Hypergraphon mean field games

(iD) **Kai Cui,** (iD) **Wasiur R. KhudaBukhsh and** (iD) **Heinz Koeppl**

COLLECTIONS

Paper published as part of the special topic on Dynamics on Networks with Higher-Order Interactions

View Online   Export Citation   CrossMark

AIP Publishing

Hypergraphon mean field games

Kai Cui,[1] 🔘 Wasiur R. KhudaBukhsh,[2] 🔘 and Heinz Koeppl[1,a)] 🔘

AFFILIATIONS

[1] Technische Universität Darmstadt, 64283 Darmstadt, Germany
[2] School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, United Kingdom

Note: This article is part of the Focus Issue, Dynamics on Networks with Higher-Order Interactions.
[a)]**Author to whom correspondence should be addressed:** heinz.koeppl@tu-darmstadt.de

ABSTRACT

We propose an approach to modeling large-scale multi-agent dynamical systems allowing interactions among more than just pairs of agents using the theory of mean field games and the notion of hypergraphons, which are obtained as limits of large hypergraphs. To the best of our knowledge, ours is the first work on mean field games on hypergraphs. Together with an extension to a multi-layer setup, we obtain limiting descriptions for large systems of non-linear, weakly interacting dynamical agents. On the theoretical side, we prove the well-foundedness of the resulting hypergraphon mean field game, showing both existence and approximate Nash properties. On the applied side, we extend numerical and learning algorithms to compute the hypergraphon mean field equilibria. To verify our approach empirically, we consider a social rumor spreading model, where we give agents intrinsic motivation to spread rumors to unaware agents, and an epidemic control problem.

Recent developments in the field of complex systems have shown that real-world multi-agent systems are often not restricted to pairwise interactions, bringing to light the need for tractable models allowing higher-order interactions. At the same time, the complexity of analysis of large-scale multi-agent systems on graphs remains an issue even without considering higher-order interactions. An increasingly popular and tractable approach of analysis is the theory of mean field games. We combine mean field games with higher-order structure by means of hypergraphons, a limiting description of very large hypergraphs. To motivate our model, we build a theoretical foundation for the limiting system, showing that the limiting system has a solution and that it approximates finite, sufficiently large systems well. This allows us to analyze otherwise intractable, large hypergraph games with theoretical guarantees, which we verify using two examples of rumor spreading and epidemics control.

I. INTRODUCTION

In recent years, there has been a surge of interest in large-scale multi-agent dynamical systems on higher-order networks due to their great generality and practical importance, e.g., in epidemiology,[1] opinion dynamics,[2,3] network synchronization,[4,5] neuroscience,[6] and more. We refer interested readers to the excellent review articles.[7–9] In addition to providing a more realistic description of the underlying processes, such large-scale systems with higher-order interactions pose interesting control problems for the reinforcement learning and control communities.[10–12] To this end, a big challenge has been to find tractable solutions.[13,14]

An increasingly popular and recent approach to the tractability issue has been to use the framework of learning in mean field games (MFGs)[15–23] and their cooperative counterpart commonly known as mean field control (MFC).[24–30] It is important to note that here, learning refers to the classical learning—i.e., iterative computation—of equilibria in game theory, as opposed to, e.g., reinforcement learning; see also, e.g., the discussion in Laurière *et al.*[31] Here are also some extensive reviews on mean field games.[32–36] Popularized by Huang *et al.*[37] and Lasry and Lions[38] in the context of differential games, mean field games and related approximations have since found application in a plethora of fields, such as transportation and traffic control,[39–41] large-scale batch processing and scheduling systems,[42–44] peer-to-peer streaming systems,[45] malware epidemics,[46] crowd dynamics and evacuation of buildings,[47–49] as well as many other applications in economics[50] and engineering.[51] Tractably finding competitive equilibria and decentralized, cooperative optimal

control solutions has been the focus of many recent works.[52–58] Since then, mean field systems have also been extended to dynamical systems on graphs, typically using the theory of large graph limits called graphons.[59,60] The graphon mean field systems can be considered either as the limit of systems with weakly interacting node state processes,[57,61] or alternatively as the result of a double limit procedure where each node constitutes a large population, or "cluster" of agents, each of which interacts with each other via inter- and intra-cluster coupling. First, infinitely many nodes are considered according to the graphon, and then infinitely many agents are considered per node (see, e.g., Caines and Huang[62,63]).

In this work, we will consider the former. The goal of our work is the synthesis of dynamical systems on hypergraphs with competitive or selfish agents. Existing analysis of hypergraph mean field systems typically remains restricted to special dynamics such as epidemiological equations[64–66] or opinion dynamics[67] on sparse graphs. In contrast, our work deals with general, agent-controlled nonlinear dynamics and equilibrium solutions. We build upon prior results for discrete-time, graph-based mean field systems[57,61,68] and extend them to incorporate higher-order hypergraphs as well as multiple layers.

Our contribution can be summarized as follows: (i) To the best of our knowledge, ours is the first general mean field game-theoretical framework for non-linear dynamics on multi-layer hypergraphs. Multi-layer networks[69] have proven extremely useful in many application areas, including infectious disease epidemiology, where different layers could be used to describe community, household, and hospital settings.[70] (ii) We prove the existence and the approximation properties of the proposed mean field equilibria. (iii) We propose and empirically verify algorithms for solving such hypergraphon mean field systems and thereby obtain a tractable approach to solving and analyzing otherwise intractable Nash equilibria on multi-layer hypergraph games. The proposed framework is of great generality, extending the recently established graphon mean field games and thereby also standard mean field games (via fully connected graphs).

After introducing some graph-theoretical preliminaries, in Sec. II, we will begin by formulating the motivating mathematical dynamical model and game on hypergraphs, as well as its more tractable mean field analog. Then, in Sec. III, we will show the existence of solutions for the mean field problem as well as quantify its approximation qualities of the finite hypergraph game, building a mathematical foundation for hypergraphon mean field games. Last, in Sec. IV, we will evaluate our model numerically for an illustrative rumor spreading game, verifying our theoretical approximation results and the obtained equilibrium behavior. All of the proofs can be found in Appendixes A and B.

Notation. On a discrete space $\mathbb{A}$, define the spaces of all (Borel) probability measures $\mathcal{P}(\mathbb{A})$ and all sub-probability measures $\mathcal{B}(\mathbb{A})$, equipped with the L_1 norm. Define the unit interval $\mathbb{I} := [0, 1]$ and its N equal-length subintervals $I_1^N, \ldots, I_N^N$ such that $\bigsqcup_{i=1}^N I_i^N = \mathbb{I}$ for any integer N, where $\bigsqcup$ denotes disjoint union and each I_i^N includes its rightmost point i/N. Denote the expectation and variance of random variables X by $\mathbb{E}[X]$, $\mathbb{V}[X]$. Define the indicator function $\mathbb{1}_A(x)$ mapping to 1 whenever $x \in A$ and 0 otherwise. For any integer k, define $[k] := \{1, \ldots, k\}$. Let $r(A, m)$ denote the set of all distinct non-empty subsets of any set A with

at most m elements and denote the set of all distinct non-empty, proper subsets by $r_<(A) := r(A, |A| - 1)$ as well as the set of all distinct non-empty subsets by $r(A) := r(A, |A|)$. To keep the notation simple, in the following, we write $r_<[k] := r_<([k])$, $r[k] := r([k])$ and identify, e.g., $r_<[k]$ with $[|r_<[k]|] := \{1, \ldots, |r_<[k]|\}$ whenever helpful. Denote the set of permutations of a set A as $\text{Sym}(A)$. Define the space of bounded, $r_<[k]$-dimensional, symmetric functions $\text{Sym}_<^{\text{ind}}[k]$ induced by permutations of the underlying set $[k]$, i.e., any bounded function $f: \mathbb{I}^{r_<[k]} \to \mathbb{R}$ is in $\text{Sym}_<^{\text{ind}}[k]$ whenever f is invariant to all permutations $\sigma \in \text{Sym}([k])$, $f(x_1, \ldots, x_k, x_{11}, x_{12}, \ldots) = f(x_{\sigma(1)}, \ldots, x_{\sigma(k)}, x_{\sigma(1)\sigma(1)}, x_{\sigma(1)\sigma(2)}, \ldots)$. Analogously, we define spaces of such functions $\text{Sym}_\leq^{\text{ind}}[k]$ and $\text{Sym}^{\text{ind}}[k]$ over $r[k]$ and $[k]$, respectively.

II. MATHEMATICAL MODEL

Before we formulate the stochastic dynamic hypergraph game and its limiting analog in Secs. II A and II B, we discuss some graph-theoretical preliminaries. A (undirected) hypergraph is defined as a pair $H = (V, E)$ of a set of vertices V and a set of hyperedges $E \subseteq 2^V \setminus \{\varnothing\}$. In contrast to edges in graphs, here hyperedges may connect an arbitrary number of vertices instead of only two. If there is no scope of confusion, we will call hyperedges of a hypergraph just edges. Denote by $V[H]$ and $E[H]$ the vertex set and edge set of a hypergraph H. The maximum cardinality of all edges of a hypergraph H is called its rank. A k-uniform hypergraph is defined as a hypergraph where all edges have cardinality k. A multi-layer hypergraph $H = (V, E^1, \ldots, E^D)$ with D layers is obtained by allowing for multiple edge sets $E^1, \ldots, E^D \subseteq 2^V \setminus \{\varnothing\}$, and we analogously write $E^d[H]$ for the dth set of edges of a multi-layer hypergraph H. We define the dth sub-hypergraph H^d of a multi-layer hypergraph H as the hypergraph with vertex set $V[H]$ and edge set $E[H^d] = E^d[H]$.

Consider any (non-uniform) hypergraph H with bounded rank $k_{\max}$. Observe the isomorphism between multi-layer uniform hypergraphs and such H by splitting hyperedges of each cardinality $k \leq k_{\max}$ into their own layer. Since this procedure can be repeated for each layer of a multi-layer hypergraph, any multi-layer hypergraph is, therefore, equivalent to a correspondingly defined multi-layer uniform hypergraph. Hence, from here on, it suffices to define and consider $[k_1, \ldots, k_D]$-uniform hypergraphs H as D-layer hypergraphs, where each layer $d = 1, \ldots, D$ is given by a k_d-uniform hypergraph with $k_d \leq k_{\max}$ (see also Fig. 1 for a visualization). For instance, in social networks, each layer could model, e.g., the k-cliques of acquaintances formed at work, friendship at school, or family relations.

To formulate the infinitely large mean field system, we define the limiting description of sufficiently dense multilayer hypergraphs as the graphs intuitively become infinite in size, called hypergraphons.[71] Here, dense means a number of edges on the order of $O(N^2)$, where N is the number of vertices, to which existing hypergraphon theory remains limited to. However, we note that an extension to more sparse models by fusing the theory of hypergraphons with L^p graphons[72–74] could be part of future work. The space of k-uniform hypergraphons $\mathbb{W}_k$ is now defined as the space of all bounded and symmetric functions $W \in \text{Sym}_<^{\text{ind}}[k]$, $W: \mathbb{I}^{r_<[k]} \to \mathbb{I}$ that are measurable. We equip $\text{Sym}_<^{\text{ind}}[k]$ with the cut (semi-)norm
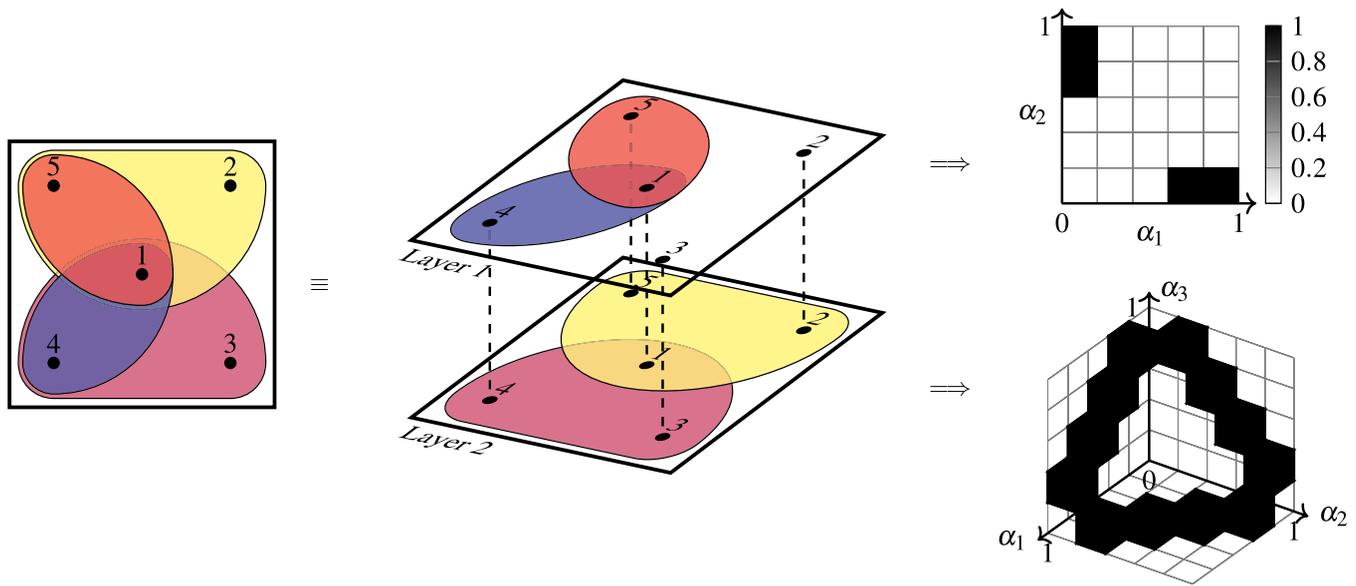
FIG. 1. An example hypergraph H is transformed into a multi-layer uniform hypergraph. On the left, a hypergraph H with nodes $V[H] = \{1, \ldots, 5\}$ and hyperedges $E[H] = \{\{1, 4\}, \{1, 5\}, \{1, 3, 4\}, \{1, 2, 5\}\}$ is depicted. An equivalent representation of H as a $[2, 3]$-uniform hypergraph H_{unif} as well as its associated hypergraphons are given, where the first and second layers each consist of edges $\{\{1, 4\}, \{1, 5\}\}$ and three-hyperedges $\{\{1, 3, 4\}, \{1, 2, 5\}\}$, respectively. The associated (step-)hypergraphons $W[H_{\text{unif}}^1]$ and $W[H_{\text{unif}}^2]$ are given as continuous versions of the (multi-dimensional) $\{0, 1\}$-valued adjacency matrices. Here, we depict only the first three coordinates for the second layer step-hypergraphon $W[H_{\text{unif}}^2]$, given by the constant 1 (black) or 0 (white). Note that while each edge corresponds to two entries in the adjacency matrix of the two-uniform case, for the three-uniform case each hyperedge corresponds to six entries, resulting in the step graphon shown (bottom right).

$\|\cdot\|_{\square^{k-1}}$ proposed by Zhao,[75] defined by

$$\|W\|_{\square^{k-1}} := \sup_{\substack{u_i : \mathbb{I}^{r[k-1]} \to \mathbb{I}, \\ u_i \in \text{Sym}^{\text{ind}}_{\leq}[k-1]}} \left| \int_{\mathbb{I}^{r<[k]}} W(\alpha) \prod_{i=1}^{k} u_i(\alpha_{r([k]\setminus\{i\})}) \, d\alpha \right|, \quad (1)$$

which (see, e.g., Lemma 8.10 of Ref. 59) coincides with the standard graphon case for $k = 2$,

$$\|W\|_{\square} = \sup_{f,g : \mathbb{I} \to \mathbb{I}} \left| \int_{\mathbb{I}^2} W(\alpha, \beta) f(\alpha) g(\beta) \, d(\alpha, \beta) \right|. \quad (2)$$

To analytically connect k-uniform hypergraphs to hypergraphons, we define the step-hypergraphons of any k-uniform hypergraph H as

$$W[H](\boldsymbol{\alpha}) = \sum_{\mathbf{m} \in [N]^k} \mathbb{1}_{E[H]}(\mathbf{m}) \cdot \prod_{i \in [k]} \mathbb{1}_{I_{m_i}^N}(\alpha_i). \quad (3)$$

For motivation, note that for any sequence of graphs with converging homomorphism densities, equivalently the step graphons converge in the cut norm to the limiting graphon, and their limiting homomorphism densities can be described by the limiting graphon.[59] Similarly, cut-norm convergence for the more general uniform hypergraphs at least implies the convergence of hypergraph homomorphism densities.[75] Accordingly, we assume hypergraph convergence in each layer of a given sequence of $[k_1, \ldots, k_D]$-uniform hypergraphs $(H_N)_{N \in \mathbb{N}}$ via the convergence of their step-hypergraphons $W_d^N := W[H_N^d]$ to a limiting hypergraphon W_d

$\in \mathbb{W}_{k_d}$ in the cut norm as visualized in Fig. 2, similar as in standard graphon mean field systems.[57,61]

Assumption 1: *The sequence of step-hypergraphons* $W^N := (W_d^N)_{d \in [D]}$ *converges on each layer in cut norm* $\|\cdot\|_{\square}$ *to some hypergraphons* $W := (W_d)_{d \in [D]} \in \times_{d \in [D]} \mathbb{W}_{k_d}$, *i.e.,*

$$\|W_d^N - W_d\|_{\square} \to 0, \quad \forall d \in [D]. \quad (4)$$

A. Finite hypergraph game

In this subsection, we will formulate a dynamical model on hypergraphs where each node is understood as an agent that is
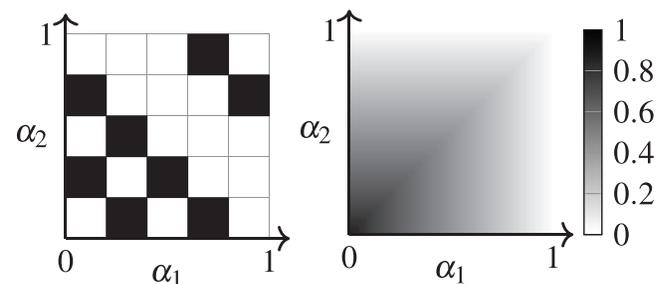
FIG. 2. Visualization of the convergence of two-dimensional step graphons to the uniform attachment graphon $W_{\text{unif}}(\alpha_1, \alpha_2) = 1 - \max(\alpha_1, \alpha_2)$.

influenced by the state distribution of all of its neighbors, according to some time-varying dynamics. Furthermore, each agent is expected to selfishly optimize its own objective, which gives rise to Nash equilibria as the solution of interest.

Consider a $[k_1, \ldots, k_D]$-uniform hypergraph and let $\mathbb{T}$ be the time index set, either $\mathbb{T} = \{0, 1, \ldots, T-1\}$ or $\mathbb{T} = \mathbb{N}_0 := \{0, 1, 2, \ldots\}$. We define N agents $i \in [N]$ each endowed with local states X_t^i and actions U_t^i from a finite state space $\mathbb{X}$ and finite action space $\mathbb{U}$, respectively. Here, $\mathbb{X}$ and $\mathbb{U}$ are assumed finite for technical reasons, though we believe that results could be extended to more general spaces in the future. States have an initial distribution $X_0^i \sim \mu_0 \in \mathscr{P}(\mathbb{X})$. For all times $t \in \mathbb{T}$ and agents $i \in [N]$, their actions are random variables following the law

$$U_t^i \sim \pi_t^i(\cdot | X_t^i), \quad (5)$$

with policy (i.e., probability distribution over actions) $\pi^i \in \Pi := \mathscr{P}(\mathbb{U})^{\mathbb{T} \times \mathbb{X}}$, that, for each node i, depends on the ith state at time t. Then, the states are random variables following the law

$$X_{t+1}^i \sim P_t(\cdot | X_t^i, U_t^i, \nu_t^{N,i}), \quad (6)$$

with transition kernels $P_t \colon \mathbb{X} \times \mathbb{U} \times \mathscr{B}(\mathbb{X}) \to \mathscr{P}(\mathbb{X})$ that, for each node i, depends on the ith state and action at time t, and $\nu_t^{N,i}$. Here, the $\times_{d=1}^D \mathscr{P}(\mathbb{X}^{k_d-1})$-valued multi-layer empirical neighborhood mean field $\nu_t^{N,i}$ is defined as

$$\nu_{t,d}^{N,i} := \frac{1}{N^{k_d-1}} \sum_{\mathbf{m} \in [N]^{k_d-1}} \mathbb{1}_{E^d[H_N]}(\mathbf{m} \cup i) \delta_{\times_{j \neq i} X_t^{m_j}}, \quad (7)$$

in its dth layer, consisting of the unnormalized state distributions of an agent i's neighbors on each layer. In other words, the state dynamics of an agent depend only on the states of nodes in their immediate neighborhood and can be influenced by the agent via its actions U_t^i.

For example, in an epidemics spread scenario, the states of each agent could model their infection status, while the actions of an agent could be to take protective measures. As a result, each agent will randomly become infected with probability depending on how many neighboring agents are infected and whether the agent is taking protective measures.

The cost functions $R_t \colon \mathbb{X} \times \mathbb{U} \times \mathscr{B}(\mathbb{X}) \to \mathbb{R}$ with discount factor $\gamma \in (0, 1)$ or in the finite horizon case $\gamma \in (0, 1]$ define the objective function for the ith agent,

$$J_i^N(\pi^1, \ldots, \pi^N) := \mathbb{E}\left[\sum_{t \in \mathbb{T}} \gamma^t R_t(X_t^i, U_t^i, \nu_t^{N,i}) \right], \quad (8)$$

which can also describe, e.g., random rewards R_t^i that are conditionally independent given $X_t^i, U_t^i, \nu_t^{N,i}$ by the law of total expectation and taking the conditional expectation, $R_t(X_t^i, U_t^i, \nu_t^{N,i}) \equiv \mathbb{E}[R_t^i \mid X_t^i, U_t^i, \nu_t^{N,i}]$.

Our goal is now to find Nash equilibria, i.e., stable policies where no agent can singlehandedly deviate and improve their own objective. Note that finding Nash equilibria in games such as the above is difficult, since (a) even existence of Nash equilibria under the above, decentralized information structure of policies is hard to show, and (b) computation of the Nash equilibria fails due to both curse of dimensionality under full observability and general complexity of computing Nash equilibria[76] (see also Saldi et al.[68] and the discussion therein).

Thus, in the finite game, we are interested in finding the following weaker notion of approximate equilibria,[17,57] where a negligible fraction of agents that remain insignificant to all other agents may remain suboptimal.

Definition 1: An (ε, p)-Nash equilibrium for $\varepsilon, p > 0$ is defined as a tuple of policies $(\pi^1, \ldots, \pi^N) \in \Pi^N$ such that for any $i \in \mathbb{J}^N$, we have

$$J_i^N(\pi^1, \ldots, \pi^N) \geq \sup_{\pi \in \Pi} J_i^N(\pi^1, \ldots, \pi^{i-1}, \pi, \pi^{i+1}, \ldots, \pi^N) - \varepsilon \quad (9)$$

for some set $\mathbb{J}^N \subseteq [N]$ of at least $\lfloor (1-p)N \rfloor$ agents.

While it may seem excessive to reduce to approximate optimality limited to a fraction of the agents, it is always possible under Assumption 1 for a finite number of agents to deviate arbitrarily from the limiting system description. Therefore, under our assumptions, it is only possible to obtain an approximate equilibrium solution for almost all agents via the mean field formulation. Although we could make stronger assumptions on the mode of convergence for hypergraphons, such a concept of convergence would be difficult to motivate from a graph theoretical perspective. Therefore, we restrict ourselves to the cut-norm convergence[75] and the above solution concept.

B. Hypergraphon mean field game

Next, we will formally let $N \to \infty$ and obtain a more tractable, reduced model consisting of any single representative agent and the distribution of agent states, the so-called mean field.

To analyze the case $N \to \infty$, however, we first introduce some preliminary definitions. We define the space of mean fields $\mathbb{M} \subseteq \mathscr{P}(\mathbb{X})^{\mathbb{T} \times \mathbb{I}}$ such that $\boldsymbol{\mu} \in \mathbb{M}$ whenever $\alpha \mapsto \mu_t^\alpha(x)$ is measurable for all $t \in \mathbb{T}, x \in \mathbb{X}$. Intuitively, a mean field is the distribution of states each of the infinitely many agents in $\mathbb{I}$ is in. Analogously, the space of policies $\boldsymbol{\Pi} \subseteq \Pi^{\mathbb{I}}$ is given by policies $\pi \in \Pi^{\mathbb{I}}$ where $\alpha \mapsto \pi_t^\alpha(u \mid x)$ is measurable for any $t \in \mathbb{T}, x \in \mathbb{X}, u \in \mathbb{U}$. Intuitively, $\pi \in \Pi^{\mathbb{I}}$ defines the behavior for each agent $\alpha \in \mathbb{I}$. For any $f \colon \mathbb{X} \times \mathbb{I} \to \mathbb{R}$ and state marginal ensemble $\boldsymbol{\mu} \in \mathscr{P}(\mathbb{X})^{\mathbb{I}}$, define

$$\boldsymbol{\mu}(f) := \int_{\mathbb{I}} \sum_{x \in \mathbb{X}} f(x, \alpha) \mu^\alpha(x) \, \mathrm{d}\alpha.$$

In the limit of $N \to \infty$, assuming that all agents follow a policy $\boldsymbol{\Pi} \subseteq \Pi^{\mathbb{I}}$, we obtain infinitely many agents $\alpha \in \mathbb{I}$, for each of whom we define the limiting hypergraphon mean field dynamics analogously to the finite hypergraph game.

The agent states have the initial distribution $X_0^\alpha \sim \mu_0 \in \mathscr{P}(\mathbb{X})$. For all times $t \in \mathbb{T}$ and agents $\alpha \in \mathbb{I}$, their actions will be random variables following the law

$$U_t^\alpha \sim \pi_t^\alpha(\cdot | X_t^\alpha), \quad (10)$$

under the policy $\pi^\alpha \in \Pi$, while their states follow the law

$$X_{t+1}^\alpha \sim P_t(\cdot | X_t^\alpha, U_t^\alpha, \nu_t^\alpha), \quad (11)$$

with the limiting, now deterministic neighborhood mean field $\nu_t^\alpha \in \times_{d=1}^D \mathscr{P}(\mathbb{X}^{k_d-1})$. Informally, by a law of large numbers, we have

replaced the distribution of finitely many neighbor states by the limiting mean field distribution ν_t^α. The dth component of this mean field is given by

$$\nu_{t,d}^\alpha := \int_{\mathbb{I}^{r<[k_d]\backslash\{1\}}} W_d(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu_t^{\beta_j}(\cdot) \, d\boldsymbol{\beta}, \qquad (12)$$

where for readability, $(\cdot)$ denotes separate coordinates of the input (the order does not matter due to symmetry). In other words, the d-layer neighborhood mean field distributions are functions

$$(x_1, \ldots, x_{k_d-1}) \mapsto \int_{\mathbb{I}^{r<[k_d]\backslash\{1\}}} W_d(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu_t^{\beta_j}(x_j) \, d\boldsymbol{\beta} \qquad (13)$$

that give the probability of random neighbors of a shared hyperedge on layer d to be in states $(x_1, \ldots, x_{k_d-1}) \in \mathbb{X}^{k_d-1}$.

Note that the same, shared $\alpha \in \mathbb{I}$ is used for all D layers, i.e., all layer neighborhood distributions of agents jointly converge to the limiting descriptions ν_t^α. This makes sense, since by Assumption 1, we assume that the agents are already ordered such that the corresponding step-hypergraphons converge to the limiting hypergraphon in cut norm on all layers jointly.

Finally, the objective will be given by

$$J_\alpha^\mu(\pi^\alpha) := \mathbb{E}\left[\sum_{t \in \mathbb{T}} \gamma^t R_t(X_t^\alpha, U_t^\alpha, \nu_t^\alpha) \right], \qquad (14)$$

which leads to the mean field counterpart of Nash equilibria. Informally, a mean field (Nash) equilibrium is given by a "consistent" tuple of policy and mean field, such that the policy is optimal under the mean field and the mean field is generated by the policy. As a result, if all agents follow the policy, they will be optimal under the generated mean field, leading to a Nash equilibrium.

More formally, we define the maps $\Phi \colon \mathbb{M} \to 2^{\boldsymbol{\Pi}}$ mapping from fixed mean field $\boldsymbol{\mu} \in \mathbb{M}$ to all optimal policies $\boldsymbol{\pi} \in \boldsymbol{\Pi} \colon \forall \alpha \in \mathbb{I}$: $\pi^\alpha \in \text{argmax}_{\tilde{\pi}} J_\alpha^\mu(\tilde{\pi})$ and similarly $\Psi \colon \boldsymbol{\Pi} \to \mathbb{M}$ mapping from policy $\boldsymbol{\pi} \in \boldsymbol{\Pi}$ to its induced mean field $\boldsymbol{\mu} \in \mathbb{M}$ such that for all $\alpha \in \mathbb{I}, t \in \mathbb{T}$ we have the initial distribution $\mu_0^\alpha = \mu_0$ and mean field evolution,

$$\mu_{t+1}^\alpha = \int_{\mathbb{X}} \int_{\mathbb{U}} P(x, u, \nu_t^\alpha) \pi_t^\alpha(du \mid x) \mu_t^\alpha(dx). \qquad (15)$$

Definition 2: *A Hypergraphon Mean Field Equilibrium (HMFE) is a pair $(\boldsymbol{\pi}, \boldsymbol{\mu}) \in \boldsymbol{\Pi} \times \mathbb{M}$ such that $\boldsymbol{\pi} \in \Phi(\boldsymbol{\mu})$ and $\boldsymbol{\mu} = \Psi(\boldsymbol{\pi})$.*

Importantly, the mean field game will be motivated rigorously in the following, and its computational complexity is independent of the number of agents. Instead, the complexity of the problem will scale with the size of agent state and action spaces $\mathbb{X}, \mathbb{U}$ and the considered time horizon in case of a finite horizon cost function, since we will solve for equilibria by repeatedly (i) computing optimal policies for discrete Markov decision processes[77] $\pi^\alpha \in \text{argmax}_{\tilde{\pi}} J_\alpha^\mu(\tilde{\pi})$ and (ii) solving the mean field evolution equation (15). In particular, mean field equilibria are guaranteed to exist, and the corresponding equilibrium policy will provide an equilibrium for large finite systems.

To obtain meaningful results, we need a standard continuity assumption (e.g., Ref. 61), since otherwise weak interaction is not guaranteed: Without continuity, a change of behavior in only one of many agents could otherwise cause arbitrarily large changes in the dynamics or rewards.

Assumption 2: *Let R_t, P_t, W each be Lipschitz continuous with Lipschitz constants $L_R, L_P, L_W > 0$.*

Remark 1: *For all but Theorem 1, we may alternatively let W be Lipschitz on finitely many disjoint hyperrectangles, i.e., let there be disjoint intervals $\{I_1, \ldots, I_Q\}$, $\cup_i I_i = \mathbb{I}$ such that $\forall i \in \{1, \ldots, Q\}$, $\forall \alpha, \tilde{\alpha} \in I_i, \forall d \in [D], \forall \boldsymbol{\beta} \in \mathbb{I}^{r<[k_d]\backslash\{1\}}$, we have*

$$|W_d(\alpha, \boldsymbol{\beta}) - W_d(\tilde{\alpha}, \boldsymbol{\beta})| \le L_W |\alpha - \tilde{\alpha}|. \qquad (16)$$

Remark 2: *Note that our model is quite general: In particular, it is also possible to model dynamics and rewards dependent on the state-action distributions instead of only state distributions, replacing $\delta_{\times_{j\neq i} X_t^{m_j}}$ by $\delta_{\times_{j\neq i}(X_t^{m_j}, U_t^{m_j})}$ in (7). This can be done by reformulating any problem as follows. Assume a problem with state and action spaces $\mathbb{X}, \mathbb{U}$ and dependence of rewards and transitions on joint state-action distributions. We can rewrite the problem as a new problem with new state space $\mathbb{X} \cup (\mathbb{X} \times \mathbb{U})$, where in the new problem, each two decision epochs $t, t+1$ correspond to a single original decision epoch, where in the first step t, we transition deterministically from $X_t^{m_j}$ to $\left(X_t^{m_j}, U_t^{m_j}\right)$ for the taken action $U_t^{m_j}$, while in the second step $t+1$, we transition and compute rewards according to the original system, ignoring any second actions taken. Choosing the square root of the discount factor and normalizing rewards will give a problem in our form that is equivalent to the original problem.*

III. THEORETICAL RESULTS

In this section, we rigorously motivate the mean field formulation by providing existence and approximation results of an HMFE. Essentially, HMFEs are guaranteed to exist and will give approximate Nash equilibria in finite hypergraph games with many agents. The reader interested primarily in applications may skip this section.

We lift the empirical distributions and policies to the continuous domain $\mathbb{I}$, i.e., for any $(\pi^1, \ldots, \pi^N) \in \Pi^N$, we define the step policy $\boldsymbol{\pi}^N \in \boldsymbol{\Pi}$ and the step empirical measures $\boldsymbol{\mu}^N \in \mathbb{M}$ by

$$\pi_t^{N,\alpha} := \sum_{i \in [N]} \mathbb{1}_{I_i^N}(\alpha) \cdot \pi_t^i, \quad \forall(\alpha, t) \in \mathbb{I} \times \mathbb{T}, \qquad (17)$$

$$\mu_t^{N,\alpha} := \sum_{i \in [N]} \mathbb{1}_{I_i^N}(\alpha) \cdot \delta_{X_t^i}, \quad \forall(\alpha, t) \in \mathbb{I} \times \mathbb{T}. \qquad (18)$$

Proofs for the results to follow can be found in Appendixes A and B and are at least structurally similar to proofs in Cui and Koeppl,[57] though they contain a number of additional considerations that we highlight in Appendixes A and B.

A. Existence of equilibria

First, we show that there exists an HMFE. We do this by rewriting the problem in a more convenient form as done in Cui and Koeppl.[57] Consider an equivalent, more standard mean field game with states $(\alpha_t, \tilde{X}_t)$, i.e., we integrate the graphon indices α into the

state. The newfound states follow the initial distribution $\tilde{X}_0 \sim \mu_0$, $\alpha_0 \sim \text{Unif}(\mathbb{I})$. Then, the actions and original state transitions follow as before, while the α_t part of the state remains fixed at all times, i.e.,

$$\tilde{U}_t \sim \tilde{\pi}_t(\cdot|\tilde{X}_t, \alpha_t),$$
$$\tilde{X}_{t+1} \sim P_t(\cdot|\tilde{X}_t, \tilde{U}_t, \tilde{\nu}_t), \quad \alpha_{t+1} = \alpha_t, \tag{19}$$

where we used the standard (non-graphical) mean field $\tilde{\mu}_t \in \mathscr{P}(\mathbb{X} \times \mathbb{I})$ (cf. Saldi *et al.*[68]) and let

$$\tilde{\nu}_{t,d} = \int_{\mathbb{I}^{r < [k_d]\backslash\{1\}}} W_d(\alpha_t, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \tilde{\mu}_t(\cdot, \beta_j) \, \mathrm{d}\boldsymbol{\beta}. \tag{20}$$

Using existing results for mean field games,[68] we obtain the existence of a potentially non-unique HMFE.

Theorem 1: *Under Assumption 2, there exists a HMFE $(\boldsymbol{\pi}, \boldsymbol{\mu}) \in \boldsymbol{\Pi} \times \mathbb{M}$.*

For uniqueness results, we refer to existing results such as the classical monotonicity condition.[38,55] However, using existing theory will not analyze the finite hypergraph structure and instead directly uses the limiting hypergraphons. In the following, we, thus, show also that the finite hypergraph games are indeed approximated well.

B. Approximation properties

Next, we will show that the finite hypergraph game and its dynamics are well approximated by the hypergraphon mean field game, which implies that the HMFE solution of the hypergraphon mean field game will give us the desired (ε, p)-Nash equilibrium in large finite hypergraph games.

To begin, we define and obtain finite N-agent system equilibria from an HMFE via the policy sharing map $\text{Id}_N(\boldsymbol{\pi}) := (\pi^1, \ldots, \pi^N) \in \Pi^N$, i.e., Id_N is defined such that each agent will act according to its position α on the hypergraphon,

$$\pi_t^i(u \mid x) := \pi_t^{\frac{i}{N}}(u \mid x), \quad \forall(i, t, x, u) \in [N] \times \mathbb{T} \times \mathbb{X} \times \mathbb{U}. \tag{21}$$

Now consider $(i, \hat{\pi})$-deviated policy tuples where the ith agent deviates from an equilibrium policy tuple to its own policy $\hat{\pi}$, i.e., policy tuples $(\pi^1, \ldots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \ldots, \pi^N)$. Note that this includes the deviation-free case as a special case. In order to obtain a (ε, p)-Nash equilibrium, we must show that for almost all i and policies $\hat{\pi}$, the $(i, \hat{\pi})$-deviated policy tuple will be approximately described by the interaction with the limiting hypergraphon mean field. For this purpose, the first step is to show the convergence of agent state distributions to the mean field.

Define for any $n \in \mathbb{N}$ the evaluation of measurable functions $f: \mathbb{X}^n \times \mathbb{I}^n \to \mathbb{R}$ under any n-dimensional product measures $\otimes^n \boldsymbol{\mu} \in P(\mathbb{X}^n)^{\mathbb{I}^n}$ as

$$\boldsymbol{\mu}(f) := \int_{\mathbb{I}^n} \sum_{x \in \mathbb{X}^n} f(x, \boldsymbol{\beta}) \prod_{i \in [n]} \mu^{\beta_i}(x_i) \, \mathrm{d}\boldsymbol{\beta}, \tag{22}$$

where $\otimes^n \boldsymbol{\mu}$ denotes the n-fold product of the measure $\boldsymbol{\mu}$, i.e., the n-dimensional distribution over agent states.

Then, our first main result is the convergence of the finite-dimensional agent state marginals to the limiting deterministic mean field, given sufficient regularity of the applied policy. For

this purpose, we introduce and optimize over a class $\boldsymbol{\Pi}_{\text{Lip}}$ of Lipschitz-continuous policies up to at most D_π discontinuities, i.e., $\boldsymbol{\pi} \in \boldsymbol{\Pi}_{\text{Lip}}$ whenever $\alpha \mapsto \pi_t^\alpha$ at any time t has at most D_π discontinuities. Note, however, that we could in principle approximate non-Lipschitz policies by classes of Lipschitz-continuous policies.

Theorem 2: *Consider a policy $\boldsymbol{\pi} \in \boldsymbol{\Pi}_{\text{Lip}}$ with associated mean field $\boldsymbol{\mu} = \Psi(\boldsymbol{\pi})$. Let $(\pi^1, \ldots, \pi^N) = \text{Id}_N(\boldsymbol{\pi})$, $\hat{\pi} \in \Pi$, $t \in \mathbb{T}$. Under the policy tuple $(\pi^1, \ldots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \ldots, \pi^N) \in \Pi^N$ and Assumption 1, we have for all finite dimensionalities $n \in \mathbb{N}$ and all measurable functions $f: \mathbb{X}^n \times \mathbb{I}^n \to \mathbb{R}$ uniformly bounded by fixed $M_f > 0$, that*

$$\mathbb{E}\left[\left|\otimes^n \boldsymbol{\mu}_t^N(f) - \otimes^n \boldsymbol{\mu}_t(f)\right|\right] \to 0, \tag{23}$$

uniformly over all possible deviations. Furthermore, the rate of convergence follows the hypergraphon convergence rate in Assumption 1 up to $O(1/\sqrt{N})$.

As a special case, by considering $n = 1$ and $f = \mathbb{1}_{\{x\}}$ for any $x \in \mathbb{X}$, we find convergence in L_1 of the empirical distribution of agent states $\frac{1}{N} \sum_{i \in [N]} \delta_{X_t^i}$ to the limiting mean field $\int_{\mathbb{I}} \mu_t^\alpha \, \mathrm{d}\alpha$.

Our second main result is the (uniform) convergence of the system for almost any agent $i \in [N]$ with deviating policy $\hat{\pi} \in \Pi$ to the system where the interaction with other agents is replaced by the interaction with the limiting deterministic mean field. Hence, we introduce new random variables for the single deviating agent, beginning with initial distribution $\hat{X}_0^{\frac{i}{N}} \sim \mu_0$. The action variables follow the deviating policy

$$\hat{U}_t^{\frac{i}{N}} \sim \hat{\pi}_t\left(\cdot|\hat{X}_t^{\frac{i}{N}}\right), \tag{24}$$

with the state transition laws

$$\hat{X}_{t+1}^{\frac{i}{N}} \sim P_t\left(\cdot|\hat{X}_t^{\frac{i}{N}}, \hat{U}_t^{\frac{i}{N}}, \nu_t^{\frac{i}{N}}\right), \tag{25}$$

i.e., we assume that all other agents act according to their corresponding equilibrium policy $\text{Id}_N(\boldsymbol{\pi})$, such that the neighborhood state distributions of most agents can be replaced by the limiting term $\nu_t^{\frac{i}{N}}$ with little error in large hypergraphs.

Theorem 3: *Consider a policy $\boldsymbol{\pi} \in \boldsymbol{\Pi}_{\text{Lip}}$ with associated mean field $\boldsymbol{\mu} = \Psi(\boldsymbol{\pi})$. Let $(\pi^1, \ldots, \pi^N) = \text{Id}_N(\boldsymbol{\pi})$, $\hat{\pi} \in \Pi$, $t \in \mathbb{T}$. Under the policy tuple $(\pi^1, \ldots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \ldots, \pi^N) \in \Pi^N$ and Assumptions 1 and 2, for any uniformly bounded family of functions $\mathbb{G}$ from $\mathbb{X}$ to $\mathbb{R}$ and any $\varepsilon, p > 0$, $t \in \mathbb{T}$, there exists $N' \in \mathbb{N}$ such that for all $N > N'$,*

$$\sup_{g \in \mathbb{G}} \left|\mathbb{E}\left[g(X_t^i)\right] - \mathbb{E}\left[g\left(\hat{X}_t^{\frac{i}{N}}\right)\right]\right| < \varepsilon \tag{26}$$

uniformly over $\hat{\pi} \in \Pi, i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N], |\mathbb{J}^N| \geq \lfloor(1-p)N\rfloor$.

Further, for any uniformly Lipschitz, uniformly bounded family of measurable functions $\mathbb{H}$ from $\mathbb{X} \times \mathscr{B}(\mathbb{X})$ to $\mathbb{R}$ and any $\varepsilon, p > 0$, $t \in \mathbb{T}$, there exists $N' \in \mathbb{N}$ such that for all $N > N'$,

$$\sup_{h \in \mathbb{H}} \left|\mathbb{E}\left[h\left(X_t^i, \nu_t^{N,i}\right)\right] - \mathbb{E}\left[h\left(\hat{X}_t^{\frac{i}{N}}, \nu_t^{\frac{i}{N}}\right)\right]\right| < \varepsilon \tag{27}$$

uniformly over $\hat{\pi} \in \Pi, i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N]$ with $|\mathbb{J}^N| \geq \lfloor(1-p)N\rfloor$.

As a corollary, we will have a good approximation of the finite hypergraph game objective through the hypergraphon mean field objective, and correspondingly the approximate Nash property of hypergraphon mean field equilibria, motivating the hypergraphon mean field game framework.

Corollary 1: *Consider a policy $\boldsymbol{\pi} \in \boldsymbol{\Pi}_{\mathrm{Lip}}$ with associated mean field $\boldsymbol{\mu} = \Psi(\boldsymbol{\pi})$. Let $(\pi^1, \ldots, \pi^N) = \mathrm{Id}_N(\boldsymbol{\pi})$, $\hat{\pi} \in \Pi$, $t \in \mathbb{T}$. Under the policy tuple $(\pi^1, \ldots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \ldots, \pi^N) \in \Pi^N$ and Assumptions 1 and 2, there exists $N' \in \mathbb{N}$ such that for all $N > N'$, we have*

$$\left| J_i^N \left(\pi^1, \ldots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \ldots, \pi^N \right) - J_{\frac{i}{N}}^{\boldsymbol{\mu}}(\hat{\pi}) \right| < \varepsilon \qquad (28)$$

uniformly over $\hat{\pi} \in \Pi$, $i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N]$ with $|\mathbb{J}^N| \geq \lceil (1-p)N \rceil$.

Corollary 2: *Consider an HMFE $(\boldsymbol{\pi}, \boldsymbol{\mu}) \in \boldsymbol{\Pi}_{\mathrm{Lip}} \times \mathbb{M}$. Under Assumptions 1 and 2, for any $\varepsilon, p > 0$ there exists N' such that for all $N > N'$, the policy $(\pi^1, \ldots, \pi^N) = \mathrm{Id}_N(\boldsymbol{\pi})$ is an (ε, p)-Nash equilibrium.*

Therefore, we find that a solution of the mean field system is a good equilibrium solution of sufficiently large finite hypergraph games.

The assumption of a class $\boldsymbol{\Pi}_{\mathrm{Lip}}$ of Lipschitz-continuous policies up to finitely many discontinuities may seem restrictive. However—similar to Theorem 5 of Ref. 57—we may discretize and partition $\mathbb{I}$ in order to solve hypergraphon mean field games to an arbitrary degree of exactness, preserving the good approximation properties on large hypergraph games.

IV. NUMERICAL EXPERIMENTS

In this section, we shall introduce an exemplary numerical problem of rumor spreading and show associated numerical solutions to demonstrate the hypergraphon mean field framework, verifying the theoretical results.

In order to learn an HMFE in our model, we shall adopt the well-founded discretization method proposed in Cui and Koeppl[57] analogous to the technique used in the proof of Theorem 1 to convert the graphon mean field game into a classical mean field game and thereby allow application of any existing mean field game algorithms such as fixed point iteration to solve for an equilibrium. In other words, we will split $\mathbb{I}$ into subintervals $I_1^N, \ldots, I_N^N$, for each of which we will pick a representing $\alpha \in I_i^N$. This α together with an agent's original state in $\mathbb{X}$ will constitute the new state. In Appendix B, we perform additional experiments for another numerical problem of epidemics control, where existing algorithms fail, pointing out potential future work.

A. Hypergraphons

In our experiments, we shall sample finite hypergraphs directly from given limiting hypergraphons, which should ensure that we obtain hypergraph sequences that fulfill Assumption 1 analogous to the standard graphon case at rate $O\left(\frac{1}{\sqrt{\log N}}\right)$ (see Lemma 10.16 of Ref. 59). To sample a k-uniform hypergraph with N nodes from a k-uniform hypergraphon W, we sample $|r_<[k]|$ uniformly distributed values from the unit interval $\{\alpha_j : \alpha_j \sim \mathrm{Unif}([0,1])\}_{j \in r_<[k]}$. Then, we add any hyperedge $B \subseteq [N]$ with probability $W(\alpha_{r_<(B)})$.

For the sake of illustration, unless otherwise noted, we will in the following consider two-layer hypergraphons, where the first layer is a two-uniform hypergraph (standard graph), while the second layer shall be a three-uniform hypergraph. For the first layer, we consider the uniform attachment graphon,

$$W_{\mathrm{unif}}(\alpha_1, \alpha_2) = 1 - \max(\alpha_1, \alpha_2),$$

the ranked attachment graphon

$$W_{\mathrm{rank}}(\alpha_1, \alpha_2) = 1 - \alpha_1 \alpha_2$$

and the flat (or p-ER) random graphon

$$W_{\mathrm{flat}} := p = 0.5.$$

In particular, the uniform attachment graphon is the limit of a random graph sequence where we iteratively add a new node N and then connect all unconnected nodes with probability $\frac{1}{N}$. Similarly, for the ranked attachment graphon, at each iteration n, we first add a new (nth) node. Before adding the node, the nodes $1, \ldots, n-1$ exist from prior iterations. The new node n is connected to all previous nodes $i = 1, \ldots, n-1$ with probability $1 - \frac{i}{n}$. Then, all other nodes that are not yet connected with each other will connect with probability $\frac{2}{n}$ (see also Chap. 11 of Ref. 59 and Fig. 3).

For the second, three-uniform layer, we similarly consider the hypergraphon resulting from converting all triangles in a standard p-ER graph into hyperedges,[75]

$$\hat{W}_{\mathrm{ind}}(\boldsymbol{\alpha}) := \mathbb{1}_{\mathbb{I}^3 \times [0,p]^3}(\boldsymbol{\alpha})$$

as well as the uniform attachment hypergraphon

$$\hat{W}_{\mathrm{unif}}(\boldsymbol{\alpha}) = 1 - \max(\alpha_1, \alpha_2, \alpha_3)$$

and its inverted version

$$\hat{W}_{\mathrm{inv-unif}}(\boldsymbol{\alpha}) = 1 - \max(1 - \alpha_1, 1 - \alpha_2, 1 - \alpha_3),$$

resulting from a similar construction as in the standard case.

B. Rumor spreading dynamics

In this section, we will describe some simple social dynamics and epidemics problems to illustrate potential applications of hypergraphon mean field games. Here, each layer could model different types of interpersonal relationships. In our particular example of two-uniform and three-uniform layers, the latter can model small
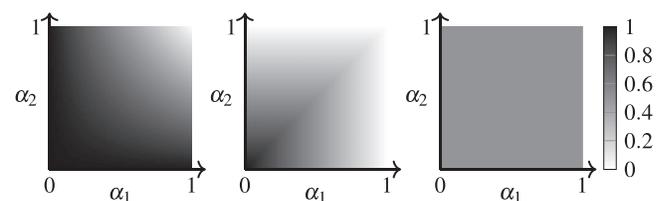
FIG. 3. Visualization of example graphons in the two-dimensional case. Left: Uniform attachment graphon; Middle: Ranked attachment graphon; Right: 0.5-ER graphon.

cliques of friends, while the former could model general acquaintanceship. We do note that social networks are typically more sparse, possessing significantly less edges than on the order of $O(N^2)$. However, our model is a first step toward rigorous limiting hypergraph models and in the future could be extended by using other graph limit theories such as L^p graphons[72–74] by extending their theory toward hypergraphons. We further imagine that similar approaches could be used, e.g., in economics[50] or engineering applications.[51]

In the classical Maki–Thompson model,[78,79] spread of rumors is modeled via three node states: ignorant, spreader, and stifler. Ignorants are unaware of the rumor, while spreaders attempt to spread the rumor. When spreaders attempt to spread to nodes that are already aware of the rumor too often, they stop spreading and become a stifler. In this work, instead of *a priori* assuming the above behavior, we will give agents an intrinsic motivation to spread or stifle rumors, giving rise to the Rumor problem. We shall consider ignorant ($\mathfrak{I}$) and aware ($\mathfrak{A}$) nodes. The behavior of aware nodes is then motivated by the gain and loss of social standing resulting from spreading rumors to ignorant and aware nodes, respectively. The possible actions $\mathbb{U} := \{\mathfrak{S}, \tilde{\mathfrak{S}}\}$ of nodes are to actively spread the rumor ($\mathfrak{S}$) or to refrain from doing so ($\tilde{\mathfrak{S}}$). The probability of an ignorant node becoming aware of the rumor at any decision epoch is then simply given by a linear combination of all layer neighborhood densities of aware, spreading nodes.

Since transition dynamics will depend on the spreading actions of neighbors, following Remark 2 we define instead the extended state space $\mathbb{X} = \{\mathfrak{I}, \mathfrak{A}\} \cup (\{\mathfrak{I}, \mathfrak{A}\} \times \mathbb{U})$. We then assume the dynamics are given at all times by

$$P((x, u) \mid x, u, \cdot) = 1, \quad P(\mathfrak{A} \mid (\mathfrak{A}, \cdot), \cdot, \boldsymbol{v}) = 1,$$

$$P(\mathfrak{A} \mid (\mathfrak{I}, \cdot), \cdot, \boldsymbol{v}) = 1 - P(\mathfrak{I} \mid (\mathfrak{I}, \cdot), \cdot, \boldsymbol{v})$$

$$= \min\left(1, \sum_{d \in [D]} \tau_d \boldsymbol{v}_d \left(\sum_{i \in [k_d]} \mathbb{1}_{\{(\mathfrak{A}, \mathfrak{S})\}}(\cdot_i)\right)\right)$$

for all $x \in \{\mathfrak{I}, \mathfrak{A}\}$, $u \in \mathbb{U}$ and similarly the rewards

$$R((\mathfrak{A}, \mathfrak{S}), \cdot, \boldsymbol{v}) = \sum_{d \in [D]} \boldsymbol{v}_d \left(\sum_{i \in [k_d]} r_d \mathbb{1}_{\{\mathfrak{I}\} \times \mathbb{U}}(\cdot_i) - c_d \mathbb{1}_{\{\mathfrak{A}\} \times \mathbb{U}}(\cdot_i)\right),$$

with $R \equiv 0$ otherwise. In other words, any aware and spreading agent obtains a reward in each layer that is proportional to the probability of a neighbor of any hyperedge sampled uniformly at random out of all connected hyperedges to be ignorant. In our experiments, we use $\tau_1 = 0.3$, $\tau_2 = 0.5$, $r_d = 0.5$, $c_d = 0.8$, $\mu_0(\mathfrak{A}) = 0.01$, and $\mathbb{T} = \{0, 1, \ldots, 49\}$.

C. Numerical results

In our experiments, we restrict ourselves to finite time horizons with $\gamma = 1$, 50 discretization points, and use backward induction with exact forward propagation to compute exact solutions. Note that a simple fixed point iteration by repeatedly computing an arbitrary optimal deterministic policy and its corresponding mean field converges to an equilibrium in the Rumor problem. In general, however, fixed point iteration (as well as more advanced state-of-the-art techniques) may fail to converge [see, e.g., the susceptible-infected-susceptible (SIS) problem in Appendix B].

In Fig. 4, we can observe that the behavior for the rumor problem is as expected. At the equilibrium, agents will continue to spread rumors until the number of aware agents reaches a critical point at which the penalty for spreading to aware agents is larger than the reward for spreading to ignorant agents. The agents with higher connectivity are more likely to be aware of the rumor. Particularly in the uniform attachment hypergraphon case, the threshold is reached at different times, since the neighborhoods of different α reach awareness at different rates depending on their connectivity. Here, a number of nodes with very low degrees will continue spreading the rumors. In Appendix B, we show additional results for inverted 3-uniform hypergraphons, which give similar results to the ones seen here. Furthermore, as can be seen in Fig. 5, the L_1
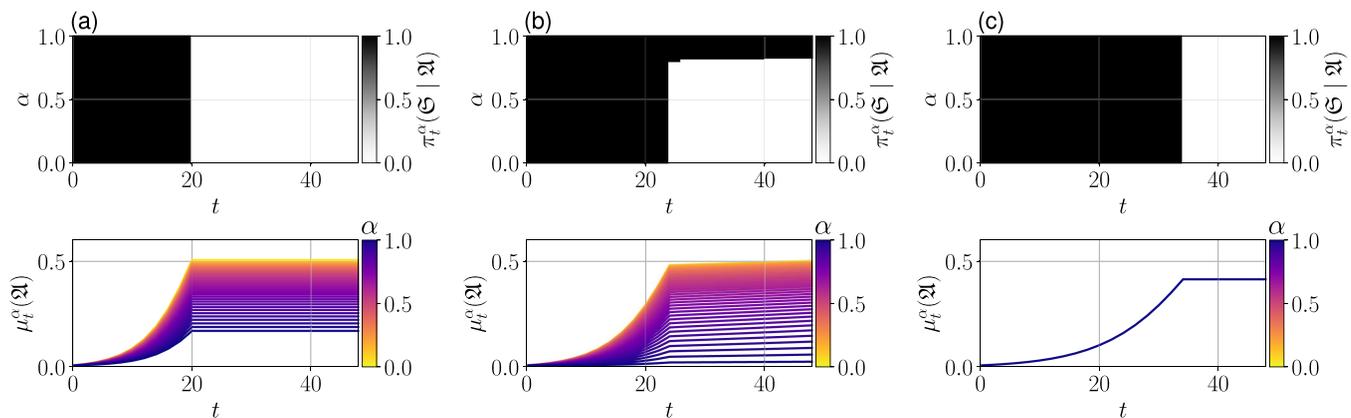
FIG. 4. Analysis of equilibrium behavior for the rumor problem. Top: The threshold policy allows spreading of rumors. It can be seen that agents spread the rumor up until a point in time where too many other agents know of the rumor. As expected, agents are more likely to hear of the rumor if they have more neighbors. (a) $(W_{\text{rank}}, \hat{W}_{\text{unif}})$; (b) $(W_{\text{unif}}, \hat{W}_{\text{unif}})$; (c) $(W_{\text{er}}, \hat{W}_{\text{ind}})$.
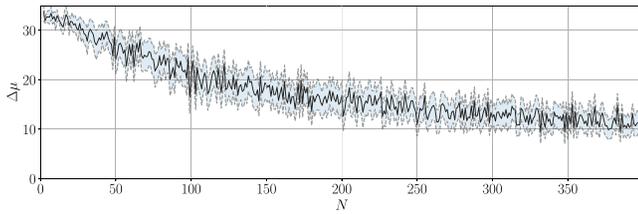
FIG. 5. Comparison between the fraction of aware nodes in the finite and mean field system under the equilibrium policy for (W_{rank}, $\hat{W}_{\text{inv-unif}}$) from Fig. 8(a) in Appendix B, averaged over 50 stochastic simulations. The shaded region depicts the 95% confidence interval at each N. It can be seen that the state distributions are increasingly well approximated by the mean field.

FIG. 7. Analysis of equilibrium behavior for the rumor problem with two-layer three-uniform hypergraphs ($\hat{W}_{\text{block}}$, $\hat{W}_{\text{unif}}$). We observe that the rumor originates in nodes with $\alpha > 0.5$, but nodes with $\alpha \leq 0.5$ eventually catch up due to their increased connectivity. (a) The equilibrium threshold policy. (b) Mean field for each α.

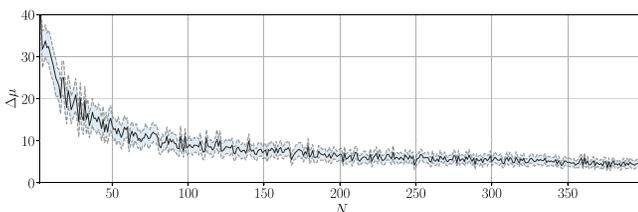error between the empirical distribution and the limiting mean field system (as vectors over time),

$$\Delta\mu = \mathbb{E}\left[\sum_{x\in\mathbb{X}}\sum_{t\in\mathbb{T}}\left|\frac{1}{N}\sum_{i\in[N]}\delta_{X_t^i}(x) - \int_{\mathbb{I}}\mu_t^\alpha(x)\,d\alpha\right|\right] \quad (29)$$

goes to zero as the number of agents increases, showing that the finite hypergraph game is well approximated by the hypergraphon mean field game for sufficiently large systems, though the error remains somewhat large due to the high variance from our sparse initialization $\mu_0(\mathfrak{I}) = 0.01$. Here, we estimated the error $\Delta\mu$ for each N over 50 realizations. Due to the $O(N^2)$ complexity of simulation and computational constraints, our experiments remain limited to the demonstrated number of agents. We repeat the experiment in Fig. 6 with a more dense initialization $\mu_0(\mathfrak{A}) = 0.1$ to reduce the aforementioned high contribution of variance from random initializations. Here, we observe that the resulting convergence is significantly faster.
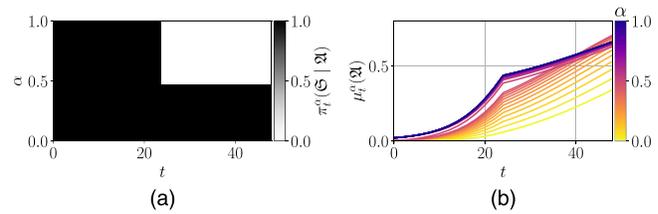
Last, in Fig. 7, we demonstrate some interesting non-linear behavior for a two-layer setting where both layers consist of three-uniform hypergraphs. Here, for the first layer, we use the block hypergraphon

$$\hat{W}_{\text{block}}(\boldsymbol{\alpha}) := \mathbb{1}_{[0,0.5]^3 \times [0,p]^3}(\boldsymbol{\alpha}) + \mathbb{1}_{(0.5,1]^3 \times [0,p]^3}(\boldsymbol{\alpha})$$

for $p = 0.5$, while for the second layer, we again use the inverted uniform attachment hypergraphon. In other words, we have a structure of two blocks on the first layer, while the second layer is more

FIG. 6. Comparison between the fraction of aware nodes in the finite and mean field system under the equilibrium policy for (W_{rank}, $\hat{W}_{\text{inv-unif}}$) as in Fig. 5, but with higher initial awareness. It can be seen that convergence is much faster, since the effect of random sparse initialization is avoided.

globally connected. Furthermore, we will initialize the rumor in the second block where $\alpha > 0.5$, i.e., $\mu_0^\alpha(\mathfrak{A}) = \mathbb{1}_{(0.5,1]}(\alpha)$. As we can see in Fig. 7, in the beginning, the rumor spreads in the second block $\alpha > 0.5$ where it originated from. After a while, however, the rumor begins to spread faster in the first block $\alpha \leq 0.5$, since nodes with low α are significantly more interconnected on the second layer.

Overall, we can see that multi-layer hypergraphon mean field games allow for more complex behavior and modeling of connections than a single-layer graphon approach.

V. CONCLUSION

In this work, we introduced a model for dynamical systems on hypergraphs that can describe agents with weak interaction via the graph structure. The model allows for a rigorous and simple mean field description that has a complexity independent of the number of agents. We verify our approach both theoretically and empirically on a rumor spreading example. By introducing game-theoretical ideas, we, thus, obtain a framework for solving otherwise intractable large-scale games on hypergraphs in a tractable manner.

We hope our work forms the basis for several future works, e.g., extensions to directed or weighted hypergraphs in order to generalize to arbitrary network motifs,[80] adaptive networks,[81] cooperative control, or consideration of edge states in addition to the vertex states we have considered in this work. Furthermore, it may be of interest to consider graph models with more adjustable clustering parameters. An extension of our rumor model and theory to continuous-time models could be fruitful. Finally, so far our work remains restricted to dense graphs and deterministic limiting graphons, while in practice, this is not always the case (e.g., preferential attachment graphs[82]). Here, L^p graphons[72–74] could provide a description for less dense cases, which are of great practical interest. We also hope that our work inspires future applications in inherently (hyper-)graphical scenarios.

in the public, commercial, or not-for-profit sectors. The authors acknowledge the Lichtenberg high performance computing cluster of the TU Darmstadt for providing computational facilities for the calculations of this research.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Kai Cui: Investigation (equal); Writing – original draft (equal). **Wasiur R. KhudaBukhsh:** Supervision (equal); Writing – review & editing (equal). **Heinz Koeppl:** Funding acquisition (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX A: PROOFS

1. Proof of Theorem 1

Proof. Under our assumptions, we can verify Assumption 1 of Ref. 68 for the equivalent standard mean field game given by (19) as in Cui and Koeppl.[57] By Theorem 3.3 of Ref. 68, there exists a mean field equilibrium $(\tilde{\pi}, \tilde{\mu})$ for (19). The policy $\tilde{\mu}$ is α-a.e. optimal under the mean field $\tilde{\mu}$ by Theorem 3.6 of Ref. 68. For all other α, there trivially exists an optimal action, i.e., we can change $\tilde{\pi}$ such that it is optimal for all α. Since the change is on a null set of $\mathbb{I}$, $(\tilde{\pi}, \tilde{\mu})$ remains a mean field equilibrium. Define the hypergraphon mean field policy $\boldsymbol{\pi}$ by $\pi_t^\alpha(u \mid x) = \tilde{\pi}_t(u \mid x, \alpha)$, then $\boldsymbol{\pi}$ is optimal under the hypergraphon mean field $\boldsymbol{\mu}$ where $\boldsymbol{\mu} = \Psi(\boldsymbol{\pi})$, since $\mu_t^\alpha = \tilde{\mu}_t(\cdot, \alpha)$ for almost every α. Finally, both $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ are measurable. Therefore, we have proven existence of the HMFE $(\boldsymbol{\pi}, \boldsymbol{\mu})$. □

2. Proof of Theorem 2

In this section, we provide the full proof of Theorem 2. In contrast to prior work, such as Cui and Koeppl,[57] we (i) extend existing mean field convergence results to n-fold products of the state distributions; and (ii) replace the state distributions by their symmetrized version, in order to obtain convergence results under the generalized cut norm (1). Propagating these changes forward, the rest of the proof is (somewhat) readily generalized and given in the following.

To begin, we introduce some notation to improve readability. Define the D-dimensional neighborhood mean fields $v_W^{\alpha, \boldsymbol{\mu}}$ with dth component

$$v_{W,d}^{\alpha, \boldsymbol{\mu}} := \int_{\mathbb{I}^{r < [k_d]\setminus\{1\}}} W_d(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu^{\beta_j}(\cdot) \, d\boldsymbol{\beta}$$

for all $\boldsymbol{\mu} \in \mathscr{P}(\mathbb{X})^{\mathbb{I}}$, $W := (W_1, \ldots, W_D) \in \times_{d=1}^D \mathbb{W}_{k_d}$ as well as the transition operator $P_W^{t, \boldsymbol{\pi}, \boldsymbol{\mu}} : \mathscr{P}(\mathbb{X})^{\mathbb{I}} \to \mathscr{P}(\mathbb{X})^{\mathbb{I}}$ such that

$$\left(\boldsymbol{\mu}' P_W^{t, \boldsymbol{\pi}, \boldsymbol{\mu}}\right)^\alpha = \sum_{x \in \mathbb{X}} \mu'^\alpha(x) \sum_{u \in \mathbb{U}} \pi^\alpha(u \mid x) P_t\left(\cdot \mid x, u, v_W^{\alpha, \boldsymbol{\mu}}\right)$$

for all $\mu' \in \mathscr{P}(\mathbb{X})^{\mathbb{I}}$, $\boldsymbol{\pi} \in \mathscr{P}(\mathbb{U})^{\mathbb{X}}$, such that, e.g.,

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t P_W^{t, \boldsymbol{\pi}_t, \boldsymbol{\mu}_t}.$$

Proof. In the following, consider arbitrary measurable functions $f : \mathbb{X} \times \mathbb{I} \to [-M_f, M_f]$, $M_f > 0$ and the telescoping sum

$$\mathbb{E}\left[\left|\otimes^n \boldsymbol{\mu}_t^N(f) - \otimes^n \boldsymbol{\mu}_t(f)\right|\right]$$

$$\leq \sum_{i=0}^{n-1} \mathbb{E}\left[\left|\otimes^{n-i} \boldsymbol{\mu}_t^N \otimes \otimes^i \boldsymbol{\mu}_t(f) - \otimes^{n-i-1} \boldsymbol{\mu}_t^N \otimes \otimes^{i+1} \boldsymbol{\mu}_t(f)\right|\right]$$

$$= \sum_{i=0}^{n-1} \mathbb{E}\left[\left|\otimes^{n-i-1} \boldsymbol{\mu}_t^N \otimes \left(\boldsymbol{\mu}_t^N - \boldsymbol{\mu}_t\right) \otimes \otimes^i \boldsymbol{\mu}_t(f)(f)\right|\right]$$

$$= \sum_{i=0}^{n-1} \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x_i \in \mathbb{X}} \int_{\mathbb{I}^{[n]\setminus\{i\}}} \sum_{\tilde{x} \in \mathbb{X}^{[n]\setminus\{i\}}} f(x, (\alpha, \boldsymbol{\beta}))\right.\right.$$
$$\cdot \prod_{j=1}^{n-i-1} \mu^{N, \beta_j}(\tilde{x}_j) \prod_{j=n-i+1}^{n} \mu^{\beta_j}(\tilde{x}_j) \, d\boldsymbol{\beta}$$
$$\left.\left.\cdot \left[\mu^{N,\alpha}(x_i) - \mu^\alpha(x_i)\right] \, d\alpha\right|\right]$$

$$= \sum_{i=0}^{n-1} \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x_i \in \mathbb{X}} g(x_i, \alpha) \left[\mu^{N,\alpha}(x_i) - \mu^\alpha(x_i)\right] \, d\alpha\right|\right],$$

where we defined $g : \mathbb{X} \times \mathbb{I} \to [-M_f, M_f]$ as

$$g(x, \alpha) := \int_{\mathbb{I}^{[n]\setminus\{i\}}} \sum_{x_{-i} \in \mathbb{X}^{[n]\setminus\{i\}}} f((x, x_{-i}), (\alpha, \boldsymbol{\beta}))$$
$$\cdot \prod_{j=1}^{n-i-1} \mu^{N, \beta_j}(x_j) \prod_{j=n-i+1}^{n} \mu^{\beta_j}(x_j) \, d\boldsymbol{\beta}.$$

Since g is a measurable function bounded by M_f, due to the prequel it suffices at any time $t \in \mathbb{T}$ to prove (23) for $n = 1$, which will imply the statement for all $n \in \mathbb{N}$.

The proof is by induction over t for $n = 1$. At $t = 0$,

$$
\mathbb{E}\left[\left|\boldsymbol{\mu}_0^N(f) - \boldsymbol{\mu}_0(f)\right|\right]
$$

$$
= \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x \in \mathbb{X}} \mu_0^{N,\alpha}(x) f(x,\alpha) - \sum_{x \in \mathbb{X}} \mu_0^\alpha(x) f(x,\alpha) \, d\alpha\right|\right]
$$

$$
= \mathbb{E}\left[\left|\sum_{i \in [N]} \left(\int_{I_i^N} f(X_0^i, \alpha) \, d\alpha - \mathbb{E}\left[\int_{I_i^N} f(X_0^i, \alpha) \, d\alpha\right]\right)\right|\right]
$$

$$
\leq \left(\mathbb{V}\left[\sum_{i \in [N]} \int_{I_i^N} f(X_0^i, \alpha) \, d\alpha\right]\right)^{\frac{1}{2}}
$$

$$
= \left(\sum_{i \in [N]} \mathbb{V}\left[\int_{I_i^N} f(X_0^i, \alpha) \, d\alpha\right]\right)^{\frac{1}{2}} \leq \frac{4M_f}{\sqrt{N}}
$$

by i.i.d. $X_0^i \sim \mu_0 = \mu_0^\alpha$ and $\mathbb{V}\left[\int_{I_i^N} f(X_0^i, \alpha) \, d\alpha\right] \leq \left(\frac{4M_f}{N}\right)^2$.

Assume that (23) holds at $t \in \mathbb{T}$. Then, at time $t + 1$, we have

$$
\mathbb{E}\left[\left|\boldsymbol{\mu}_{t+1}^N(f) - \boldsymbol{\mu}_{t+1}(f)\right|\right]
$$

$$
\leq \mathbb{E}\left[\left|\boldsymbol{\mu}_{t+1}^N(f) - \boldsymbol{\mu}_t^N P_{W^N}^{t,\pi_t^N,\mu_t^N}(f)\right|\right]
$$

$$
+ \mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_{W^N}^{t,\pi_t^N,\mu_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\pi_t^N,\mu_t^N}(f)\right|\right]
$$

$$
+ \mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\pi_t^N,\mu_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\pi_t,\mu_t^N}(f)\right|\right]
$$

$$
+ \mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\pi_t,\mu_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\pi_t,\mu_t}(f)\right|\right]
$$

$$
+ \mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\pi_t,\mu_t}(f) - \boldsymbol{\mu}_{t+1}(f)\right|\right]
$$

and in the following, we will analyze each term.

For the first term, observe first that by definition,

$$
\int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} W_d^N(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu_t^{N,\beta_j}(\cdot) \, d\boldsymbol{\beta}
$$

$$
= \frac{1}{N^{k_d-1}} \sum_{\mathbf{m} \in [N]^{k_d-1}} \mathbb{1}_{E^d[H_N]}(\mathbf{m} \cup i) \delta_{\times_{j \neq i} X_t^{m_j}},
$$

and, therefore,

$$
\boldsymbol{\mu}_t^N P_{W^N}^{t,\pi_t^N,\mu_t^N}(f) = \mathbb{E}\left[\int_{I_i^N} f(X_{t+1}^i, \alpha) \, d\alpha \,\middle|\, \mathbf{X}_t\right]
$$

such that we again obtain

$$
\mathbb{E}\left[\left|\boldsymbol{\mu}_{t+1}^N(f) - \boldsymbol{\mu}_t^N P_{W^N}^{t,\pi_t^N,\mu_t^N}(f)\right|\right]
$$

$$
= \mathbb{E}\left[\left|\sum_{i \in [N]} \left(g(X_{t+1}^i) - \mathbb{E}\left[g(X_{t+1}^i) \mid \mathbf{X}_t\right]\right)\right|\right]
$$

$$
\leq \left(\mathbb{E}\left[\left(\sum_{i \in [N]} \left(g(X_{t+1}^i) - \mathbb{E}\left[g(X_{t+1}^i) \mid \mathbf{X}_t\right]\right)\right)^2\right]\right)^{\frac{1}{2}}
$$

$$
= \left(\sum_{i \in [N]} \mathbb{E}\left[\left(g(X_{t+1}^i) - \mathbb{E}\left[g(X_{t+1}^i) \mid \mathbf{X}_t\right]\right)^2\right]\right)^{\frac{1}{2}} \leq \frac{4M_f}{\sqrt{N}},
$$

where $g(x) := \int_{I_i^N} f(x,\alpha) \, d\alpha$, $|g| \leq \frac{M_f}{N}$, by using the law of total expectation and conditional independence of $\{X_{t+1}^i\}_{i \in [N]}$ given $\mathbf{X}_t := \{X_t^i\}_{i \in [N]}$.

For the second term, first note that we can replace the distributional terms by their symmetrized version: For any $k \in \mathbb{N}$, any $W \in \text{Sym}_<^{\text{ind}}[k]$ and any step empirical measure or mean field $\boldsymbol{\mu} \in \mathscr{P}(\mathbb{X})^{\mathbb{I}}$, we have by symmetry that the associated neighborhood probabilities are invariant to all permutations $\sigma \in \text{Sym}([k-1])$ of states $\mathbb{X}^{k-1}$, i.e., for any $x \in \mathbb{X}^{k-1}, \alpha \in \mathbb{I}$

$$
\int_{\mathbb{I}^{r<[k]\setminus\{1\}}} W(\alpha, \beta) \prod_{i=1}^{k-1} \mu^{\beta_i}(x_i)) \, d\beta
$$

$$
= \frac{1}{(k-1)!} \sum_{\sigma \in \text{Sym}([k-1])} \int_{\mathbb{I}^{r<[k]\setminus\{1\}}} W(\alpha, \beta) \prod_{i=1}^{k-1} \mu^{\beta_i}(x_{\sigma(i)})) \, d\beta
$$

$$
= \int_{\mathbb{I}^{r<[k]\setminus\{1\}}} W(\alpha, \beta) \underbrace{\frac{1}{(k-1)!} \sum_{\sigma \in \text{Sym}([k-1])} \prod_{i=1}^{k-1} \mu^{\beta_i}(x_{\sigma(i)}))}_{u_1 \in \text{Sym}_\leq^{\text{ind}}[k-1]} \, d\beta,
$$

and, hence, Assumption 1 implies that

$$
\int_{\mathbb{I}} \left|\int_{\mathbb{I}^{r<[k]\setminus\{1\}}} W(\alpha, \beta) u_1\left(\beta_{[k]\setminus\{1\}}\right) \, d\beta\right| \, d\alpha
$$

$$
\leq \int_{\mathbb{I}[k]} \left|\int_{\mathbb{I}^{r<[k]\setminus[k]}} W(\alpha, \beta) u_1\left(\alpha_{[k]\setminus\{1\}}\right) \, d\beta\right| \, d\alpha
$$

$$
\leq \sup_{\substack{u_1: \mathbb{I}^{r[k-1]} \to \mathbb{I}, \\ u_1 \in \text{Sym}_\leq^{\text{ind}}[k-1]}} \int_{\mathbb{I}[k]} \left|\int_{\mathbb{I}^{r<[k]\setminus[k]}} W(\alpha, \beta) \, d\beta u_1\left(\alpha_{[k]\setminus\{1\}}\right)\right| \, d\alpha
$$

$$
= \sup_{\substack{u_1: \mathbb{I}^{r[k-1]} \to \mathbb{I}, \\ u_1 \in \text{Sym}_\leq^{\text{ind}}[k-1]}} \int_{\mathbb{I}[k]} \int_{\mathbb{I}^{r<[k]\setminus[k]}} W(\alpha, \beta) \, d\beta u_1\left(\alpha_{[k]\setminus\{1\}}\right) \, d\alpha
$$

$$
\leq \|W\|_{\square^{k-1}} \to 0
$$

for any $x \in \mathbb{X}^{k-1}$, $\alpha \in \mathbb{I}$ by letting $u_1(\alpha_{r_{<}([k]\setminus\{i\})}) := \frac{1}{(k-1)!}$ $\sum_{\sigma \in \mathrm{Sym}([k-1])} \prod_{i=1}^{k-1} \mu^{\beta_i}(x_{\sigma(i)}))$ in (1). Therefore,

$$\mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_{W^N}^{t,\boldsymbol{\pi}_t^N,\boldsymbol{\mu}_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}^{N_t},\boldsymbol{\mu}_t^N}(f)\right|\right]$$

$$= \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x \in \mathbb{X}} \mu_t^{N,\alpha}(x) \sum_{u \in \mathbb{U}} \pi_t^{N,\alpha}(u \mid x) \sum_{x' \in \mathbb{X}} f(x',\alpha)\right.\right.$$

$$\left.\left. \cdot \left[P_t\left(x' \mid x, u, v_{W^N}^{\alpha,\boldsymbol{\mu}_t^N}\right) - P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t^N}\right)\right] \mathrm{d}\alpha\right|\right]$$

$$\leq |\mathbb{X}| M_f L_P \, \mathbb{E}\left[\int_{\mathbb{I}} \left\|v_{W^N}^{\alpha,\boldsymbol{\mu}_t^N} - v_W^{\alpha,\boldsymbol{\mu}_t^N}\right\| \mathrm{d}\alpha\right]$$

$$\leq |\mathbb{X}| M_f L_P \, \mathbb{E}\left[\int_{\mathbb{I}} \sum_{d \in [D]} \sum_{x \in \mathbb{X}^{k_d-1}} \left|\int_{\mathbb{I}^{r_{<}[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{N,\beta_j}(x_j)\right)\right.\right.$$

$$\left.\left. \cdot \left[W_d^N(\alpha,\boldsymbol{\beta}) - W_d(\alpha,\boldsymbol{\beta})\right] \mathrm{d}\boldsymbol{\beta}\right| \mathrm{d}\alpha\right]$$

$$\leq |\mathbb{X}| M_f L_P |\mathbb{X}| \sum_{d \in [D]} \|W_d^N - W_d\|_{\square^{k_d-1}} \to 0$$

by Assumption 1, and at rate $O(1/\sqrt{N})$ if (4) converges at rate $O(1/\sqrt{N})$.

For the third term, we have

$$\mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}_t^N,\boldsymbol{\mu}_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}_t,\boldsymbol{\mu}_t^N}(f)\right|\right]$$

$$= \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x \in \mathbb{X}} \mu_t^{N,\alpha}(x) \sum_{u \in \mathbb{U}} \left[\pi_t^{N,\alpha}(u \mid x) - \pi_t^\alpha(u \mid x)\right]\right.\right.$$

$$\left.\left. \cdot \sum_{x' \in \mathbb{X}} P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t^N}\right) f(x',\alpha) \mathrm{d}\alpha\right|\right]$$

$$\leq M_f \, \mathbb{E}\left[\int_{\mathbb{I}} \left|\pi_t^{N,\alpha}(u \mid x) - \pi_t^\alpha(u \mid x)\right| \mathrm{d}\alpha\right]$$

$$= M_f \, \mathbb{E}\left[\sum_{j \in [N]\setminus\{i\}} \int_{I_j^N} \left|\pi_t^{\frac{[N\alpha]}{N}}(u \mid x) - \pi_t^\alpha(u \mid x)\right| \mathrm{d}\alpha\right]$$

$$+ M_f \, \mathbb{E}\left[\int_{I_i^N} \left|\hat{\pi}_t(u \mid x) - \pi_t^\alpha(u \mid x)\right| \mathrm{d}\alpha\right]$$

$$\leq M_f \cdot \frac{L_\pi}{N} + M_f \cdot \frac{2|D_\pi|}{N} + M_f \cdot \frac{2}{N}$$

by $\boldsymbol{\pi} \in \boldsymbol{\Pi}_{\mathrm{Lip}}$ with Lipschitz constant L_π and up to D_π discontinuities, where we bound the integrands by 2.

For the fourth term, we find that

$$\mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}_t,\boldsymbol{\mu}_t^N}(f) - \boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}_t,\boldsymbol{\mu}_t}(f)\right|\right]$$

$$= \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x \in \mathbb{X}} \mu_t^{N,\alpha}(x) \sum_{u \in \mathbb{U}} \pi_t^\alpha(u \mid x) \sum_{x' \in \mathbb{X}} f(x',\alpha)\right.\right.$$

$$\left.\left. \cdot \left[P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t^N}\right) - P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t}\right)\right] \mathrm{d}\alpha\right|\right]$$

$$\leq |\mathbb{X}| M_f L_P \, \mathbb{E}\left[\int_{\mathbb{I}} \left\|v_W^{\alpha,\boldsymbol{\mu}_t^N} - v_W^{\alpha,\boldsymbol{\mu}_t}\right\| \mathrm{d}\alpha\right]$$

$$\leq |\mathbb{X}| M_f L_P \, \mathbb{E}\left[\int_{\mathbb{I}} \sum_{d \in [D]} \sum_{x \in \mathbb{X}^{k_d-1}} \left|\int_{\mathbb{I}^{r_{<}[k_d]\setminus\{1\}}} W_d(\alpha,\boldsymbol{\beta})\right.\right.$$

$$\left.\left. \cdot \left[\prod_{j=1}^{k_d-1} \mu^{N,\beta_j}(x_j) - \prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j)\right] \mathrm{d}\boldsymbol{\beta}\right| \mathrm{d}\alpha\right]$$

$$= |\mathbb{X}| M_f L_P \int_{\mathbb{I}} \sum_{d \in [D]} \sum_{x \in \mathbb{X}^{k_d-1}} \mathbb{E}\left[\left|\int_{\mathbb{I}^{[k_d]\setminus\{1\}}} \int W_d(\alpha,\boldsymbol{\beta},\boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\zeta}\right.\right.$$

$$\left.\left. \cdot \left[\prod_{j=1}^{k_d-1} \mu^{N,\beta_j}(x_j) - \prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j)\right] \mathrm{d}\boldsymbol{\beta}\right|\right] \mathrm{d}\alpha$$

$$= |\mathbb{X}| M_f L_P \int_{\mathbb{I}} \sum_{d \in [D]} \sum_{x \in \mathbb{X}^{k_d-1}}$$

$$\mathbb{E}\left[\left|\otimes^{k_d-1} \boldsymbol{\mu}_t^N(f_{x,\alpha}) - \otimes^{k_d-1} \boldsymbol{\mu}_t(f_{x,\alpha})\right|\right] \mathrm{d}\alpha \to 0$$

at the rate in the induction assumption, by applying the induction assumption (23) for $n = k_d - 1$ to the functions

$$f_{x,\alpha}(x',\boldsymbol{\beta}) = \int_{\mathbb{I}^{r_{<}[k_d]\setminus[k_d]}} W_d(\alpha,\boldsymbol{\beta},\boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\zeta} \cdot \mathbb{1}_{\{x\}}(x')$$

for any $(x,\alpha) \in \mathbb{X}^{k_d-1} \times \mathbb{I}$.

For the fifth term, we analogously obtain

$$\mathbb{E}\left[\left|\boldsymbol{\mu}_t^N P_W^{t,\boldsymbol{\pi}_t,\boldsymbol{\mu}_t}(f) - \boldsymbol{\mu}_t P_W^{t,\boldsymbol{\pi}_t,\boldsymbol{\mu}_t}(f)\right|\right]$$

$$= \mathbb{E}\left[\left|\int_{\mathbb{I}} \sum_{x \in \mathbb{X}} \left[\mu_t^{N,\alpha}(x) - \mu_t^\alpha(x)\right] \sum_{u \in \mathbb{U}} \pi_t^\alpha(u \mid x)\right.\right.$$

$$\left.\left. \cdot \sum_{x' \in \mathbb{X}} P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t}\right) f(x',\alpha) \mathrm{d}\alpha\right|\right]$$

$$= \mathbb{E}\left[\left|\boldsymbol{\mu}_t^N(f') - \boldsymbol{\mu}_t(f')\right|\right] \to 0$$

at the rate in the induction assumption, by applying the induction assumption (23) to

$$f'(x,\alpha) = \sum_{u \in \mathbb{U}} \pi_t^\alpha(u \mid x) \sum_{x' \in \mathbb{X}} P_t\left(x' \mid x, u, v_W^{\alpha,\boldsymbol{\mu}_t}\right) f(x',\alpha).$$

This concludes the proof by induction. $\qquad \square$

3. Proof of Theorem 3

The proof of Theorem 3 mirrors the proof in Ref. 57 apart from propagating the multidimensional convergence results forward, and we give the entire proof for completeness and convenience. Again, we introduce some notation to improve readability. For any $\alpha \in \mathbb{I}$, $d \in [D]$, define maps $\nu_d^\alpha: \mathscr{P}(\mathbb{X})^{\mathbb{I}} \to \mathscr{P}(\mathbb{X})$ and $\nu_{N,d}^\alpha: \mathscr{P}(\mathbb{X})^{\mathbb{I}} \to \mathscr{P}(\mathbb{X})$ as

$$\nu_d^\alpha(\boldsymbol{\mu}) := \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} W_d(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu^{\beta_j}(\cdot) \, \mathrm{d}\boldsymbol{\beta},$$

$$\nu_{N,d}^\alpha(\boldsymbol{\mu}) := \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} W_d^N(\alpha, \boldsymbol{\beta}) \prod_{j=1}^{k_d-1} \mu^{\beta_j}(\cdot) \, \mathrm{d}\boldsymbol{\beta},$$

with D-dimensional shorthands

$$\nu^\alpha(\boldsymbol{\mu}) := (\nu_d^\alpha(\boldsymbol{\mu}))_{d\in[D]},$$

$$\nu_N^\alpha(\boldsymbol{\mu}) := (\nu_{N,d}^\alpha(\boldsymbol{\mu}))_{d\in[D]},$$

such that by definition $\nu_t^\alpha = \nu^\alpha(\boldsymbol{\mu}_t)$ and $\nu_t^{N,i} = \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N)$.

Proof. To begin, we prove (26) $\implies$ (27) at any fixed time t. Define the uniform bound M_h and uniform Lipschitz constant L_h of functions in $\mathbb{H}$. For any $h \in \mathbb{H}$, we have

$$\left| \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N)) \right] - \mathbb{E}\left[h(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|$$

$$= \left| \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N)) \right] - \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|$$

$$+ \left| \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] - \mathbb{E}\left[h(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|$$

$$+ \left| \mathbb{E}\left[h(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] - \mathbb{E}\left[h(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|,$$

which we will analyze as $N \to \infty$.

For the first term, we obtain

$$\left| \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N)) \right] - \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|$$

$$\leq \mathbb{E}\left[\mathbb{E}\left[\left| h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N)) - h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right| \mid X_t^i \right] \right]$$

$$\leq L_h \mathbb{E}\left[\left\| \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N) - \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right\| \right]$$

$$= L_h \mathbb{E}\left[\sum_{d\in[D]} \sum_{x\in\mathbb{X}^{k_d-1}} \left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} W_d^N(\alpha, \boldsymbol{\beta}) \right. \right.$$

$$\left. \left. \cdot \left[\prod_{j=1}^{k_d-1} \mu^{N,\beta_j}(x_j) - \prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right] \mathrm{d}\boldsymbol{\beta} \right| \right] \to 0$$

uniformly by applying Theorem 2 to the functions

$$f_{N,i,x}(x', \boldsymbol{\beta}) = \int_{\mathbb{I}^{r<[k_d]\setminus[k_d]}} W_d^N(\frac{i}{N}, \boldsymbol{\beta}, \boldsymbol{\zeta}) \, \mathrm{d}\boldsymbol{\zeta} \cdot \mathbb{1}_{\{x\}}(x').$$

For the second term, we analogously have

$$\left| \mathbb{E}\left[h(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] - \mathbb{E}\left[h(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t)) \right] \right|$$

$$\leq L_h \| \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t) - \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \|_1$$

$$\leq L_h \sum_{d\in[D]} \sum_{x\in\mathbb{X}^{k_d-1}} \left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d^N(\frac{i}{N}, \boldsymbol{\beta}) - W_d(\frac{i}{N}, \boldsymbol{\beta}) \right] \mathrm{d}\boldsymbol{\beta} \right|$$

$$\leq L_h \sum_{d\in[D]} \sum_{x\in\mathbb{X}^{k_d-1}} \left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d^N(\frac{i}{N}, \boldsymbol{\beta}) - N \int_{I_i^N} W_d(\alpha, \boldsymbol{\beta}) \, \mathrm{d}\alpha \right] \mathrm{d}\boldsymbol{\beta} \right|$$

$$+ L_h \sum_{d\in[D]} \sum_{x\in\mathbb{X}^{k_d-1}} \left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[N \int_{I_i^N} W_d(\alpha, \boldsymbol{\beta}) \, \mathrm{d}\alpha - W_d(\frac{i}{N}, \boldsymbol{\beta}) \right] \mathrm{d}\boldsymbol{\beta} \right|,$$

where for the former (finite) sum, we have

$$\left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d^N(\frac{i}{N}, \boldsymbol{\beta}) - N \int_{I_i^N} W_d(\alpha, \boldsymbol{\beta}) \, \mathrm{d}\alpha \right] \mathrm{d}\boldsymbol{\beta} \right|$$

$$= \left| N \int_{I_i^N} \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d^N(\alpha, \boldsymbol{\beta}) - W_d(\alpha, \boldsymbol{\beta}) \right] \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha \right|$$

$$\leq N \int_{I_i^N} \left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d^N(\alpha, \boldsymbol{\beta}) - W_d(\alpha, \boldsymbol{\beta}) \right] \mathrm{d}\boldsymbol{\beta} \right| \mathrm{d}\alpha$$

$$=: I_i^N,$$

since by definition of the step-hypergraphon, $W_d^N(\frac{i}{N}, \boldsymbol{\beta}) = W_d^N(\alpha, \boldsymbol{\beta})$ over $\alpha \in I_i^N$. Therefore,

$$\frac{1}{N} \sum_{i=1}^N I_i^N = \int_{\mathbb{I}^{r<[k_d]}} \left[W_d^N(\beta) - W_d(\beta) \right] \prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \, \mathrm{d}\beta \to 0$$

as in the proof of Theorem 2 by Assumption 1. Fix $\varepsilon, p > 0$. As N becomes sufficiently large, there must exist $\mathbb{J}_1^N, |\mathbb{J}_1^N| \geq \lfloor (1-p)N \rfloor$ such that

$$I_i^N < \varepsilon, \quad \forall i \in \mathbb{J}_1^N.$$

We prove this by contradiction: Assume there does not exist such $\mathbb{J}_1^N$, then there exist at least $\lceil pN \rceil$ agents where $I_i^N \geq \varepsilon$. Since $I_i^N \geq 0$, it follows that $\frac{1}{N} \sum_{i=1}^N I_i^N \geq \frac{1}{N} \lceil pN \rceil \varepsilon \geq \varepsilon p$, which contradicts the convergence to zero of $\frac{1}{N} \sum_{i=1}^N I_i^N$. Repeating the argument for each $d \in [D], x \in \mathbb{X}^{k_d-1}$ bounds the first sum.

For the latter (finite) sum, we have

$$\left| \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[N \int_{I_i^N} W_d(\alpha, \beta) \, \mathrm{d}\alpha - W_d\left(\frac{i}{N}, \boldsymbol{\beta} \right) \right] \mathrm{d}\beta \right|$$

$$= \left| N \int_{I_i^N} \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left(\prod_{j=1}^{k_d-1} \mu^{\beta_j}(x_j) \right) \right.$$

$$\left. \cdot \left[W_d(\alpha, \boldsymbol{\beta}) - W_d\left(\frac{\lceil N\alpha \rceil}{N}, \beta \right) \right] \mathrm{d}\beta \, \mathrm{d}\alpha \right|$$

$$\leq N \int_{I_i^N} \int_{\mathbb{I}^{r<[k_d]\setminus\{1\}}} \left| W_d(\alpha, \boldsymbol{\beta}) - W_d\left(\frac{\lceil N\alpha \rceil}{N}, \beta \right) \right| \mathrm{d}\beta \, \mathrm{d}\alpha$$

$$\leq N \frac{1}{N} \cdot \frac{L_W}{N} = \frac{L_W}{N} \to 0$$

by Assumption 2. Alternatively, under only block-wise Lipschitz W as in (16), the same result is obtained by first separating out finitely many i (at most $Q - 1$) for which Lipschitzness fails, trivially bounding their terms by $\frac{2(Q-1)}{N}$. For all other i, there exists $k \in \{1, \dots, Q\}$ such that $I_i^N \times I_j \subseteq I_k \times I_j$, i.e., the Lipschitz bound applies.

For the third term, again fix $\varepsilon, p > 0$. Then, by our initial assumption of (26), for sufficiently large N, there exists a set $\mathbb{J}_2^N$, $|\mathbb{J}_2^N| \geq \lfloor (1-p)N \rfloor$ such that

$$\left| \mathbb{E}\left[h\left(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] - \mathbb{E}\left[h\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right| < \varepsilon, \quad \forall i \in \mathbb{J}_2^N$$

independent of $\hat{\pi} \in \Pi$.

This completes the proof of (26) $\implies$ (27) at any time t, since by the prequel, the intersection of all correspondingly chosen, finitely many sets $\mathbb{J}_i^N$ for sufficiently large N has at least $N - \sum_i \lceil p_i N \rceil$ elements, which is always larger than $N - \lceil pN \rceil$ for any $p > 0$ by choosing p_i sufficiently small.

Finally, we show (26) at all times t using the prequel by induction, which will imply (27). By definition for $t = 0$, $\hat{X}_t^{\frac{i}{N}} \sim \mu_0$ and $X_t^i \sim \mu_0$ imply

$$\left| \mathbb{E}\left[g(X_0^i) \right] - \mathbb{E}\left[g\left(\hat{X}_0^{\frac{i}{N}} \right) \right] \right| = 0.$$

For the induction step, define the uniform bound M_g of functions in $\mathbb{G}$. Observe that

$$\left| \mathbb{E}\left[g\left(X_{t+1}^i \right) \right] - \mathbb{E}\left[g\left(\hat{X}_{t+1}^{\frac{i}{N}} \right) \right] \right|$$

$$= \left| \mathbb{E}\left[l_{N,t}\left(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N) \right) \right] - \mathbb{E}\left[l_{N,t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right|$$

using the uniformly bounded and Lipschitz functions

$$l_{N,t}(x, \nu) := \sum_{u \in \mathbb{U}} \hat{\pi}_t(u \mid x) \sum_{x' \in \mathbb{X}} P_t(x' \mid x, u, \nu) g(x'),$$

with bound M_g and Lipschitz constant $|\mathbb{X}|M_g L_P$. By induction assumption (26) and (26) $\implies$ (27), there exists $N' \in \mathbb{N}$ such that for all $N > N'$, we have

$$\left| \mathbb{E}\left[l_{N,t}\left(X_t^i, \nu_N^{\frac{i}{N}}(\boldsymbol{\mu}_t^N) \right) \right] - \mathbb{E}\left[l_{N,t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right| < \varepsilon$$

uniformly over $\hat{\pi} \in \Pi, i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N]$ with $|\mathbb{J}^N| \geq \lfloor (1-p)N \rfloor$. This concludes the proof by induction. $\square$

4. Proof of Corollary 1

Proof. The result follows more or less directly from Theorem 3. Consider first the finite horizon case $\mathbb{T} = \{0, 1, \dots, T-1\}$. Define

$$R_t^{\hat{\pi}}(x, \nu) := \sum_{u \in \mathbb{U}} R_t(x, u, \nu) \hat{\pi}_t(u \mid s),$$

with uniform bound M_R and Lipschitz constant $|U|L_R$. Therefore, by choosing the maximum over all N' for all finitely many times $t \in \mathbb{T}$ via Theorem 3, there exists $N' \in \mathbb{N}$ such that for all $N > N'$, we have

$$\left| J_i^N\left(\pi^1, \dots, \pi^{i-1}, \hat{\pi}, \pi^{i+1}, \dots, \hat{\pi} \right) - J_{\frac{i}{N}}^{\boldsymbol{\mu}}(\hat{\pi}) \right|$$

$$\leq \sum_{t=0}^{T-1} \left| \mathbb{E}\left[R_t^{\hat{\pi}_t}\left(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] - \mathbb{E}\left[R_t^{\hat{\pi}_t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right| < \varepsilon$$

uniformly over $\hat{\pi} \in \Pi, i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N]$ with $|\mathbb{J}^N| \geq \lfloor (1-p)N \rfloor$.
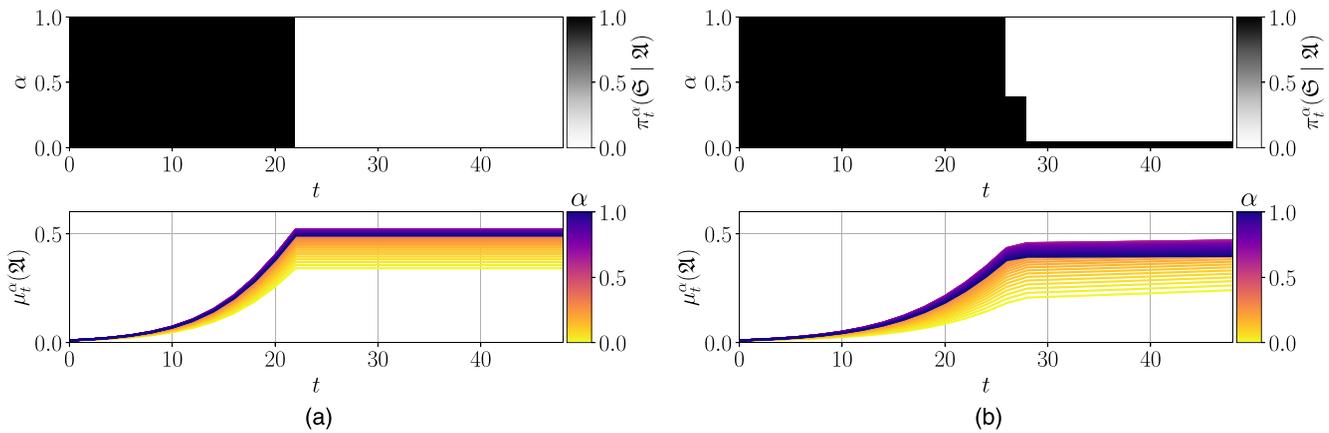
FIG. 8. Analysis of equilibrium behavior for the rumor problem on additional hypergraphon configurations. (a) $(W_{\text{rank}}, \hat{W}_{\text{inv-unif}})$; (b) $(W_{\text{unif}}, \hat{W}_{\text{inv-unif}})$.

For the infinite horizon problem $\mathbb{T} = \mathbb{N}_0$, we first pick some time $T' > \frac{\log \frac{\varepsilon(1-\gamma)}{4M_R}}{\log \gamma}$ such that

$$\sum_{t=0}^{T'-1} \gamma^t \left| \mathbb{E}\left[r_{\hat{\pi}_t}\left(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] - \mathbb{E}\left[r_{\hat{\pi}_t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right|$$

$$+ \gamma^{T'} \sum_{t=T'}^{\infty} \gamma^{t-T'} \left| \mathbb{E}\left[r_{\hat{\pi}_t}\left(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] - \mathbb{E}\left[r_{\hat{\pi}_t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right|$$

$$< \sum_{t=0}^{T'-1} \gamma^t \left| \mathbb{E}\left[r_{\hat{\pi}_t}\left(X_t^i, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] - \mathbb{E}\left[r_{\hat{\pi}_t}\left(\hat{X}_t^{\frac{i}{N}}, \nu^{\frac{i}{N}}(\boldsymbol{\mu}_t) \right) \right] \right| + \frac{\varepsilon}{2}$$

and again apply Theorem 3 to the remaining finite sum. □

5. Proof of Corollary 2

Proof. The result follows directly from Corollary 1. Let $\varepsilon > 0$, then by Corollary 1, there exists $N' \in \mathbb{N}$ such that for all $N > N'$, we have

$$\max_{\pi \in \Pi} \left(J_i^N\left(\pi^1, \ldots, \pi^{i-1}, \pi, \pi^{i+1}, \ldots, \pi^N \right) - J_i^N\left(\pi^1, \ldots, \pi^N \right) \right)$$

$$\leq \max_{\pi \in \Pi} \left(J_i^N\left(\pi^1, \ldots, \pi^{i-1}, \pi, \pi^{i+1}, \ldots, \pi^N \right) - J_{\frac{i}{N}}^{\boldsymbol{\mu}}(\pi) \right)$$

$$+ \max_{\pi \in \Pi} \left(J_{\frac{i}{N}}^{\boldsymbol{\mu}}(\pi) - J_{\frac{i}{N}}^{\boldsymbol{\mu}}\left(\pi^{\frac{i}{N}} \right) \right)$$

$$+ \left(J_{\frac{i}{N}}^{\boldsymbol{\mu}}\left(\pi^{\frac{i}{N}} \right) - J_i^N\left(\pi^1, \ldots, \pi^N \right) \right)$$

$$< \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2} = \varepsilon$$
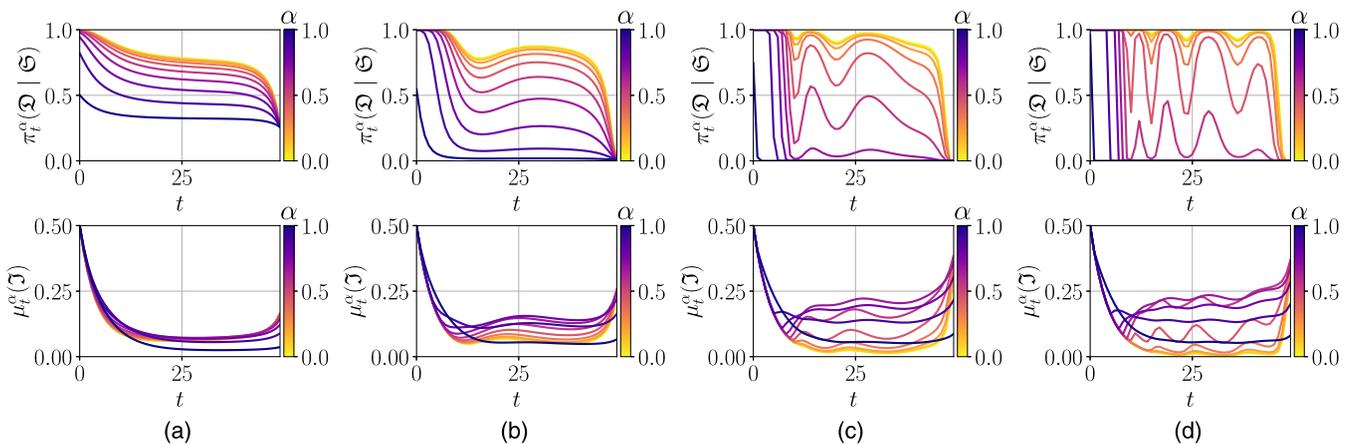

FIG. 9. The solution policy and corresponding mean field for graphons $(W_{\text{unif}}, \hat{W}_{\text{unif}})$ from Fig. 10 at different iterations n. It can be observed that in the SIS problem, the solution oscillates between taking precautions and not taking precautions. (a) $n = 20$; (b) $n = 100$; (c) $n = 500$; (d) $n = 1500$.

uniformly over $i \in \mathbb{J}^N$ for some $\mathbb{J}^N \subseteq [N]$ with $|\mathbb{J}^N| \geq \lfloor (1-p)N \rfloor$, where by definition of equilibrium optimality, we obtained

$$\max_{\pi \in \Pi} \left(J^{\boldsymbol{\mu}}_{\frac{i}{N}}(\pi) - J^{\boldsymbol{\mu}}_{\frac{i}{N}}\left(\pi^{\frac{i}{N}} \right) \right) = 0.$$

This concludes the proof. □

APPENDIX B: ADDITIONAL EXPERIMENTS

In Fig. 8, we show additional results for the Rumor problem and inverted three-uniform hypergraphons. There, we find almost inverted results as in Fig. 4, indicating that the influence of connections from the second layer are more important under the given problem parameters. However, we note that surprisingly, the highest awareness is reached for intermediate α.

As an additional example, in the timely SIS problem, we assume that there exists an epidemic that spreads to neighboring nodes according to the classical SIS dynamics (see, e.g., Ref. 83). Analogously, we may consider extensions to arbitrary variations of the SIS model, such as susceptible-infected-recovered (SIR) or susceptible-exposed-infected-recovered (SEIR). Each healthy (or susceptible, $\mathfrak{S}$) agent can take costly precautions ($\mathfrak{P}$) to avoid becoming infected ($\mathfrak{I}$), or ignore ($\bar{\mathfrak{P}}$) precautions at no further cost. Since being infected itself is costly, an equilibrium solution must balance the expected cost of infections against the cost of taking precautions. Formally, we define the state space $\mathbb{X} = \{\mathfrak{S}, \mathfrak{I}\}$ and action space $\mathbb{U} = \{\bar{\mathfrak{P}}, \mathfrak{P}\}$ such that

$$P(\mathfrak{I} \mid \mathfrak{S}, \bar{\mathfrak{P}}, \boldsymbol{\nu}) = \min \left(1, \sum_{d \in [D]} \tau_d \boldsymbol{\nu}_d \left(\mathbb{1}_{\{\mathfrak{I}\}} \right) \right),$$

$$P(\mathfrak{I} \mid \mathfrak{S}, \mathfrak{P}, \cdot) = 0, \quad P(\mathfrak{S} \mid \mathfrak{I}, \cdot, \cdot) = \delta,$$

with infection rates $\tau_d > 0$, $\sum_d \tau_d \leq 1$, recovery rate $\delta \in (0,1)$, and rewards $R(x, u, \cdot) = c_P \mathbb{1}_{\{\mathfrak{P}\}}(u) + c_I \mathbb{1}_{\{\mathfrak{I}\}}(x)$ with infection and precaution costs $c_P > 0$, $c_I > 0$. In our experiments, we will use $\tau_d = 0.8$, $\delta = 0.2$, $c_P = 0.5$, $c_I = 2$, $\mu_0(\mathfrak{I}) = 0.5$, and $\mathbb{T} = \{0, 1, \dots, 49\}$.

Existing state-of-the-art approaches such as online mirror descent (OMD)[55] (and similarly fictitious play, see, e.g., Ref. 19) as depicted in Figs. 9 and 10 for ten discretization points did not converge to an equilibrium in the considered 2000 iterations, though we expect that the methods will converge when running for significantly more iterations—e.g., 400 000 iterations as in Ref. 58—which we could not verify here due to the computational complexity. We expect that existing standard results using monotonicity conditions[38,55] can be extended to the hypergraphon case in order to guarantee convergence of aforementioned learning algorithms. However, this remains outside the scope of our work. In particular, for the ranked attachment graphon and hypergraphon, the final behavior as seen in Fig. 9 remains with an average final exploitability ΔJ of above 0.25, which is defined as

$$\Delta J(\pi) = \int_{\mathbb{I}} \sup_{\pi^* \in \Pi} J^{\Psi(\pi)}_\alpha(\pi^*) - J^{\Psi(\pi)}_\alpha(\pi) \, d\alpha$$

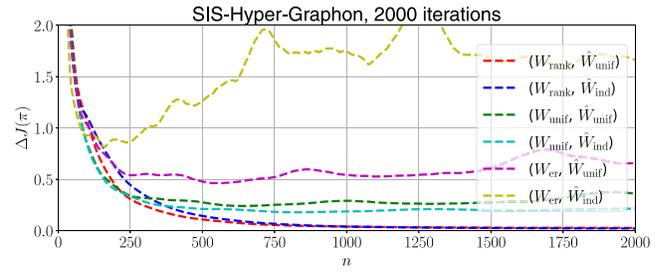and must be zero for an exact equilibrium.

FIG. 10. Average exploitability over iterations n of online mirror descent[55] on the SIS problem. It can be observed that for some configurations, the method fails to converge to an equilibrium.

REFERENCES

[1] L. Horstmeyer and C. Kuehn, Phys. Rev. E **101**, 022305 (2020).
[2] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Nat. Commun. **10**, 1 (2019).
[3] X.-J. Xu, S. He, and L.-J. Zhang, Chaos **32**, 023125 (2022).
[4] P. S. Skardal, L. Arola-Fernández, D. Taylor, and A. Arenas, Phys. Rev. Res. **3**, 043193 (2021).
[5] M. S. Anwar and D. Ghosh, Chaos **32**, 033125 (2022).
[6] C. Ziegler, P. S. Skardal, H. Dutta, and D. Taylor, Chaos **32**, 023128 (2022).
[7] M. A. Porter, in *Emerging Frontiers in Nonlinear Science* (Springer, 2020), pp. 131–159.
[8] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Phys. Rep. **874**, 1 (2020).
[9] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, arXiv:2104.11329 (2021).
[10] K. Zhang, Z. Yang, and T. Başar, *Handbook of Reinforcement Learning and Control* (Springer, Cham, 2021), p. 321.
[11] G. Qu and N. Li, in *2019 IEEE 58th Conference on Decision and Control (CDC)* (IEEE, 2019), pp. 6479–6486.
[12] H. Gu, X. Guo, X. Wei, and R. Xu, arXiv:2108.02731 (2021).
[13] Y. Lin, G. Qu, L. Huang, and A. Wierman, Adv. Neural Inf. Process. Syst. **34**, 7825–7837 (2021).
[14] G. Qu, A. Wierman, and N. Li, in *Learning for Dynamics and Control* (PMLR, 2020), pp. 256–266.
[15] P. Cardaliaguet and S. Hadikhanloo, ESAIM: Control, Optim. Calc. Var. **23**, 569 (2017).
[16] X. Guo, A. Hu, R. Xu, and J. Zhang, Adv. Neural Inf. Process. Syst. **32**, 4967–4977 (2019).
[17] R. Elie, J. Perolat, M. Laurière, M. Geist, and O. Pietquin, in *Proceedings of the AAAI Conference on Artificial Intelligence* (AAAI Press, 2020), pp. 7143–7150.
[18] X. Guo, A. Hu, R. Xu, and J. Zhang, arXiv:2003.06069 (2020).
[19] K. Cui and H. Koeppl, in *International Conference on Artificial Intelligence and Statistics* (PMLR, 2021), pp. 1909–1917.
[20] Y. Chen, J. Liu, and B. Khoussainov, arXiv:2104.14654 (2021).
[21] J. Bonnans, P. Lavigne, and L. Pfeiffer, arXiv:2109.05785 (2021).
[22] B. Anahtarci, C. D. Kariksiz, and N. Saldi, in *2021 60th IEEE Conference on Decision and Control (CDC)* (IEEE, 2021), pp. 3048–3053.
[23] X. Guo, R. Xu, and T. Zariphopoulou, "Entropy regularization for mean field games with learning," Math. Operations Res. (published online 2022).
[24] J. Arabneydi and A. Mahajan, in *53rd IEEE Conference on Decision and Control* (IEEE, 2014), pp. 1669–1674.
[25] H. Pham and X. Wei, ESAIM: Control, Optim. Calc. Var. **24**, 437 (2018).
[26] H. Gu, X. Guo, X. Wei, and R. Xu, arXiv:2002.04131 (2020).
[27] W. U. Mondal, M. Agarwal, V. Aggarwal, and S. V. Ukkusuri, arXiv:2109.04024 (2021).
[28] K. Cui, A. Tahir, M. Sinzger, and H. Koeppl, in *2021 IEEE 60th Conference on Decision and Control (CDC)* (IEEE, 2021), pp. 5239–5246.
[29] W. U. Mondal, V. Aggarwal, and S. V. Ukkusuri, arXiv:2203.00035 (2022).
[30] R. Carmona, M. Laurière, and Z. Tan, arXiv:1910.12802 (2019).

[31] M. Laurière, S. Perrin, M. Geist, and O. Pietquin, arXiv:2205.12944 (2022).

[32] Y. Achdou and I. Capuzzo-Dolcetta, SIAM J. Numer. Anal. **48**, 1136 (2010).

[33] O. Guéant, J.-M. Lasry, and P.-L. Lions, *Paris-Princeton Lectures on Mathematical Finance 2010* (Springer, 2011), pp. 205–266.

[34] A. Bensoussan, J. Frehse, and P. Yam, *Mean Field Games and Mean Field Type Control Theory* (Springer, 2013), Vol. 101.

[35] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I-II* (Springer, 2018).

[36] P. E. Caines, in *Encyclopedia of Systems and Control* (Springer, 2021), pp. 1197–1202.

[37] M. Huang, R. P. Malhamé, and P. E. Caines, Commun. Inf. Syst. **6**, 221 (2006).

[38] J.-M. Lasry and P.-L. Lions, Japanese J. Math. **2**, 229 (2007).

[39] T. Tanaka, E. Nekouei, A. R. Pedram, and K. H. Johansson, IEEE Trans. Autom. Control **66**, 880 (2020).

[40] T. Cabannes, M. Lauriere, J. Perolat, R. Marinier, S. Girgin, S. Perrin, O. Pietquin, A. M. Bayen, E. Goubault, and R. Elie, arXiv:2110.11943 (2021).

[41] K. Huang, X. Chen, X. Di, and Q. Du, Transp. Res. Part C: Emerg. Technol. **128**, 103189 (2021).

[42] A. F. Hanif, H. Tembine, M. Assaad, and D. Zeghlache, IEEE/ACM Trans. Netw. **24**, 624 (2015).

[43] S. Kar, R. Rehrmann, A. Mukhopadhyay, B. Alt, F. Ciucu, H. Koeppl, C. Binnig, and A. Rizk, Perform. Eval. **144**, 102142 (2020).

[44] W. R. KhudaBukhsh, S. Kar, B. Alt, A. Rizk, and H. Koeppl, IEEE Trans. Parallel Distrib. Syst. **31**, 2594 (2020).

[45] W. R. KhudaBukhsh, J. Rückert, J. Wulfheide, D. Hausheerv, and H. Koeppl, in *2016 IFIP Networking Conference (IFIP Networking) and Workshops* (IFIP Open Digital Library, 2016), pp. 386–394.

[46] S. Eshghi, M. H. R. Khouzani, S. Sarkar, and S. S. Venkatesh, IEEE/ACM Trans. Netw. **24**, 283 (2016).

[47] A. Tcheukam, B. Djehiche, and H. Tembine, in *2016 35th Chinese Control Conference (CCC)* (IEEE, 2016), pp. 9218–9223.

[48] B. Djehiche, A. Tcheukam, and H. Tembine, IEEE Trans. Autom. Control **62**, 5154 (2017).

[49] A. Aurell and B. Djehiche, SIAM J. Control Optim. **56**, 434 (2018).

[50] R. Carmona, arXiv:2012.05237 (2020).

[51] B. Djehiche, A. Tcheukam, and H. Tembine, AIMS Electronics Electr. Eng. **1**, 18 (2017).

[52] S. Gao and P. E. Caines, in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)* (IEEE, 2017), pp. 1052–1057.

[53] J. Subramanian and A. Mahajan, in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, 2019), pp. 251–259.

[54] S. Perrin, M. Laurière, J. Pérolat, R. Élie, M. Geist, and O. Pietquin, arXiv:2109.09717 (2021).

[55] J. Perolat, S. Perrin, R. Elie, M. Laurière, G. Piliouras, M. Geist, K. Tuyls, and O. Pietquin, arXiv:2103.00623 (2021).

[56] D. Vasal, R. Mishra, and S. Vishwanath, in *2021 American Control Conference (ACC)* (IEEE, 2021), pp. 730–736.

[57] K. Cui and H. Koeppl, in *International Conference on Learning Representations (ICLR, 2022)*.

[58] M. Laurière, S. Perrin, S. Girgin, P. Muller, A. Jain, T. Cabannes, G. Piliouras, J. Pérolat, R. Élie, O. Pietquin *et al.*, arXiv:2203.11973 (2022).

[59] L. Lovász, *Large Networks and Graph Limits* (American Mathematical Society, 2012). Vol. 60.

[60] D. Glasscock, Not. AMS **62**, 46–48 (2015).

[61] E. Bayraktar, S. Chakraborty, and R. Wu, arXiv:2003.13180 (2020).

[62] P. E. Caines and M. Huang, in *2018 IEEE Conference on Decision and Control (CDC)* (IEEE, 2018), pp. 4129–4134.

[63] P. E. Caines and M. Huang, in *2019 IEEE 58th Conference on Decision and Control (CDC)* (IEEE, 2019), pp. 286–292.

[64] Á Bodó, G. Y. Katona, and P. L. Simon, Bull. Math. Biol. **78**, 713 (2016).

[65] N. W. Landry and J. G. Restrepo, Chaos **30**, 103117 (2020).

[66] D. J. Higham and H.-L. de Kergorlay, arXiv:2108.05451 (2021).

[67] J. Noonan and R. Lambiotte, Phys. Rev. E **104**, 024316 (2021).

[68] N. Saldi, T. Basar, and M. Raginsky, SIAM J. Control Optim. **56**, 4256 (2018).

[69] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, J. Complex Netw. **2**, 203 (2014).

[70] K. A. Jacobsen, M. G. Burch, J. H. Tien, and G. A. Rempała, J. Biol. Dyn. **12**, 746 (2018).

[71] G. Elek and B. Szegedy, Adv. Math. **231**, 1731 (2012).

[72] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao, Ann. Probab. **46**, 337 (2018).

[73] C. Borgs, J. Chayes, H. Cohn, and Y. Zhao, Trans. Am. Math. Soc. **372**, 3019 (2019).

[74] C. Fabian, K. Cui, and H. Koeppl, arXiv:2209.03880 (2022).

[75] Y. Zhao, Random Struct. Algorith. **47**, 205 (2015).

[76] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou, SIAM J. Comput. **39**, 195 (2009).

[77] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, 2014).

[78] D. Maki and M. Thompson, *Mathematical Models and Applications: With Emphasis on the Social, Life, and Management Sciences* (Prentice-Hall, 1973).

[79] V. V. Junior, P. M. Rodriguez, and A. Speroto, J. Stat. Phys. **181**, 1204 (2020).

[80] K. Cui, W. R. KhudaBukhsh, and H. Koeppl, arXiv:2201.04999 (2022).

[81] F. Garbe, J. Hladk`y, M. Šileikis, and F. Skerman, arXiv:2201.12272 (2022).

[82] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi, Eur. J. Comb. **32**, 985 (2011).

[83] I. Z. Kiss, J. C. Miller, P. L. Simon, *Mathematics of Epidemics on Networks - From Exact to Approximate Models* (Springer, Cham, 2017), Vol. 598.