



Finite State Graphon Games with Applications to Epidemics

Alexander Aurell¹ · René Carmona¹ · Gökçe Dayanıklı¹ · Mathieu Laurière¹

Accepted: 25 October 2021 / Published online: 1 January 2022
© The Author(s) 2021

Abstract

We consider a game for a continuum of non-identical players evolving on a finite state space. Their heterogeneous interactions are represented with a graphon, which can be viewed as the limit of a dense random graph. A player's transition rates between the states depend on their control and the strength of interaction with the other players. We develop a rigorous mathematical framework for the game and analyze Nash equilibria. We provide a sufficient condition for a Nash equilibrium and prove existence of solutions to a continuum of fully coupled forward-backward ordinary differential equations characterizing Nash equilibria. Moreover, we propose a numerical approach based on machine learning methods and we present experimental results on different applications to compartmental models in epidemiology.

Keywords Graphon games · Epidemiological models · Machine learning

1 Introduction

With the recent pandemic of COVID-19, the importance of management of large populations in order to control the evolution of the disease has been recognized globally. How a pandemic plays out is a consequence of the interplay of many complex processes, e.g., disease specific

This article is part of the topical collection “Modeling and Control of Epidemics” edited by Quanyan Zhu, Elena Gubar and Eitan Altman.

This work was done with the support of NSF DMS-1716673, ARO W911NF-17-1-0578, and AFOSR # FA9550-19-1-0291.

✉ René Carmona
rcarmona@princeton.edu

Alexander Aurell
aaurell@princeton.edu

Gökçe Dayanıklı
gokced@princeton.edu

Mathieu Laurière
lauriere@princeton.edu

¹ Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA

spread mechanisms, the network of social interactions, and society-wide efforts to stop or slow the spread. As individuals, we have all made choices during the ongoing pandemic about the extent to which we minimize our personal risk of being infected. However, there is a trade-off between being careful and the pursuit of happiness. As we all have learned by now, our risk is not only determined by our own vigilance but also by others' choices and our environment.

In the framework of rational agents¹, each individual anticipates the action of their neighbours, neighbours' neighbours, etc., and any other external influence, then selects their action as a best response to others' actions. In other words, they want to optimize their private outcome while taking into account their surrounding environment, which includes other agents' actions. This type of strategic interaction is a non-cooperative game. The communication between agents in the game may be restricted by geography, social circles, and other factors. Moreover, people interact with different intensity, depending on their occupation and personality. Hence, the agents in the game, each with their own predisposition for risk, will act in a wide variety of ways and thus naturally form a heterogeneous crowd.

Consider the game discussed above where all agents anticipate the others' action and then selfishly plays a best response. Strategy profiles (the collection of all players' actions) consistent with such behavior are Nash equilibria, *i.e.*, profiles such that no player profits from a unilateral deviation. Computing Nash equilibria in games with large number of players is particularly hard; under some specific assumptions, approximate equilibria can be found by using the mean field game approach developed independently by Lasry-Lions [45,46] and Huang-Malhamé-Caines [35]. The approach has found many practical applications, examples include the mathematical modeling of price movement in energy markets; pedestrian crowd motion and evacuation; and epidemic disease spread.

One of the fundamental assumptions in mean field game theory is that agents are indistinguishable and interact homogeneously. However, in some real-world applications such as the modeling of epidemics, the diversity of individuals and the variation of their interactions are important factors to consider. Examples of such include the effects of travel restrictions, multiple age groups with distinct social behavior and risk profiles, and the spectrum of pre-existing health conditions. The aspects listed above require the game to have a non-uniform underlying network structure. Games with a large number of non-identical players can be analyzed with so-called graphon games whenever the network specifying the interactions is dense².

Epidemics are driven by the spread of the disease from infected to susceptible agents. The set of susceptible agents is not necessarily the whole non-infected population, for certain diseases immunity is gained after exposure. The evidence suggests that the COVID-19 virus mainly spreads through close contact³. Fortunately, the disease transmission probability can be decreased by efforts of the individual. For example, an individual can choose to avoid public and closed spaces, wear a protective mask or do their shopping online. When two people meet the disease transmission probability depends on both sides' effort. The disease is less likely to occur if both parts wear protective masks than if just one part does. However, the decrease in risk of transmission is not additive in the interacting agents' efforts. Using this intuition, we assume in this paper that disease transmission likelihood depends on efforts

¹ We will interchangeably use the words agent, player, and individual.

² Loosely speaking, a dense network or graph is one in which the number of edges is close to the maximal number of edges. For such graphs, there is a functional limit of the adjacency matrix interpreted as a step function. See [48] for an exposé of the theory of such limits.

³ <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html> (last accessed on 10 April, 2021)

in a multiplicative way. The effort of the agents will therefore be given the name “contact factor control”, in line with the game-based model introduced in [2].

Before giving the full description of the epidemiological graphon game model, how it can be approached numerically, and all technical details, we expand upon the heuristics of graphon-based interaction with contact factor control and review the literature of related fields of research in the next sections.

1.1 The SIR Graphon Game with Contact Factor Control

The most famous compartmental model in epidemics is arguably the classical Susceptible-Infected-Removed (SIR) model⁴. In this section, we take advantage of its wide familiarity and compact formulation to further motivate for the concept of contact factor control and graphon-type interaction.

In order to give a description of the rate at which agents become infected, we need to first introduce some notation. Consider N individuals who transition between the states Susceptible (S), Infected (I), and Removed (R). An individual in state R has either gained immunity or deceased. Denote the state of agent $j \in \{1, \dots, N\}$ at time t by $X_t^{j,N}$. A susceptible individual might encounter infected individuals, resulting in disease transmission. Encounters occur pairwise and randomly throughout the population with intensity β . The number of encounters with infected agents in a short time interval $[t - \Delta t, t)$ is approximately proportional to the the share of the population in state I at t . Between each agent pair (j, k) , we set the interaction strength to $w(x_j, x_k)$, where w is a graphon (see Definition 1 for its definition) and x_j, x_k are random variables uniformly distributed on $[0, 1]$. Hence agent k 's transition rate from state S to I is scaled by w and of the form $\beta \frac{1}{N} \sum_{k=1}^N w(x_j, x_k) \mathbf{1}_I(X_{t-}^{k,N})$. Upon infection, an individual starts the path to recovery. The jump from state I to R happens after an exponentially distributed time with rate γ . The state R is absorbing.

Denoting $Z_t^{j,N} := \frac{1}{N} \sum_{k=1}^N w(x_j, x_k) \mathbf{1}_I(X_{t-}^{k,N})$, the transition rate matrix for player $j \in \{1, \dots, N\}$ is

$$Q(Z_t^{j,N}) = \begin{bmatrix} -\beta Z_t^{j,N} & \beta Z_t^{j,N} & 0 \\ 0 & -\gamma & \gamma \\ 0 & 0 & 0 \end{bmatrix}. \tag{1}$$

Here we use the order S, I, R for the columns and the rows. For instance, the term $\beta Z_t^{j,N}$ encodes the rate at which an agent arrives in I coming from S. As explained above, this rate is proportional to the weighted average $Z_t^{j,N}$ of infected agents interacting with player j . At this stage, the distinguishability of the players is seen in the aggregate variables $(Z_t^{j,N})_{j=1}^N$ which in general differ in value. As $N \rightarrow \infty$, we expect the probability distribution flow of player j to converge to the solution of the ordinary differential equation (ODE)

$$\dot{p}^{x_j}(t) = p^{x_j}(t)Q(Z_t^{x_j}), \quad p^{x_j}(0) = p_0^{x_j}, \tag{2}$$

where $p_0^{x_j}$ is some given initial distribution over the states and

$$Z_t^{x_j} := \int_I w(x_j, y)p^y(t, I)dy. \tag{3}$$

Here $p^{x_j}(t)$ is understood as a vector of length 3 whose coordinates correspond to the probability for player x_j of being in state S, I and R, respectively at time t . Equation (2)

⁴ Depending on the modeling, R can refer to “Removed” or “Recovered”.

encodes the evolution of these probabilities. $Z_t^{x_j}$ is the average number of infected agents around player x_j , weighted by the pairwise interaction strength $w(x_j, y)$. Scaling $p^x(t)$, $x \in [0, 1]$, by a population size N , $Np^x(t) =: (S^x(t), I^x(t), R^x(t))$, we retrieve a formulation of the compartmental SIR model with graphon-based interactions

$$\begin{aligned} \dot{S}^x(t) &= -\frac{\beta}{N} \left(\int_I w(x, y) I^y(t) dy \right) S^x(t), & S^x(0) &= Np_0^x(S), \\ \dot{I}^x(t) &= \frac{\beta}{N} \left(\int_I w(x, y) I^y(t) dy \right) S^x(t) - \gamma I^x(t), & I^x(0) &= Np_0^x(I), \\ \dot{R}^x(t) &= \gamma I^x(t), & R^x(0) &= Np_0^x(R). \end{aligned}$$

So far we have considered a model without action. Assuming agents choose a “contact factor” in order to decrease their risk of getting infected, which enters the model in line with the discussion above, we express this new feature mathematically as follows: Given that the meeting frequency is β , pairing is random, disease spreads from infected agents to susceptible, and the spread probability is scaled by the efforts of the individuals that meet in a multiplicative way, the transition rate for individual j from S to the I is

$$\beta \alpha_t^j \frac{1}{N} \sum_{k=1}^N w(x_j, x_k) \alpha_t^k \mathbf{1}_I(X_{t-}^{k,N}), \tag{4}$$

where α_t^k denotes the (contact factor) action of individual $k \in \{1, \dots, N\}$ at time t , selected from set of actions A . Along the lines of the heuristics of mean field game theory, we anticipate that in an appropriate approximation of our interacting system in the limit $N \rightarrow \infty$, agent x transitions from susceptible to infected with rate

$$\beta \alpha_t^x \int_I w(x, y) \left(\int_A a \rho_t^y(da, l) \right) dy, \tag{5}$$

where ρ_t^y is the joint distribution of action and state of player y at time t and α_t^x is the action of individual x .

1.2 Related Literature

1.2.1 Graphon Games and Finite State Space Mean Field Games

Challenges related to large populations arise in the game theory, and mean field game (MFG) theory presents a toolbox to compute equilibria in these games with large number of players. MFGs were first developed for continuous state space models [35,45,46] and later for a finite state space models [32,33,42]. The theory for finite state MFGs has been extended in many directions, with contributions including a probabilistic approach [14], the master equation approach [4], minor-major player games [12], and extended games⁵ [13]. Further, finite state mean field control, risk-sensitive control, and zero-sum games are treated in [17–19] which cover cases of unbounded jump intensities. Graphon games [3,7–9,23,30,49] have recently been receiving an increasing research interest. The motivation is the study of strategic decision making in the face of a large non-complete dense networks of distinguishable agents. The

⁵ In MFG theory “extended” refers to formulations where interaction is modelled with the joint distribution of action and state of players.

graphon game's rising popularity stems from its ability to handle heterogeneity of agents to an extent far beyond the MFG theory.

1.2.2 Mean Field Games and Related Models for Epidemics

Decision making in compartmental models (*e.g.*, the SIR model) have been studied intensively for a long time, with an increasing interest recently with the COVID-19 pandemic. In the form of games and optimal control problems, disease-combating efforts ranging from strategies for social contracts to vaccination have been analyzed in the literature. Here, we focus on work relying on the graphon game theory and the mean-field approach.

During an epidemic where the disease is prone to close-contact transmission (one example being COVID-19) control of the contact rate or other social distancing protocols are go-to solutions in the fight against the disease. Such strategies and related variations have been studied in the context of mean field games [16,28], mean field optimal control [47], and mean field type games [55]. In reality the population needs to be tested for the disease in order to accurately assess the risks in the decision making process. Two recent papers studying optimal testing policies are [15,36]. The effects of vaccination and accompanying decision making problems are not studied in this paper; they have been analyzed in MFG-related settings since before the COVID-19 pandemic [27,31,37,43,44,50]. Recently, the interplay between a population in which a disease spreads and a regulator has been studied in the form of Stackelberg games. In such models, the members in the populations are taking actions based on policies issued by the regulator, while the regulator anticipates the population's reaction and optimizes the policy. The case of a cooperative population has been studied in [36], while in [2] a population of selfish agents with contact factor control has been studied.

An example of a deterministic optimal control problem for centralized decision making during a pandemic in a society with multiple communicating subpopulations is given in [29]. The subpopulations interact over a non-uniform graph. A central planner wants to flatten the (global) curve of infections, leading to the optimal control problem. Sending the number of subpopulations to infinity, we anticipate a limit where each interaction is weighted by a graphon and the limit model would be reminiscent of the interacting system of Kolmogorov equations studied in this paper.

1.2.3 Networks and Graphons in Epidemiology

There is a vast body of literature on epidemiology modeling with network interactions. A review of the studies that use idealized networks⁶ in epidemiology models can be found in [39]. More closely related to the ideas in this paper, there are recent contributions connecting epidemic models and graphons. In [57] a sensitivity analysis on the graphon SIS epidemic model is conducted. An infinite dimensional SIS model with application to targeted vaccination is considered in [24]. The paper [40] proposes a model with local-density dependent Markov processes interacting through a graphon structure, and considers applications to epidemiology. In a similar but more general setting to the SIR model with graphon interaction, [1] studies convergence of a stochastic particle system to an SIR-like system of PDEs with spatial interaction. We note [1] and its continuation [25,26] may be relevant for a future study of the convergence of N -player Nash equilibria to the equilibria in the finite state graphon

⁶ Some examples of idealized networks given in this work are Small-World Networks, Scale-Free Networks, Exponential Random Graph Models.

game. The works mentioned in this section only consider the dynamics of the population without taking the agents' decision making into account.

1.3 Contributions and Paper Structure

This paper is, to the best of our knowledge, the first to address the analysis and numerical resolution of graphon games that are time-dependent and with a discrete state space. The application to epidemiology model departs from the traditional literature on epidemiology models and graphon models by the incorporation of a game theoretical aspect: here we go beyond dynamic graphon systems and find Nash equilibria for rational agents. We construct a probabilistic particle model for a continuum of interacting agents and prove that graphon aggregates must be deterministic (as in *e.g.* (5)) under a set of natural conditions on the strategies and transition rates. This motivates the study of the asymptotic deterministic model formulation and gives a transparent interpretation of the agent's control in the applied context. We derive theoretical results for the deterministic model: a verification theorem and an existence theorem for the coupled continuum of forward-backward ordinary differential equations (FBODEs) that characterize the finite state graphon game at equilibrium are proven. This is reminiscent of the mean field game framework, except that here the population is heterogeneous due to the graphon-based interactions. This makes the computation of solutions much more challenging. We then propose a machine learning method to solve the FBODE system. Finally, we consider a graphon game model for epidemic disease spread. Multiple test cases are solved with the proposed numerical method and the experimental results are discussed.

The outline of the rest of the paper is as follows: In Sect. 2, we introduce the model and analyze its deterministic formulation. In Sect. 3, we introduce the numerical approach and give experiment results. In Sect. 4, a theoretical framework for the model's probabilistic framework is presented and we rigorously define the graphon game. For the sake of conciseness, the proofs are postponed to the appendices.

2 Model

2.1 Setup and Preliminaries

Let $n \in \mathbb{N}$ and let E be the finite set $\{1, \dots, n\}$. For each $e \in E$ define the difference operator Δ_e acting on functions on E by the formula $[\Delta_e \phi](e') = \phi(e') - \phi(e)$. We identify the set of probability measures on E , $\mathcal{P}(E)$, with the simplex $\Delta(E) := \{x = (x_1, \dots, x_n) \in \mathbb{R}_+^n : \sum_i x_i = 1\}$ and endow it with the Euclidean distance. Throughout the paper, the notation $\mathcal{P}(\cdot)$ will be used to denote the set of Borel probability measures.

Let $T > 0$ be a finite time horizon. A process $(f_t)_{t \in [0, T]}$ will be denoted with its bold letter symbol f . Let $\mathcal{C} := C([0, T]; \mathbb{R})$ be the space of continuous real-valued functions from $[0, T]$, $\mathcal{D} := D([0, T]; \mathbb{R})$ be the space of real-valued functions from $[0, T]$ càdlàg at $t \in [0, T)$ and continuous at $t = T$. We denote the uniform norm by $\|x\|_T := \sup_{s \in [0, T]} |x(s)|$, $x \in \mathcal{D}$. We note that $(\mathcal{C}, \|\cdot\|_T)$ and $(\mathcal{D}, \|\cdot\|_T)$ are both Banach spaces, only the former is separable. Let $\mathcal{D}_E \subset \mathcal{D}$ be the set of functions in $f \in \mathcal{D}$ such that $f([0, T]) \subset E$. Since \mathcal{D}_E is a closed subset of \mathcal{D} , $(\mathcal{D}_E, \|\cdot\|_T)$ is a complete metric space.

Let I be the unit interval equipped with the Euclidean distance. We denote by λ_I and $\mathcal{B}(I)$ the Lebesgue measure and Borel σ -field on I , respectively. The set I is indexing the

continuum of players in the graphon game. Throughout the paper, we will employ the notation $\underline{\phi} := (\phi(x))_{x \in I}$ for functions with domain I . Furthermore, in most cases we will denote the index argument with a superscript: $\phi^x := \phi(x)$.

This paper studies games with heterogeneous interactions. When the players in the game interact, the weight they give to each others' action is parameterized by their indices. This weighted averaging is captured as the integration with respect to a graphon kernel function (see Section 1 for the discussion). Here, we give the most prevalent definition of a graphon.

Definition 1 A graphon is a symmetric Borel-measurable function, $w : I \times I \rightarrow [0, 1]$.

The graphon induces an operator W from $L^2(I)$ to itself: for any $\underline{\phi} \in L^2(I)$,

$$[W\underline{\phi}]^x := \int_I w(x, y)\phi^y dy. \tag{6}$$

Remark 1 The definition of a graphon varies somewhat in the literature. Some authors require neither symmetry nor a non-negative range. As mentioned in the introduction, the graphon as defined here can be used to represent the limit of a sequence of dense random graphs as the number of nodes (players) goes to infinity. For example, the constant graphon $w(x, y) = p$ is in a sense the limit of a sequence of Erdős-Rényi graphs with parameter $p \in [0, 1]$. Conversely, random graphs can be sampled from a graphon in at least two different ways: either we sample points in I and construct a weighted graph whose weights are given by the graphon, or we sample points in I and then sample edges with probabilities given by the graphon.

A Q -matrix with real-valued entries $q_{i,j}$, $i, j \in E$, is an $n \times n$ matrix with non-negative off-diagonal entries such that:

$$q_{i,i} = - \sum_{j=1:j \neq i}^n q_{i,j}, \quad i \in E. \tag{7}$$

In this paper, we will consider controlled Q -matrices with entries that depend on population aggregates. More specifically, we let for each $x \in I$, $q_{i,j}^x : A \times \mathbb{R} \mapsto \mathbb{R}$, $i, j \in E$, be bounded measurable functions such that $Q^x(\alpha, z) := [q_{i,j}^x(\alpha, z)]_{i,j=1}^n$ is a Q -matrix for all $(\alpha, z) \in A \times \mathbb{R}$. We are going to work under the following assumption on the rates $q_{i,j}^x$:

Condition 1 (i) There is a finite constant $q_{\max} > 0$ such that for all $x \in I$, $(i, j) \in E^2$, $a \in A$, and $z \in \mathbb{R}$:

$$|q_{i,j}^x(a, z)| \leq q_{\max}. \tag{8}$$

(ii) There is a finite constant $C > 0$, possibly depending on n , such that for $p = 1, 2$ and for all $x \in I$, $t \in [0, T]$, $\alpha \in A$, $(i, j) \in E^2$, and $z, z' \in \mathbb{R}$

$$|q_{i,k+i}^x(\alpha, z) - q_{j,k+j}^x(\alpha, z')| \leq C (\mathbf{1}_{\{i \neq j\}} + |z - z'|^p).$$

Since the state space is finite and the rates are assumed to be uniformly bounded in Condition 1.(i), the Hölder-continuity assumption of Condition 1.(ii) is less restrictive than it otherwise would have been.

2.2 The Finite State Graphon Games Model for Epidemiology

In this section, we give a descriptive introduction to the epidemiological graphon game without going in to all the technical details. A rigorous mathematical motivation, built on the

theory of Fubini extensions to accommodate for a continuum of independent jump processes, is presented in Sect. 4.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a continuum of E -valued pure jump processes $X^x = (X_t^x)_{t \in [0, T]}$ indexed over $x \in I$. That is, for each $\omega \in \Omega$, $X^x(\omega) \in \mathcal{D}_E$. The stochastic process X^x models the state trajectory of player x . The initial state X_0^x is sampled from a distribution $p_0^x \in \mathcal{P}(E)$. Each player implements a strategy α^x , a process taking values in the compact interval $A \subset \mathbb{R}$ described in more detail below. The players interact and each player’s state trajectory is potentially influenced by the whole strategy profile $(\alpha^x)_{x \in I}$. To emphasize dependence we denote the state trajectory of player x as $X^{\alpha, x}$ given a strategy profile $\alpha = (\alpha^x)_{x \in I}$. For all $x \in I$, $X^{\alpha, x}$ is a E -valued pure jump process with rate matrix $Q^x(\alpha_t^x, Z_t^{\alpha, x})$ at time $t \in [0, T]$. The rate matrix is controlled by α^x and influenced by the population aggregate $Z^{\alpha, x}$. The aggregates we consider are averages weighted by a graphon w , more specifically of the form $[WK(\alpha_t, X_{t-}^\alpha)]$ (cf. (6)) for some function $K : A \times E \rightarrow \mathbb{R}$. In Sect. 4 we prove that the aggregate is a deterministic function of time, henceforth we write

$$Z_t^{\alpha, x} = \int_I w(x, y) \mathbb{E} [K(\alpha_t^y, X_{t-}^{\alpha, y})] dy, \tag{9}$$

where w is a graphon.

K will be called the *impact function* since it quantifies how much a player’s joint state and control distribution impacts the aggregate variable. One example is the impact function in Sect. 1.1 where $K(\alpha, e) = \alpha \mathbf{1}_{(e=1)}$, where the interpretation being that the aggregate is the averaged contact factor control of infected players. We have the following assumption on K :

Condition 2 *There exist finite constants $L_K, C_K > 0$ such that for all $a, a' \in A$ and $e \in E$:*

$$|K(a, e) - K(a', e)| \leq L_K |a - a'|, \quad |K(a, e)| \leq C_K.$$

Our (for now formal) probabilistic definition of the interacting system of players is complete. The rigorous analysis of the system, the construction of a continuum of state trajectories, and conditions under which the aggregate is deterministic, *i.e.*, of the form (9), is treated in detail in Sect. 4.

The key feature of the graphon game is that the aggregate variable is in general not the same for two distinct players. The players are therefore distinguishable and there is no “representative agent”, as in MFGs⁷. As a direct consequence, there is no flow of player state distributions common to all players. Instead, each player has their private flow. Denote by $p^{\alpha, x}(t, e)$ the probability that player $x \in I$ that is in state $e \in I$ at time $t \in [0, T]$, given that the population plays the strategy profile α . We shall argue that player x ’s state distribution flow $p^{\alpha, x}$ solves the Kolmogorov forward equation

$$\frac{d}{dt} p^{\alpha, x}(t) = p^{\alpha, x}(t) Q^x(\alpha_t^x, Z_t^{\alpha, x}), \quad p^{\alpha, x}(t) := (p^{\alpha, x}(t, e))_{e \in E} \tag{10}$$

with initial condition $p^{\alpha, x}(0) = p_0^x$ and the player’s aggregate variable $Z_t^{\alpha, x}$ is

$$Z_t^{\alpha, x} = \int_I w(x, y) \left(\int_{A \times E} K(a, e) \rho_t^{\alpha, y}(da, de) \right) dy, \tag{11}$$

⁷ This is of course not the case if the graphon is constant. In this case, the graphon game is in fact equivalent to an MFG, so there is a representative agent. There are a few more examples of this kind, such as the piecewise constant graphon yielding a game equivalent to a multipopulation MFG with one representative agent for each subpopulation.

with $\rho_t^{\alpha,y}$ being the joint probability law of control and state, $(\alpha_t^y, X_t^{\alpha,y})$.

We turn our focus to the players’ actions. We will make three standing assumptions that directly affect which strategies the players will choose. The first is that the environment that endogenously affects the players (but is known to the players) is varying smoothly over time, with no abrupt changes for example in lockdown penalties or expected recovery time. Secondly, if the players can impact their environment with their control, then the environment varies smoothly with their control too. For example, the risk of infection depends continuously on the agent’s level of cautiousness. Finally, the players’ strategies are decentralized, *i.e.*, are unaffected by the transition of any agent other than themselves. Under these circumstances, the players have no apparent reason to discontinuously change their action over time except at times of transition between states. Such strategies (A -valued; decentralized; continuous in time between changes of the player’s own state) will be called admissible and the set of admissible strategies denoted by \mathbb{A} . The setting is further discussed in Sect. 4.

In this paper, we focus on the finite horizon problem where the cost is composed of two components: a running cost and a terminal cost. For each player $x \in I$, the assumptions on the conditions that the running and terminal cost functions, $f^x : [0, T] \times E \times \mathbb{R} \times A \rightarrow \mathbb{R}$ and $g^x : E \times \mathbb{R} \rightarrow \mathbb{R}$, satisfy are given later in the text together with the theoretical results. The total expected cost to player x for playing the strategy $\sigma \in \mathbb{A}$ while the population plays the strategy profile $\underline{\alpha}$ is

$$\mathcal{J}^x(\sigma; \underline{\alpha}) = \mathbb{E} \left[\int_0^T f^x(t, X_t^{\alpha,x}, Z_t^{\alpha,x}, \sigma_t) dt + g^x(X_T^{\alpha,x}, Z_T^{\alpha,x}) \right]. \tag{12}$$

As we shall see, a change in player x ’s control has no effect on the aggregate. Hence, the expected cost depends on the strategy profile only indirectly through the value of the aggregate variable. See Sect. 4 for the details. Therefore, hereinafter we shall use the notation $J^x(\sigma; \mathbf{Z}^{\alpha,x})$ for the right-hand side of (12). In light of this, we employ the following definition of a Nash equilibrium in the graphon game:

Definition 2 The strategy profile $\underline{\alpha}$ is a Nash equilibrium if it is admissible and no player can gain from a unilateral deviation, *i.e.*,

$$J^x(\alpha^x; \mathbf{Z}^{\alpha,x}) \leq J^x(\sigma; \mathbf{Z}^{\alpha,x}), \quad \forall x \in I, \forall \sigma \in \mathbb{A}.$$

2.3 Analysis of Finite State Graphon Games

By Definition 2, an admissible strategy profile $\hat{\alpha}$ is a Nash equilibrium if there exists an aggregate profile $\hat{\mathbf{Z}} = (\hat{Z}^x)_{x \in I}$ such that

- for all $x \in I$, $\hat{\alpha}^x$ minimizes $J^x(\cdot; \hat{\mathbf{Z}}^x)$;
- for all $x \in I$, \hat{Z}^x is the aggregate perceived by player x if the population uses strategy profile $\hat{\alpha}$.

This alternative formulation has the advantage to split the characterization of the equilibrium into two parts and in the first part, the optimization problem faced by a single agent is performed with while the aggregate is fixed.

With a flow \mathbf{Z}^x being fixed, we define the value function of player $x \in I$ as, for $t \in [0, T]$ and $e \in E$,

$$u^x(t, e) := \inf_{\sigma \in \mathbb{A}} \mathbb{E} \left[\int_t^T f^x(s, X_s^{\sigma, Z^x, x}, Z_s^x, \sigma_s) dt + g^x(X_T^{\sigma, Z^x, x}, Z_T^x) \mid X_t^{\sigma, Z^x, x} = e \right]$$

where $X^{\sigma, Z^x, x}$ is an E -valued pure jump process with transition rate matrix $Q^x(\sigma_t, Z_t^x)$ at time $t \in [0, T]$ and initial distribution p_0^x .

To derive optimality conditions, we introduce the Hamiltonian of player x :

$$H^x(t, e, z, h, a) := \vec{\mathbf{1}}_e Q^x(a, z)h + f^x(t, e, z, a), \tag{13}$$

where $h \in \mathbb{R}^n$ and $\vec{\mathbf{1}}_e$ is the coordinate (row) vector in direction e in \mathbb{R}^n . We assume that $A \ni \alpha \mapsto H^x(t, e, z, h, \alpha)$ admits a unique measurable minimizer $\hat{a}_e^x(t, z, h)$ for all (t, z, h) and define the minimized Hamiltonian of player x :

$$\hat{H}^x(t, e, z, h) := H^x(t, e, z, h, \hat{a}_e^x(t, z, h)). \tag{14}$$

The dynamic programming principle of optimal control leads to the HJB equation for u^x that reads

$$\dot{u}^x(t, e) + \hat{H}^x(t, e, Z_t^x, u^x(t, \cdot)) = 0, \quad u^x(T, e) = g^x(e, Z_T^x), \tag{15}$$

where $\dot{u}^x(t, e)$ denotes the time derivative of $u^x(t, e)$. Noting that $\vec{\mathbf{1}}_e Q(a, z)h = \vec{\mathbf{1}}_e Q(a, z)\Delta_e h$, (15) can be equivalently written as

$$\dot{u}^x(t, e) + \hat{H}^x(t, e, Z_t^x, \Delta_e u^x(t, \cdot)) = 0, \quad u^x(T, e) = g^x(e, Z_T^x). \tag{16}$$

In the following theorem, we verify that the solution of the HJB equation indeed is the value function of the infinitesimal agent’s control problem, and we provide an expression for an optimal Markovian control in terms of this value function and the aggregate.

Theorem 1 *If $u^x : [0, T] \times E \ni (t, e) \mapsto u^x(t, e) \in \mathbb{R}$ is a continuously differentiable solution to the HJB equation (15), then u^x is the value function of the optimal control problem when the flow Z^x is given. Moreover, the function*

$$\hat{\phi}^x(t, e) = \hat{a}_e^x(t, Z_t^x, u^x(t, \cdot)) \tag{17}$$

gives an optimal Markovian control.

Next, we prove the existence of a solution to the coupled Kolmogorov-HJB system at equilibrium. For that purpose we place the following condition:

- Condition 3** (i) *There exist two functions Q_1 and Q_2 with Q_2 locally Lipschitz such that $Q(a, z) = Q_1(z) + aQ_2(z)$ for all $a \in A, z \in \mathbb{R}$.*
 (ii) *$a \mapsto f^x(t, e, z, a)$ is continuously differentiable, and as a function of a it is strongly convex, uniformly in (t, e, z) with constant λ ; $(t, z) \mapsto \partial_a f^x(t, e, z, a)$ is locally Lipschitz continuous, uniformly in $a \in A, e \in E$.*
 (iii) *f and g are uniformly bounded and $z \mapsto f^x(t, e, z, a)$ is continuous.*

As a consequence of Condition 3(i) and 3(ii) $a \mapsto H^x(t, e, z, h, a)$ is continuously differentiable and strictly convex, and $(t, z, h) \mapsto \hat{a}_e(t, z, h)$ is locally Lipschitz continuous (see Lemma 1 in Appendix A.2). We denote Lipschitz constant of $[-c, c] \ni z \mapsto \hat{a}(t, z, h)$ by $L_{\hat{a}}(c)$, which can be bounded from above using smoothness properties of Q^x and f^x . More specifically, $L_{\hat{a}}$ depends on the local Lipschitz coefficients of $z \mapsto \partial_a Q^x(a, z)$ and $(t, z, h) \mapsto \partial_a f^x(t, e, z, a)$, see the proof of Lemma 1). Recall that C_K denotes the uniform upper bound of the impact function K guaranteed by Condition 2.

Theorem 2 *Assume Conditions 1, 2, 3 hold. If $\|w\|_{L^2(I \times I)} L_K L_{\hat{a}}(C_K) < 1$, then the coupled Kolmogorov–HJB forward-backward system at equilibrium*

$$\begin{cases} \dot{p}^x(t, e) = \sum_{e' \in E} q_{e',e}^x(\hat{\phi}^x(t, e'), Z_t^x) p^x(t, e'), \quad \forall e \in E, t \in [0, T] \\ \dot{u}^x(t, e) = -\hat{H}^x(t, e, Z_t^x, \Delta_e u^x(t, \cdot)), \quad \forall e \in E, t \in [0, T] \\ Z_t^x = \int_I w(x, y) \left(\sum_{e \in E} K(\hat{\phi}^y(t, e), e) p^y(t, e) \right) dy \\ u^x(T, e) = g^x(e, Z_T^x), \quad p^x(0, e) = p_0^x(e), \\ \hat{\phi}^x(t, e) = \hat{a}_e^x(t, Z_t^x, u^x(t, \cdot)) \end{cases} \quad (18)$$

admits a bounded solution (u, p) in $C([0, T]; L^2(I \times E) \times L^2(I \times E))$ such that for each $(t, x) \in [0, T] \times I$, $p^x(t, \cdot)$ is a probability mass function on E .

2.4 The Finite State Graphon Game for SIR with Contact Factor Control

Here, we introduce a model which we shall use as a test bed for the numerical algorithm presented in Sect. 3. It is inspired by the first example scenario in [2] and builds on the case discussed in Sect. 1.1. It is a compartmental model with four possible states: (S)usceptible, (I)nfected, (R)ecovered and (D)eceased. The agents choose their level of contact factor. A regulator (government or health care authority) recommends state-dependent contact factor levels to the agents, denoted by $\lambda^{(e)}$, $e \in \{S, I, R\}$. For enforcement purposes, it also sets penalties for deviation from these levels. The cost has 3 components: The first component penalizes the agent for not following the regulator’s recommended contact factor level, the second is the cost of treatment for an infected agent (this cost can be player specific due to individual differences in health care plan coverage, etc.), and the last one is the cost for being deceased. In this setting, the running cost is written as

$$\begin{aligned} f^x(t, e, z, \alpha) &= \frac{c_\lambda}{2} \left(\lambda^{(S)}(t) - \alpha \right)^2 \mathbf{1}_{(e=S)} + \frac{1}{2} \left(\lambda^{(R)}(t) - \alpha \right)^2 \mathbf{1}_{(e=R)} \\ &\quad + \left(\frac{1}{2} \left(\lambda^{(I)}(t) - \alpha \right)^2 + c_I(x) \right) \mathbf{1}_{(e=I)} + c_D(x) \mathbf{1}_{(e=D)}, \end{aligned} \quad (19)$$

where c_I and c_D are nonnegative costs functions (of the player index). We set the terminal cost to be identically zero, $g^x(e, z) = 0$ for all $(e, z) \in E \times \mathbb{R}$. The transition rate matrix for player x is given as:

$$Q^x(\alpha, z) = \begin{matrix} & \begin{matrix} S & I & R & D \end{matrix} \\ \begin{matrix} S \\ I \\ R \\ D \end{matrix} & \begin{bmatrix} \dots & \beta(x)\alpha z & 0 & 0 \\ 0 & \dots & \rho(x)\gamma(x) & (1 - \rho(x))\gamma(x) \\ \kappa(x) & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots \end{bmatrix} \end{matrix}, \quad (20)$$

where $\beta, \gamma, \kappa, \rho$ are nonnegative parameter functions, $0 \leq \rho \leq 1$, determining the rates of infection, recovery, reinfection, and decrease, and where in each row, the diagonal entry is the negative of the sum of the other terms on the same row. In line with the discussion

in Sect. 1.1, the transition rate from state S to I depends on the agent's own decision and the aggregate variable⁸. Furthermore, when an infected agent transitions, she goes to state R with probability ρ and to state D with probability $(1 - \rho)$. Then, for player x the optimality conditions yield

$$\begin{aligned}\hat{\phi}^x(t, S) &= \lambda^{(S)}(t) + \frac{\beta(x)}{c_\lambda} Z_t^x (u^x(t, S) - u^x(t, I)) \\ \hat{\phi}^x(t, I) &= \lambda^{(I)}(t) \\ \hat{\phi}^x(t, R) &= \lambda^{(R)}(t)\end{aligned}$$

and the forward-backward graphon ODE system reads:

$$\begin{aligned}\dot{p}^x(t, \cdot) &= p^x(t, \cdot) Q^x(\hat{\phi}^x(t, S), Z_t^x) \\ \dot{u}^x(t, S) &= \beta(x) \hat{\phi}^x(t, S) Z_t^x (u^x(t, S) - u^x(t, I)) - \frac{c_\lambda}{2} (\lambda^{(S)}(t) - \hat{\phi}^x(t, S))^2, \\ \dot{u}^x(t, I) &= \rho(x) \gamma(x) (u^x(t, I) - u^x(t, R)) \\ &\quad + (1 - \rho(x)) \gamma(x) (u^x(t, I) - u^x(t, D)) - c_I(x), \\ \dot{u}^x(t, R) &= \kappa(x) (u^x(t, R) - u^x(t, S)), \\ \dot{u}^x(t, D) &= -c_D(x), \\ u^x(T, e) &= 0, \quad p^x(0, e) = p_0^x(e), \quad e \in \{S, I, R\}, \\ Z_t^x &= \int_I w(x, y) \hat{\phi}^y(t, I) p^y(t, I) dy, \quad t \in [0, T], \quad x \in I.\end{aligned}$$

We note that with a careful choice of $\beta(x)$, c_λ , and $(\lambda^{(e)})_{e \in E}$ this system will satisfy the sufficient condition for existence from Theorem 2.

3 Numerical Approach

We rewrite the continuum of FBODEs (18) as the solution of a minimization problem: minimize

$$\mathbb{J}(\theta) = \int_I \sum_{e \in E} |u_\theta^x(T, e) - g^x(e, Z_{\theta, T}^x)|^2 dx, \quad (21)$$

where (p_θ, u_θ) solve the forward-forward continuous system of ODEs:

$$\begin{cases} \dot{p}_\theta^x(t, e) = \sum_{e' \in E} q_{e', e}^x(Z_{\theta, t}^x, \hat{\alpha}^x(t, e', Z_{\theta, t}^x, u_\theta^x(t, \cdot))) p_\theta^x(t, e'), \\ \dot{u}_\theta^x(t, e) = -\hat{H}^x(t, e, Z_{\theta, t}^x, \Delta_e u_\theta^x(t, \cdot)), \\ Z_{\theta, t}^x = \int_I w(x, y) \left(\sum_{e \in E} K(\hat{\alpha}^y(t, e, Z_{\theta, t}^y, u_\theta^y(t, \cdot)), e) p_\theta^y(t, e) \right) \lambda(dy) \\ u_\theta^x(0, e) = \varphi_\theta^x(e), \quad p^x(0, e) = p_0^x(e), \quad e \in E, \quad t \in [0, T]. \end{cases} \quad (22)$$

⁸ The rate is presented in (20) as $\beta(x)\alpha z$. For an arbitrary impact function K this would lead to a violation of Condition 1.1. Here we are however considering contact factor control and aggregates of the form (11), which are bounded by Condition 2 and the compactness of A . The Q -matrix (20) can be modified without changing the model so that Condition 1.1 is satisfied; we defer from this for the sake of presentation.

This “shooting” strategy is reminiscent of the one used *e.g.* in [20,21,41] for stochastic optimal control problem and *e.g.* in [2, 11] for mean field games in a numerical context. However, here we deal with a continuum of ODEs rather than a finite number of stochastic differential equations. Here θ is the parameter in the function φ replacing the initial condition of u . Typically, θ is a real-valued vector of dimension the number of degrees of freedom in the parametric function φ . In general, the true initial condition for u^x is a nonlinear function with a potentially complicated shape. So we need to choose a rich enough class of parametric functions. In the implementation, we used a deep neural network with a feedforward architecture. See, *e.g.*, [10] for a description of the feedforward neural network architecture we used in the implementation.

Our strategy to find θ is to run a gradient-descent based method. To alleviate the computational cost and to introduce some randomness, at each iteration we replace the above cost \mathbb{J} by an empirical average over a finite set of indices x , which is also used to approximate the value of the aggregate quantities. More precisely, for a finite set \mathbf{S} of indices, we introduce

$$\mathbb{J}^N(\theta, \mathbf{S}) = \frac{1}{N} \sum_{x \in \mathbf{S}} \sum_{e \in E} |u_{\theta, \mathbf{S}}^x(T, e) - g^x(e, Z_{\theta, \mathbf{S}, T}^x)|^2, \tag{23}$$

where $(p_{\theta, \mathbf{S}}, u_{\theta, \mathbf{S}})$ solves the forward-forward (finite) system of ODEs:

$$\begin{cases} \dot{p}_{\theta, \mathbf{S}}^x(t, e) = \sum_{e' \in E} q_{e', e}^x(Z_{\theta, \mathbf{S}, t}^x, \hat{a}^x(t, e', Z_{\theta, t}^x, u_{\theta, \mathbf{S}}^x(t, \cdot))) p_{\theta, \mathbf{S}}^x(t, e'), \\ \dot{u}_{\theta, \mathbf{S}}^x(t, e) = -\hat{H}^x(t, e, Z_{\theta, \mathbf{S}, t}^x, \Delta_e u_{\theta, \mathbf{S}}^x(t, \cdot)), \\ Z_{\theta, \mathbf{S}, t}^x = \frac{1}{N} \sum_{y \in \mathbf{S}} w(x, y) \left(\sum_{e \in E} K(\hat{a}^y(t, e, Z_{\theta, t}^y, u_{\theta, \mathbf{S}}^y(t, \cdot)), e) p_{\theta, \mathbf{S}}^y(t, e) \right), \\ u_{\theta, \mathbf{S}}^x(0, e) = \varphi_{\theta}^x(e), \quad p^x(0, e) = p_0^x(e), \quad e \in E, \quad x \in \mathbf{S}. \end{cases} \tag{24}$$

Algorithm 1: DeepGraphonGame Algorithm

- 1 **Input:** Initial parameter $\theta^{(0)}$ for the neural network φ_{θ} playing the role of the initial condition for u ; number of iterations K ; sequence $(\beta_k)_{k=0, \dots, K-1}$ of learning rates; coefficients of the problem; a number of indices N
 - 2 **Output:** Approximation of θ^* minimizing \mathbb{J}^N , see (21)
 - 1: Initialize the parameter θ of φ_{θ} with $\theta^{(0)}$
 - 2: **for** $k = 0, 1, 2, \dots, K - 1$ **do**
 - 3: Sample $\mathbf{S} = (x^i)_{i \in \llbracket N \rrbracket}$
 - 4: Solve the FFODE finite system (24) with initial condition $u_{\theta, \mathbf{S}}(0, \cdot) \equiv \varphi_{\theta(k)}$
 - 5: Compute the gradient $\nabla_{\theta} \mathbb{J}^N(\theta^{(k)}, \mathbf{S})$ of $\mathbb{J}^N(\theta^{(k)}, \mathbf{S})$ defined in (23)
 - 6: Set $\theta^{(k+1)} = \theta^{(k)} - \beta_k \nabla_{\theta} \mathbb{J}^N(\theta^{(k)}, \mathbf{S})$
 - 7: **end for**
 - 8: **return** $\theta^{(K)}$
-

3.1 Piecewise Constant Graphon

Let m^1, m^2, \dots, m^K be non-negative numbers such that $\sum_{k=1}^K m^k = 1$. We divide the player population into K groups, B^1, \dots, B^K , placing all players with index $x \in [0, m^1)$ into

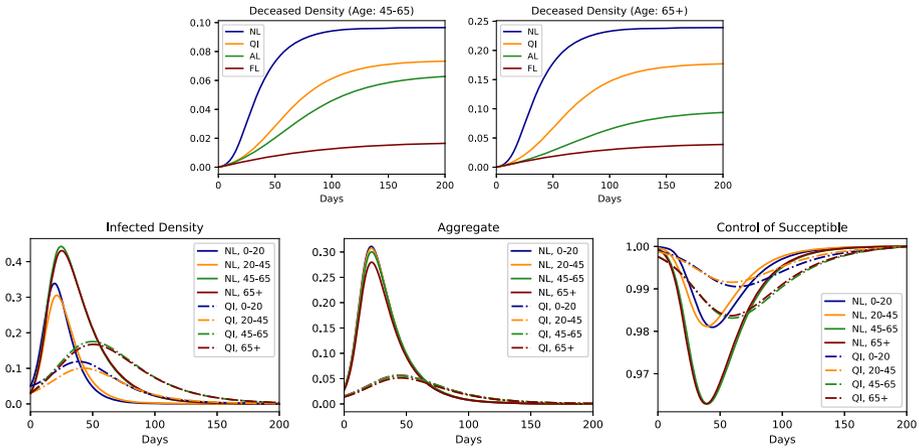


Fig. 1 *Top:* Density of deceased people for age groups 45-65 (*left*) and 65+ (*right*) under different policies where NL: No Lockdown, QI: Quarantine for Infected, AL: Age Specific Lockdown, FL: Full Lockdown. *Bottom:* Comparison Plots for No Lockdown and Quarantine for Infected policies: Density of infected people (*left*), Aggregate Z (*middle*), Control of susceptible people (*right*) are plotted for each 4 age groups under both policies

group B^1 , etc. We assume that players belonging to the same group are indistinguishable. For example, all players within a group must have the same recovery rate γ and if $x, x' \in I$ are the indices of two players in the same group, then $w(x, y) = w(x', y)$ for all $y \in I$. In this situation, we only need to specify the graphon’s values on each block of indices corresponding to a group, since the graphon is a constant on each block. Let us identify the group B^i with its index block (or set). We can compactly represent the interaction weights between the blocks with a connection matrix $[w_{ij}]_{i,j=1}^K$, where w_{ij} is the connection strength between players in block B^i and players in block B^j . Then, for all players x in block B^i

$$Z_i^x = \int_I w(x, y)\phi^y(t, l)p^y(t, l)dy = \sum_{k=1}^K w_{ik}\lambda^{(l,k)}p^k(t, l)m^k, \tag{25}$$

which is constant over $x \in B^i$. This is a feature of that we sometimes see when the piecewise constant graphon is used. Furthermore, we assume β, γ, κ constant over each block but can differ between blocks. It opens up the possibility for us to solve the graphon game with classical numerical methods⁹ and a way of evaluating the DeepGraphonGame algorithm.

Turning to the remaining example set up, we specify the cost structure as a particular case of the general formulation of Sect. 2.4. We assume that the regulator has set $\lambda^{(S,k)}, \lambda^{(l,k)}, \lambda^{(R,k)}$, $k = 1, \dots, K$, different for each block.

In this scenario, we first study the policy effects on the death ratio in the age groups. The policies compared are no lockdown (NL); quarantine for infected (QI); age specific lockdown (AL); full lockdown (FL). We can see in Fig. 1 that the death ratio decreases nearly 30% if infected individuals are quarantined, compared to no lockdown. Furthermore, if an age specific lockdown is implemented, we see that even more lives are saved while not deteriorating the economy. Zooming in on the comparison of no lockdown and quarantine

⁹ These methods are referring to solving an ODE system similar to the one given in Sect. 2.4 for a finite (and possibly a small) number of x ’s.

Table 1 Parameters of the experiment with different age groups

Age	0–20	20–45	45–65	65+	Age	β	γ	ρ	$p_0(S)$	m
0-20	1.0	0.9	0.8	0.7	0-20	0.4	0.1	1.0	0.95	0.27
20-45	0.9	0.9	0.8	0.8	20-45	0.3	0.1	1.0	0.97	0.33
45-65	0.8	0.8	0.9	0.8	45-65	0.3	0.05	0.9	0.97	0.27
65+	0.7	0.8	0.8	0.8	65+	0.3	0.05	0.75	0.97	0.13

Table 2 Parameters used in the experiments with age-groups specific lockdowns (NL: No Lockdown, QI: Quarantine for Infected, AL: Age Specific Lockdown, FL: Full Lockdown): λ^S and λ^I vary between the age groups.

Parameters	T	λ^S	λ^I	λ^R	c_λ	c_I	c_D
NL	200	[1.0, 1.0, 1.0, 1.0]	[1.0, 1.0, 1.0, 1.0]	1.0	10	1	1
QI	200	[1.0, 1.0, 1.0, 1.0]	[0.5, 0.5, 0.5, 0.5]	1.0	10	1	1
AL	200	[0.5, 1.0, 1.0, 0.5]	[0.5, 0.5, 0.5, 0.5]	1.0	10	1	1
FL	200	[0.5, 0.5, 0.5, 0.5]	[0.5, 0.5, 0.5, 0.5]	1.0	10	1	1

The lockdown is imposed by decreasing the λ^S and λ^I values of the corresponding age groups

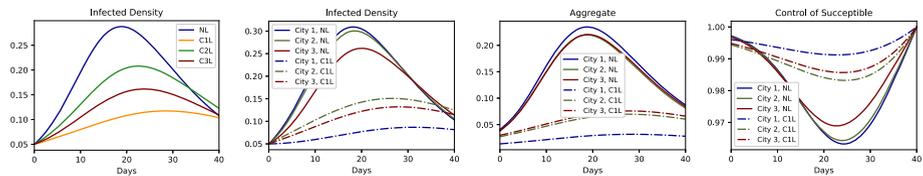


Fig. 2 Density of infected people in the whole population (including all cities) under 4 different policies, NL: No Lockdown, C1L: City 1 Lockdown, C2L: City 2 Lockdown, C3L: City 3 Lockdown (left). Comparison Plots for No Lockdown and City 1 Lockdown policies: Density of infected people in each city (middle left), Aggregate Z in each city (middle right), Control of susceptible people in each city (right)

for infected (second row of Fig. 1), we note that susceptible individuals are using a smaller contact factor when there is no quarantine in place. They optimize their risk and hence want to be more cautious, since the risk of getting infected is higher (Tables 1 and 2).

Secondly, in the same scenario, we model multiple cities with different attributes and study the effects of the travel restrictions. In this experiment, we compare a universal no-travel policy to the policy where traveling in or out of one of the cities is restricted (totaling four policies in the comparison). City 1 is a highly populated city with a more contagious virus variant, city 2 also has this variant; however, it is a small city. City 3 is a highly populated city but with a less contagious virus variant. For visual simplification, we assume that there are no deaths (*i.e.* $\rho = 0$). In Fig. 2, we can see that the infected-density curve can be flattened the most if city 1 has travel restrictions. The reason is the existence of the more contagious variant and the large size of the city 1. We note that when this restriction is implemented the susceptible individuals feel relieved and increase their contact factor control (Tables 3 and 4).

Table 3 Connection matrix (left) and Parameters (right) used in the experiment with different cities

Block	City 1	City 2	City 3	Block	β	$p_0(S)$	m
City 1	0.3	0.3	0.3	City 1	0.4	0.95	0.4
City 2	0.3	1.0	0.7	City 2	0.4	0.95	0.2
City 3	0.3	0.7	1.0	City 3	0.3	0.95	0.4

The connection matrix shown here is the case where there is a lockdown in City 1. When there is no lockdown the interaction weights are as follows: $w(1, 1) = 1.0$, $w(1, 2) = 0.9$ and $w(1, 3) = 0.8$

Table 4 Parameters used in the experiment with different cities

Parameters	T	γ	λ^S	λ^I	λ^R	c_λ	c_I	c_D
All policies	40	0.1	1.0	0.9	1.0	10	1	1

3.1.1 Sanity Check for the Numerical Approach

Here, we test the `DeepGraphonGame` algorithm by comparing its solution to the solution obtained by solving the ODE system for the cities-example when city 1 has travel restrictions (Table 4). As can be seen in Fig. 3, the `DeepGraphonGame` algorithm approximates the exact result well. A plot of the function $x \mapsto u^x(0, S)$, where u^x is the numerically computed value function, can be seen on the right side of the bottom row in Fig. 3. We can clearly see that agents in the same block have the same $u^x(0, S)$ values. From this we infer that the `DeepGraphonGame` algorithm is performing well when learning this piecewise constant function.

3.2 General Graphon

To show scalability of the proposed numerical approach now we focus on the second example in [2] with the SEIRD model where the state (E)xposed is added. An individual is in state E when infected but is not yet infectious. Hence, the agents evolve from S to E and then to I, and the infection rate from S to E depends on the proportion of the infected agents. The diagram of the dynamics can be seen in Fig. 4. The cost structure is similar to the one used in Sect. 2.4. After introducing the state E, we set

$$f^x(t, e, z, \alpha) = \frac{c_\lambda}{2} \left(\lambda^{(S)}(t) - \alpha \right)^2 \mathbf{1}_{(e=S)} + \frac{c_\lambda}{2} \left(\lambda^{(E)}(t) - \alpha \right)^2 \mathbf{1}_{(e=E)} + \left(\frac{1}{2} \left(\lambda^{(I)}(t) - \alpha \right)^2 + c_I(x) \right) \mathbf{1}_{(e=I)} + \frac{1}{2} \left(\lambda^{(R)}(t) - \alpha \right)^2 \mathbf{1}_{(e=R)} + c_D(x) \mathbf{1}_{(e=D)}.$$

In this example, we focus on an application where the agents are not homogeneous over blocks. The interaction strength between individual x and y is now given by the power law graphon: $w(x, y) = (xy)^{-g}$ where $-\infty < g \leq 0$ is a constant¹⁰. Intuitively, the power law graphon models interactions in a population where a small number of individuals are responsible for a large number of the interactions. For example, a population with superspreaders¹¹ can be modeled with this graphon. The model with an underlying power law graphon interaction requires us to solve a continuum of coupled ODEs which is not computationally feasible.

¹⁰ We can realize that if $g = 0$, the setting is equivalent to a mean field game.

¹¹ A superspreader is an infected person who is able to transmit the disease to a disproportionately high number of people.

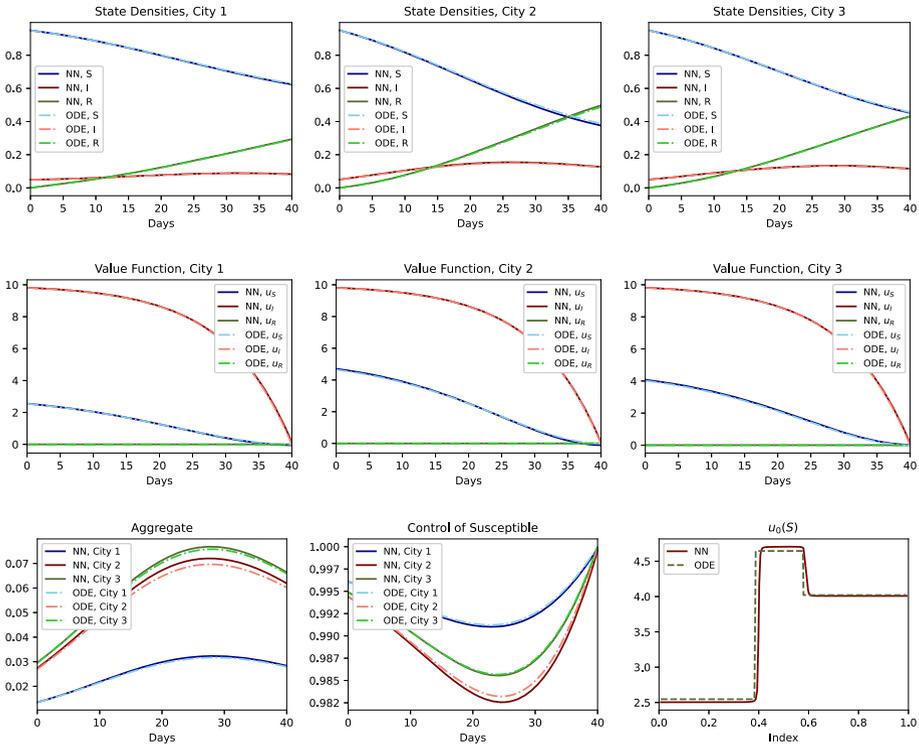
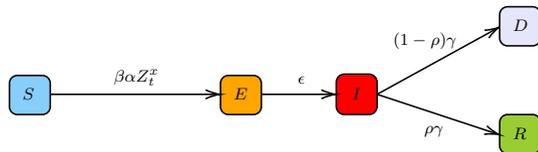


Fig. 3 Comparison of the ODE and NN results when there is a lockdown for City 1: *Top*: State densities in City 1 (left), City 2 (middle) and City 3 (right). *Middle*: Value functions given each state in City 1 (left), City 2 (middle) and City 3 (right). *Bottom*: Aggregate Z (left), Control of susceptible people (middle) and value function at time 0 given state is susceptible as a function of index, $u^x(0, S)$ (right)

Fig. 4 Diagram of SEIRD model for individual x



However, by using the DeepGraphonGame algorithm, the solution can be learned by using simulated particles, *i.e.* agents.

According to CDC, COVID-19 reinfection is very rare¹²; therefore, we assume that there is no reinfection (*i.e.* $\kappa = 0$). Furthermore, the recovery duration is around 10 days since the symptom onset¹³. For this reason, we assume that $\gamma = 0.1 \text{ days}^{-1}$. According to the recent study conducted by Lauer *et al.* [56], an exposed person begins to show symptoms after around 5 days. Based on this observation, we choose $\epsilon = 0.2 \text{ days}^{-1}$. Finally, the Basic

¹² <https://www.cdc.gov/coronavirus/2019-ncov/your-health/reinfection.html> (last accessed on 10 April, 2021)

¹³ <https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html> (last accessed on 10 April, 2021)

Table 5 Parameters in the SEIRD model experiments with power law graphon

Param.	T	β	γ	ϵ	ρ	c_λ	c_I	c_D	g	λ^S	λ^E	λ^I	λ^R
Values	40	0.2	0.1	0.2	0.95	10	1	1	-0.2	1.0	1.0	0.9	1.0

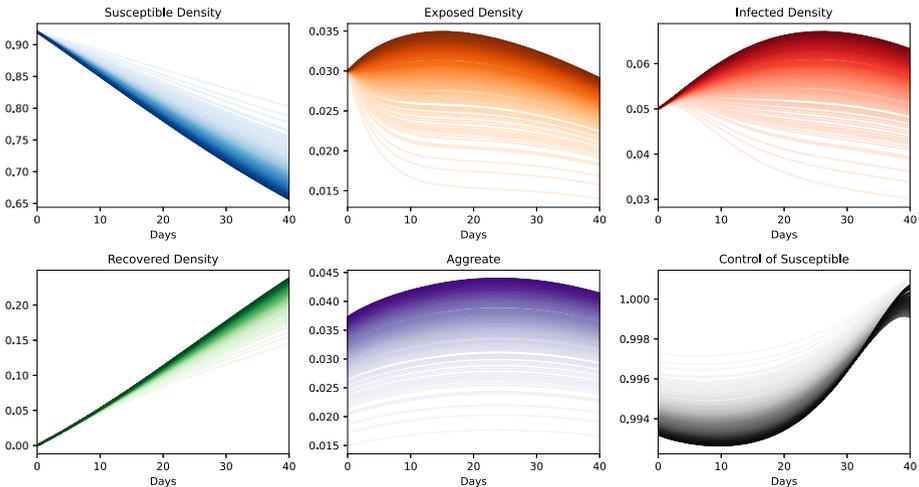


Fig. 5 Results of agents from a sampled population. In each plot, colors are chosen from a continuous colormap to represent the index of the agents with the following convention: If the index x of a player is higher, the color of the line is darker. *Top*: Probability of being susceptible (*left*), exposed (*middle*) and infected (*right*). *Bottom*: Probability of being recovered (*left*), Aggregate Z (*middle*), Control at the susceptible state (*right*)

Reproduction Number estimate $R_0 = 2$ used by CDC¹⁴ leads us to set $\beta = R_0 \times \gamma = 0.2$ in our simulations (Table 5).

The experiment results for a sampled finite subset of the agent population are presented in Fig. 5. In the figure, each line corresponds to one agent and the color of the plot gets darker as the index of the agent (*i.e.* x) increases. Our first observation is that as the index of the agent increases, the aggregate also increases. In response to this high aggregate, the agent lower its contact rate, in order to protect itself. However, this protection is not enough to neutralize the effects of the high levels of the aggregate and the probability of the agent to be infected is still elevated.

4 The Probabilistic Approach to Finite State Graphon Games

This section contains a closer study of the continuum of interacting jump processes that constitute the graphon game dynamics. Going back to the informal discussion in the introductory Sect. 1.1, it is not clear that there would be an adequate law of large numbers so that (4) converges to (5) since the averaged random variables are dependent. A common approach in economic theory for this situation is to consider the continuum limit and average the continuum of random variables with respect to a non-atomic probability measure over I . Using the example from Sect. 1.1 again, a continuum limit of (4) is

¹⁴ <https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html> (last accessed on 10 April, 2021)

$$\beta \alpha_t^{x_j} \int_I w(x_j, y) \alpha_t^y \mathbf{1}_I(X_{t-}^y) dy.$$

However, the integral in the expression above is ill-defined. There is an issue of constructing a continuum of independent random variables (here, that would be the driving Poisson noise) that are jointly measurable in the sample and the index. If the construction is done in the usual way via the Kolmogorov construction, then almost all random variables are essentially equal to an arbitrarily given function on the index space (*i.e.*, as random variables they are constants). Hence, in any interesting case the function $y \mapsto X_{t-}^y$ will not be measurable with respect to the Lebesgue measure. One solution proposed by economists is to extend the usual probability space to a so-called a Fubini extension [52], a probability space over $\Omega \times I$ where Fubini’s theorem holds. On the Fubini extension a continuum of random variables can be constructed that are essentially pairwise independent (e.p.i.; see Theorem 3 for the definition) and jointly measurable in sample and index. Moreover, there is hope for an exact law of large numbers [53], justifying our assumption about the determinism of the aggregate variable in the previous sections. We will construct a Fubini extension that carries a continuum of e.p.i. Poisson random measures. Then, each player path will be defined in the representation of the counting processes associated to the pure jump process as a stochastic integral with respect to a family of independent Poisson random measures with Lebesgue mean measure on $\mathbb{R} \times [0, T]$ as suggested in Skorokhod [51] and Grigelionis [34].

4.1 Theoretical Background and Definitions

4.1.1 Poisson Random Measures

Let us denote by $(\Gamma, \mathcal{G}, \mathbb{G})$ the measure space $(\mathbb{R} \times [0, T], \mathcal{B}(\mathbb{R} \times [0, T]), \text{Leb}_{\mathbb{R} \times [0, T]})$. We first recall the definition of a Poisson random measure. A family $(N_\omega(G))_{G \in \mathcal{G}}$ of random variables defined on some probability space $(H, \mathcal{H}, \mathbb{H})$ is said to be a Poisson random measure with measure \mathbb{G} if

- for all $G \in \mathcal{G}$ such that $\mathbb{G}(G) < \infty$, $N_\omega(G)$ is a Poisson random variable with rate $\mathbb{G}(G)$;
- the random variables $N_\omega(G_1), \dots, N_\omega(G_n)$ are mutually independent whenever the sets $G_1, \dots, G_n \in \mathcal{G}$ have finite \mathbb{G} -measures and are disjoint;
- for all $\omega \in H$, $N_\omega(\cdot)$ is a measure on (Γ, \mathcal{G}) .

We define $\tilde{M} := \{\mu \mid \mu \text{ is a } \sigma\text{-finite non-negative measure on } (\Gamma, \mathcal{G})\}$, and the subset of locally finite measures by

$$M := \{\mu \in \tilde{M} \mid \mu(\tilde{G}) < \infty \text{ for all bounded } \tilde{G} \in \mathcal{G}\}.$$

For all bounded measurable \tilde{G} , define the mappings $I_{\tilde{G}} : \tilde{M} \ni \mu \mapsto \mu(\tilde{G}) \in \mathbb{R}$. Let \tilde{M} and M be the σ -algebras induced by the mappings $I_{\tilde{G}}$ on \tilde{M} and M respectively. For us, a random measure will be a measurable function from $(H, \mathcal{H}, \mathbb{H})$ into (\tilde{M}, \tilde{M}) that almost surely takes values in (M, M) . We shall also use the fact that M (equipped with the vague topology) is a Polish space [22]. We denote the law of N on (M, M) by \mathcal{N} .

The Poisson random measure N has an accompanying martingale. For all bounded $A \in \mathcal{B}(\mathbb{R})$, $\hat{N}(A, t) := N(A \times [0, t]) - \text{Leb}(A)t$ is a square integrable zero-mean martingale.

4.1.2 The Fubini Extension

In the model, state dynamics are given as E -valued jump processes. We will construct such a process from $2|E| - 1 = 2n - 1$ independent Poisson random measures. The possible jumps will be those so that the state process jumps between two integers in E , at most $n - 1$ steps up or down. The initial state of player $x \in I$ is randomly sampled from a pre-selected distribution $p_0^x \in \mathcal{P}(E)$.

In order to model idiosyncratic random shocks affecting the dynamics of the individual states, we use the framework of Fubini extensions. It allows us to capture a form of independence for a continuum of random variables, while preserving joint measurability.

Definition 3 If $(\Omega, \mathcal{F}, \mathbb{P})$ and $(I, \mathcal{I}, \lambda)$ are probability spaces, a probability space $(\Omega \times I, \mathcal{W}, \mathbb{Q})$ extending the usual product space $(\Omega \times I, \mathcal{F} \otimes \mathcal{I}, \mathbb{P} \otimes \lambda)$ is said to be a Fubini extension if for any real-valued \mathbb{Q} -integrable function f on $(\Omega \times I, \mathcal{W})$

- (i) the two functions $f_x : \omega \mapsto f(\omega, x)$ and $f_\omega : x \mapsto f(\omega, x)$ are integrable, respectively, on $(\Omega, \mathcal{F}, \mathbb{P})$ for λ -a.e. $x \in I$, and on $(I, \mathcal{I}, \lambda)$ for \mathbb{P} -a.e. $\omega \in \Omega$;
- (ii) $\int_\Omega f_x(\omega) d\mathbb{P}$ and $\int_I f_\omega(x) d\lambda(x)$ are integrable, respectively, on $(I, \mathcal{I}, \lambda)$ and $(\Omega, \mathcal{F}, \mathbb{P})$, with the Fubini property

$$\begin{aligned} \int_{\Omega \times I} f(\omega, x) d\mathbb{Q}(\omega, x) &= \int_I \left(\int_\Omega f_x(\omega) d\mathbb{P}(\omega) \right) d\lambda(x) \\ &= \int_\Omega \left(\int_I f_\omega(x) d\lambda(x) \right) d\mathbb{P}(\omega). \end{aligned}$$

The following theorem summarizes the results by Sun and collaborators, see for example [53] and [54], which we use as a foundation for our model.

Theorem 3 *There exists a probability space $(I, \mathcal{I}, \lambda)$ extending $(I, \mathcal{B}_I, \lambda_I)$, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a Fubini extension $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, \mathbb{P} \boxtimes \lambda)$ such that for any measurable mapping $\underline{\phi}$ from $(I, \mathcal{I}, \lambda)$ to $\mathcal{P}(E \times M^{2n-1})$ there is an $\mathcal{F} \boxtimes \mathcal{I}$ -measurable process $\underline{f} : \Omega \times I \rightarrow E \times M^{2n-1}$ such that the random variables $f^x = \underline{f}(\cdot, x)$ are essentially pairwise independent (e.p.i.), i.e., for λ -a.e. $x \in I$, f^x is independent of f^y for λ -a.e. $y \in I$, and $\mathbb{P} \circ (f^x)^{-1} = \phi^x$ for all $x \in I$.*

We set $\phi^x = p_0^x \otimes (\otimes_{k=-n+1}^{n-1} \mathcal{N})$, where \mathcal{N} is the probability law of the Poisson random measure introduced above, and p_0^x is the initial distribution of player x . By Theorem 3 (which holds for $\underline{\phi}$ since E and M are Polish spaces) there exists a collection of random variables $(\underline{\xi}, \underline{N}_k; k = -n + 1, \dots, n - 1)$ on a Fubini extension $(\Omega \times I, \mathcal{F} \boxtimes \mathcal{I}, \mathbb{P} \boxtimes \lambda)$, that are e.p.i. and ϕ^x -distributed for all $x \in I$. With the model in Sect. 2 in mind, we assume that the mapping $x \mapsto p_0^x$ is Lebesgue-measurable (this assumption is however not necessary for the analysis that follows).

We denote by $L^2_{\boxtimes}(\Omega \times I; \mathcal{D})$ the Bochner space of all (equivalence classes of) strongly $(\mathbb{P} \boxtimes \lambda, \mathcal{B}(\mathcal{D}))$ -measurable functions $\underline{f} : \Omega \times I \rightarrow \mathcal{D}$ for which

$$\mathbb{E}^{\boxtimes} \left[\|\underline{f}\|_T^2 \right] = \int_{\Omega \times I} \|f^x(\omega)\|_T^2 \mathbb{P} \boxtimes \lambda(d\omega, dx) < +\infty.$$

We define $L^2_{\boxtimes}(\Omega \times I; \mathcal{C})$ in the same way, with \mathcal{C} replacing \mathcal{D} above. By e.g. [38, Ch. 1.2.b], $L^2_{\boxtimes}(\Omega \times I; \mathcal{C})$ and $L^2_{\boxtimes}(\Omega \times I; \mathcal{D})$ are Banach spaces.

For later reference, we define also the set \mathcal{L}_E as the subset of $L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$ of $\mathbb{P} \otimes \lambda$ -a.e. \mathcal{D}_E -valued functions. One can show that \mathcal{L}_E is a closed subset of $L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$, hence $(\mathcal{L}_E, \|\cdot\|_{L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})})$ is a complete metric space.

4.1.3 The Set of Admissible Strategies

We can now give a rigorous definition of the set of admissible strategy profiles. Recall that A is a compact subset of \mathbb{R} .

Definition 4 We define the set $\underline{\mathbb{A}}$ of admissible strategy profiles in feedback form as the set of A -valued, $\mathcal{I} \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(E)$ -measurable functions $\underline{\alpha}$ on $I \times [0, T] \times E$ such that $\underline{\alpha}(x, \cdot, e)$ is a continuous function on $[0, T]$ for every $(x, e) \in I \times E$.

We will sometimes use the same notation $\underline{\mathbb{A}}$ for the set of admissible control processes associated to an admissible strategy profile in feedback form $\underline{\alpha}$. At time t , for player $x \in I$, the value of such a control process is the action $\underline{\alpha}(x, t, X_{t-}^{x,\alpha})$ where $X_{t-}^{x,\alpha}$ is the state of player x just before time t . These control processes are predictable with respect to the filtration generated by the player’s private state, and decentralized since they do not depend directly on other players’ states.

The continuity in time is a strong assumption and prohibits the player from immediately reacting to abrupt changes in their environment. However, if a player transitions between two states at time t their control can be discontinuous at that time (as a multivariate function of time and state). In Sect. 2, a rationale was given for restricting our attention to such controls.

4.2 The Finite State Graphon Game in the Fubini Extension

We begin by describing an interacting system of a continuum of particles. First, we define the decoupled system where the aggregate variable vector has been “frozen”. Then, we define the aggregate with a fixed point argument. Finally, we prove that the aggregate is in fact deterministic.

Consider the pure jump stochastic integral equation (here written formally)

$$X_t^{\alpha, \underline{z}, x} = \xi^x + \sum_{k=-n+1}^{n-1} k \int_{\mathbb{R} \times (0, t]} \mathbf{1}_{[0, \kappa^x(X_{s-}^{\alpha, \underline{z}, x}, k, \alpha_s^x, z_{s-}^x)]}(y) N_k^x(dy, ds), \tag{26}$$

where $x \in I, t \in [0, T], \underline{\alpha}(\omega) = (\alpha^x(\omega))_x := (\alpha(x, t, X_{t-}^{\alpha, \underline{z}, x}(\omega)))_{t,x}$ for some admissible strategy profile $\underline{\alpha}, \underline{z} \in L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$, for any $k \in \mathbb{Z}, x \in I, s \in [0, T], a \in A, z \in \mathbb{R}, i \in E$

$$\kappa^x(i, k, a, z) := \begin{cases} q_{i, i+k}^x(a, z), & i+k \in E, \\ 0, & i+k \notin E, \end{cases} \tag{27}$$

is the rate of jumps from state i to state $i+k$ given the action $a \in A$ and the aggregate value z . The proposition below asserts that (26) has a unique solution in \mathcal{L}_E , the subset of $L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$ defined in the end of Sect. 4.1.2.

Proposition 1 Assume that Condition 1 holds. Let $\underline{z} \in L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$ and $\underline{\alpha} \in \underline{\mathbb{A}}$ be fixed. Then there is a unique strong solution $\underline{X}^{\alpha, \underline{z}} \in \mathcal{L}_E$ to (26), i.e., a $\mathbb{P} \otimes \lambda$ -a.e. \mathcal{D}_E -valued process satisfying (26) $\mathbb{P} \otimes \lambda$ -a.s.

Note that the quantity

$$M_k^x(t) = \int_{\mathbb{R} \times (0,t]} \mathbf{1}_{[0, \kappa^x(X_{s-}^{\alpha, z, x}, k, \alpha_s^x, z_{s-}^x)]}(y) N_k^x(dy, ds),$$

appearing in the right-hand side of (26), is a counting process with intensity $\kappa^x(X_{t-}^{\alpha, z, x}, k, \alpha_t^x, z_{t-}^x)$ at time $t \in [0, T]$ so by construction the solution to (26) (granted by Proposition 1) is almost surely an E -valued pure jump process with intensity matrix $Q^x(\alpha_t^x(\omega), z_{t-}^x(\omega))$ at time $t \in [0, T]$.

For a fixed admissible strategy profile $\underline{\alpha}$, consider now the coupled system

$$\begin{aligned} X_t^{\alpha, x} &= \xi^x + \sum_{k=-n+1}^{n-1} k \int_{\mathbb{R} \times (0,t]} \mathbf{1}_{[0, \kappa_s^x(X_{s-}^{\alpha, x}, k, \alpha_s^x, Z_{s-}^{\alpha, x})]}(y) N_k^x(dy \otimes ds), \\ Z_t^{\alpha, x} &= \int_I w(x, y) K(\alpha_t^y, X_{t-}^{\alpha, y}) \lambda(dy). \end{aligned} \tag{28}$$

The next theorem proves that (28) is well-posed with a unique solution in $L^2_{\mathbb{F}}$ -sense. It further specifies the regularity of the solution: the aggregate variable \underline{Z}^α must $\mathbb{P} \otimes \lambda$ -a.s. be a deterministic and a continuous function of time.

Theorem 4 *Let Condition 1 and 2 hold, and let $\underline{\alpha} \in \underline{\mathbb{A}}$.*

- (i) *There exists a unique solution $\underline{X}^\alpha \in \mathcal{L}_E$ to (28). The corresponding aggregate \underline{Z}^α is a random variable in $L^2_{\mathbb{F}}(\Omega \times I; \mathcal{C})$.*
- (ii) *The aggregate \underline{Z}^α is $\mathbb{P} \otimes \lambda$ -a.s. equal to a deterministic (i.e., constant in ω) function in $L^2_{\mathbb{F}}(\Omega \times I; \mathcal{C})$.*
- (iii) *There is a unique pair $\check{\underline{X}}^\alpha$ and $\check{\underline{Z}}^\alpha$ of versions of \underline{X}^α and \underline{Z}^α , respectively, solving (28) for all $x \in I$ in the standard L^2 -sense. Moreover $\check{\underline{Z}}^\alpha$ is deterministic and continuous in time for all $x \in I$.*

Theorem 4 justifies working with a model defined for all $x \in I$ with a deterministic, continuous-in-time aggregate in Sect. 2. From here on, we will represent the $L^2_{\mathbb{F}}$ -elements solving system (28) with the version defined for all $x \in I$ and drop the check in the notation.

Remark 2 If admissible strategy profiles did not have the prescribed continuity property we could not expect the aggregate to be a continuous function of time. One example of such a case is found in [2] where a regulator imposes a penalty that is discontinuous in time, resulting in equilibrium controls and aggregates discontinuous in time. We leave the analysis of the more general case to future work.

We now turn to the notion of player costs and equilibrium. Denote by $\underline{\mathbb{A}}$ the set of A -valued and $\mathcal{B}([0, T]) \otimes \mathcal{B}(E)$ -measurable functions on $[0, T] \times E$, continuous $t \in [0, T]$ for every $e \in E$ that. If the player population plays according to an admissible strategy profile $\underline{\alpha} \in \underline{\mathbb{A}}$ and player $x \in I$ decides to play strategy $\sigma = (\sigma(t, X_{t-}^{(\underline{\alpha}^{-x}, \sigma), x}))_{t \in [0, T]}$ where $\sigma \in \underline{\mathbb{A}}$ and

$$(\underline{\alpha}^{-x}, \sigma)^y := \begin{cases} \alpha^y, & \text{if } y \neq x \\ \sigma, & \text{if } y = x, \end{cases}$$

then $(\underline{\alpha}^{-x}, \sigma)$ is an admissible strategy profile and the player’s expected cost for using σ is

$$\mathcal{J}^x(\sigma; \underline{\alpha}) := \mathbb{E} \left[\int_0^T f^x(t, X_t^{(\underline{\alpha}^{-x}, \sigma), x}, \beta_t, Z_t^{\alpha, x}) dt + h^x(X_T^{(\underline{\alpha}^{-x}, \sigma), x}, Z_T^{\alpha, x}) \right].$$

In fact, \mathbb{A} is the set of strategies a player can deviate to without destroying the admissibility of the strategy profile. Therefore, we say that if $\hat{\alpha} = (\hat{\alpha}^x)_{x \in I} \in \underline{\mathbb{A}}$ satisfies

$$\mathcal{J}^x(\hat{\alpha}^x; \hat{\alpha}) \leq \mathcal{J}^x(\sigma; \hat{\alpha}), \quad \sigma \in \mathbb{A}, \quad x \in I,$$

the $\hat{\alpha}$ is a Nash equilibrium of the graphon game. The dependence of the cost on the whole strategy profile is unnecessarily complicated, as the following reasoning shows. Notice that $\underline{Z}^{(\alpha^{-x}, \sigma), x} = \underline{Z}^{\alpha, x}$ since $(\alpha^{-x}, \sigma) = \alpha$ for λ -a.e. $x \in I$. The other players' actions appear in player x 's cost indirectly, through the aggregate $\underline{Z}^{\alpha, x}$, which is unaffected if one specific player changes control (it is an integral with respect to a non-atomic measure). Thus, we write $\mathcal{J}^x(\sigma; \alpha)$ as $J^x(\sigma; \underline{Z}^{\alpha, x})$ a function taking an admissible strategy and an aggregate variable trajectory:

$$\mathbb{A} \times \mathcal{C} \ni (\sigma, \zeta) \mapsto J^x(\sigma; \zeta) \in \mathbb{R}.$$

In light of this, an equivalent definition of the Nash equilibrium is that a strategy profile $\hat{\alpha} = (\hat{\alpha}^x)_{x \in I}$ is a Nash equilibrium in the graphon game if it satisfies

$$J^x(\alpha^x; \underline{Z}^{\alpha, x}) \leq J^x(\sigma; \underline{Z}^{\alpha, x}), \quad \sigma \in \mathbb{A}, \quad x \in I,$$

further justifying the game setup in Sect. 2.

5 Conclusion and outlook

In this paper, we introduced stochastic graphon games in which the agents evolve in a finite state space. We provided optimality conditions in the form of a continuum of forward-backward ODE system, for which we established existence of solutions. We proposed a numerical method based on a neural network approximation of the initial condition of the FBODE. We then applied our theoretical framework and numerical method to a class of models from epidemiology and we provided several test cases. From here, several directions can be considered for future work. An interesting aspect would be to incorporate a regulator with a Stackelberg type model as was done in [2] without graphon structure. The theoretical analysis would probably rely on a combination of the tools developed in the present work together with tools from optimal contract theory. However there would be some important challenges depending on the class of controls that are admissible for the regulator. Their controls could indeed lead to discontinuities in the incentives to the population, which would raise subtle measurability questions. This is left for future work. Another direction is to consider more realistic epidemiological models (e.g., with more compartments). Such models would be more complex and we expect our proposed machine learning numerical method to be helpful from this point of view. Furthermore, to be able to use graphon games to make epidemiological predictions, it would be interesting to investigate further how to use real data in the model and in the numerical method.

Acknowledgements The authors would like to thank Boualem Djehiche and Yeneng Sun for helpful discussions.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Proofs for Section 2

A.1 Theorem 1

First, assume player x uses control σ and, using the HJB equation (15) and the definition of the minimized Hamiltonian (14), note that

$$d \left(u^x(t, X_t^{\sigma, Z^x, x}) + \int_0^t f(s, X_s^{\sigma, Z^x, x}, Z_s^x, \sigma_s) ds \right) = M_t \\ + H^x(t, X_t^{\sigma, Z^x, x}, Z_t^x, u^x(t, \cdot), \sigma_t) - H^x(t, X_t^{\sigma, Z^x, x}, Z_t^x, u^x(t, \cdot), \hat{a}_{X_t^{\sigma, Z^x, x}}^x(t, Z_t^x, u^x(t, \cdot))),$$

where M is a zero-mean martingale. Hence, we write

$$\mathbb{E} \left[u^x(t, X_t^{\sigma, Z^x, x}) + \int_0^t f(s, X_s^{\sigma, Z^x, x}, Z_s^x, \sigma_s) ds \right] \\ = \mathbb{E} \left[u^x(0, X_0^{\sigma, Z^x, x}) \right] + \int_0^t \mathbb{E} \left[H^x(s, X_s^{\sigma, Z^x, x}, Z_s^x, u^x(s, \cdot), \sigma_s) \right. \\ \left. - H^x(s, X_s^{\sigma, Z^x, x}, Z_s^x, u^x(s, \cdot), \hat{a}_{X_s^{\sigma, Z^x, x}}^x(s, Z_s^x, u^x(s, \cdot))) \right] ds.$$

Note that the second expectation in the right hand side is non-negative since $\hat{a}_e^x(t, z, h)$ minimizes $A \ni \alpha \mapsto H^x(t, e, z, h, \alpha)$ by assumption. In fact, since it is the unique minimizer, this term is strictly positive unless $\sigma_t = \hat{a}_{X_t^{\sigma, Z^x, x}}^x(t, Z_t^x, u^x(t, e))$.

By taking the expectation and by recalling the terminal condition we deduce that:

$$J^x(\sigma; Z^x) = \mathbb{E} \left[u^x(T, X_T^{\sigma, Z^x, x}) + \int_0^T f(s, X_s^{\sigma, Z^x, x}, Z_s^x, \sigma_s) ds \right] \\ \geq \mathbb{E} \left[u^x(0, X_0^x) \right] = J^x(\hat{\phi}^x; Z^x),$$

where, for the last equality, we used the interpretation of u^x as player x 's value function. Furthermore the inequality above is an equality if and only if $\sigma_t = \hat{a}_{X_t^{\sigma, Z^x, x}}^x(t, Z_t^x, u^x(t, e))$.

A.2 Regularity of the Optimal Control

Here we show that \hat{a} is continuous in (t, z, h) .

Lemma 1 *Assume Conditions 3.3 and 3.3 hold. For every $x \in I$ and $e \in E$, $(t, z, h) \mapsto \hat{a}_e^x(t, z, h)$ defined by the Hamiltonian minimizer in (14) is continuous. Moreover $(t, z, h) \mapsto \hat{a}_e^x(t, z, h)$ is locally Lipschitz continuous, i.e., for every positive constants C_z and C_h , for every $e \in E$, $a \in A$, the function $(t, z, h) \mapsto \hat{a}_e^x(t, z, h)$ is Lipschitz continuous on $[0, T] \times [-C_z, C_z] \times [-C_h, C_h]^{|E|}$ (with a Lipschitz constant possibly depending on T, C_z, C_h).*

Proof By the assumption on the dependence of H^x on a , $\hat{a}_e^x(t, z, h) = \hat{a}^x(t, e, z, h)$ is the unique solution of the variational inequality (with unknown a):

$$\forall b \in A, \quad (b - a)\partial_a H^x(t, e, z, h, a) \geq 0.$$

Let C_z and C_h be positive constants. Let $\theta = (t, e, z, h)$ and $\theta' = (t', e, z', h')$ with $(t, z, h), (t', z', h') \in [0, T] \times [-C_z, C_z] \times [-C_h, C_h]^E$. Let $a = \hat{a}^x(\theta)$ and $a' = \hat{a}^x(\theta')$. We deduce from the above inequality that:

$$(a' - a)[\partial_a H^x(\theta, a') - \partial_a H^x(\theta, a)] \leq (a' - a)[\partial_a H^x(\theta, a') - \partial_a H^x(\theta', a')].$$

By the assumption on the strict convexity of f , we have

$$\lambda|a' - a|^2 \leq (a' - a)(\partial_a f^x(t, e, z, a') - \partial_a f^x(t, e, z, a)).$$

So, combining the above inequalities and the property $|\partial_a Q^x(a, z) - \partial_a Q^x(a', z)| = 0$, we get

$$\begin{aligned} \lambda|a' - a|^2 &\leq (a' - a)(\partial_a f^x(t, e, z, a') - \partial_a f^x(t, e, z, a)) \\ &= (a' - a) [\partial_a H^x(\theta, a') - \partial_a H^x(\theta, a)] \\ &\leq (a' - a) [\partial_a H^x(\theta, a') - \partial_a H^x(\theta', a')] \\ &\leq C|a' - a| \left(\sup_{e' \in E} \left| \vec{\mathbf{1}}_e \partial_a Q^x(a', z)_{e'} - \vec{\mathbf{1}}_{e'} \partial_a Q^x(a', z')_{e'} \right| + \sup_{e' \in E} |h_{e'} - h'_{e'}| \right) \\ &\quad + |a' - a| |\partial_a f^x(t, e, z, a') - \partial_a f^x(t', e, z', a')|, \end{aligned}$$

where C can depend on q_{\max} , C_z , and C_h . We conclude by using the continuity and local Lipschitz continuity properties of $\vec{\mathbf{1}}_e \partial_a Q^x$ and $\partial_a f^x$. □

A.3 Theorem 2

Step 1: Definition of the solution space

We start by letting, for every $C_1 > 0$, \mathcal{K}_{C_1} be the closed ball of continuous functions (u, p) from $[0, T]$ into $L^2(I \times E) \times L^2(I \times E)$ such that for all $(t, x) \in [0, T] \times I$, $p^x(t, e) \geq 0$ for $e \in E$ and $\sum_{e \in E} p^x(t, e) = 1$, and such that (u, p) is bounded by C_1 in uniform norm. In other words, \mathcal{K}_{C_1} is the subset of $C([0, T]; L^2(I \times E) \times L^2(I \times E))$ for which the second component is a probability on E and for which the uniform norm is bounded by C_1 , whose value will be fixed in Step 5 below.

Step 2: Definition of the aggregate mapping

For each $(u, p) \in \mathcal{K}_{C_1}$ we define the map $\Phi^{(u,p)}$ which takes $(Z_t^x)_{0 \leq t \leq T, x \in I}$ into

$$\Phi^{(u,p)}((Z_t^x)_{0 \leq t \leq T, x \in I}) = \left(\int_I w(x, y) \sum_{e \in E} K(\hat{a}(t, e, Z_t^y, u^y(t, \cdot)), e) p^y(t, e) dy \right)_{0 \leq t \leq T, x \in I}.$$

We now prove that, if we choose the space of aggregates properly, $\Phi^{(u,p)}$ has a unique fixed point, say $(\hat{Z}_t^x)_{0 \leq t \leq T, x \in I}$, which depends continuously in $(u, p) \in \mathcal{K}_{C_1}$. Indeed, let $z, \tilde{z} \in \mathcal{Z}$ where

$$\mathcal{Z} := \{f \in C([0, T]; L^2(I; \mathbb{R})) : |f_t^x| \leq C_K, t \in [0, T], \text{ a.e. } x \in I\}.$$

In light of Condition 2, $\Phi^{(u,p)}(\mathcal{Z}) \subset \mathcal{Z}$. Moreover, \mathcal{Z} is a closed subset of the Banach space $C([0, T]; L^2(I; \mathbb{R}))$, hence a complete metric space. Using Cauchy-Schwarz inequality and the Lipschitz continuity of K and \hat{a} (given by Condition 2 and Lemma 1), we get

$$\sup_{t \in [0, T]} \int_I |\Phi^{(u,p)}(z)^y - \Phi^{(u,p)}(\tilde{z})^y|^2 dy \leq C_\Phi(C_K) \sup_{t \in [0, T]} \int_I |z_t^y - \tilde{z}_t^y|^2 dy,$$

where we used the fact that $|I| = 1$ and $C_\Phi(\cdot) := \|w\|_{L^2(I \times I)} L_K L_{\hat{a}}(\cdot)$. Recall that we assume $C_\Phi(C_K) < 1$. With Banach fixed point theorem we conclude that there exists a unique fixed point in \mathcal{Z} to $\Phi^{(u,p)}$. We denote it $\hat{Z}^{(u,p)} = (\hat{Z}_t^{(u,p),x})_{0 \leq t \leq T, x \in I}$.

Step 3: Solving the Kolmogorov equation

Given $\hat{Z}^{(u,p)}$ and u , we solve the Kolmogorov equation and we get the solution \hat{p} . Existence and uniqueness of the solution \hat{p} is provided by the Cauchy-Lipschitz-Picard theorem; see, e.g., [5, Theorem 7.3] (viewing q as a linear operator acting on the Banach space $L^2(I \times E)$). Furthermore, given Condition 1.1, the time derivative of \hat{p} is bounded. Therefore, we conclude that \hat{p} is equicontinuous.

Step 4: Solving the HJB equation

Given $\hat{Z}^{(u,p)}$ and \hat{p} , we solve the HJB equation and we get the solution \hat{u} . Here again, existence and uniqueness of the solution \hat{u} is provided by the Cauchy-Lipschitz-Picard theorem (viewing \hat{H} as a Lipschitz operator acting on the Banach space $L^2(I \times E)$). Furthermore, there is a uniform bound on the time derivative of \hat{u} since the Hamiltonian \hat{H} is bounded given Condition 3.3 and Condition 1.1. Hence \hat{u} is equicontinuous.

Step 5: Application of Schauder’s theorem

Let us call Ψ the mapping constructed by the above steps, namely, $\Psi : \mathcal{K}_{C_1} \ni (u, p) \mapsto (\hat{u}, \hat{p})$. By steps 3 and 4 above, if we choose C_1 large enough, Ψ maps \mathcal{K}_{C_1} onto \mathcal{K}_{C_1} , so it is well-defined. Furthermore, by the same steps and the Arzela-Ascoli theorem $\Psi(\mathcal{K}_{C_1})$ is compact. Finally we argue the continuity as follows:

We first show the continuity of $\hat{Z}^{(u,p)}$ in u and p . Consider a sequence $(u^n, p^n)_n$ such that $(u^n, p^n) \in \mathcal{K}_{C_1}$ for every n , and $\lim_{n \rightarrow \infty} (u^n, p^n) = (u, p)$. We denote $Z^n := Z^{(u^n, p^n)}$ and prove below that $\lim_{n \rightarrow \infty} \hat{Z}^n = \hat{Z}$. By Lipschitz continuity of K and \hat{a} :

$$\begin{aligned} & \sup_{t \in [0, T]} \int_I |\hat{Z}_t^{n,x} - \hat{Z}_t^x|^2 dx \leq \\ & \sup_{t \in [0, T]} \int_I \left(C_\Phi(C_K) |\hat{Z}_t^{n,x} - \hat{Z}_t^x|^2 + C_p |p^{n,x}(t, \cdot) - p^x(t, \cdot)|^2 + C_u |u^{n,x}(t, \cdot) - u^x(t, \cdot)|^2 \right) dx, \end{aligned}$$

for some constants $C_p, C_u \geq 0$, and where $C_\Phi(C_K) = \|w\|_{L^2(I \times I)} L_K L_{\hat{a}}(C_K)$. Hence

$$\begin{aligned} & \sup_{t \in [0, T]} \int_I |\hat{Z}_t^{n,x} - \hat{Z}_t^x|^2 dx \\ & \leq \sup_{t \in [0, T]} \int_I \frac{1}{1 - C_\Phi(C_K)} \left(C_p |p^{n,x}(t, \cdot) - p^x(t, \cdot)|^2 + C_u |u^{n,x}(t, \cdot) - u^x(t, \cdot)|^2 \right) dx, \end{aligned}$$

which tends to 0 as $n \rightarrow \infty$.

Next, we study the continuity of \hat{p} . We have, for $s \in [0, T]$:

$$\begin{aligned} & \int_I |\hat{p}^{n,x}(s, \cdot) - \hat{p}^x(s, \cdot)|^2 dx \\ & \leq \int_I \int_0^s \sum_{e, e' \in E} \left(|q_{e',e}^x(\hat{a}(t, e', \hat{Z}_t^{n,x}, u^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right. \\ & \quad \left. - q_{e',e}^x(\hat{a}(t, e', \hat{Z}_t^x, u^x(t, \cdot)), \hat{Z}_t^x) \right|^2 |\hat{p}^{n,x}(t, e')|^2 \\ & \quad \left. + |q_{e',e}^x(\hat{a}(t, e', \hat{Z}_t^x, u^x(t, \cdot)), \hat{Z}_t^x)|^2 |\hat{p}^{n,x}(t, e') - \hat{p}^x(t, e')|^2 \right) dt dx \end{aligned}$$

$$\begin{aligned} &\leq C \left(\int_0^s \sum_{e, e' \in E} \int_I \left| q_{e', e}^x(\hat{a}(t, e', \hat{Z}_t^{n,x}, u^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) - q_{e', e}^x(\hat{a}(t, e', \hat{Z}_t^x, u^x(t, \cdot)), \hat{Z}_t^x) \right|^2 dx dt \right. \\ &\quad \left. + \int_0^s \int_I \left| \hat{p}^{n,x}(t, \cdot) - \hat{p}^x(t, \cdot) \right|^2 dx dt \right), \end{aligned}$$

where we used Condition 1.1 and a uniform (*i.e.*, independent of n) bound on $\hat{p}^{n,x}(t, \cdot)$ (since q is bounded independently of n). By Grönwall’s inequality, we obtain

$$\begin{aligned} &\int_I \left| \hat{p}^{n,x}(s, \cdot) - \hat{p}^x(s, \cdot) \right|^2 dx \\ &\leq C \int_0^s \sum_{e, e' \in E} \int_I \left| q_{e', e}^x(\hat{a}(t, e', \hat{Z}_t^{n,x}, u^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) - q_{e', e}^x(\hat{a}(t, e', \hat{Z}_t^x, u^x(t, \cdot)), \hat{Z}_t^x) \right|^2 dx dt \end{aligned}$$

which tends to 0 as $n \rightarrow \infty$, and we concluded by using the continuity of q (given by Condition 1.1).

Finally, we study the continuity of \hat{u} . For $s \in [0, T]$, we have:

$$\begin{aligned} &\int_I \left| \hat{u}^{n,x}(s, \cdot) - \hat{u}^x(s, \cdot) \right|^2 dx \\ &\leq \int_I \int_s^T \left| -H^x(t, e, \hat{Z}_t^{n,x}, \Delta_e \hat{u}^{n,x}(t, \cdot)) + H^x(t, e, \hat{Z}_t^x, \Delta_e \hat{u}^x(t, \cdot)) \right|^2 dt dx \\ &\leq \int_s^T \int_I \sum_{e' \in E} \left| q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right|^2 \left| \hat{u}^x(t, \cdot) - \hat{u}^{n,x}(t, \cdot) \right|^2 dx dt \\ &\quad + \int_I \int_s^T \left| q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot)), \hat{Z}_t^x) - q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right|^2 \left| \hat{u}^x(t, \cdot) \right|^2 dt dx \\ &\quad + \int_I \int_s^T \left| -f^x(t, e, \hat{Z}_t^{n,x}, \hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot))) + f^x(t, e, \hat{Z}_t^x, \hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot))) \right|^2 dt dx \\ &\leq \int_s^T \int_I \sum_{e' \in E} \left| q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right|^2 \left| \hat{u}^x(t, \cdot) - \hat{u}^{n,x}(t, \cdot) \right|^2 dx dt \\ &\quad + C \int_I \int_s^T \left| q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot)), \hat{Z}_t^x) - q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right|^2 dt dx \\ &\quad + \int_I \int_s^T \left| -f^x(t, e, \hat{Z}_t^{n,x}, \hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot))) + f^x(t, e, \hat{Z}_t^x, \hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot))) \right|^2 dt dx. \end{aligned}$$

Using the boundedness of \hat{u}^x and $\hat{u}^{n,x}$ (given by Condition 3.3) and the boundedness of q (given by Condition 1.1), we deduce with Grönwall’s inequality:

$$\begin{aligned} &\int_I \left| \hat{u}^{n,x}(s, \cdot) - \hat{u}^x(s, \cdot) \right|^2 dx \\ &\leq C \left(\int_I \int_s^T \left| q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot)), \hat{Z}_t^x) - q_{e, e'}(\hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot)), \hat{Z}_t^{n,x}) \right|^2 dt dx \right. \\ &\quad \left. + \int_I \int_s^T \left| -f^x(t, e, \hat{Z}_t^{n,x}, \hat{a}(t, e, \hat{Z}_t^{n,x}, \hat{u}^{n,x}(t, \cdot))) + f^x(t, e, \hat{Z}_t^x, \hat{a}(t, e, \hat{Z}_t^x, \hat{u}^x(t, \cdot))) \right|^2 dt dx \right), \end{aligned}$$

which tends to zero as $n \rightarrow \infty$ by the continuity of f and q (given by Condition 3.3 and Condition 1.1 respectively) and the fact that $\lim_{n \rightarrow \infty} \hat{Z}^n = \hat{Z}$.

We conclude the proof by applying Schauder’s theorem, see *e.g.*, [5, Exercice 6.26], which yields the existence of a fixed point $(\hat{u}, \hat{p}) \in \mathcal{K}_C$ to Ψ .

B Proofs for Sect. 4

B.1 Proposition 1

The proof is inspired by [17, Thm. 3.2] where the authors study problems of optimal control of McKean-Vlasov-type pure jump processes and propose the use of the identities found between (29) and (30) below.

Consider the mapping $\Psi : \mathcal{L}_E \rightarrow \mathcal{L}_E$ defined by

$$\Psi(\underline{\chi}) := \underline{\xi} + \sum_{k=-n+1}^{n-1} k \int_{(0, \cdot]} \mathbf{1}_{[0, \kappa_s(\underline{\chi}_{s-}, k, \alpha_s, z_{s-})]}(y) \underline{N}_k(dy, ds), \quad \underline{\chi} \in \mathcal{L}_E.$$

We note that Ψ is well-defined. Indeed, for any $\underline{\chi} \in \mathcal{L}_E$, $\Psi(\underline{\chi})$ is a linear combination of the initial condition $\underline{\xi}$ and the Poisson random measures evaluated at measurable sets, hence $\mathcal{F} \otimes \mathcal{I}$ -measurable. By construction $\Psi(\underline{\chi})$ is $\mathbb{P} \otimes \lambda$ -a.e. \mathcal{D}_E -valued which implies the integrability.

To conclude the proof, we show that Ψ has the contraction property, *i.e.*, that there exists a constant $C < 1$ such that

$$\mathbb{E}^{\otimes} \left[\|\Psi(\underline{\chi}) - \Psi(\underline{\tilde{\chi}})\|_T^2 \right] \leq C \mathbb{E}^{\otimes} \left[\|\underline{\chi} - \underline{\tilde{\chi}}\|_T^2 \right], \quad \underline{\chi}, \underline{\tilde{\chi}} \in \mathcal{L}_E.$$

By independence of the Poisson measures $(N_k^x)_{k=1}^n$ for any fixed $x \in I$, the compensated martingales $(\tilde{N}_k^x)_{k=1}^n$ (cf. Sect. 4.1.1) are orthogonal. Hence, for any $\underline{\chi} \in \mathcal{L}_E$ and $t \in [0, T]$,

$$\begin{aligned} & \xi^x(\omega) + \sum_{k=-n+1}^{n-1} k \int_{(0, t]} \mathbf{1}_{[0, \kappa_s(\chi_s^x(\omega), k, \alpha_s^x(\omega), z_s^x(\omega))]}(y) N_k^x(\omega, dy, ds) \\ &= \xi^x(\omega) + \int_0^t \sum_{k=-n+1}^{n-1} k \kappa_s^x(\chi_s^x(\omega), k, \alpha_s^x(\omega), z_s^x(\omega)) ds + M_t^{\alpha, z, x}(\omega), \quad \mathbb{P} \otimes \lambda\text{-a.e.} \end{aligned}$$

where $M^{\alpha, z}$ is a zero-mean stochastic integral, and we have that

$$\begin{aligned} & |\chi_t^x(\omega) - \tilde{\chi}_t^x(\omega)|^2 \\ & \leq C \left(\int_0^t \sum_{k=-n+1}^{n-1} |k| |\kappa_s^x(\chi_s^x(\omega), k, \alpha_s^x(\omega), z_s^x(\omega)) - \kappa_s^x(\tilde{\chi}_s^x(\omega), k, \alpha_s^x(\omega), z_s^x(\omega))| ds \right)^2 \\ & \quad + C |M_t^{\alpha, z, x}(\omega) - \tilde{M}_t^{\alpha, z, x}(\omega)|^2, \quad \mathbb{P} \otimes \lambda\text{-a.e.} \end{aligned} \tag{29}$$

where the process $\tilde{M}^{\alpha, z}$ is the equivalent of $M^{\alpha, z}$ but for $\tilde{\chi}$. The predictable quadratic variation of the process $M^{\alpha, z} - \tilde{M}^{\alpha, z}$ is

$$\begin{aligned} & \langle M^{\alpha, z, x} - \tilde{M}^{\alpha, z, x} \rangle_t = \\ & \sum_{k=-n+1}^{n-1} k^2 \int_{\mathbb{R} \times (0, t]} (\mathbf{1}_{[0, \kappa_s^x(\chi_s^x, k, \alpha_s^x, z_s^x)]}(y) - \mathbf{1}_{[0, \kappa_s^x(\tilde{\chi}_s^x, k, \alpha_s^x, z_s^x)]}(y))^2 dy \otimes ds, \quad \mathbb{P} \otimes \lambda\text{-a.e.} \end{aligned}$$

where the ω -dependence has been suppressed in the notation. Expanding the square and integrating, using the identities

$$(\mathbf{1}_{[0, a]} - \mathbf{1}_{[0, b]})^2 = \mathbf{1}_{[0, a]} - 2\mathbf{1}_{[0, \min(a, b)]} + \mathbf{1}_{[0, b]},$$

and $a - 2 \min(a, b) + b \leq |a - b|$ for $a, b \geq 0$, we obtain

$$\langle M^{\alpha, z, x} - \tilde{M}^{\alpha, z, x} \rangle_t \leq \int_0^t \sum_{k=-n+1}^{n-1} k^2 |\kappa_s^x(X_s^x, k, \alpha_s^x, z_s^x) - \kappa_s^x(\tilde{X}_s^x, k, \alpha_s^x, z_s^x)| ds, \mathbb{P} \boxtimes \lambda\text{-a.e.}$$

From the Lipschitz continuity imposed by Condition 1.1, we get

$$\begin{aligned} \langle M^{\alpha, z, x}(\omega) - \tilde{M}^{\alpha, z, x}(\omega) \rangle_t &\leq C \int_0^t \mathbf{1}_{\{X_s^x(\omega) \neq \tilde{X}_s^x(\omega)\}} ds \\ &\leq C \int_0^t \|X^x(\omega) - \tilde{X}^x(\omega)\|_s^2 ds, \mathbb{P} \boxtimes \lambda\text{-a.e.} \end{aligned} \tag{30}$$

After taking expectation of (29), we get using Doob’s inequality and (30) that

$$\mathbb{E}^{\boxtimes} \left[\|\Psi(\underline{X}) - \Psi(\tilde{X})\|_T^2 \right] \leq C \int_0^T \mathbb{E}^{\boxtimes} \left[\|\underline{X} - \tilde{X}\|_s^2 \right] ds.$$

Iterating the inequality, we get for any $N \in \mathbb{N}$ that

$$\mathbb{E}^{\boxtimes} \left[\|(\Psi^N)(\underline{X}) - \Psi(\tilde{X})^{(N)}\|_T^2 \right] \leq \frac{(CT)^N}{N!} \mathbb{E}^{\boxtimes} \left[\|\underline{X} - \tilde{X}\|_T^2 \right],$$

where (Ψ^N) denotes the N -fold composition of Ψ . Thus, for some N large enough, (Ψ^N) is a contraction and it follows from the Banach fixed-point theorem for iterated mappings (see, e.g. [6]) that Ψ has a unique fixed point in the set \mathcal{L}_E . The fixed point is the unique (up to $\mathbb{P} \boxtimes \lambda$ -modification) strong solution that was sought.

B.2 Theorem 4

The first step of the proof is to show that the aggregate variable is well-defined as a fixed point to the mapping

$$\begin{aligned} U^\alpha &: L^2_{\boxtimes}(\Omega \times I; \mathcal{D}) \rightarrow L^2_{\boxtimes}(\Omega \times I; \mathcal{D}), \\ \underline{z} &\mapsto [U^\alpha \underline{z}] : (\omega, x) \mapsto \left(\int_I w(x, y) K \left(\alpha_t^y(\omega), X_{t-}^{\alpha, z, y}(\omega) \right) \lambda(dy) \right)_{t \in [0, T]} \end{aligned} \tag{31}$$

where $\underline{X}^{\alpha, z} \in \mathcal{L}_E$ is the solution to (26) characterized in Proposition 1.

Lemma 2 *Let Condition 1 hold. For each $\alpha \in \underline{\mathbb{A}}$ the mapping U^α has a unique fixed point in $L^2_{\boxtimes}(\Omega \times I; \mathcal{D})$.*

Denoting the fixed point by \underline{Z}^α , the next lemma uses the Exact Law of Large Numbers [53] to guarantee that \underline{Z}^α is \mathcal{C} -valued $\mathbb{P} \boxtimes \lambda$ -a.s. for each $\alpha \in \underline{\mathbb{A}}$.

Lemma 3 *Let Conditions 1 and 2 hold, and let $\alpha \in \underline{\mathbb{A}}$. Then $[U^\alpha \underline{z}] \in L^2_{\boxtimes}(\Omega \times I; \mathcal{C})$ for each $\underline{z} \in L^2_{\boxtimes}(\Omega \times I; \mathcal{D})$.*

This proves the part (i) of the theorem. Part (ii) and (iii) can be shown along the same lines of proof as is used in [3, Thm. 2]

B.2.1 Proof of Lemma 2

Let $\underline{z}, \tilde{z} \in L^2_{\mathbb{F}}(\Omega \times I; \mathcal{D})$. We have for $(\omega, x) \in \Omega \times I$:

$$\|U^{\alpha}_{\underline{z}}(\omega, x) - U^{\alpha}_{\tilde{z}}(\omega, x)\|_T^2 \leq C \int_I \|X^{\alpha, \underline{z}, y}(\omega) - X^{\alpha, \tilde{z}, y}(\omega)\|_T^2 \lambda(dy). \quad (32)$$

Following similar lines of proof as in Proposition 1, get the initial $\mathbb{P} \otimes \lambda$ -a.e. estimates

$$\begin{aligned} |X_t^{\alpha, \underline{z}, y}(\omega) - X_t^{\alpha, \tilde{z}, y}(\omega)|^2 &\leq C \left(\int_0^t \sum_{k=-n+1}^{n-1} |k| |\kappa_s^y(X_{s-}^{\alpha, \underline{z}, y}(\omega), k, \alpha_s^y(\omega), z_{s-}^y(\omega)) \right. \\ &\quad \left. - \kappa_s^y(X_{s-}^{\alpha, \tilde{z}, y}(\omega), k, \alpha_s^y(\omega), \tilde{z}_{s-}^y(\omega))| ds \right)^2 + C |M_t^{\alpha, \underline{z}, y}(\omega) - M_t^{\alpha, \tilde{z}, y}(\omega)|^2 \end{aligned} \quad (33)$$

and

$$\begin{aligned} &\langle M^{\alpha, \underline{z}, x}(\omega) - M^{\alpha, \tilde{z}, x}(\omega) \rangle_t \\ &\leq \int_0^t \sum_{k=-n+1}^{n-1} |k|^2 |\kappa_s^y(X_{s-}^{\alpha, \underline{z}, y}(\omega), k, \alpha_s^y(\omega), z_{s-}^y(\omega)) - \kappa_s^y(X_{s-}^{\alpha, \tilde{z}, y}(\omega), k, \alpha_s^y(\omega), \tilde{z}_{s-}^y(\omega))| ds. \end{aligned}$$

In view of Condition 1

$$\begin{aligned} &|X_t^{\alpha, \underline{z}, x}(\omega) - X_t^{\alpha, \tilde{z}, x}(\omega)|^2 \\ &\leq C \int_0^t \left(\|X^{\alpha, \underline{z}, x}(\omega) - X^{\alpha, \tilde{z}, x}(\omega)\|_s^2 + |z_{s-}^x(\omega) - \tilde{z}_{s-}^x(\omega)|^2 \right) ds, \quad \mathbb{P} \otimes \lambda\text{-a.s.} \end{aligned} \quad (34)$$

Taking expectation in (32) we get (using the Fubini property) that

$$\mathbb{E}^{\otimes} [\|U^{\alpha}_{\underline{z}} - U^{\alpha}_{\tilde{z}}\|_T^2] \leq C \mathbb{E}^{\otimes} [\|\underline{X}^{\alpha, \underline{z}} - \underline{X}^{\alpha, \tilde{z}}\|_T^2].$$

Then, by (33), Doob's inequality, and (34) we get

$$\mathbb{E}^{\otimes} [\|\underline{X}^{\alpha, \underline{z}} - \underline{X}^{\alpha, \tilde{z}}\|_T^2] \leq C \int_0^T \mathbb{E}^{\otimes} [\|\underline{X}^{\alpha, \underline{z}} - \underline{X}^{\alpha, \tilde{z}}\|_s^2 + |z_{s-} - \tilde{z}_{s-}|^2] ds.$$

Hence, after one use of Gronwall's inequality, we have that

$$\mathbb{E}^{\otimes} [\|U^{\alpha}_{\underline{z}} - U^{\alpha}_{\tilde{z}}\|_T^2] \leq C \int_0^T \mathbb{E}^{\otimes} [\|\underline{z} - \tilde{z}\|_s^2] ds$$

and we conclude the proof in the same way as in Proposition 1.

B.2.2 Proof of Lemma 3

Consider the function

$$[0, T] \ni t \mapsto [U^{\alpha}_{\underline{z}}]_t(\omega, x) := \int_I w(x, y) K(a_t^y, X_{t-}^{\alpha, \underline{z}, y}) \lambda(dy), \quad (\omega, x) \in \Omega \times I.$$

Recall that $\alpha \in \underline{\mathbb{A}}$ means that $\alpha_t^y = \alpha(y, t, X_{t-}^{\alpha, \underline{z}, y})$ for $y \in I$ and $t \in [0, T]$. It follows from the Lipschitz assumption on K , compactness of A , and definition of $\underline{\mathbb{A}}$ that

$$\begin{aligned}
 & |K(\alpha_t^y, X_{t-}^{\alpha, z, y}) - K(\alpha_s^y, X_{s-}^{\alpha, z, y})| \\
 & \leq C \left(|\underline{\alpha}(y, t, X_{s-}^{\alpha, z, y}) - \underline{\alpha}(y, s, X_{s-}^{\alpha, z, y})| + |X_{t-}^{\alpha, z, y} - X_{s-}^{\alpha, z, y}| \right), \quad s, t \in [0, T].
 \end{aligned}$$

Let $t_j \in [0, T]$, $j \in \mathbb{N}$, be a sequence converging to $t^* \in [0, T]$. Without loss of generality, assume that the sequence is non-decreasing. Recall that q_{\max} denotes the uniform upper bound for the intensity rates, see Condition 1.1. For $(\omega, x) \in (\Omega \times I)$,

$$\begin{aligned}
 & \left| \int_I w(x, y) \left(X_{t_j-}^{\alpha, z, y}(\omega) - X_{t^*-}^{\alpha, z, y}(\omega) \right) \lambda(dy) \right| \\
 & \leq C \int_I \sum_{k=-n+1}^{n-1} \left| \int_{\mathbb{R} \times [t_j, t^*)} 1_{[0, \kappa_s^y(X_{s-}^{\alpha, z, y}(\omega), k, \alpha_s^y(\omega), z_s^y(\omega))]}(u) N_k^y(\omega, du \otimes ds) \right| \lambda(dy) \\
 & \leq C \int_I \tilde{N}_j^y(\omega) \lambda(dy),
 \end{aligned}$$

where $\tilde{N}_j^y := \sum_{k=-n+1}^{n-1} N_k^y([0, q_{\max}] \times [t_j, t^*))$ is for each $y \in I$ a Poisson-distributed random variable with intensity $(2n - 1)q_{\max}|t^* - t_j|$, since the summands $N_k^y([0, q_{\max}] \times [t_j, t^*))$, $k = -n + 1, \dots, n - 1$ are independent Poisson-distributed with intensity $q_{\max}|t^* - t_j|$. Moreover, $(\tilde{N}_j^y)_{y \in I}$ are e.p.i., so by the Exact Law of Large Numbers

$$\int_I \tilde{N}_j^y(\omega) \lambda(dy) = \int_I \mathbb{E} \left[\tilde{N}_j^y \right] \lambda(dy) = (2n - 1)q_{\max}|t^* - t_j|, \quad \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Hence, by the boundedness of A , it holds for all $x \in I$ and \mathbb{P} -a.e. $\omega \in \Omega$ that

$$\begin{aligned}
 & \lim_{j \rightarrow \infty} |[U^{\alpha, z}]_{t_j}(\omega, x) - [U^{\alpha, z}]_{t^*}(\omega, x)| \\
 & \leq \lim_{j \rightarrow \infty} C \int_I \left(|\underline{\alpha}(y, t_j, X_{t_j-}^{\alpha, z, y}, y) - \underline{\alpha}(y, t^*, X_{t_j-}^{\alpha, z, y})| + \tilde{N}_j^y(\omega) \right) \lambda(dy) \rightarrow 0,
 \end{aligned}$$

and in particular $[U^{\alpha, z}](\omega, x) \in \mathcal{C}$ for $\mathbb{P} \boxtimes \lambda$ -a.e. $(\omega, x) \in \Omega \times I$. Measurability and square-integrability of $U^{\alpha, z}$ follows by the same lines of proof as Lemma 1 in [3].

References

1. The Incubation Period of Coronavirus Disease (2019) (COVID-19) (2020) From publicly reported confirmed cases: estimation and application. *Ann Intern Med* 172(9):577–582
2. Andersson H, Djehiche B (1995) Limit theorems for multitype epidemics. *Stochast Process Appl* 56(1):57–75
3. Alexander A, Rene C, Gokce D, Mathieu L (2020) Optimal incentives to mitigate epidemics: a Stackelberg mean field game approach. [arXiv:2011.03105](https://arxiv.org/abs/2011.03105)
4. Alexander A, Rene C, Mathieu L (2021) Stochastic graphon games: II. the linear-quadratic case. [arXiv:2105.12320](https://arxiv.org/abs/2105.12320)
5. Bayraktar E, Cohen A (2018) Analysis of a finite state many player game using its master equation. *SIAM J Control Opt* 56(5):3538–3568
6. Brezis H (2011) *Functional analysis. Sobolev spaces and partial differential equations*. Universitext, Springer, New York
7. Bryant VW (1968) A remark on a fixed-point theorem for iterated mappings. *Am Math Monthly* 75:399–400
8. Caines PE, Huang M (2019) Graphon Mean field games and the GMFG equations: ϵ -nash equilibria. In: 2019 IEEE 58th conference on decision and control (CDC), pp 286–292
9. Caines PE, Huang M (2018) Graphon mean field games and the GMFG equations. 2018 IEEE conference on decision and control (CDC)

10. Carmona R, Cooney D, Graves C, Lauriere M (2021) Stochastic graphon games: I. The static case. To appear in *Math Oper Res*
11. Carmona R, Laurière M (2021) Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games I: the ergodic case. *SIAM J Numer Anal* 59(3):1455–1485
12. Carmona R., Laurière M (2021) Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II—the finite horizon case. [arXiv:1908.01613](https://arxiv.org/abs/1908.01613)
13. Carmona R, Wang P (2016) Finite state mean field games with major and minor players. [arXiv:1610.05408](https://arxiv.org/abs/1610.05408)
14. Carmona R, Wang P (2021) A probabilistic approach to extended finite state mean field games. *Math Oper Res*
15. Cecchin A, Fischer M (2018) Probabilistic approach to finite state mean field games. *Appl Math Opt*, pp 1–48
16. Charpentier A, Elie R, Laurière M, Tran VC (2020) Covid-19 pandemic control: balancing detection policy and lockdown intervention under icu sustainability. *Math Modell Nat Phenomena* 15:57
17. Cho S (2020) Mean-field game analysis of sir model with social distancing. [arXiv:2005.06758](https://arxiv.org/abs/2005.06758)
18. Choutri SE, Djehiche B (2019) Mean-field risk sensitive control and zero-sum games for Markov chains. *Bulletin des Sciences Mathématiques* 152:1–39
19. Choutri SE, Djehiche B, Tembine H (2019) Optimal control and zero-sum games for Markov chains of mean-field type. *Math Control Related Fields* 9(3):571
20. Choutri SE, Hamidou T (2018) A stochastic maximum principle for Markov chains of mean-field type. *Games* 9(4):84
21. Cvitanic J, Possamaï D, Touzi N (2018) Dynamic programming approach to principal-agent problems. *Finance Stoch* 22(1):1–37
22. Jakša Cvitanic, Jianfeng Zhang (2013) *Contract theory in continuous-time models*. Springer Finance, Springer, Heidelberg
23. Andrew DD, Edwin P (2012) *Superprocesses at Saint-Flour*. Springer
24. Delarue F (2017) Mean field games: a toy model on an Erdős-Renyi graph. *ESAIM: Procs*, 60:1–26
25. Delmas JF, Dronnier D , Zitt P (2021) Targeted Vaccination Strategies for an Infinite-Dimensional SIS Model. [arXiv:2103.10330](https://arxiv.org/abs/2103.10330)
26. Djehiche Boualem, Kaj Ingemar (1995) The rate function for some measure-valued jump processes. *The annals of probability*, pages 1414–1438
27. Djehiche B, Alexander Schied (1998) Large deviations for hierarchical systems of interacting jump processes. *J Theor Prob* 11(1):1–24
28. Doncel J, Gast N, Gaujal B(2020) A mean-field game analysis of SIR dynamics with vaccination. *Prob Eng Inf Sci*, pp 1–18
29. Romuald Elie, Emma Hubert, Gabriel Turinici (2020) Contact rate epidemic control of COVID-19: an equilibrium view. *Math Modell Nat Phenomena* 15:35
30. Gao S, Caines PE (2019) Spectral representations of graphons in very large network systems control. 2019 IEEE 58th conference on decision and Control (CDC)
31. Gao S, Caines PE, Huang M (2021) LQG graphon mean field games. [arXiv:2004.00679](https://arxiv.org/abs/2004.00679)
32. Gaujal B, Doncel J, Gast N (2020) Vaccination in a large population: mean field equilibrium versus social optimum. In: *netgcoop'20*, Cargèse, France, September
33. Gomes AD, Mohr J, Rigao Souza Rafael (2010) Discrete time, finite state space mean field games. *Journal de mathématiques pures et appliquées* 93(3):308–328
34. Gomes Diogo A, Joana Mohr, Rigao Souza Rafael (2013) Continuous time finite state mean field games. *Appl Math Opt* 68(1):99–143
35. Grigelionis B (1971) On representation of integer-valued random measures by means of stochastic integrals with respect to the Poisson measure. *Lithuanian Math J* 1:93–108
36. Huang M, Malhamé RP, Caines Peter E et al (2006) Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun Inf Syst* 6(3):221–252
37. Hubert E, Mastrolia T, Possamaï D, Warin X (2020) Incentives, lockdown, and testing: from Thucydides’s analysis to the COVID-19 pandemic. [arXiv:2009.00484](https://arxiv.org/abs/2009.00484)
38. Hubert E, Turinici G (2018) Nash-MFG equilibrium in a SIR model with time dependent newborn vaccination. *Ricerche di Matematica* 67(1):227–246
39. Hytönen T, Van Neerven J, Veraar M, Weis L (2016) *Anal Banach spaces*, vol 12. Springer
40. Keeling JM, Eames TDK (2005) Networks and epidemic models. *J R Soc Interface* 2(4):295–307
41. Keliger D, Horvath I, Takacs B (2020) Local-density dependent Markov processes on graphons with epidemiological applications. [arXiv:2008.08109](https://arxiv.org/abs/2008.08109)
42. Kohlmann M, Zhou XY (2000) Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach. *SIAM J Control Optim* 38(5):1392–1407

43. Kolokoltsov VN (2012) Nonlinear Markov games on a finite state space (mean-field and binary interactions). *Int J Stat Prob* 1(1):77–91
44. Laguzet L, Turinici G (2015) Individual vaccination as Nash equilibrium in a SIR model with application to the 2009–2010 influenza A (H1N1) epidemic in France. *Bull Math Biol* 77(10):1955–1984
45. Laguzet L, Turinici G, Yahiaoui G (2016) Equilibrium in an individual-societal SIR vaccination model in presence of discounting and finite vaccination capacity. In: *New trends in differential equations, control theory and optimization: proceedings of the 8th congress of Romanian mathematicians*, pp 201–214. World Scientific
46. Lasry JM, Lions PL (2006) Jeux à champ moyen. i-le cas stationnaire. *Comptes Rendus Mathématique* 343(9):619–625
47. Lasry JM, Lions PL (2006) Jeux à champ moyen. ii-horizon fini et contrôle optimal. *Comptes Rendus Mathématique* 343(10):679–684
48. Lee W, Liu S, Tembine H, Li W, Osher S (2020) Controlling propagation of epidemics via mean-field games. *arXib:2006.01249*
49. László Lovász (2012) *Large networks and graph limits*, vol 60. American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI
50. Parise F, Ozdaglar AE (2019) Graphon games: a statistical framework for network games and interventions. *SSRN Electron J*
51. Francesco Salvarani, Gabriel Turinici (2018) Optimal individual strategies for influenza vaccines with imperfect efficacy and durability of protection. *Mathematical Biosciences & Engineering* 15(3):629
52. Skorokhod AV (1982) *Studies in the theory of random processes*, vol 7021. Courier Dover Publications
53. Sun Y (1998) A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN. *Math. Econom* 29(4):419–503
54. Sun Y (2006) The exact law of large numbers via Fubini extension and characterization of insurable risks. *J Econom Theory* 126(1):31–69
55. Sun Y, Zhang Y (2009) Individual risk and Lebesgue extension without aggregate uncertainty. *J Econom Theory* 144(1):432–443
56. Tembine T (2020) Covid-19: data-driven mean-field-type game perspective. *Games* 11(4):51
57. Vizuete R, Frasca P, Frasca F (2020) Graphon-based sensitivity analysis of SIS epidemics. *IEEE Control Syst Lett* 4(3):542–547

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.