}_k^\alpha)=\mathbb{E}_k[P_k(\boldsymbol{x}_k^\alpha)^2+2(\boldsymbol{x}_k^\alpha)\boldsymbol{s}_k^\alpha+m_k^\alpha], \tag{40}$$
$$k=\{0,...,T\},$$

where P_k^ρ is an positive scalar, and $\boldsymbol{s}_k$ and m_k^α are $L_2([0,1])$ valued functions for all $k=\{0,...,T\}$ derived from the following backwards recurrence relations,

$$F_k^\rho=\rho(R+\rho P_{k+1}^\rho b^2)^{-1}P_{k+1}^\rho ab, \tag{41}$$

$$G_k^\rho=\rho(R+\rho P_{k+1}^\rho b^2)^{-1}P_{k+1}^\rho bc, \tag{42}$$

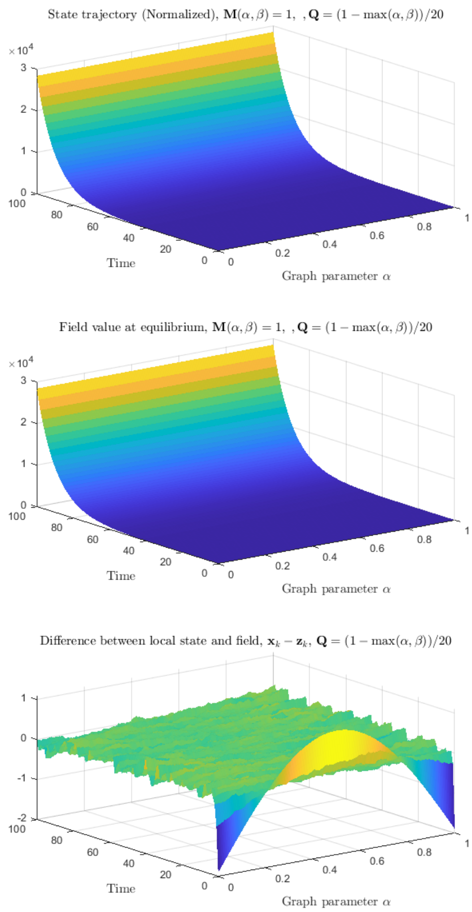

Figure 5: Top: The state trajectory of the system x_k when the graphon field is given by the Erdos-Renyi graphon $M=1$. As the system is unstable ($c=0.1$), the state of each agent tends to infinity. Middle: The associated graphon field. Bottom: The difference between each agent's state and field. Despite the controlled system being fundamentally unstable, the state of each agent very closely tracks the graphon field.

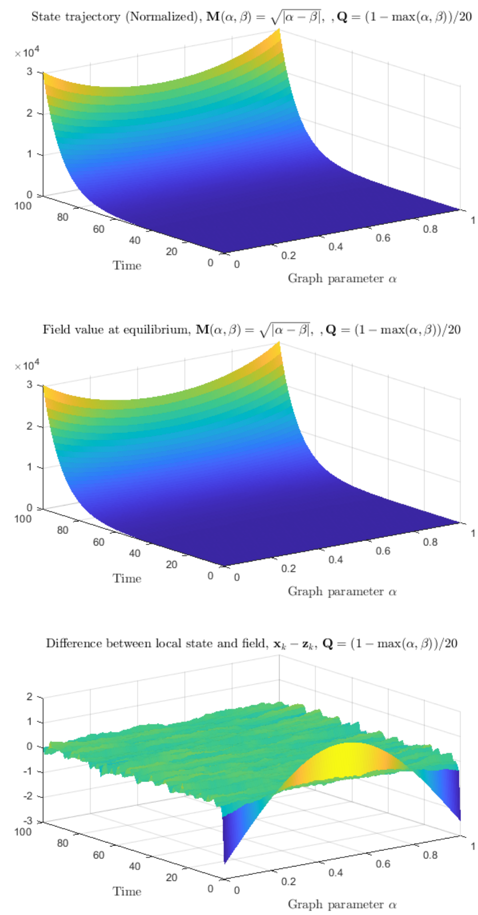

Figure 6: Top: The state trajectory of the system x_k when the graphon field is given by the graphon $M=\sqrt{|\alpha-\beta|}$ after normalization by its maximal eigenvalue. As the system is unstable ($c=0.1$), the state of each agent tends to infinity. Middle: The associated graphon field. Bottom: The difference between each agent's state and field. Despite the controlled system being fundamentally unstable, the state of each agent very closely tracks the graphon field.

$$
\begin{aligned}
H_k^\rho =&\rho(R+\rho P_{k+1}^\rho b^2)^{-1}b, && (43)\\
P_k^\rho =&S+R(F_k^\rho)^2+\rho^2(a-bF_k^\rho)^2 P_{k+1} && (44)\\
\boldsymbol{s}_k^\alpha =&-S\boldsymbol{d}_k^\alpha+F_k^\rho R(G_k^\rho \boldsymbol{d}_k^\alpha+H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]) && (45)\\
&+\rho(a-bF_k^\rho)P_{k+1}^\rho\\
&\quad\times\left[(c-bG_k^\rho)\boldsymbol{d}_k^\alpha-bH_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\
&+\rho(a-bF_k^\rho)\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha],\\
m_k^\alpha =&\boldsymbol{d}_k^{\alpha *}S\boldsymbol{d}_k^\alpha+\rho\Big[(G_k^\rho\boldsymbol{d}_k^\alpha+H_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])^*R && (46)\\
&\quad\times(G_k^\rho\boldsymbol{d}_k^\alpha+H_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])\\
&+\left[(c-bG_k^\rho)\boldsymbol{d}_k^\alpha-bH_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]P_{k+1}^\rho\\
&\quad\times\left[(c-bG_k^\rho)\boldsymbol{d}_k^\alpha-BH_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\
&+2\left[(c-bG_k^\rho)\boldsymbol{d}_k^\alpha-BH_k^\rho\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\\
&+\mathbf{Q}(\alpha,\alpha)+\mathbb{E}_k[m_{k+1}^\alpha]\Big],
\end{aligned}
$$

with the terminal conditions

$$P_T^\rho = S, \tag{47}$$
$$\boldsymbol{s}_T^\alpha = -S\boldsymbol{d}_T^\alpha, \tag{48}$$
$$m_T^\alpha = S||\boldsymbol{d}_T^\alpha||^2. \tag{49}$$

Further, the optimal control is given by

$$u_k^{o,\alpha} = -\rho(R+\rho b^2 P_{k+1}^\rho)^{-1}[P_{k+1}^\rho(a\boldsymbol{x}_k^\alpha + c\boldsymbol{d}_k^\alpha) + \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]] \tag{50}$$
$$=: -F_k^\rho \boldsymbol{x}_k^\alpha - G_k^\rho \boldsymbol{d}_k^\alpha - H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]. \tag{51}$$

As in the non-discounted case, the process to be tracked is given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$ at time k. Let Γ_k and Ψ_k be $L_2([0,1])$ operators defined by the backwards recursion equations

$$\Psi_k^\rho = -S\mathbb{I} + F_k^\rho R(G_k^\rho \mathbb{I} + H_k^\rho \Psi_{k+1}^\rho \Gamma_k^\rho) + \rho(a-bF_k^\rho)^* P_{k+1}^\rho \big[(c-bG_k^\rho)\mathbb{I} - b\Psi_{k+1} H_k \Gamma_k\big] + \rho(a-bF_k)^* \Psi_{k+1}\Gamma_k, \tag{52}$$
$$\Gamma_k^\rho = (\mathbb{I} + bH_k^\rho \boldsymbol{M}\Psi_{k+1})^{-1}[(a-bF_k^\rho)\mathbb{I} + (d-bG_k^\rho)\boldsymbol{M}] \tag{53}$$

with the terminal conditions

$$\Psi_T^\rho = -S\mathbb{I}. \tag{54}$$

Assume that for all $k = \{0,...,T-1\}$, the inverse $(\mathbb{I} + bH_k\boldsymbol{M}\Psi_{k+1})^{-1}$ exists. Then, for the system (15) as with the non-discounted game,

$$\mathbb{E}_k[\boldsymbol{z}_{k+1}] = \Gamma_k^\rho \boldsymbol{z}_k, \tag{55}$$
$$\boldsymbol{s}_k = \Psi_k^\rho \boldsymbol{z}_k, \tag{56}$$

and the trajectory is generated by

$$\boldsymbol{u}_k^\rho = -F_k^\rho \boldsymbol{x}_k - (G_k^\rho \mathbb{I} + H_k^\rho \Psi_{k+1}\Gamma_k^\rho)\boldsymbol{z}_k \tag{57}$$

gives the optimal tracking trajectory for each α.

It is assumed that the infinite horizon feedback solution (when it exists) is given by the fixed point to the following algebraic Riccati and operator equations:

$$F_\infty^\rho = \rho(R+\rho b^2 P_\infty^\rho)^{-1} P_\infty^\rho ab, \tag{58}$$
$$G_\infty^\rho = \rho(R+\rho b^2 P_\infty^\rho)^{-1} P_\infty^\rho bc, \tag{59}$$
$$H_\infty^\rho = \rho(R+\rho b^2 P_\infty^\rho b)^{-1} b, \tag{60}$$
$$P_\infty^\rho = S + F_\infty^{\rho *} R F_\infty^\rho + \rho(a-bF_\infty^\rho)^* P_\infty^\rho (a-bF_\infty^\rho), \tag{61}$$
$$\Psi_\infty^\rho = -S\mathbb{I} + F_\infty^\rho R(G_\infty^\rho \mathbb{I} + H_\infty^\rho \Psi_\infty^\rho \Gamma_\infty^\rho) + \rho(a-bF_\infty^\rho)^* P_\infty^\rho \big[(c-bG_\infty^\rho)\mathbb{I} - bH_\infty \Psi_\infty \Gamma_\infty\big] + (a-bF_\infty^\rho)^* \Psi_\infty \Gamma_\infty^\rho, \tag{62}$$
$$\Gamma_\infty^\rho = (\mathbb{I} + bH_\infty^\rho \boldsymbol{M}\Psi_\infty)^{-1}[(a-bF_\infty^\rho)\mathbb{I} + (c-bG_\infty^\rho)\boldsymbol{M}]. \tag{63}$$

Under the condition that $c = 0$ (that is, the states of each agent evolve with strictly local state and controls) and that the graphon $\boldsymbol{M}$ has finite rank K, then the operators can be explicitly solved with $(K \times K) + 1$ equations.

Define the orthogonal subspaces $\mathcal{S}^{\boldsymbol{M}}$ to be the the linear subspace spanned by the basis of the system graphon $\boldsymbol{M}$, and $\breve{\mathcal{S}}$ to be the orthogonal complement of $\mathcal{S}^{\boldsymbol{M}}$ such that

$$L_2[0,1] =: \mathcal{S}^{\boldsymbol{M}} \oplus \breve{\mathcal{S}}. \tag{64}$$

Then, define the $K \times K$ matrix M to be

$$M_{i,j} := \langle \boldsymbol{M}\phi_i, \phi_j \rangle, \quad i, j \in (1, ..., K) \tag{65}$$

and the complement identity operator $\breve{\mathbb{I}} : L_2[0,1] \to \breve{\mathcal{S}}$ to be the projection operator satisfying

$$\breve{\mathbb{I}}\boldsymbol{d} = \boldsymbol{d} - \sum_{k=1}^{K} \langle \boldsymbol{d}, \phi_k \rangle \phi_k, \quad \forall \boldsymbol{d} \in L_2[0,1]. \tag{66}$$

Theorem 2 (Finite Rank Closed Form Feedback)**.** *If the state of each node is a scalar, $c = 0$, and the system graphon $\boldsymbol{M}$ is of rank $K < \infty$ with associated orthonormal basis $\{\phi_k\}_{k=1}^K$, then Ψ_∞ and Γ_∞ is a solution of the quadratic operator equation*

$$BH_\infty \boldsymbol{M} \Psi_\infty^2 + (C\mathbb{I} + SBH_\infty^\rho \boldsymbol{M})\Psi_\infty^\rho + S\mathbb{I} = 0, \tag{67}$$

where

$$C := 1 - (a - bF_\infty^\rho)[F_\infty^\rho R - \rho(a - bF_\infty^\rho)P_\infty^\rho b \tag{68}$$

$$- \rho(a - bF_\infty)^2] \tag{69}$$

Then, in the low rank case Eq. (67) is solved by the orthogonal operator equations

$$\Psi_\infty := \breve{\Psi}_\infty + \sum_{i=1}^{K}\sum_{j=1}^{K} [\Psi_\infty^{\boldsymbol{M}}]_{ij} \langle \phi_i, \cdot \rangle \tag{70}$$

$$c\breve{\Psi}_\infty + S\breve{\mathbb{I}} = 0, \quad \breve{\Psi}_\infty : L_2[0,1] \to \breve{\mathcal{S}} \tag{71}$$

$$bH_\infty M(\Psi_\infty^{\boldsymbol{M}})^2 + (C + bSH_\infty)\Psi_\infty^{\boldsymbol{M}} + SI_K = 0, \quad \Psi_\infty^{\boldsymbol{M}} \in \mathcal{R}^{K \times K} \tag{72}$$

where I_K is the K-dimensional identity operator.

Proof. Note that if $c = 0$, then $G_\infty = 0$. Eq. (67) is attained by substituting the definition of Γ_∞ into Eq. (A42), and multiplying on the right by $\mathbb{I} + bH_\infty \boldsymbol{M}\Psi_\infty$. The equation can then be solved explicitly in two orthogonal subspaces $\mathcal{S}^{\boldsymbol{M}}$ and $\breve{\mathcal{S}}$ using the definitions provided in the lemma. □

Note 1. When $\boldsymbol{M}$ is rank one, then Eq. (72) is a simple quadratic equation in one variable. Thus, there are two solutions due to the square root. In simulations, when the system parameters are changed, Eqs (52)–(54) may converge to either. In particular, the operator may converge to the solution with the positive radical when the eigenvalue associated with the eigenfunction of $\boldsymbol{M}$ is less than one, and to the solution with the negative radical when the associated eigenfunction equals one. Further, while the solution found through Eqs (52)–(54) is necessarily a Nash equilibrium for the finite-horizon game, this may not be the case for the other solution of the infinite-horizon game. For $\boldsymbol{M}$ of rank larger than one, a matrix quadratic formula is required, for which there are 2^K solutions.

As will be explored in the following numerical simulations, there are typically multiple solutions for the operator equations. Given initial conditions aligned with the eigenfunctions of the graphon $\boldsymbol{M}$, some of the generated solutions are clearly worse for all players.

5.1 Numerical simulation

The following numerical solutions demonstrate the existence of multiple solutions of the infinite time horizon discounted problem, and that the finite horizon discounted problem can converge to both of them depending on whether or not the underlying graphon operator is normalized.

For each of the simulations, the discount factor $\rho = 0.95$ is used, with system parameters $a = 1$, $b = 1$, $c = 0$ and costs $S = 1$ and $R = 1$ and a 200 node discretization of the unit interval. The rank one graphon $\boldsymbol{M}(\alpha, \beta) = (\alpha^2 - 1)(\beta^2 - 1)$ was used to generate the graphon field. In the first set of three simulations, the graphon is non-normalized (contracting), and for the second set of three simulations the graphon is normalized. For all agents in all simulations, the initial condition $\boldsymbol{x}_0(\alpha) = 1,\ \alpha \in [0, 1]$ was used.

Fig. 7 shows the state trajectory of the finite time discounted game when the graphon is non-normalized.

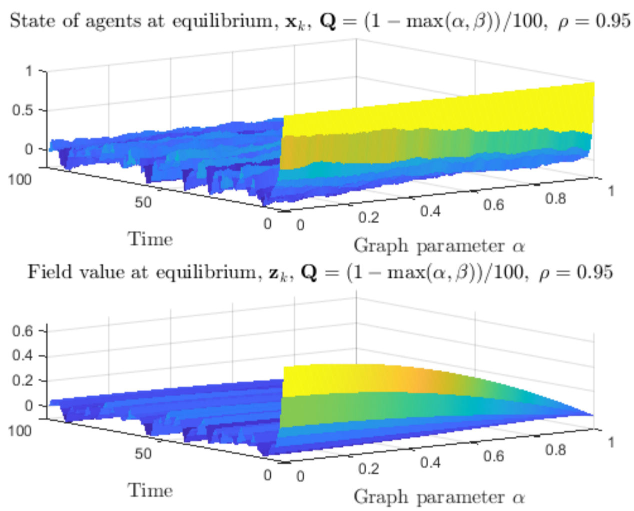

Figure 7: Finite time horizon discounted game with a contracting graphon. As expected, the state trajectory of the finite time horizon controlled system approaches zero.

Fig. 8 demonstrates that when the positive root solution of Eq. (72) is used to calculate Ψ^{ρ}_{∞} and Γ^{ρ}_{∞}, the solution very closely tracks the controlled state trajectory shown in Fig. 7. Indeed, the operator norm distance of $\Psi^{\rho}_{2}\Gamma^{\rho}_{1}$ and $\Psi^{\rho}_{\infty}\Gamma^{\rho}_{\infty}$ was on the order of machine precision, indicating that the sequence converged.

Meanwhile, Fig. 9 shows that when the negative root solution of Eq. (72) is used, the controlled state trajectory is exceptionally unstable.

Next, consider the case where the graphon $\boldsymbol{M}$ is normalized by its eigenvalue. As with the nondiscounted game, the solution of the finite time horizon discounted game approaches a scaled eigenfunction of the graphon, in this case, $\phi(\alpha) = \alpha^2 - 1$. Fig. 10 shows this behavior.

However, unlike the previous case, the positive root solution of Eq. 72 does not create a trajectory that closely matches the finite time horizon game. It instead creates a trajectory that approaches zero for all agents, even though that seems sub-optimal, shown in Fig. 11.

When the graphon $\boldsymbol{M}$ is normalized, the finite horizon discounted game solution instead converges to negative root solution. A calculated trajectory of the controlled system using the negative root solution is shown in Fig. 12, which approaches an eigenfunction of the graphon $\boldsymbol{M}$ in a similar manner to the finite horizon discounted game trajectory.

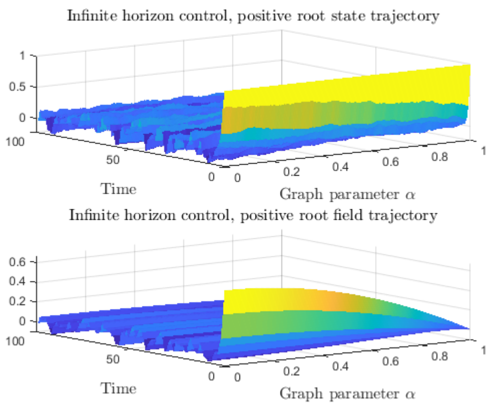

Figure 8: The controlled state trajectory of the game when the positive root solution of Eq. 72 is used, which closely resembles the finite time horizon discounted game trajectory.

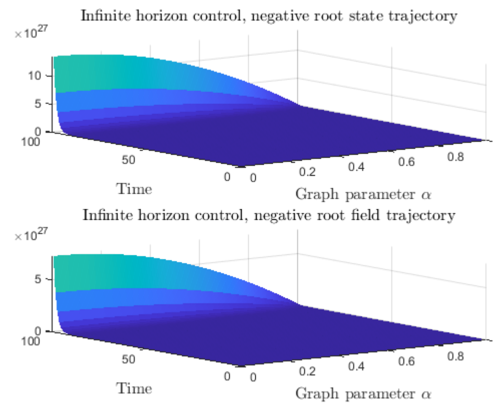

Figure 9: The system is unstable using the negative root solution of Eq. 72.

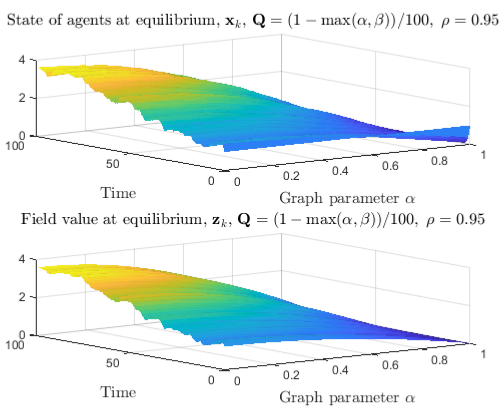

Figure 10: The controlled trajectory of the finite horizon discounted game closely tracks the eigenfunction of the system, even under noise, when M is normalized by its largest eigenvalue.

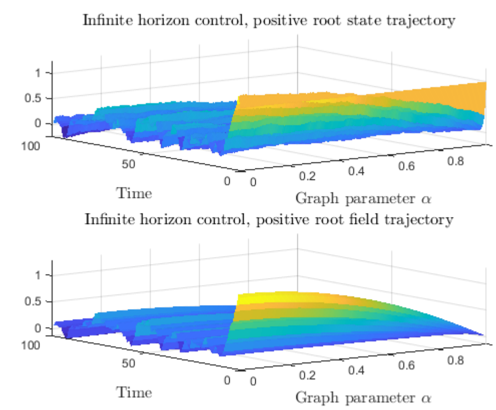

Figure 11: The controlled state trajectory of the normalized system using the positive root solution approaches zero for all agents, instead of approaching an eigenfunction.

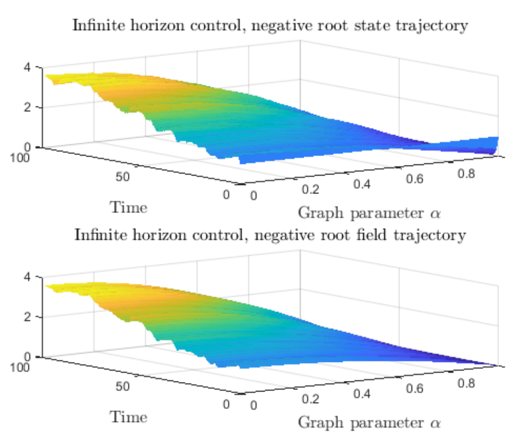

Figure 12: When M is normalized, the backwards recursion instead converges to the negative root solution.

6 Future work

There are some clear directions for future research. First, the work should be extended to limit graphs embedded in metric spaces using the embedded graph limit theory developed in Caines [2, 6]. This theory generalizes the concept used implicitly in the numerical simulations above, where each node in the graph is located uniformly at a point on the unit interval. Embedded graph limit theory is a method for describing graph limits that exist in geometric spaces more general than the unit interval, for instance those where each node is located in $\mathcal{R}^2$ or $\mathcal{R}^3$. In the dense graph case, this is a straightforward generalization, but may not have a direct solution analog for sparse graphs.

The precise criteria for the convergence of the operators Ψ_k^ρ and Γ_k^ρ to Ψ_∞^ρ and Γ_∞^ρ is a topic for open investigation.

This article considered Nash equilibria with full state information. This will be expanded to other information sets, such as those where each agent has only local information and, hence, estimation of the status of the overall graphon field may be of value. In particular, applying it to the case where each agent uses Kalman filtering to estimate the overall graphon field would be of interest. Further research in this area would require the consideration of common noise [7], where, unlike in the full information case, the local state $\boldsymbol{x}_k^\alpha$ would not be conditionally independent of the full state $\{\boldsymbol{x}_k^\beta,\ \beta \in [0,1]\}$.

Appendix

A.1 Multivariate state notation

Let each agent $\alpha \in [0,1]$ have n local states and m local controls, and let $A \in \mathcal{R}^{n\times n}, D \in \mathcal{R}^{n\times n},\ B \in \mathcal{R}^{n\times m}$. Then, define the state evolution equation as

$$\boldsymbol{x}_{k+1}^\alpha = (A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{z}_k^\alpha) + \boldsymbol{g}_k^\alpha, \tag{A1}$$

where the graphon field $\boldsymbol{z}_k^\alpha$ is an n dimensional real vector for all α defined by the following blockwise form, with graphons $\boldsymbol{M}_{11}, \boldsymbol{M}_{12}, ...\boldsymbol{M}_{nn}$

$$\boldsymbol{z}_k^\alpha := \begin{bmatrix} (\boldsymbol{z}_k^\alpha)_1 \\ (\boldsymbol{z}_k^\alpha)_2 \\ \vdots \\ (\boldsymbol{z}_k^\alpha)_1 \end{bmatrix} \tag{A2}$$

$$= \begin{bmatrix} \int_0^1 \big[\boldsymbol{M}_{11}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_1 + \boldsymbol{M}_{12}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_2 \\ + \cdots + \boldsymbol{M}_{1n}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_n\big] d\beta \\ \int_0^1 \big[\boldsymbol{M}_{21}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_1 + \boldsymbol{M}_{22}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_2 \\ + \cdots + \boldsymbol{M}_{2n}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_n\big] d\beta \\ \vdots \\ \int_0^1 \big[\boldsymbol{M}_{n1}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_1 + \boldsymbol{M}_{n2}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_2 \\ + \cdots + \boldsymbol{M}_{nn}(\alpha,\beta)(\boldsymbol{x}_k^\beta)_n\big] d\beta \end{bmatrix} \tag{A3}$$

$$= \begin{bmatrix} \boldsymbol{M}_{11} & \boldsymbol{M}_{12} & \cdots & \boldsymbol{M}_{1n} \\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22} & \cdots & \boldsymbol{M}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{M}_{n1} & \boldsymbol{M}_{n2} & \cdots & \boldsymbol{M}_{nn} \end{bmatrix} \begin{bmatrix} (_k)_1 \\ (_k)_2 \\ \vdots \\ (_k)_1 \end{bmatrix} \tag{A4}$$

By abuse of notation, we denote this blockwise form

$$\boldsymbol{z}_k := \boldsymbol{M}\boldsymbol{x}_k \tag{A5}$$

as in the single-state-per-node case.

Following this, the relevant $\mathcal{R}^m$ dimensional controls for each $\alpha \in [0,1]$ would be defined by

$$\begin{aligned} F_k =&(R + B^* P_{k+1} B)^{-1} B^* P_{k+1} A, &\text{(A6)}\\ G_k =&(R + B^* P_{k+1} B)^{-1} B^* P_{k+1} D, &\text{(A7)}\\ H_k =&(R + B^* P_{k+1} B)^{-1} B^*, &\text{(A8)}\\ P_k =&S + F_k^* R F_k + (A - BF_k)^* P_{k+1}(A - BF_k), &\text{(A9)}\\ \boldsymbol{s}_k^\alpha =& - S\boldsymbol{z}_k^\alpha + F_k^* R(G_k \boldsymbol{z}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]) &\text{(A10)}\\ &+ (A - BF_k)^* P_{k+1}\\ &\quad\times \left[(D - BG_k)\boldsymbol{z}_k^\alpha - BH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\ &+ (A - BF_k)^* \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha],\\ m_k^\alpha =&\boldsymbol{z}_k^{\alpha *} S \boldsymbol{z}_k^\alpha + (G_k \boldsymbol{z}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])^* R &\text{(A11)}\\ &\quad\times (G_k \boldsymbol{z}_k^\alpha + H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])\\ &+ \left[(D - G_k)\boldsymbol{z}_k^\alpha - BH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]^* P_{k+1}\\ &\quad\times \left[(D - G_k)\boldsymbol{z}_k^\alpha - BH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]\\ &+ 2\left[(D - G_k)\boldsymbol{z}_k^\alpha - BH_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\\ &+ \mathbf{Q}(\alpha,\alpha) + \mathbb{E}_k[m_{k+1}^\alpha], \end{aligned}$$

with the terminal conditions

$$\begin{aligned} P_T =&S, &\text{(A12)}\\ \boldsymbol{s}_T^\alpha =& - S\boldsymbol{z}_T^\alpha, &\text{(A13)}\\ m_T^\alpha =&S||\boldsymbol{z}_T^\alpha||_2^2. &\text{(A14)} \end{aligned}$$

Further, the optimal control is given by

$$\begin{aligned} u_k^{o,\alpha} =& - (R + B^* P_{k+1} B)^{-1} B^* [P_{k+1}(A\boldsymbol{x}_k^\alpha + D\boldsymbol{z}_k^\alpha) &\text{(A15)}\\ &\quad + \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]]\\ =:& - F_k \boldsymbol{x}_k^\alpha - G_k \boldsymbol{z}_k^\alpha - H_k \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]. &\text{(A16)} \end{aligned}$$

Similarly, Γ_k and Ψ_k would be defined by

$$\begin{aligned} \Psi_k =& - S\mathbb{I} + F_k R(G_k \mathbb{I} + H_k \Psi_{k+1}\Gamma_k) &\text{(A17)}\\ &+ (A - BF_k)^* P_{k+1}\left[(D - BG_k)\mathbb{I}\right.\\ &\left. - B\Psi_{k+1} H_k \Gamma_k\right]\\ &+ (A - BF_k)^* \Psi_{k+1}\Gamma_k,\\ \Gamma_k =&(\mathbb{I} + BH_k \boldsymbol{M}\Psi_{k+1})^{-1}[(A - BF_k)\mathbb{I} &\text{(A18)}\\ &+ (D - BG_k)\boldsymbol{M}] \end{aligned}$$

with the terminal condition

$$\Psi_T = - S\mathbb{I}. \tag{A19}$$

The infinite time horizon discounted problem has a similar formulation:

By the same proof approach to the finite time non-discounted game, this is associated with the sequence of value functions

$$\begin{aligned} V_k^\alpha(\mathcal{F}_k^\alpha) =&\mathbb{E}_k[(\boldsymbol{x}_k^\alpha)^* P_k (\boldsymbol{x}_k^\alpha) + 2(\boldsymbol{x}_k^\alpha)^* \boldsymbol{s}_k^\alpha + m_k^\alpha], &\text{(A20)}\\ &k = \{0, ..., T\}, \end{aligned}$$

where P_k^ρ is an positive scalar, and $\boldsymbol{s}_k$ and m_k^α are $L_2([0,1])$ valued functions for all $k = \{0,...,T\}$ derived from the following backwards recurrence relations,

$$F_k^\rho = \rho(R + \rho B^* P_{k+1}^\rho)^{-1} B^* P_{k+1}^\rho A, \tag{A21}$$

$$G_k^\rho = \rho(R + \rho B^* P_{k+1}^\rho)^{-1} B^* P_{k+1}^\rho D, \tag{A22}$$

$$H_k^\rho = \rho(R + \rho B^* P_{k+1}^\rho)^{-1} B^*, \tag{A23}$$

$$P_k^\rho = S + F_k^{\rho *} R F_k^\rho + \rho(A - BF_k^\rho)^* P_{k+1}^\rho (A - BF_k^\rho), \tag{A24}$$

$$\begin{aligned} \boldsymbol{s}_k^\alpha = & - S\boldsymbol{d}_k^\alpha + F_k^{\rho *} R(G_k^\rho \boldsymbol{d}_k^\alpha + H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]) \\ & + \rho(A - BF_k^\rho)^* P_{k+1}^\rho \\ & \quad \times \left[(D - BG_k^\rho)\boldsymbol{d}_k^\alpha - BH_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] \\ & + \rho(A - BF_k^\rho)^* \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha], \end{aligned} \tag{A25}$$

$$\begin{aligned} m_k^\alpha = & \boldsymbol{d}_k^{\alpha *} S \boldsymbol{d}_k^\alpha + \rho \Bigg[(G_k^\rho \boldsymbol{d}_k^\alpha + H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha])^* R \\ & \quad \times (G_k^\rho \boldsymbol{d}_k^\alpha + H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]) \\ & + \left[(D - BG_k^\rho)\boldsymbol{d}_k^\alpha - BH_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right]^* P_{k+1}^\rho \\ & \quad \times \left[(D - G_k^\rho)\boldsymbol{d}_k^\alpha - BH_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] \\ & + 2\left[(D - G_k^\rho)\boldsymbol{d}_k^\alpha - BH_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]\right] \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha] \\ & + \mathbf{Q}(\alpha,\alpha) + \mathbb{E}_k[m_{k+1}^\alpha] \Bigg], \end{aligned} \tag{A26}$$

with the terminal conditions

$$P_T^\rho = S, \tag{A27}$$

$$\boldsymbol{s}_T^\alpha = - S\boldsymbol{d}_T^\alpha, \tag{A28}$$

$$m_T^\alpha = S||\boldsymbol{d}_T^\alpha||^2. \tag{A29}$$

Further, the optimal control is given by

$$\begin{aligned} u_k^{o,\alpha} = & - \rho(R + \rho B^* P_{k+1}^\rho)^{-1} B^* [P_{k+1}^\rho (A\boldsymbol{x}_k^\alpha + D\boldsymbol{d}_k^\alpha) \\ & + \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]] \end{aligned} \tag{A30}$$

$$=: - F_k^\rho \boldsymbol{x}_k^\alpha - G_k^\rho \boldsymbol{d}_k^\alpha - H_k^\rho \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]. \tag{A31}$$

Let the process to be tracked be given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$ for time k. Let Γ_k and Ψ_k be $L_2([0,1])$ operators which are defined by the backwards recursion equations

$$\begin{aligned} \Psi_k^\rho = & - S\mathbb{I} + F_k^\rho R(G_k^\rho \mathbb{I} + H_k^\rho \Psi_{k+1}^\rho \Gamma_k^\rho) \\ & + \rho(A - BF_k^\rho)^* P_{k+1}^\rho \left[(D - BG_k^\rho)\mathbb{I} \right. \\ & \left. - B\Psi_{k+1} H_k \Gamma_k\right] \\ & + \rho(A - BF_k)^* \Psi_{k+1} \Gamma_k, \end{aligned} \tag{A32}$$

$$\begin{aligned} \Gamma_k^\rho = & (\mathbb{I} + BH_k^\rho \boldsymbol{M} \Psi_{k+1})^{-1} [(A - BF_k^\rho)\mathbb{I} \\ & + (D - BG_k^\rho)\boldsymbol{M}] \end{aligned} \tag{A33}$$

with the terminal conditions

$$\Psi_T^\rho = - S\mathbb{I}. \tag{A34}$$

Assume that for all $k = \{0,...,T-1\}$, the inverse $(\mathbb{I} + BH_k \boldsymbol{M}\Psi_{k+1})^{-1}$ exists. Then, as with the non-discounted game,

$$\mathbb{E}_k[\boldsymbol{z}_{k+1}] = \Gamma_k^\rho \boldsymbol{z}_k, \tag{A35}$$

$$\boldsymbol{s}_k = \Psi_k^\rho \boldsymbol{z}_k, \tag{A36}$$

and the trajectory is generated by

$$\boldsymbol{u}_k^\rho = -F_k^\rho \boldsymbol{x}_k - (G_k^\rho \mathbb{I} + H_k^\rho \Psi_{k+1}\Gamma_k^\rho)\boldsymbol{z}_k \tag{A37}$$

gives the optimal tracking trajectory for each α.

Assume that the infinite horizon feedback solution (when it exists) is given by the fixed point to the following algebraic Riccati and operator equations:

$$F_\infty^\rho = \rho(R + \rho B^* P_\infty^\rho B)^{-1} B^* P_\infty^\rho A, \tag{A38}$$

$$G_\infty^\rho = \rho(R + \rho B^* P_\infty^\rho B)^{-1} B^* P_\infty^\rho D, \tag{A39}$$

$$H_\infty^\rho = \rho(R + \rho B^* P_\infty^\rho B)^{-1} B^*, \tag{A40}$$

$$P_\infty^\rho = S + F_\infty^{\rho *} R F_\infty^\rho + \rho(A - BF_\infty^\rho)^* P_\infty^\rho (A - BF_\infty^\rho), \tag{A41}$$

$$\begin{aligned}\Psi_\infty^\rho = &- S\mathbb{I} + F_\infty^\rho R(G_\infty^\rho \mathbb{I} + H_\infty^\rho \Psi_\infty^\rho \Gamma_\infty^\rho) \\ &+ \rho(A - BF_\infty^\rho)^* P_\infty^\rho \big[(D - BG_\infty^\rho)\mathbb{I} \\ &\quad - BH_\infty \Psi_\infty \Gamma_\infty\big] \\ &+ (A - BF_\infty^\rho)^* \Psi_\infty \Gamma_\infty^\rho,\end{aligned} \tag{A42}$$

$$\begin{aligned}\Gamma_\infty^\rho = &(\mathbb{I} + BH_\infty^\rho \boldsymbol{M}\Psi_\infty)^{-1}[(A - BF_\infty^\rho)\mathbb{I} \\ &+ (D - BG_\infty^\rho)\boldsymbol{M}].\end{aligned} \tag{A43}$$

A.2 Proof of Lemma 1

The dynamic programming principle is applied to find the optimal control. From the terminal condition

$$V_T^\alpha(\boldsymbol{x}_k) = ||\boldsymbol{x}_T^\alpha - \boldsymbol{d}_T^\alpha||_S^2. \tag{A44}$$

Then, $P_T = S$, $\boldsymbol{s}_T^\alpha = -S\boldsymbol{d}_T^\alpha$, and $m_T^\alpha = ||\boldsymbol{d}_T^\alpha||_S^2$.

By the dynamic programming assumption,

$$V_k^\alpha(\boldsymbol{x}_k) = \min_u \mathbb{E}_k \left[||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||u||_R^2 + V_{k+1}^\alpha(\boldsymbol{x}_{k+1})\right] \tag{A45}$$

$$= \min_u ||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||u||_R^2 + \mathbb{E}_k[V_{k+1}^\alpha(\boldsymbol{x}_{k+1})] \tag{A46}$$

$$\begin{aligned}= &\min_u ||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ \mathbb{E}_k[P_{k+1}(\boldsymbol{x}_{k+1}^\alpha)^2 + 2(\boldsymbol{x}_{k+1}^\alpha)\boldsymbol{s}_{k+1}^\alpha \\ &\quad + m_{k+1}^\alpha]\end{aligned} \tag{A47}$$

$$\begin{aligned}= &\min_u ||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ \mathbb{E}_k[(a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha + \boldsymbol{g}_k^\alpha)^2 P_{k+1}] \\ &+ 2\mathbb{E}_k[a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha + \boldsymbol{g}_k^\alpha]\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha] \\ &+ \mathbb{E}_k[m_{k+1}^\alpha]\end{aligned} \tag{A48}$$

$$\begin{aligned}= &\min_u ||\boldsymbol{x}_k^\alpha - \boldsymbol{d}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ (a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha)P_{k+1} \\ &\cdot (a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha) + \mathbf{Q}(\alpha, \alpha) \\ &+ 2(a\boldsymbol{x}_k^\alpha + b\boldsymbol{u}_k^\alpha + c\boldsymbol{d}_k^\alpha)\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha] + \mathbb{E}_k[m_{k+1}^\alpha].\end{aligned} \tag{A49}$$

Note that the right-hand expression of (A49) is differentiable and convex in u, and hence the optimal control is

$$u_k^{o,\alpha} = -(R + b^2 P_{k+1})^{-1} b[P_{k+1}(a\boldsymbol{x}_k^\alpha + c\boldsymbol{d}_k^\alpha) \tag{A50}$$

$$
\begin{aligned}
&+ \mathbb{E}_k[\boldsymbol{s}^{\alpha}_{k+1}]] \\
=: &- F_k\boldsymbol{x}^{\alpha}_k - G_k\boldsymbol{d}^{\alpha}_k - H_k\mathbb{E}_k[\boldsymbol{s}^{\alpha}_{k+1}]. \qquad \text{(A51)}
\end{aligned}
$$

Applying the optimal control to the value function and re-arranging terms gives equations (A6–A26) as required. □

A.3 Proof of Lemma 2

First, recall that by definition $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$, and hence when applying the optimal control at time $k = T-1$,

$$
\begin{aligned}
\mathbb{E}_{T-1}[\boldsymbol{z}_T] &= \mathbb{E}_{T-1}[\boldsymbol{M}\boldsymbol{x}_T] && \text{(A52)}\\
&= \mathbb{E}_{T-1}[\boldsymbol{M}(a\boldsymbol{x}_{T-1} + b\boldsymbol{u}_{T-1} + c\boldsymbol{z}_{T-1} + \boldsymbol{g}_{T-1})]] && \text{(A53)}\\
&= \boldsymbol{M}[a\boldsymbol{x}_{T-1} + b(-F_{T-1}\boldsymbol{x}_{T-1} && \text{(A54)}\\
&\quad - G_{T-1}\boldsymbol{z}_{T-1} - H_{T-1}\mathbb{E}_{T-1}[\boldsymbol{s}_T]) + c\boldsymbol{z}_{T-1}] \\
&= \boldsymbol{M}[(a - bF_{T-1})\boldsymbol{x}_{T-1} + (c - bG_{T-1})\boldsymbol{z}_{T-1} && \text{(A55)}\\
&\quad - bG_{T-1}\mathbb{E}_{T-1}[\boldsymbol{s}_T]] \\
&= (a - bF_{T-1})\boldsymbol{M}\boldsymbol{x}_{T-1} && \text{(A56)}\\
&\quad + (c - bG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} - bH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[\boldsymbol{s}_T] \\
&= (a - bF_{T-1})\boldsymbol{z}_{T-1} + (c - bG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} && \text{(A57)}\\
&\quad - bH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[\boldsymbol{s}_T].
\end{aligned}
$$

Then, applying the terminal condition $\boldsymbol{s}_T = -S\boldsymbol{z}_T$,

$$
\begin{aligned}
\mathbb{E}_{T-1}[\boldsymbol{z}_T] &= (a - bF_{T-1})\boldsymbol{z}_{T-1} + (c - bG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} && \text{(A58)}\\
&\quad + bH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[S\boldsymbol{z}_T] \\
\mathbb{E}_{T-1}[\boldsymbol{z}_T] - bH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[S\boldsymbol{z}_T] &= (a - bF_{T-1})\boldsymbol{z}_{T-1} && \text{(A59)}\\
&\quad + (c - bG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathbb{E}_{T-1}[\boldsymbol{z}_T] &= (\mathbb{I} - SBH_{T-1}\boldsymbol{M})^{-1}[(A - BF_{T-1})\mathbb{I} && \text{(A60)}\\
&\quad + (c - bG_{T-1})\boldsymbol{M}]\boldsymbol{z}_{T-1} && \text{(A61)}\\
&=: \Gamma_{T-1}\boldsymbol{z}_{T-1}. && \text{(A62)}
\end{aligned}
$$

Observing this, make the following inductive hypothesis:

$$
\begin{aligned}
\mathbb{E}_k[\boldsymbol{z}_{k+1}] &= \Gamma_k\boldsymbol{z}_k, && \text{(A63)}\\
\boldsymbol{s}_k &= \Psi_k\boldsymbol{z}_k, && \text{(A64)}
\end{aligned}
$$

where Ψ_k and Γ_k are $L_2([0,1])$ operators for each $k \in \{0, ..., T\}$. Applying the inductive hypotheses to the expectation of $\boldsymbol{z}_{k+1}$,

$$
\begin{aligned}
\mathbb{E}_k[\boldsymbol{z}_{k+1}] &= [(a - bF_k)\mathbb{I} + (c - bG_k)\boldsymbol{M}]\boldsymbol{z}_k && \text{(A65)}\\
&\quad - bH_k\boldsymbol{M}\mathbb{E}_k[\boldsymbol{s}_{k+1}] \\
&= [(a - bF_k)\mathbb{I} + (c - bG_k)\boldsymbol{M}]\boldsymbol{z}_k && \text{(A66)}\\
&\quad - bH_k\boldsymbol{M}\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}] \\
&= (\mathbb{I} + bH_k\boldsymbol{M}\Psi_{k+1})^{-1} && \text{(A67)}\\
&\quad \cdot [(a - bF_k)\mathbb{I} + (c - bG_k)\boldsymbol{M}]\boldsymbol{z}_k \\
&=: \Gamma_k\boldsymbol{z}_k, && \text{(A68)}
\end{aligned}
$$

which shows Eq. (A63). Applying the inductive hypotheses to the recursion for $\boldsymbol{s}_k$,

$$\begin{aligned}\boldsymbol{s}_k =& - S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}]) \\ & + (a - bF_k)P_{k+1} \\ & \cdot \left[(c - bH_k)\boldsymbol{z}_k - bH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}]\right] \\ & + (a - bF_k)^*\mathbb{E}_k[\boldsymbol{s}_{k+1}] \qquad \text{(A69)} \\ =& - S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]) \\ & + (a - bF_k)P_{k+1}\left[(c - bG_k)\boldsymbol{z}_k \right.\\ & \left. - bH_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]\right] \\ & + (a - bF_k)^*\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}] \qquad \text{(A70)} \\ =& - S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\Psi_{k+1}\mathbb{E}_k[\boldsymbol{z}_{k+1}]) \\ & + (a - bF_k)^*P_{k+1}\left[(c - bG_k)\boldsymbol{z}_k \right.\\ & \left. - b\Psi_{k+1}H_k\mathbb{E}_k[\boldsymbol{z}_{k+1}]\right] \\ & + (a - bF_k)^*\Psi_{k+1}\mathbb{E}_k[\boldsymbol{z}_{k+1}] \qquad \text{(A71)} \\ =& - S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\Psi_{k+1}\Gamma_k\boldsymbol{z}_k) \\ & + (a - bF_k)P_{k+1}\left[(c - bG_k)\boldsymbol{z}_k \right.\\ & \left. - B\Psi_{k+1}H_k\Gamma_k\boldsymbol{z}_k\right] \\ & + (a - bF_k)^*\Psi_{k+1}\Gamma_k\boldsymbol{z}_k \qquad \text{(A72)} \\ =:& \Psi_k\boldsymbol{z}_k. \qquad \text{(A73)}\end{aligned}$$

Then, the optimal control $\boldsymbol{u}_k^o$ is given in

$$\begin{aligned}\boldsymbol{u}_k^o =& - F_k\boldsymbol{x}_k - G_k\boldsymbol{z}_k - H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}] \qquad \text{(A74)} \\ =& - F_k\boldsymbol{x}_k - G_k\boldsymbol{z}_k - H_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}] \qquad \text{(A75)} \\ =& - F_k\boldsymbol{x}_k - (G_k\mathbb{I} + H_k\Psi_{k+1}\Gamma_k)\boldsymbol{z}_k. \qquad \text{(A76)}\end{aligned}$$

□

References

[1] Alain Bensoussan, Jens Frehse, and Phillip Yam. Mean Field Games and Mean Field Type Control Theory. SpringerBriefs in Mathematics. Springer, New York, NY, 2013.

[2] Peter E. Caines. Embedded Vertexon-Graphons and Embedded GMFG Systems. In IEEE Conference on Decision and Control, December 2022.

[3] Peter E. Caines and Minyi Huang. Graphon Mean Field Games and the GMFG Equations. In 2018 IEEE Conference on Decision and Control (CDC), pages 4129–4134, December 2018.

[4] Peter E. Caines and Minyi Huang. Graphon Mean Field Games and the GMFG Equations: Epsilon-Nash Equilibria. In 2019 IEEE 58th Conference on Decision and Control (CDC), pages 286–292, December 2019.

[5] Peter E. Caines and Minyi Huang. Graphon Mean Field Games and Their Equations. SIAM Journal on Control and Optimization, 59(6):4373–4399, January 2021.

[6] Peter E. Caines and Minyi Huang. Mean Field Games on Dense and Sparse Networks: The Graphexon MFG Equations. In 2024 American Control Conference (ACC), pages 4230–4235, July 2024.

[7] Rene Carmona and François Delarue. Probabilistic Theory of Mean Field Games with Applications II. 2018.

[8] Alex Dunyak and Peter E. Caines. Linear Stochastic Graphon Systems with Q-Space Noise. In 2022 IEEE 61st Conference on Decision and Control (CDC), pages 3926–3932, December 2022.

[9] Alex Dunyak and Peter E. Caines. Graphon Field Tracking Games with Discrete Time Q-noise. In 2023 62nd IEEE Conference on Decision and Control (CDC), pages 8194–8199, Singapore, Singapore, December 2023. IEEE.

[10] Shuang Gao and Peter E. Caines. Graphon Control of Large-Scale Networks of Linear Systems. IEEE Transactions on Automatic Control, 65(10):4090–4105, October 2020.

[11] Shuang Gao, Rinel Foguen Tchuendom, and Peter E. Caines. Linear quadratic graphon field games. Communications in Information and Systems, 21(3):341–369, June 2021. International Press of Boston.

[12] Minyi Huang, Peter E. Caines, and Roland P. Malhame. The NCE (Mean Field) Principle With Locality Dependent Cost Interactions. IEEE Transactions on Automatic Control, 55(12):2799–2805, December 2010.

[13] Frank L. Lewis, Draguna Vrabie, and Vassili Syrmos. The Tracking Problem and Other LQR Extensions. In Optimal Control, pages 177–212. John Wiley & Sons, Ltd, 2012.

[14] László Lovász. Large Networks and Graph Limits. 2012.