# VLASOV EQUATIONS ON DIRECTED HYPERGRAPH MEASURES

CHRISTIAN KUEHN [1,2] AND CHUANG XU [3]

Abstract. In this paper we propose a framework to investigate the mean field limit (MFL) of interacting particle systems on directed hypergraphs. We provide a non-trivial measure-theoretic viewpoint and make extensions of directed hypergraphs as directed hypergraph measures (DHGMs), which are measure-valued functions on a compact metric space. These DHGMs can be regarded as hypergraph limits which include limits of a sequence of hypergraphs that are sparse, dense, or of intermediate densities. Our main results show that the Vlasov equation on DHGMs are well-posed and its solution can be approximated by empirical distributions of large networks of higher-order interactions. The results are applied to a Kuramoto network in physics, an epidemic network, and an ecological network, all of which include higher-order interactions. To prove the main results on the approximation and well-posedness of the Vlasov equation on DHGMs, we robustly generalize the method of [Kuehn, Xu. Vlasov equations on digraph measures, arXiv:2107.08419, 2021] to higher-dimensions. In particular, we successfully extend the arguments for the measure-valued functions $f\colon X\to\mathcal{M}_+(X)$ to those for $f\colon X\to\mathcal{M}_+(X^{k-1})$, where $X$ is the vertex space of DHGMs and $k\in\mathbb{N}\setminus\{1\}$ is the *cardinality* of the DHGM.

## 1. Introduction

Many science phenomena in diverse areas including epidemiology [30], ecology [43], physics [40], social science [25], communication systems [17], etc., can be described as an interacting particle system (IPS), or equivalently, a dynamical system on networks [39, 40]. Classical works mainly focused on the dynamics of the IPS with interactions of rather simple types coupled on an all-to-all graph [39]. More recently, studies have concentrated more on the complex dynamics caused by (i) the interaction types and (ii) the heterogeneity of the underlying network/graph [2]. In particular, certain IPS incorporating indirect interactions beyond pairwise interactions between two particles have been considered [6, 7]. For instance, in a deterministic chemical reaction network, the concentration of a species A depends on all reactions where A appears either in the reactant or in the product [20]. Such interactions are called *higher order (or polyadic) interactions* [2, 49, 8].

It is in general difficult to analyze an IPS analytically or numerically, due to the large size (the total species counts in a *chemical reaction network* [20] or the total population size [37]) of the IPS. One popular way to study these IPS is via the so-called *mean field analysis* [46]. For example, given a large system of *indistinguishable* oscillators coupled on a given graph, one asks how to characterize the dynamics of a typical oscillator? It turns out that one can first sample the individual behavior of the oscillators on each node independently to derive a empirical distribution of the first $N$ oscillators, and then consider a *weak limit* of the sequence of empirical distributions as $N\to\infty$. Such a limit, if it exists, is the so-called *mean field*

[1] Department of Mathematics, Technical University of Munich, Munich, Garching bei München, 85748, Germany.

[2] Munich Data Science Institute (MDSI), Technical University of Munich, Munich, Garching bei München, 85748, Germany.

3 Department of Mathematics, University of Hawai'i at Mānoa, Honolulu, Hawai'i, 96822, USA

*E-mail addresses*: ckuehn@ma.tum.de (Christian Kuehn), chuangxu@hawaii.edu (Chuang Xu).

*Date*: 9th May 2023.

*Key words and phrases.* Mean field limit, higher-order interaction, hypergraphs, Kuramoto model, epidemic dynamics, Lotka-Volterra systems, sparse networks.

*limit* (MFL) [46, 23]. MFLs generally relate to a transport type PDE, the so-called Vlasov equation (VE), whose weak solution is the density of the MFL. Hence to explore the dynamics of the IPS of a large size, one can turn to study the dynamics of the VE.

Let us review some recent works, where the emphasis is the influence of the heterogeneity of the underlying graph on the dynamics of the network, in which case particles are distinguishable. Among all types of networks in epidemiology, ecology, social science, one key benchmark example is the *Kuramoto network*, a network of oscillators used to model diverse phenomena in e.g., physics, neuroscience, etc. [40]. For deterministic oscillator dynamics (particularly for deterministic Kuramoto network coupled on deterministic or random (e.g., Erdős-Rényi) graphs), the well-posedness and approximation of the MFL has been investigated [28, 15] for a Kuramoto network coupled on a sequence of heterogeneous graphs whose limit is a *graphon* [35]. Based on the VE, one can then further study the bifurcations in the Kuramoto network [15]. Later, the results were generalized to the situation, where the graph limit may not necessarily be a graphon, e.g., a *graphop* [22], a *digraph measure* [31], or a limit of a sequence of sparse graphs–a generalized graphon [26]. We point out that the technical approach in [26] is different from that in [31]. In stochastic settings, MFL results were established for IPS based on systems of SDEs coupled also on heterogeneous graphs (particularly sparse graphs due to practical considerations) [34, 33, 42].

There are many network models with higher-order interactions, e.g., the Kuramoto-Sakaguchi network [5], mathematical models based on random replicator dynamics [2], three-player games [2], Chua oscillator models [21], among many others. Effectively almost all of these models have been only been studied by direct numerical simulation or formal analytical techniques [2]. In contrast, rigorous theoretical results for MFLs of IPS with higher-order interactions are extremely rare. The underlying topological structure for such systems are hypergraphs (see Figure 1 for the illustration of a hypergraph) [38]. In [6], the MFL of a Kuramoto network coupled on *complete $k$-partite hypergraphs* for finite $k \in \mathbb{N}$ was analyzed. In [29], *quenched mean field approximation* (QMFA) of continuous time Markov chains on directed hypergraphs of a uniformly bounded cardinality (for the definition of cardinality, see Definition 2.6 in the next section) was analyzed. It is noteworthy that the QMFA is also called $N$-intertwined mean field approximation (NIMFA). Such mean field approximation neglects *dynamical correlation between nodes of the hypergraph*, which is different from the mean field limit. Our results on MFLs may help provide a deeper insight, how dynamics of a network of higher-order interactions depend on the heterogeneity of the underlying coupling graph/hypergraph.

Indeed, even from the viewpoint of graph theory, there exists a much smaller literature on hypergraph limits (a special case of hypergraph limits, is called *ultralimit hypergraph* in [19], corresponding to the limit of dense hypergraphs) [47, 24, 27, 19, 35, 13, 51], in comparison to the graph limits [4, 36, 35, 32, 1].

In this paper, we aim to propose a general framework to study the MFL of a class of IPS coupled on heterogeneous hypergraphs. For the ease of exposition, here we only consider dynamical systems on one hypergraph. Straightforward extension to a finitely many hypergraphs can be done, e.g., in a similar manner as in [31].

Consider the following network dynamical system on an $r$-layer directed hypergraph of rank $\kappa$:

$$\dot{\phi}_i^N = h_i^N(t,\phi_i^N) + \sum_{\ell=1}^{r} \frac{1}{N^{k_\ell-1}} \sum_{j_1=1}^{N} \cdots \sum_{j_{k_\ell-1}=1}^{N} W^{(\ell,N)}_{i,j_1,\ldots,j_{k_\ell-1}} g_\ell(t,\phi_i^N,\phi_{j_1}^N,\ldots,\phi_{j_{k_\ell-1}}^N), \tag{1.1}$$

where for $\ell = 1,\ldots,r$, $g_\ell$ is the coupling function of layer $\ell$, $W^{(\ell,N)} = (W^{(\ell,N)}_{i,j_1,\ldots,j_{k_\ell-1}})$ is the weight matrix associated with the $k_\ell$-uniform hypergraph $\mathcal{H}^{\ell,N} = (\mathcal{V}^N, \mathcal{E}^{\ell,N})$ of layer $\ell$ with the set of nodes $V^N = [N]$ consisting of integers from 1 to $N$ and the set $\mathcal{E}^{\ell,N}$ of (hyper)-edges; $\phi_i^N \in \mathbb{R}^d$ stands for the dynamics on the $i$-th node, and $h_i^N$ is the vector field for node $i$, for $i = 1, \dots, N$. Note that a directed edge $(i, j_1, \dots, j_{k_\ell - 1}) \in \mathcal{E}^{\ell,N}$ if and only if $W^{\ell,N}_{i,j_1,\dots,j_{k_\ell-1}} > 0$. Hence one can regard $\mathcal{H}^N = (\mathcal{V}^N, \mathcal{E}^N)$ with $\mathcal{E} = \cup_{\ell=1}^r \mathcal{E}^{\ell,N}$, is a hypergraph of rank $\kappa := \max\{k_{\ell\colon \ell=1,\dots,,r}\}$. We assume w.l.o.g. that $\{k_\ell\}_{\ell=1}^r$ consists of distinct numbers. Indeed, one can always take the uinon of all uniform hypergraphs of the same cardinality as a new hypergraph. We emphasize that $\kappa$ (and therefore $r$) is independent of the graph size $N$. Although such framework may not be suitable for modelling networks on a sequence of simplicial complexes with growing cardinalities [41, 44], it already covers several important network dynamical models from different disciplines of sciences [2]. Furthermore, one can easily extend the results of this paper to investigate the MFL of IPS on Riemannian manifolds with higher-order interactions (see Section 6.1).

To better reveal the essence of our main result (Theorem 5.6), we illustrate it here for a one-dimensional Kuramoto network coupled on a *higher-dimensional ring* which can be decomposed into two uniform hypergraphs, where in (1.1), $r = 2$, the adjacency matrices associated with a 2-uniform and a 3-uniform hypergraphs for $\ell = 1, 2$ are

$$W^{(1,N)}_{ij} = \begin{cases} 1, & \text{if} \quad |i-j| = 1 \bmod N, \\ 0, & \text{else,} \end{cases} \qquad W^{(2,N)}_{ijk} = \begin{cases} 1, & \text{if} \quad |i-j| = |j-k| = 1 \bmod N,\ k \neq i, \\ 0, & \text{else.} \end{cases}$$

This is a *sparse* hypergraph of cardinality 3. Assume $h_i^N(t,u) \equiv h(t,u)$ for some $h \in \mathcal{C}^1([0,T] \times \mathbb{T})$ and $T > 0$, for all $i = 1, \dots, N$ and $N \in \mathbb{N}$. Let $X = \mathbb{T}$. It is readily verified that the weak limit of $W^{(\ell,N)}$ is a measure-valued function $\eta_\ell\colon x \mapsto \eta_\ell^x$ defined as follows for $\ell = 1, 2$:

$$\eta_1^x = 2\delta_x, \quad \eta_2^x = 2\delta_{(x,x)}, \quad \text{for} \quad x \in X.$$

Then the VE (i.e., the mean field equation) is given by

$$\begin{aligned} &\frac{\partial \rho}{\partial t}(t,x,\phi) + \mathrm{div}_\phi \left( \rho(t,x,\phi) \widehat{V}[\eta, \rho(\cdot), h](t,x,\phi) \right) = 0, t \in (0,T],\ x \in X,\ \mathfrak{m}\text{-a.e. } \phi \in \mathbb{T}, \\ &\rho(0,\cdot) = \rho_0(\cdot), \end{aligned}$$

where

$$\begin{aligned} \widehat{V}[\eta, \rho(\cdot), h](t,x,\phi) = &\int_X \int_{\mathbb{T}} g_1(t,\phi,\psi_1) \rho(t,y_1,\psi_1) \mathrm{d}\psi_1 \mathrm{d}\eta_1^x(y_1) + \int_{X^2} \int_{\mathbb{T}} \int_{\mathbb{T}} g_2(t,\phi,\psi_1,\psi_2) \\ &\rho(t,y_2,\psi_2) \rho(t,y_1,\psi_1) \mathrm{d}\psi_2 \mathrm{d}\psi_1 \mathrm{d}\eta_2^x(y_1,y_2) + h(t,x,\phi). \end{aligned}$$

As we will see in Section 6.1, one may obtain well-posedness as well as approximation of the MFL as an application of Theorem 5.6. In fact, our most general result applies to the model

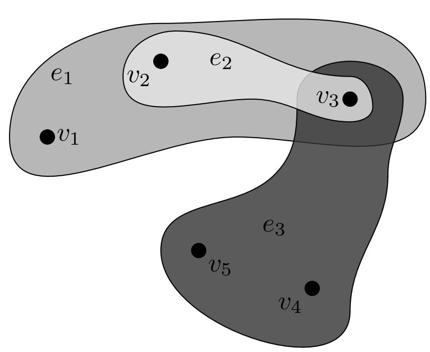


FIGURE 1. A hypergraph of rank 3, one hyper-edge of cardinality 2 ($e_2$), and two hyper-edges of cardinality 3 ($e_1$ and $e_3$). The set of vertices is $V = \{v_1, \dots, v_5\}$, and the set of hyper-edges is $E = \{e_1, e_2, e_3\}$.

(1.1) on hypergraph limits using a suitable *equation of characteristics* involving generalized measure-valued functions $\eta_\ell$, see Section 3.

In the following, we describe the methods we use to obtain the well-posedness as well as approximation in the context of model (1.1). First, we propose *directed hypergraph measures* (abbreviated as DHGM), which are measure-valued functions defined on a compact metric space. The definition of DHGM is motivated by the so-called *digraph measure* proposed in [31], corresponding to the limit of a sequence of directed graphs. We point out that such a generalization to hypergraphs is different from the hypergraph limits defined in the literature of combinatorics and graph theory [19, 35]. In particular, we compare the difference between two DHGMs using uniform bounded Lipschitz metric, in contrast to the classical *homeomorphism density* [19, 26] used in the contexts of both graph limits and hypergraph limits. Resting on the proposed notion of DHGM, we propose a generalized model of (1.1), which can be regarded as the *fiberized equation of characteristics* associated with (1.1). We mention that such type of equation of characteristics indexed by vertices of a graph limit seemed to first appear in [28]. Moreover, we then prove the well-posedness of the equation of characteristics. Using the flow generated by this equation of characteristics, we define the generalized VE (or, fixed point equation), which is a time-dependent measure-valued equation (i.e., the measure at time $t$ is the push-forward of the initial measure under the flow of the equation of characteristics). Then, under rather mild conditions for the DHGM (basically, uniform boundedness of the DHGM), we obtain the well-posedness of the VE (a transport type PDE whose *uniform weak* solution is the density of the MFL). To obtain the approximation of the MFL by empirical measures, we rely on the recently developed result of uniform approximation of probability measures on Euclidean spaces [50, 14, 3] (see also [31]), which is different from the Martingale Convergence Theorem for the approximation of integrable functions exmployed in [28]. For the approximation results to hold, we need some continuity of the DHGM.

We like to point out that for combinatorial reasons, it is known to be *inappropriate* to view hypergraph limits as a function defined on the space of vertices [35], provided e.g. the underlying metric is the *homomorphism density.* Nevertheless, for analytical reasons, we here use a *uniform* bounded Lipschitz metric, which may induce a uniform weak topology on the space of measure-valued functions [31]. To work under this metric, we successfully obtain the results on the MFL. This also demonstrates the robustness of the idea initiated in [31] to study MFLs of IPS on heterogeneous graph limits.

Next, we outline the structure of the paper. In Section 2, we introduce the notation, provide some basics in measure theory, and define *directed hypergraph measures* (DHGM) with ample examples of independent interest. In Section 4, using DHGM, we propose a generalized model of (1.1), which serves as the *fiberized equation of characteristics.* In Section 5, we prove approximation of the MFL by a sequence of ODEs coupled on hypergraphs, based on several lemmas from [31]. Finally in Section 6.1, we apply our theoretical results to three network models of higher-order interactions in physics, epidemiology, and ecology.

## 2. Preliminaries

**Notation.** Let $\mathbb{R}_+$ be the set of nonnegative real numbers. For $i=1,2$, let $X_i$ be a complete subspace of a finite dimensional Euclidean space endowed with the metric $d_i$ induced by the $\ell_1$-norm $|\cdot|$. For instance $X_i$ can be a sphere, a torus, or any other compact set of a Euclidean space. For any $k\in\mathbb{N}$ and $x\in\mathbb{R}^k$, let $[k]\coloneqq\{j\}_{j=1}^k$, $\delta_x$ denote the Dirac measure at $x$, and $\mathfrak{m}$ be the Lebesgue measure on $\mathbb{R}^k$; here we omit the explicit dependence of $\mathfrak{m}$ on the dimension $k$. For any measurable subset $A$ in a Euclidean space endowed with norm $|\cdot-\cdot|$, let $\dim_H(A)$ denote the Hausdorff dimension of $A$, $\mathrm{Diam}A\coloneqq\sup_{x,y\in A}|x-y|$ be its diameter; by convention, $\mathrm{Diam}\,A=0$ if the cardinality $\#A\le 1$. We use $\lambda|_A$ to denote the uniform (probability) measure over $A$ whenever appropriate (e.g., when $A$ is either finite or Lebesgue measurable with a finite Lebesgue measure). Let $\mathbb{1}_A$ be the indicator function on $A$. Let $X$ be a Borel measurable space. For $k \in \mathbb{N}$, and for any Borel set $A \subseteq X^k$, let $\pi_1 A = \{z_1 \in X : z \in A\}$ denote the projection of $A$ onto the first coordinate in $X$; let $A_x = \{(z_2, \ldots, z_k) : z_1 = x, z \in A\}$ be the *vertical slice* of $A$ through a point in $A$ with the first coordinate $x \in X$. For any $a \in A$, let $\delta_x$ be the Dirac measure at the point $a \in A$. For two real-valued functions $f$ and $g$, we denote $f \precsim g$ if there exists $C > 0$ such that $\frac{f(x)}{g(x)} \le C$, and $f \sim g$ if both $f \precsim g$ and $g \precsim f$.

**Measure-valued functions.** Let $\mathcal{M}_+(X_2)$ be the set of all finite Borel positive measures on $X_2$ and $\mathcal{P}(X_2)$ the space of all Borel probability measures on $X_2$. Given a reference measure $\mu_{X_2} \in \mathcal{P}(X_2)$, let $\mathcal{M}_{+,\mathsf{abs}}(X_2) \subseteq \mathcal{M}_+(X_2)$ be the space of all absolutely continuous finite positive measures w.r.t. $\mu_{X_2}$.

Let $\mathcal{B}(X_1, X_2)$ ($\mathcal{C}(X_1, X_2)$, $\mathcal{C}_{\mathsf{b}}(X_1, X_2)$, respectively) be the space of bounded measurable functions (continuous functions, bounded continuous functions, respectively) from $X_1$ to $X_2$ equipped with the same uniform metric:

$$d(f, g) = \sup_{x \in X_1} d_2(f(x), g(x)).$$

Let $\mathcal{L}(X_1, X_2) \coloneqq \{g \in \mathcal{C}(X_1, X_2) : \mathcal{L}(g) \coloneqq \sup_{x \neq y} \frac{d_2(g(x), g(y))}{d_1(x, y)} < \infty\}$ be the space of Lipschitz continuous functions from $X_1$ to $X_2$. Hence $\mathcal{BL}(X_1, X_2) = \mathcal{B}(X_1, X_2) \cap \mathcal{L}(X_1, X_2)$ denotes the space of bounded Lipschitz continuous functions. In particular, when $X_2 = \mathbb{R}$, we suppress $X_2$ in $\mathcal{B}(X_1, X_2)$ and simply write $\mathcal{B}(X_1)$. Similarly, we write $\mathcal{C}(X_1)$ for $\mathcal{C}(X_1, \mathbb{R})$, etc. In this case, $\mathcal{B}(X_1)$, $\mathcal{C}(X_1)$ and $\mathcal{C}_{\mathsf{b}}(X_1)$ are all Banach spaces with the supremum norm

$$\|f\|_\infty \coloneqq \sup_{x \in X_1} |f(x)|.$$

Let $\mu_{X_1} \in \mathcal{P}(X_1)$ be the reference measure on $X_1$. Let $\mathcal{BL}_1(X_2) = \{g \in \mathcal{BL}(X_2) : \mathcal{BL}(g) \coloneqq \|g\|_\infty + \mathcal{L}(g) \le 1\}$.

Define the bounded Lipschitz norm (on the space of all finite signed Borel measures):

$$\|\nu\|_{\mathsf{BL}} \coloneqq \sup_{f \in \mathcal{BL}_1(X_2)} \int_{X_1} f \mathrm{d}\nu, \quad \nu \in \mathcal{M}_+(X_2),$$

which induces the bounded Lipschitz distance: For $\nu^1,\ \nu^2 \in \mathcal{M}_+(X_2)$,

$$d_{\mathsf{BL}}(\nu^1, \nu^2) = \sup_{f \in \mathcal{BL}_1(X_2)} \int f(x) \mathrm{d}(\nu^1(x) - \nu^2(x)).$$

Recall that $(\mathcal{M}_+(X_2), d_{\mathsf{BL}})$ is a complete metric space [10]. Define the *uniform bounded Lipschitz metric*:

$$d_\infty(\eta_1, \eta_2) = \sup_{x \in X} d_{\mathsf{BL}}(\eta_1^x, \eta_2^x), \quad \text{for} \quad \eta_1, \eta_2 \in \mathcal{B}(X_1, \mathcal{M}_+(X_2)). \tag{2.1}$$

Hence $\mathcal{B}(X_1, \mathcal{M}_+(X_2))$ and $\mathcal{C}_{\mathsf{b}}(X_1, \mathcal{M}_+(X_2))$ equipped with the uniform bounded Lipschitz metric are complete.

For every $\eta \in \mathcal{M}_+(X_2)$, let

$$\|\eta\|_{\mathsf{TV}} = \sup_{A \in \mathfrak{B}(X_2)} \eta(A) = \eta(X_2)$$

be the total variation *norm* of $\eta$, where $\mathfrak{B}(X_2)$ is the Borel sigma algebra of $X_2$. For $\mathcal{B}(X_1, \mathcal{M}_+(X_2)) \ni \eta : x \mapsto \eta^x$, let

$$\|\eta\| \coloneqq \sup_{x \in X_1} \|\eta^x(X_2)\|$$

Let

$$\mathcal{B}_*(X_1, \mathcal{M}_+(X_2)) \coloneqq \left\{\eta \in \mathcal{B}(X_1, \mathcal{M}_+(X_2)) : \int_{X_1} \eta^x(X_2) \mathrm{d}\mu_{X_1}(x) = 1\right\},$$

$$\mathcal{C}_*(X_1,\mathcal{M}_+(X_2)) \coloneqq \mathcal{B}_*(X_1,\mathcal{M}_+(X_2)) \cap \mathcal{C}(X_1,\mathcal{M}_+(X_2)),$$

Let $\mathcal{I} \subseteq \mathbb{R}$ be a compact interval and $k \in \mathbb{N}$. For $h \in \mathcal{C}(\mathcal{I} \times X_1 \times X_2, \mathbb{R}^k)$, let

$$\|h\|_{\infty,\mathcal{I}} \coloneqq \sup_{\tau\in\mathcal{I}} \sup_{u\in X_1} \sup_{v\in X_2} |h(\tau,u,v)|.$$

For $\eta_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}(X_1,\mathcal{M}_+(X_2)))$, let

$$\|\eta_\cdot\| = \sup_{t\in\mathcal{I}} \sup_{x\in X_1} \|\eta_t^x\|_{\mathsf{TV}}$$

be the norm of the function $\eta_\cdot$.

To construct solutions to Vlasov equations on DHGMs, we need to first define "weak continuity" of measure-valued functions [31].

**Definition 2.1.** For $i = 1, 2$, let $X_i$ be a complete subspace of a finite dimensional Euclidean space. Assume $(X_1, \mathfrak{B}(X_1), \mu_{X_1})$ is a compact probability space. For

$$\mathcal{B}_*(X_1,\mathcal{M}_+(X_2)) \ni \eta\colon \begin{cases} X_1 \to \mathcal{M}_+(X_2), \\ x \mapsto \eta^x, \end{cases}$$

$\eta$ is *weakly $x$-continuous* (or *weakly continuous* if the variable is clear from the context) if for every $f \in \mathcal{C}_{\mathsf{b}}(X_2)$,

$$\mathcal{C}(X_1) \ni \eta(f)\colon \begin{cases} X_1 \to \mathbb{R}, \\ x \mapsto \eta^x(f) \coloneqq \int_{X_2} f \mathrm{d}\eta^x. \end{cases}$$

**Definition 2.2.** For $i = 1, 2$, let $X_i$ be a complete subspace of a finite dimensional Euclidean space. Assume $(X_1, \mathfrak{B}(X_1), \mu_{X_1})$ is a compact probability space. Let $\mathcal{I} \subseteq \mathbb{R}$ be a compact interval. For

$$\eta_\cdot\colon \begin{cases} \mathcal{I} \to \mathcal{B}(X_1,\mathcal{M}_+(X_2)), \\ t \mapsto \eta_t, \end{cases}$$

$\eta_\cdot$ is *$x$-uniformly weakly $t$-continuous* (or *uniformly weakly $t$-continuous* if the spatial variable $x$ is clear from the context) if for every $f \in \mathcal{C}_{\mathsf{b}}(X_2)$, $t \mapsto \eta_t^x(f)$ is continuous in $t$ uniformly in $x \in X_1$.

Recall the following proposition from [31, Proposition 3.9] on the relation between continuity and (uniform) weak continuity. This proposition will be used the proofs of the main results.

**Proposition 2.3.** *For $i = 1, 2$, let $X_i$ be a subspace of a finite dimensional Euclidean space. Assume $X_1$ and $X_2$ are both compact, and $(X_1, \mathfrak{B}(X_1), \mu_{X_1})$ is a probability space. Let $\mathcal{I} \subseteq \mathbb{R}$ be a compact interval.*

(i) *Let $\eta_\cdot\colon \mathcal{I} \to \mathcal{B}_*(X_1,\mathcal{M}_+(X_2))$. Then $\eta_\cdot$ is $x$-uniformly weakly $t$-continuous if and only if $\eta_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X_1,\mathcal{M}_+(X_2)))$.*
(ii) *Assume $\eta_\cdot,\ \xi_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X_1,\mathcal{M}_+(X_2)))$, then $\|\eta_\cdot\| < \infty$ and $t \mapsto d_\infty(\eta_t, \xi_t)$ is continuous.*
(iii) *Assume $\eta \in \mathcal{C}(X_1,\mathcal{M}_+(X_2))$. Then $\eta$ is weakly $x$-continuous.*
(iv) *Assume $\eta \in \mathcal{C}(X_1,\mathcal{M}_+(X_2))$. Let $m \in \mathbb{N}$,*

$$\eta(m)\colon X_1^m \ni (x_1,\ldots,x_m) \mapsto \otimes_{j=1}^m \eta^{x_j} \in \mathcal{M}_+(X_2^m).$$

*Then $\eta(m)$ is weakly $x$-continuous.*

*Proof.* Items (ii)-(iii) follow directly from [31, Proposition 3.9]. The last statement (iv) can be proved similarly to (iii). We only prove (i).

Step I. $x$-uniform weak $t$-continuity implies continuity. Assume $\eta_\cdot$ is $x$-uniformly weakly $t$-continuous. Fix $t \in \mathcal{I}$ and $\mathcal{I} \ni t_j \to t$. Since $\eta_\cdot$ is $x$-uniformly weakly $t$-continuous, for every $f \in \mathcal{C}_{\mathsf{b}}(X_2)$,

$$(\eta_{t_j})_x(f) \to (\eta_t)_x(f) \quad \text{uniformly in } x.$$

Since $X_2$ is a complete, separable metric space, $d_{\mathsf{BL}}$ metrizes the weak-$*$ topology of $\mathcal{M}_+(X_2)$ [10, Thm. 8.3.2]. Using the supremum representation of $d_{\mathsf{BL}}$ and note that $\mathcal{BL}_1(X_2) \subseteq \mathcal{C}_{\mathsf{b}}(X_2)$, we have

$$\lim_{j\to\infty} d_{\mathsf{BL}}((\eta_{t_j})_x, (\eta_t)_x) = 0,$$

and the convergence is uniform in $x \in X_1$. This means that

$$\lim_{j\to\infty} \sup_{x\in X_1} d_{\mathsf{BL}}((\eta_{t_j})_x, (\eta_t)_x) = 0,$$

i.e.,

$$\lim_{j\to\infty} d_\infty(\eta_{t_j}, \eta_t) = 0,$$

which shows $\eta_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_{\mu_{X_1},1}(X_1, \mathcal{M}_+(X_2)))$.

Step II. Continuity implies $x$-uniform weak $t$-continuity. Assume $\eta_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_{\mu_{X_1},1}(X_1, \mathcal{M}_+(X_2)))$. For every fixed $t \in \mathcal{I}$ and $\mathcal{I} \ni t_j \to t$, we have

$$\lim_{j\to\infty} d_\infty(\eta_{t_j}, \eta_t) = 0. \tag{2.2}$$

Let $f \in \mathcal{C}_{\mathsf{b}}(X_2)$. Since $X_2$ is compact, by [16, Corollary 6.2.2] for every $\varepsilon > 0$, there exists $M \ge 1$ and $\tilde{f} \in \mathcal{BL}(X_2)$ such that $\mathcal{BL}(\tilde{f}) \le M$ and

$$\|f - \tilde{f}\|_\infty < \frac{\varepsilon}{3(1 + \|\eta_\cdot\|)} \tag{2.3}$$

It follows from (2.2) that there exists $J \in \mathbb{N}$ such that for all $j \ge J$,

$$\sup_{g\in\mathcal{BL}_1(X_2)} |\eta^x_{t_j}(g) - \eta^x_t(g)| < \frac{\varepsilon}{3M} \tag{2.4}$$

Note that $\frac{\tilde{f}}{M} \in \mathcal{BL}_1(X_2)$. Hence

$$|\eta^x_{t_j}(\tilde{f}) - \eta^x_t(\tilde{f})| = M|\eta^x_{t_j}(\tilde{f}/M) - \eta^x_t(\tilde{f}/M)| < M \cdot \frac{\varepsilon}{3M} = \frac{\varepsilon}{3}, \quad \text{for all} \quad x \in X$$

Using triangle inequality, (2.3) and (2.4),

$$\begin{aligned} |\eta^x_{t_j}(f) - \eta^x(f)| \le& |\eta^x_{t_j}(f) - \eta^x_{t_j}(\tilde{f})| + |\eta^x_{t_j}(\tilde{f}) - \eta^x(\tilde{f})| + |\eta^x(\tilde{f}) - \eta^x(f)| \\ \le& \|f - \tilde{f}\|_\infty \|\eta_\cdot\| + M|\eta^x_{t_j}(\tilde{f}/M) - \eta^x(\tilde{f}/M)| + \|f - \tilde{f}\|_\infty \|\eta_\cdot\| \\ \le& \frac{\varepsilon}{3(1 + \|\eta_\cdot\|)} + M\frac{\varepsilon}{3M} + \frac{\varepsilon}{3(1 + \|\eta_\cdot\|)} < \varepsilon \end{aligned}$$

This shows that $\eta_\cdot$ is $x$-uniformly weakly $t$-continuous.

□

**Definition 2.4.** For $i = 1, 2$, let $X_i$ be a complete subspace of a finite dimensional Euclidean space. Assume that $(X_1, \mathfrak{B}(X_1), \mu_{X_1})$ is a compact probability space. Let $\mathcal{I} \subseteq \mathbb{R}$ be a compact interval and $\alpha > 0$. For $\nu^1_\cdot, \nu^2_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X_1, \mathcal{M}_+(X_2)))$, let

$$d_\alpha(\nu^1_\cdot, \nu^2_\cdot) = \sup_{t\in\mathcal{I}} \mathrm{e}^{-\alpha t} d_\infty(\nu^1_t, \nu^2_t)$$

be a *weighted* uniform metric. In particular,

$$d_0(\nu^1_\cdot, \nu^2_\cdot) = \sup_{t\in\mathcal{I}} d_\infty(\nu^1_t, \nu^2_t).$$

**Proposition 2.5.** [31, Proposition 3.11] *For $i = 1, 2$, let $X_i$ be a complete subspace of a finite-dimensional Euclidean space. Assume that $(X_1, \mathfrak{B}(X_1), \mu_{X_1})$ is a compact probability space. Let $\mathcal{I} \subseteq \mathbb{R}$ be a compact interval and $\alpha > 0$. Then $(\mathcal{C}(\mathcal{I}, \mathcal{B}_*(X_1, \mathcal{M}_+(X_2))), d_\alpha)$ and $(\mathcal{C}(\mathcal{I}, \mathcal{C}_*(X_1,\ \mathcal{M}_+(X_2))), d_\alpha)$ are both complete.*

2.1. **Hyper-digraph measures.** In this subsection, we recall some basics of hypergraphs and define their generalizations with illustrative examples.

**Definition 2.6.** Let $\kappa \in \mathbb{N}$. A *directed hypergraph* $\mathcal{H}$ is a pair $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of $N \in \mathbb{N}$ vertices, and $\mathcal{E} \subseteq \cup_{k \in I} \mathcal{V}^k$ with $I \subseteq [\kappa]$, consists of hyper-edges (or edges for short, so long as no confusion will arise in the context) identified as a vector without having repetitions in the coordinates. We assume w.l.o.g. that $\kappa \in I$ (otherwise one can choose $\max I$ to replace $\kappa$), and $\kappa$ is called the *rank* of $\mathcal{H}$. For every directed hyper-edge $e = (v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ for some $k \in I$, $v_{i_1}$ and $v_{i_k}$ are the *head* and the *tail* of the hyper-edge, respectively; $k$ is the *cardinality* of the edge $e$. In particular, if $I = \{k\}$ is a singleton, then $\mathcal{H}$ is a *$k$-uniform directed hypergraph* with $k$ being the cardinality of each directed edge. A $k$-uniform directed hypergraph $(\mathcal{V}^N, \mathcal{E}^N)$ is *$[k]$-permutation invariant* if $\mathcal{E}^N = \{((j_{\sigma(1)}, j_{\sigma(2)}, \ldots, j_{\sigma(k)}))\colon (j_1, j_2, \ldots, j_k) \in \mathcal{E}^N\}$ for any permutation $\sigma\colon [k] \to [k]$. A $[k]$-permutation invariant directed hypergraph is a *$k$-uniform hypergraph*. A $k$-uniform hypergraph consisting of all possible hyper-edges of cardinality $k$ is called a *complete $k$-uniform hypergraph*, and denoted $\mathcal{K}_N^k$. for a directed hyper-edge $e \in \mathcal{E}$ of cardinality $k$. If every directed hyper-edge $e \in \mathcal{E}$ is assigned a positive weight $a_e$, then the directed hypergraph $\mathcal{H}$ is called a *weighted directed hypergraph*.[1] A directed hypergraph $\mathcal{H}$ of a finite rank $\kappa$ can be decomposed into a finite union of $k$-uniform hypergraphs $\mathcal{H}_k$ for $2 \le k \le \kappa$ of the same set of vertices; in this case, each subhypergraph $\mathcal{H}_k$ is called a *layer* of $\mathcal{H}$, and $\mathcal{H}$ is a *multi-layer directed hypergraph* if it contains more than one layer.

*Remark* 2.7.
- The concept of layer of a hypergraph of a finite rank comes from [18].
- We point out that hypergraph defined in the literature (e.g.,[19]) is *undirected* in the sense that each edge is regarded as a set of distinct vertices rather than a vector. For example, let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = [4]$ be a directed 3-hypergraph with $\mathcal{E}$ consisting of all hyper-edges of cardinality 3. Hence $(1, 2, 3)$ and $(1, 3, 2)$ stand for two different directed hyper-edges.
- The density of a sequence of $k$-uniform hypergraphs $\{\mathcal{H}^N = (\mathcal{V}^N, \mathcal{E}^N)\}_{N \in \mathbb{N}}$ is characterized by the asymptotics of the sequence of degrees of the hypergraphs:

$$D(\mathcal{H}^N) = \frac{\#\mathcal{E}^N}{\binom{\#\mathcal{V}^N}{k}}. \tag{2.5}$$

  The sequence $\{\mathcal{H}^N\}_{N \in \mathbb{N}}$ is *dense* if $D(\mathcal{H}^N) \sim 1$; it is *sparse* if $\#\mathcal{E}^N \precsim \#\mathcal{V}^N$; otherwise, it is neither dense nor sparse.

**Example 2.8.** Consider the hypergraph given in Figure 1. It is a 2-layer hypergraph $\mathcal{H} = (\mathcal{E}, \mathcal{V})$ consisting of in total of 2 $k$-uniform hypergraphs with $\mathcal{V} = \{v_i\}_{i=1}^5$:
- a 2-uniform hypergraph $\mathcal{H}_2 = (\mathcal{V}, \mathcal{E}_2)$ with $\mathcal{E}_2 = \{e_2\}$;
- a 3-uniform hypergraph $\mathcal{H}_3 = (\mathcal{V}, \mathcal{E}_3)$ with $\mathcal{E}_3 = \{e_1, e_3\}$;

When $\mathcal{H}$ is regarded as a weighted directed hypergraph, we can assign a unit weight to each hyper-edge, i.e., $a_{e_j} = 1$ for $j = 1, 2, 3$.

Now we propose one candidate as a generalization of directed $k$-uniform hypergraph.

**Definition 2.9.** For $k \in \mathbb{N} \setminus \{1\}$, a measure-valued function in $\mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$ is called a *$k$-uniform directed hypergraph measure* (abbreviated as "*$k$-uniform DHGM*"). A $k$-uniform

[1] Throughout we will omit "weighted" since we will always refer to weighted hypergraphs.

DHGM $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$ is a $k$-HGM if $\xi := \mu_X \otimes \eta^x$ as a measure in $\mathcal{M}_+(X^k)$[2] is $[k]$*-permutation invariant* for any permutation $\sigma\colon [k] \to [k]$:

$$\mathrm{d}\xi(x_1, \ldots, x_k) = \mathrm{d}\xi(\sigma(x_1), \ldots, \sigma(x_k)), \quad (x_1, \ldots, x_k) \in X^k.$$

A measure-valued function $\eta$ in $\mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$ is called a *directed hypergraph measure of rank* $\kappa$ if it is a finite sum of $k_\ell$-uniform directed hypergraph measure $\eta_\ell$ for $1 \le \ell \le m$ such that $\eta = \sum_{\ell=1}^m \eta_{k_\ell}$ with $\kappa = \max_{1 \le \ell \le m} k_\ell$. In this case, $\eta$ is a *multi-layer DHGM*, and each $k_\ell$-uniform directed hypergraph measure $\eta_\ell$ is called a *layer* of $\eta$.

**Definition 2.10.** A DHGM $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$ is called a *directed hyper-graphon* w.r.t. a reference measure $\mu_X \in \mathcal{P}(X)$ if $\mu_X \otimes \eta^x \in \mathcal{M}_+(X^k)$ is absolutely continuous w.r.t. $\otimes_{j=1}^k \mu_X$. A $[k]$-permutation invariant directed hyper-graphon is a *hyper-graphon*.

Based on the density characterization of a sequence of $k$-uniform hypergraphs in Remark 2.7, we propose the criterion for density of a $k$-uniform DHGM.

**Definition 2.11.** A $k$-uniform DHGM $\eta$ is *dense* if

$$\inf_{x \in X} \dim_H(\operatorname{supp}\ \eta^x) = \dim_H(X) > 0.$$

Hence a hypergraphon is dense. A $k$-uniform DHGM $\eta$ is *sparse* if

$$\sup_{x \in X} |\operatorname{supp} \eta^x| < \infty.$$

A sparse DHGM is called a *directed hypergraphing.*

*Remark* 2.12*.* In contrast to a directed multi-layer hypergraph of a finite rank, a DHGM may not always be decomposed into finitely many $k$-uniform DHGMs. Instead it may be decomposed into countably infinite uniform DHGMs of *unbounded* ranks (e.g., simplicial complexes of unbounded ranks [44]). In this case, this directed hypergraph limit may be represented as an object even more complicated than a measure-valued function in $\mathcal{B}(X, \mathcal{M}_+(X^\infty))$.

**Representation of finite hypergraphs.** There are many ways of representing a finite (hyper)graph: by extended real-valued functions or measurable sets (*graphons* [36], *extended graphons* [26], or *hypergrahons* in the sense of [19]), by positive linear operators (*graphop* [1]), or by measure-valued functions (so-called *digraph measures* [31]). We now provide two ways of representing a fintie hypergraph here (assuming $X = [0,1]$ with $\mu_X$ being the Lebesgue measure on $X$, for the ease of exposition).

Let $\mathcal{H}^N = ([N], \mathcal{E}^N)$ be a weighted $k$-uniform directed hypergraph. Let $\{I_i^N\}$ be an equipartition of $X$ with

$$I_i^N = \left[\tfrac{i-1}{N}, \tfrac{i}{N}\right[, \quad 1 \le i \le N-1, \quad \text{and} \quad I_N^N = \left[\tfrac{N-1}{N}, 1\right] \tag{2.6}$$

Then $\{I_{i_1,\ldots,i_k}^N\}$ is a uniform partition of $X^k$ with cubes $I_{i_1,\ldots,i_k}^N := \otimes_{j=1}^k I_{i_j}^N$ for $\{i_1, \ldots, i_k\} \in [N]^k$. We associate with $\mathcal{E}^N$ an $N^k$-dimensional non-negative adjacency matrix $(a_{i_1\ldots i_k}^N)$ such that $a_{i_1\ldots i_k}^N > 0$ if and only if $(i_1 \ldots i_k) \in \mathcal{E}^N$.

(i) Function representation. Define a simple function $W^N \in L^p(X^k)$:

$$W^N = \sum_{i_1=1}^N \cdots \sum_{i_k=1}^N a_{i_1,\ldots,i_k}^N \mathbb{1}_{I_{i_1,\ldots,i_k}^N} \tag{2.7}$$

[2] $\xi$ is understood as:

$$\xi(A) := \int_{\pi_1 A} \eta^x(A_x) \mathrm{d}\mu_X(x), \quad \text{for all measurable set} \quad A \subseteq X^k$$

Define a piecewise uniform measure valued function $\eta_N \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\frac{\mathrm{d}\eta_N^x(y_1,\ldots,y_{k-1})}{\mathrm{d}(\otimes_{\ell=1}^{k-1}\mu_X)(y_1,\ldots,y_{k-1})} = W^N(x,y_1,\ldots,y_{k-1}), \quad x \in X$$

where $\otimes_{\ell=1}^{k-1}\mu_X$ is the product measure on $X^{k-1}$ satisfying

$$(\otimes_{\ell=1}^{k-1}\mu_X)(\prod_{\ell=1}^{k-1} A_\ell) = \otimes_{\ell=1}^{k-1}\mu_X(A_\ell),$$

and $A_\ell \subseteq X$ for $\ell = 1,\ldots,k-1$ are Borel measurable sets.

(ii) Atomic mesaure representation. Define $\tilde{\eta}_N \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\tilde{\eta}_N^x = N^{-(k-1)} \sum_{i_2=1}^{N} \cdots \sum_{i_k=1}^{N} a^N_{i_1,\ldots,i_k} \delta_{\left(\frac{2i_2-1}{2N},\ldots,\frac{2i_k-1}{2N}\right)}, \quad x \in I^N_{i_1}$$

Indeed, assume that:

**(H)** $\max_{1\le i\le N} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N a^N_{i,j_1,\ldots,j_{k-1}} = o(N^k)$

Then these two representations are *asymptotically the same.*

**Proposition 2.13.** *Assume (**H**). Then* $\lim_{N\to\infty} d_\infty(\eta_N, \tilde{\eta}_N) = 0$.

*Proof.*

$$\begin{aligned}
d_\infty(\eta_N,\tilde{\eta}_N) &= \sup_{x\in X} d_{\mathsf{BL}}(\eta_N^x, \tilde{\eta}_N^x)\\
&= \max_{1\le i\le N} \sup_{x\in I_i^N} \sup_{f\in\mathcal{BL}_1(X^{k-1})} \int_{X^{k-1}} f(y_1,\ldots,y_{k-1})\mathrm{d}(\eta_N^x - \tilde{\eta}_N^x)\\
&= \max_{1\le i\le N} \sup_{x\in I_i^N} \sup_{f\in\mathcal{BL}_1(X^{k-1})} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N \int_{\prod_{\ell=1}^{k-1} I^N_{j_\ell}} f(y_1,\ldots,y_{k-1})\mathrm{d}(\eta_N^x - \tilde{\eta}_N^x)\\
&= \max_{1\le i\le N} \sup_{x\in I_i^N} \sup_{f\in\mathcal{BL}_1(X^{k-1})} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N a^N_{i,j_1,\ldots,j_{k-1}}\\
&\quad \cdot \int_{\prod_{\ell=1}^{k-1} I^N_{j_\ell}} \big(f(y_1,\ldots,y_{k-1}) - f(\tfrac{2i_2-1}{2N},\ldots,\tfrac{2i_k-1}{2N})\big)\mathrm{d}(\otimes_{\ell=1}^{k-1}\mu_X)(y_1,\ldots,y_{k-1})\\
&\le \max_{1\le i\le N} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N a^N_{i,j_1,\ldots,j_{k-1}}\\
&\quad \cdot \int_{\prod_{\ell=1}^{k-1} I^N_{j_\ell}} |(y_1,\ldots,y_{k-1}) - (\tfrac{2i_2-1}{2N},\ldots,\tfrac{2i_k-1}{2N})|\mathrm{d}(\otimes_{\ell=1}^{k-1}\mu_X)(y_1,\ldots,y_{k-1})\\
&\le \max_{1\le i\le N} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N a^N_{i,j_1,\ldots,j_{k-1}}\\
&\quad \cdot \int_{\prod_{\ell=1}^{k-1} I^N_{j_\ell}} \big(|y_1 - \tfrac{2i_2-1}{2N}| + \ldots + |y_{k-1} - \tfrac{2i_k-1}{2N}|\big)\mathrm{d}(\otimes_{\ell=1}^{k-1}\mu_X)(y_1,\ldots,y_{k-1})\\
&\le \frac{k-1}{4} \max_{1\le i\le N} N^{-k} \sum_{j_1=1}^N \cdots \sum_{j_{k-1}=1}^N a^N_{i,j_1,\ldots,j_{k-1}} \to 0, \quad \text{as} \quad N\to\infty
\end{aligned}$$

$\square$

*Remark* 2.14. (i) Proposition 2.13 implies that with weights of appropriate scales, the two sequences of measure-valued functions representing the same sequence of finite graphs, diverge simultaneously, or converge to the same limit, i.e., whether the sequence of hypergraphs viewed as DHGMs converges or not is independent of the choice of the representation.

(ii) Note that

$$\|\eta^N\| = \|\tilde{\eta}^N\| = \max_{1\leq i\leq N} N^{-k+1} \sum_{j_1=1}^{N} \cdots \sum_{j_{k-1}=1}^{N} a^N_{i,j_1,\dots,j_{k-1}}$$

and (**H**) is equivalent to $\|\tilde{\eta}^N\| = o(N)$, which is a necessary condition for $\{\eta^N\}_{N\in\mathbb{N}}$ as well as $\{\tilde{\eta}^N\}_{N\in\mathbb{N}}$ to be convergent. Moreover, if assuming $\{\eta^N\}_{N\in\mathbb{N}}$ as well as $\{\tilde{\eta}^N\}_{N\in\mathbb{N}}$ converges, then from the proof we know that $d_\infty(\eta_N, \tilde{\eta}_N)$ vanishes with a rate no slower than $N^{-1}$.

(iii) This proposition also rules out the illusion that the atomic measure representation of finite (hyper)graphs implies the sequence of the (hyper)graphs are sparse. See Example 2.15 below.

(iv) As will be seen below, $X$ is not always chosen to be $[0,1]$. Nevertheless, one can always represent a finite hypergraph in two ways analogously defined above.

**Example 2.15.** Let $k \in \mathbb{N}\setminus\{1\}$ and $\{\mathcal{K}^N_k\}_{N\geq k}$ be the sequence of complete $k$-uniform hypergraphs [35]. Then $\{\mathcal{K}^N_k\}_{N\geq k}$, represented as DHGMs (in either of the two ways aforementioned), converges to the *complete $k$-uniform hypergraphon* $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\eta^x \equiv \lambda|_{X^{k-1}} \quad \text{for all} \quad x \in X.$$

**Example 2.16.** Consider the graph of $N$ vertices with the adjacency matrix

$$a^N_{ij} = \begin{cases} N^\alpha, & \text{if} \quad j = i+1,\ i < N \\ 0, & \text{else} \end{cases}$$

for some constant $\alpha > 0$. Hence

$$\max_{1\leq i\leq N} \sum_{j=1}^{N} a^N_{ij} = N^\alpha.$$

For $0 \leq \alpha < 2$, (**H**) is satisfied, and there is no difference in the two representations in the convergence of the graph sequence. With the atomic mesaure representation $\xi_N \in \mathcal{B}(X, \mathcal{M}_+(X))$:

$$\tilde{\eta}^x_N = N^{\alpha-1}\delta_{(2i+1)/(2N)}, \quad x \in I^N_i, \quad i = 1,\dots,N-1,$$

where $\{I^N_i\}_{i=1}^N$ is given in (2.6). The sequence converges to $\xi(\alpha) \in \mathcal{B}(X, \mathcal{M}_+(X))$ given by

$$\xi^x(\alpha) = \begin{cases} \delta_x, & \text{for } x \in X,\ \alpha = 1 \\ 0, & \text{for } x \in X,\ \alpha < 1 \end{cases} \tag{2.8}$$

which is a non-trivial graph limit if and only if $\alpha = 1$. This shows that the sequence no matter represented in either way, diverges in $\mathcal{B}(X, \mathcal{M}_+(X))$ for $1 < \alpha < 2$. In contrast, using the function representation, one can construct a 2-uniform DHGM (a so-called *digraph measure* (DGM) [31]) $\eta_N \in \mathcal{B}(X, \mathcal{M}_+(X))$ with the density at each $x \in X$ given by $W^N(x, \cdot)$:

$$W^N(x,y) = \frac{\mathrm{d}\eta^x_N(y)}{\mathrm{d}\mu_X(y)} = \begin{cases} N^\alpha & \text{if} \quad x \in I^N_i, \quad y \in I^N_{i+1}, \quad i = 1,\dots,N-1 \\ 0 & \text{else} \quad \text{a.e. } y \in [0,1] \end{cases}$$

It is obvious that

$$\|W^N\|_{L^p(X^2)} = \begin{cases} N^{\alpha-2/p}(N-1)^{1/p} & \text{if } 0 < p < \infty \\ N^\alpha & \text{if } p = \infty \end{cases}$$

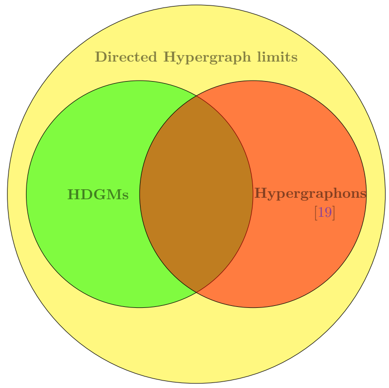


FIGURE 2. Venn diagram for different hypergraph limits.

Hence $\{W^N\}_{N\in\mathbb{N}}$ converges to $0 \in L^p(X^2)$ for $0 < p < \frac{1}{\alpha}$ while diverges in $L^p(X^2)$ for $\frac{1}{\alpha} \le p \le \infty$. Hence in a space of integrable graphons, the sequence of graphs either converges to a trivial limit or diverges. Despite the cut metric is different from the metric induced by the $L^1$-norm induced, the convergence of $\{W^N\}_{N\in\mathbb{N}}$ in $L^1$-norm coincides with the convergence in the cut metric.

*Remark* 2.17.

- The topology induced by the total variation distance is stronger than the weak topology (note that for absolutely continuous measures of the same mass, the distance between two measures in total variation norm is equivalent to the distance of their densities in $L^p$-norm), which results in a fact that the $L^p$ graphons are a proper subset of the space of positive measures in general. This also explains why DHGMs may contain non-dense graph limits.
- Here we provide some explanation on why our definition of hypergraphon differs from the standard one defined in [19], pertaining the dimension of the vertex space. In [19], in order to define limit object of $k$-uniform hypergraphs w.r.t. the topology induced by *homomorphism density*, one needs to regard the limit object as a measurable subset of $[0,1]^{2^k}$, where $2^k$ is the number of subsets of $[k]$. Since the empty set plays no role [35], one can also barely consider measurable subsets of $[0,1]^{2^k-1}$, by excluding the coordinate of a point indexed by the empty set. The reason of considering the limit of hypergraphs as such sets rather than a function is due to the topology induced by some generalized *homomorphism density* [19]. For $k=2$, homomorphism density from graph $\mathcal{F}$ to graph $\mathcal{G}$ refers to as the probability of a random mapping from the vertex set of $\mathcal{F}=([N],\mathcal{E})$ to that of $\mathcal{G}$, whose definition can be extended as

$$t(\mathcal{F},\mathcal{W}) = \int_{[0,1]^N} \prod_{(i,j)\in\mathcal{E}} \mathcal{W}(x_i,x_j)dx_1\dots dx_N, \tag{2.9}$$

representing the homomorphism density from a finite graph $\mathcal{F}$ to a limit object (graphon)–a measurable function $\mathcal{W}\colon [0,1]^2 \to [0,1]$. The convergence characterized in terms of the homomorphism density can also be characterized by the *cut distance* [12, Theorem 3.8]. However, in order to ensure the existence of the limit for a convergent sequence of hypergraphs w.r.t. the distance defined below analogous to (2.9) [35]:

$$t(\mathcal{F},\mathcal{W}) = \int_{[0,1]^N} \prod_{(i_1,\dots,i_k)\in\mathcal{E}} \mathcal{W}(x_{i_1},\dots,x_{i_k})dx_1\dots dx_N, \tag{2.10}$$

for any $k$-uniform test hypergraph $\mathcal{F} = ([N], \mathcal{E})$, a *naïve* straightforward extension to limit of hypergraphs defined as measurable functions on $[0,1]^k$, is *impossible*. The interested reader may refer to [35] for an example. Hence an appropriate generalized definition of homomorphism density as a probability when regarding hypergraph limit as measurable subsets of $[0,1]^{2^k}$, is proposed [19]. The existence of limit theorem, confined to the family of such measurable sets as the space of hypergraph limits, holds w.r.t. this generalized definition of homomorphism density [19]. In short, *such a definition of hypergraph limit is for the completeness of the limit objects under homomorphism density.* In contrast, we treat these limits from the measure-valued function viewpoint, and the completeness is naturally guaranteed due to very basic properties of space of continuous/bounded functions on a (compact) complete metric space. Hence, the set of hypergraph limits as HGMs may not coincide with that defined as hypergraphons in [19]. For instance for $k = 2$, despite the uniform weak topology induced by the uniform bounded Lipschitz metric is stronger than the topology induced by homomophism density (essentially due to the "uniformity" $\sup_{x\in X}$), the space of graph measures [31] is *not* a proper subset of any space of integral graphons as it contains non-absolutely continuous measure-valued functions, evidenced by Example 2.16 as well as examples below. Hence they may share a non-trivial intersection of *dense* graph limits (see Figure 2). In addition, due to the smaller dimension of the underlying space the functions defined in, the set of hypergraphons in the sense of Definition 2.10 should be a subset of that of hypergraphons defined in [19]. It would be of great interest to identify the non-empty (evidenced by, e.g., the counterexample in [35]) complement set and compare the differences, in order to see how narrow our definition can be, so that we may have a chance to further generalize the definition. Conversely, two DHGMs, when viewed as measurable functions on $X^k$, may coincide so long as they differ on a measure zero set. For the same reason, non-dense graph limits may be represented as *non-trivial* DHGMs; but not as (hyper)graphons, they are identified as zero.

- We would like to point out that, graph limits have been defined as measures on the square $X^2$ [32] in the literature, which particularly are used to characterize sparse graphs. Nevertheless, the DGMs [31] as well as HDGMs defined here and in [31] are measure-valued *functions* on $X$, and the motivation of such a definition rather than product measures or kernels (graphons) on the product space owes to the *fiberized characteristic equation* indexed by vertices of graph limits. Nevertheless, though rare in the literature, it is noteworthy that similar topics have been dealt with under the more combinatorial cut metric in a very elegant way in [26]. The results in [26] only pertain sparse graph limits. It will be interesting to see if the approach in [26] can be further generalized to cover non-sparse (hyper)graph limits.

Next, we provide several examples of $k$-uniform HDGMs which are not dense.

**Example 2.18.** *Torical graph measure.* For every $N \in \mathbb{N}$, let $\mathcal{G}^N = (\mathcal{V}^N, \mathcal{E}^N)$ with $\mathcal{V}^N = [N]^2$ and $\mathcal{E}^N = \{((i_1,i_2),(j_1,j_2)) \in (\mathcal{V}^\mathcal{N})^2 \colon i_1 - j_1 = \lfloor N/2 \rfloor \mod N\}$ with the weights

$$a_{(i_1,i_2),(j_1,j_2)} = \begin{cases} 1 & \text{if } j_1 - i_1 = \lfloor N/2 \rfloor \mod N \\ 0 & \text{otherwise} \end{cases}$$

Now we represent $\mathcal{G}^N$ as a *graph measure* [31] $\eta_N \in \mathcal{B}(X, \mathcal{M}_+(X))$ with $X = \mathbb{T}^2$ given as follows:

$$\eta_N^x = \lambda_{\mathbb{T}_{\left(\frac{2i-1}{2N}+\frac{1}{2}\right) \bmod 1}}, \quad x \in \tilde{I}_i^N \times \tilde{I}_j^N, \quad i,j = 1,\dots,N,$$

where

$$\tilde{I}_i^N = [\tfrac{i-1}{N}, \tfrac{i}{N}[ \quad \text{for } i = 1,\dots,N, \tag{2.11}$$

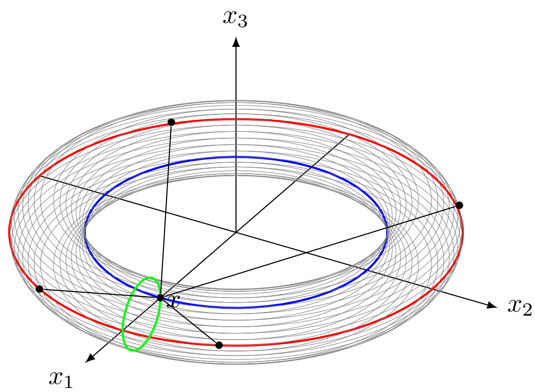


FIGURE 3. Example 2.18. Torical graph measure. Every point on the blue circle connects to every point on the red circle.

and $\mathbb{T}_a \coloneqq \{a\} \times \mathbb{T}$ for every $a \in \mathbb{T}$. Then it is readily shown that $\eta_N$ converges to $\eta \in \mathcal{B}(X, \mathcal{M}_+(X))$ with

$$\eta^x = \lambda_{\mathbb{T}_{(x_1+\frac{1}{2}) \bmod 1}}, \quad x \in X$$

which is an absolutely continuous measure supported on a subset of $X$ of a lower Hausdorff dimension. Geometrically, on the generalized graph $\eta$, every vertex $x = (x_1, x_2)$ on the circle $\mathbb{T}^{x_2} \coloneqq \{y \in X : y_2 = x_2\}$ connects to every vertex on the circle $\mathbb{T}_{(x_1+1/2) \mathrm{mod}\ 1}$ perpendicular to $\mathbb{T}^{x_2}$. We call $\eta$ *torical graph measure.*

**Example 2.19.** *Sparse triangle* 3*-uniform hypergraphing.* Consider a sequence of hypergraphs $\{\mathcal{H}^N\}_N$ of rank 3 with vertex sets being triangularization points of $X$:

$$\mathcal{V}^N = \Big\{\big(\tfrac{i}{2N}, \tfrac{\sqrt{3}}{2} \cdot \tfrac{j}{2N}\big) \colon j \le \min\{2i, 4N - 2i\}, \quad i, j = 0, \ldots, 2N\Big\},$$

$$\mathcal{E}^N = \Big\{\big(\big(\tfrac{i}{2N}, \tfrac{\sqrt{3}}{2} \cdot \tfrac{j}{2N}\big), \big(\tfrac{i+N}{2N}, \tfrac{\sqrt{3}}{2} \cdot \tfrac{j}{2N}\big), \big(\tfrac{i+N}{2N}, \tfrac{\sqrt{3}}{2} \cdot \tfrac{j+N}{2N}\big)\big) \colon \\ j \le \min\{2i, 2N - 2i\}, \quad i, j = 0, \ldots, N\Big\}$$

Let $a^N_{v_1,v_2,v_3} = \begin{cases} 1 & \text{if } (v_1, v_2, v_3) \in \mathcal{E}^N \\ 0 & \text{otherwise} \end{cases}$ We represent $\mathcal{H}^N$ as a 3-uniform HGM $\eta_N \in \mathcal{B}(X, \mathcal{M}_+(X^2))$:

$$\eta^x_N = \begin{cases} \delta_{\left(\frac{2i-1}{4N}, \frac{\sqrt{3}}{2} \cdot \frac{2j-1}{4N}\right)} & \text{if} \quad x \in A^N_i \times B^N_j, \quad j \le \min\{2i, 2N - 2i\}, \quad i, j = 0, \ldots, N \\ 0 & \text{otherwise} \end{cases}$$

Let $X$ be the triangle with vertices $(0, 0)$, $(0, \frac{\sqrt{3}}{2})$, and $(\frac{1}{2}, \frac{1}{2})$. Define $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^2))$:

$$\eta^x = \begin{cases} \delta_{\left(x + \left(\frac{1}{2}, 0\right), x + \left(\frac{1}{4}, \frac{\sqrt{3}}{4}\right)\right)} & \text{if} \quad x \in \frac{1}{2}X \\ 0 & \text{otherwise} \end{cases}$$

where $\frac{1}{2}X = \{y \in X : 2y \in X\}$. It can be readily shown that $d_\infty(\eta_N, \eta) \to 0$ as $N \to \infty$. Note that $\eta$ is a 3-uniform hypergraphing.

Below we provide more examples of $k$-uniform DHGM for general $k$.

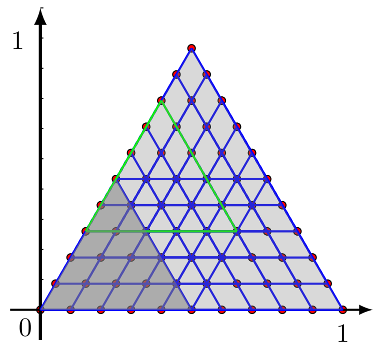


FIGURE 4. 3-uniform triangular hypergraph $\mathcal{H}^5$ in Example 2.19. A typical hyperedge consisting of three nodes as vertices of the triangle is colored in green (all others are blue). The darker shaded area consists of all vertices which belong to one "triangle" hyperedge located in the lower left corner.

**Example 2.20.** Let $k \in \mathbb{N}$. Define $\eta \colon X \to \mathcal{M}_+(X^{k-1})$ by

$$\eta^x = f(x)\delta_{g(x)}, \quad x \in X,$$

where $f \in \mathcal{C}(X, \mathbb{R}_+)$ and $g \in \mathcal{C}(X, X^{k-1})$. Such DHGM can be a $k$-uniform directed hypergraphing.

**Example 2.21.** Let $X = \mathbb{T}$, $\mu_X = \lambda|_{\mathbb{T}}$, and $k \in \mathbb{N} \setminus \{1\}$. Define the *circular $k$-DHGM* $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\eta^x = \otimes_{j=1}^{k-1}(\delta_{(x+1/4) \bmod 1} + \delta_{(x+3/4) \bmod 1}), \quad x \in \mathbb{T}.$$

It is readily verified that $\#\operatorname{supp} \eta^x = 2^{k-1}$, and hence the circular $k$-DHGM is sparse.

**Example 2.22.** Let $X = [0,1]$, $\mu_X = \lambda|_X$, and $k \in \mathbb{N} \setminus \{1\}$. Let $\Delta \coloneqq \{(x_1, \ldots, x_k) \in \mathbb{R}_+^k \colon \sum_{j=1}^k x_j \le 1\}$ and $\Delta_{x_1}$ be the slice of $\Delta$ for $x_1 \in X$. Define $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\eta^{x_1} = k(1-x_1)^{k-1}\lambda|_{\Delta_{x_1}}, \quad \text{for} \quad x_1 \in X.$$

Then it is straightforward to show that $\mu_X \otimes \eta^{x_1}$ can be viewed as the uniform (probability) measure on $\Delta$, which is a dense $k$-DHGM. See Figure 5.

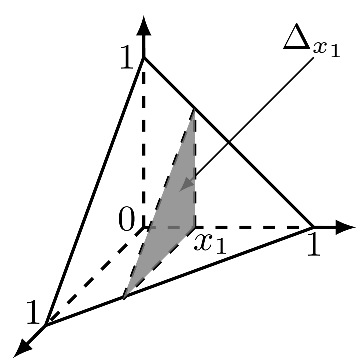


FIGURE 5. Example 2.22 for $k = 3$.

**Example 2.23.** Let $k \in \mathbb{N}\setminus\{1\}$. Consider the following sequence of dense *directed $k$-uniform directed hypergraphs* $\{\mathcal{H}^N\}_{N\in\mathbb{N}}$ with $\mathcal{H}^N = (\mathcal{V}^N, \mathcal{E}^N)$ with $\mathcal{V}^N = [N]$ and

$$\mathcal{E}^N = \left\{(i, j_1, \ldots, j_{k-1})\colon i + \left\lfloor \tfrac{N}{4} \right\rfloor \le j_1, \ldots, j_{k-1} \le i + \left\lfloor \tfrac{3N}{4} \right\rfloor \mod N, \quad 1 \le i \le N\right\}$$

Let $X = \mathbb{T}$, $\mu_X = \lambda|_{\mathbb{T}}$. We can represent this hypergraph sequence as $\eta_N \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\eta_N^x = \mathbb{1}_{\prod_{\ell=1}^{k-1} \tilde{I}^N_{j_\ell}}(y_1, \ldots, y_{k-1}), \quad x \in \tilde{I}^N_i, \quad (i, j_1, \ldots, j_{k-1}) \in \mathcal{E}^N,$$

where $\tilde{I}^N_\ell$ is defined in (2.11). Then it is readily verified that $\eta^N$ converges in the uniform bounded Lipschitz metric to the $k$-uniform DHGM $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$

$$\eta^x = \otimes_{j=1}^{k-1} \lambda|_{[x+1/4, x+3/4[ \bmod 1}, \quad \text{for} \quad x \in X.$$

Note that $\eta$ is also a dense $k$-DHGM since $\inf_{x\in X} \dim_H(\operatorname{supp}\eta^x) = k-1 = \dim_H X^{k-1}$.

**Example 2.24.** Let $k \in \mathbb{N} \setminus \{1, 2\}$ and $X = \mathbb{T}$. Define $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$:

$$\eta^x = \otimes_{j=1}^{k-1} \eta_j^x, \quad x \in X$$

where for $j = 1, \ldots, k-1$,

$$\eta_j^x = \begin{cases} \delta_{(x+1/4) \bmod 1} + \delta_{(x+3/4) \bmod 1} & \text{if } j \text{ is odd}, \\ \lambda|_{[x+1/4, x+3/4[ \bmod 1} & \text{if } j \text{ is even}. \end{cases}$$

Then by Definition 2.11, $\eta$ is neither dense nor sparse.

**Example 2.25.** Let $X = \mathbb{S}^d$ for some $d \in \mathbb{N}$ and $2 \le k \le d$ be an integer. Let $\mathcal{E} = \{(x_1, \ldots, x_k) \in X^k\colon x_i \cdot x_j = 0, \quad \text{for any} \quad i \ne j, \quad 1 \le i, j \le k\}$ be the set of hyperedges. Then $(X, \mathcal{E})$ can be regarded as a generalized $k$-uniform hypergraph. Next, we define $\eta \in \mathcal{B}(X, \mathcal{M}_+(X^{k-1}))$ to represent this generalized hypergraph:

$$\eta^{x_1} = \lambda|_{S(x_1)}, \quad x_1 \in X,$$

where

$$S(x_1) \coloneqq \{(x_2, \ldots, x_k)\colon x_2 \in x_1^\perp,\ x_3 \in x_1^\perp \cap x_2^\perp, \ldots,\ x_\ell \in \cap_{j=1}^{\ell-1} x_j^\perp, \ldots,\ x_k \in \cap_{j=1}^{k-1} x_j^\perp\},$$

where $a^\perp \coloneqq \{y \in X^k\colon a \cdot y = 0\}$. Note that

$$0 < \dim_H(S(x_1)) = \sum_{j=1}^{k-1}(d-j) = \tfrac{(2d-k)(k-1)}{2} < \dim_H(X^{k-1}) = d(k-1)$$

Hence $\eta$ is neither dense nor sparse. We call this $\eta$ *$d$-dimensional spherical $k$-uniform HGM*.

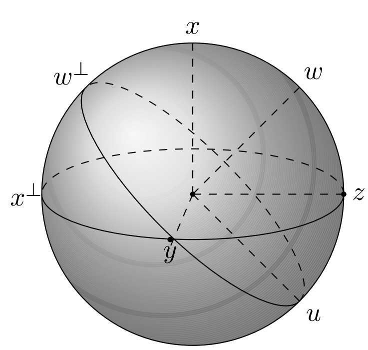


FIGURE 6. 2-dimensional spherical 3-uniform HGM. $X = \mathbb{S}^2$. $\eta^x = \lambda|_{S(x)}$ with $S(x) = \{(y, z) \in X^2\colon y \in x^\perp,\ z \in x^\perp \cap y^\perp\}$. Note that $\dim_H(S(x)) = 1$.

## 3. Generalized interacting particle system

To state a generalized IPS on hypergraphs, we first collect some standing assumptions.

### 3.1. Assumptions.

(**A1**) $(X, \mathfrak{B}(X), \mu_X)$ is a compact Polish probability space equipped with the metric induced by the 1-norm of $\mathbb{R}^{r_1} \supseteq X$.

(**A2**) For $\ell = 1, \ldots, r$, $(t, u) \mapsto g_\ell(t, u) \in \mathbb{R}^{r_2(k_\ell - 1)}$ is continuous in $t \in \mathcal{I}$, and locally Lipschitz continuous in $u \in \mathbb{R}^{k_\ell r_2}$ uniformly in $t$, i.e., for every $u \in \mathbb{R}^{k_\ell r_2}$, there exists a neighbourhood $\mathcal{N} \subseteq \mathbb{R}^{k_\ell r_2}$ of $u$ such that

$$\sup_{t\in\mathcal{I}} \sup_{\substack{u_1 \neq u_2, \\ u_1, u_2 \in \mathcal{N}}} \frac{|g_\ell(t, u_1) - g_\ell(t, u_2)|}{|u_1 - u_2|} < \infty,$$

for $\ell = 1, \ldots, r$.
(**A3**) $(t, x, \phi) \mapsto h(t, x, \phi) \in \mathbb{R}^{r_2}$ is continuous in $t \in \mathcal{I}$, and locally Lipschitz continuous in $\phi \in \mathbb{R}^{r_2}$ uniformly in $(t, x)$, i.e., for every $\phi \in \mathbb{R}^{r_2}$ for some $r_2 \in \mathbb{N}$, there exists a neighbourhood $\mathcal{N} \subseteq \mathbb{R}^{r_2}$ of $\phi$ such that

$$\sup_{t\in\mathcal{I}} \sup_{x\in X} \sup_{\substack{\phi_1 \neq \phi_2, \\ \phi_1, \phi_2 \in \mathcal{N}}} \frac{|h(t, x, \phi_1) - h(t, x, \phi_2)|}{|\phi_1 - \phi_2|} < \infty.$$

(**A4**) $\eta_\ell \in \mathcal{B}(X, \mathcal{M}_+(X^{k_\ell - 1}))$, for $\ell = 1, \ldots, r$.

(**A4**)$'$ $\eta_\ell \in \mathcal{C}(X, \mathcal{M}_+(X^{k_\ell - 1}))$, for $\ell = 1, \ldots, r$.

(**A5**) $\nu_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_+(\mathbb{R}^{r_2})))$ is uniformly compactly supported in the sense that there exists a compact set $E_{\nu_\cdot} \subseteq \mathbb{R}^{r_2}$ such that $\cup_{t\in\mathbb{R}} \cup_{x\in X} \operatorname{supp} \nu_t^x \subseteq E_{\nu_\cdot}$.

(**A6**) There exists a convex compact set $Y \subseteq \mathbb{R}^{r_2}$ such that for all $\nu_\cdot$ satisfying (**A5**) uniformly supported within $Y$, the following inequality holds:

$$V[\eta, \nu_\cdot, h](t, x, \phi) \cdot \upsilon(\phi) \le 0, \quad \text{for all } t \in \mathcal{I},\ x \in X,\ \phi \in \partial Y,$$

where $\partial Y = Y \cap \overline{\mathbb{R}^{r_2} \setminus Y}$, $\upsilon(\phi)$ is the outer normal vector at $\phi$, and

(3.1)
$$V[\eta, \nu_\cdot, h](t, x, \phi) = \sum_{\ell=1}^{r} \int_{X^{k_\ell - 1}} \underbrace{\int_{\mathbb{R}^{r_2}} \cdots \int_{\mathbb{R}^{r_2}}}_{k_\ell - 1} g_\ell(t, \phi, \psi_1, \ldots, \psi_{k_\ell - 1}) \mathrm{d}\nu_t^{y_{k_\ell - 1}}(\psi_{k_\ell - 1}) \cdots \mathrm{d}\nu_t^{y_1}(\psi_1)$$
$$\cdot \mathrm{d}\eta_\ell^x(y_1, \ldots, y_{k_\ell - 1}) + h(t, x, \phi), \quad t \in \mathcal{I}, x \in X,\ \phi \in \mathbb{R}^{r_2}$$

(**A7**) $(t, x, \phi) \mapsto h(t, x, \phi) \in \mathbb{R}^{r_2}$ is continuous in $x$ uniformly in $\phi$:

$$\lim_{|x - x'| \to 0} \sup_{\phi \in Y} |h(t, x, \phi) - h(t, x', \phi)| = 0, \quad t \in \mathcal{I},$$

where $Y$ is the compact set given in (**A6**). Moreover, $h$ is integrable uniformly in $x$:

$$\int_0^T \int_Y \sup_{x\in X} |h(t, x, \phi)| \mathrm{d}\phi \mathrm{d}t < \infty.$$

Under (**A1**)-(**A5**), the *Vlasov operator* $V$ given in (3.1) is well defined.

Now we provide some intuitive explanation for these assumptions. These assumptions can be regarded as analogues of those for digraph measures in [31], in the context of DHGMs. Assumption (**A1**) means that the underlying generalized directed hyper-digraphs (DHGMs) have the same compact vertex space $X$. Such compactness is crucial in establishing discretization of DHGMs. Assumptions (**A2**)-(**A3**) are the standard Lipschitz conditions for the well-posedness of (non-local) ODE models. Assumption (**A4**) means that we interpret the hyper-graphs as measure-valued functions (see also [31]); note that we can think of $\eta_\ell^x$ as the local hyper-edge density or connectivity near vertex $x$. Next, we need the assumption for the approximation of the VE (i.e., the mean field equation for the IPS) that the family of DHGMs $\eta^x$ are *continuous* in the vertex variable $x$, which is encoded in assumption $(\mathbf{A4})'$ (essentially used in Lemma 5.3). As mentioned in [31], $(\mathbf{A4})'$ is sufficient but not necessary for the approximation results. For instance, one can relax this assumption by allowing $x \mapsto \eta_\ell^x$ $(\ell = 1, \ldots, r)$ to have finitely many discontinuity points.

Next, we propose a generalized network on $r$ limits of sequences of directed hypergraphs of cardinality $k_\ell$, for $\ell = 1, \ldots, r$.

$$\begin{aligned} \frac{\mathrm{d}\phi(t,x)}{\mathrm{d}t} =& V[\eta, \nu_\cdot, h](t, x, \phi(t,x)), \quad t \in \mathcal{I}, \quad x \in X, \\ \phi(0,x) =& \varphi(x), \qquad\qquad\qquad\qquad x \in X. \end{aligned} \tag{3.2}$$

Indeed, a special case of the network (3.2) is the following discrete set of ODEs:

(3.3)
$$\dot{\phi}_i^N(t) = h_i^N(t, \phi_i^N(t)) + \sum_{\ell=1}^r \frac{1}{N^{k_\ell-1}} \sum_{j_1=1}^N \cdots \sum_{j_{k_\ell-1}=1}^N W_{i,j_1,\ldots,j_{k_\ell-1}}^{\ell,N} g_\ell(t, \phi_i^N(t), \phi_{j_1}^N(t), \ldots, \phi_{j_{k_\ell-1}}^N(t)).$$

To verify that (3.3) is a special case of (3.2), let $X = [0,1]$, and $\{I_i^N\}_{i=1}^N$ be an equipartition of $X$ with $I_i^N$ being defined in Example 2.6, $\mu_X \in \mathcal{P}(X)$ be the reference measure,

$$\nu_t^x = \delta_{\phi_i^N(t)}, \quad x \in I_i^N, \quad i = 1, \ldots, N,$$

$$\frac{\mathrm{d}\eta_{\ell,N}^x(y_1, \ldots, y_{k_\ell-1})}{\prod_{j=1}^{k_\ell-1} \mathrm{d}\mu_X(y_j)} = W_{i_1,\ldots,i_{k_\ell-1}}^{\ell,N}, \quad (y_1, \ldots, y_{k_\ell-1}) \in I_{i_1}^N \times \ldots \times I_{i_{k_\ell-1}}^N, \quad \ell = 1, \ldots, r,$$

$$\phi(t,x) = \phi_i^N(t), \quad h(t, x, \phi(t,x)) = h_i^N(t, \phi_i^N(t)), \quad x \in I_i^N.$$

Substituting the expressions above into (3.2) simply yields (3.3).

The following well-posedness of the network (3.3) is a standard result of ODE theory [48].

**Proposition 3.1.** *Assume $h_i^N$ (for $i = 1, \ldots, N$) and $g_\ell$ (for $\ell = 1, \ldots, N$) are locally Lipschitz. Then there exists a local solution to the IVP of (3.3). In particular, if there exists a compact positively invariant set for (3.3), then the solution is global.*

The IVP of (3.2) confined to a finite time interval $\mathcal{I}$ is the so-called *fiberized equation of characteristics* (or *fiberized characteristic equation*) [28, 31]. When the underlying space $X$ is finite, and the measures $\nu_t^x$ and $\eta_\ell^x$ for all $x \in X$ are finitely supported, (3.2) becomes a system of ODEs as (3.3) coupled on a finite set of directed graphs in terms of $\{\eta_\ell\}_{1 \le \ell \le r}$. Hence, the fiberized characteristic equation connects a finite-dimensional IPS and the VE, while maintaining the information about both systems. Analogous to Proposition 3.1, the well-posedness of (3.2) is also a standard result from ODE theory [48].

**Theorem 3.2.** *Assume (**A1**)-(**A5**). Let $\phi_0 \in \mathcal{B}(X, \mathbb{R}^{r_2})$. Then for every $x \in X$ and $t_0 \in \mathcal{I}$, there exists a solution $\phi(t,x)$ to the IVP of (3.2) with $\phi(t_0, x) = \phi_0(x)$ for all $t \in (T_{\min}^{x,t_0}, T_{\max}^{x,t_0}) \cap \mathcal{I}$ with $(T_{\min}^{x,t_0}, T_{\max}^{x,t_0}) \subseteq \mathbb{R}$ being a neighbourhood of $t_0$ such that*

*(i) either (i-a) $T_{\max}^{x,t_0} > T$ or (i-b) $T_{\max}^{x,t_0} \le T$ and $\lim_{t \uparrow T_{\max}^{x,t_0}} |\phi(t,x)| = \infty$ holds, and*

*(ii) either (ii-a) $T_{\min}^{x,t_0} < 0$ or (ii-b) $T_{\min}^{x,t_0} \ge 0$ and $\lim_{t \downarrow T_{\min}^{x,t_0}} |\phi(t,x)| = \infty$ holds.*

*In addition, assume* (**A6**) *and* $\nu.$ *is uniformly supported within* $Y$*, then* $(T^{x,t_0}_{\min}, T^{x,t_0}_{\max}) \cap \mathcal{I} = \mathcal{I}$ *for all* $x \in X$*, and there exists a fiberized flow of the vector field* $V[\eta, \nu., h]$ *such that*

$$\begin{aligned}\frac{d}{dt}\Phi^x_{t,0}[\eta,\nu.,h](\phi) =& V[\eta,\nu.,h](t,x,\Phi^x_{t,0}[\eta,\nu.,h](\phi)), \quad t\in\mathcal{I},\\ \Phi^x_{0,0}[\eta,\nu.,h](\phi) =& \phi\end{aligned}$$

## 4. Well-posedness of the Vlasov equation

From Theorem 3.2, we have for all $x \in X$,

$$\Phi^x_{t,0}[\eta,\nu.,h])^{-1} = \Phi^x_{0,t}[\eta,\nu.,h], \quad t \in \mathcal{I}.$$

The pushforward under the flow $\Phi^x_{t,0}[\eta,\nu.,h]$ of an initial measure $\nu^x_0 \in \mathcal{B}_*(X, \mathcal{M}_+(Y))$ defines another time-dependent measure in $\mathcal{B}_*(X, \mathcal{M}_+(Y))$ via the following *fixed point equation*

$$\nu^x_t = (\mathcal{A}[\eta,h]\nu)^x_t, \quad t \in \mathcal{I}. \tag{4.1}$$

where $(\mathcal{A}[\eta,h]\nu)^x_t := \Phi^x_{0,t}[\eta,\nu.,h]_\sharp \nu^x_0$ denotes the push-forwards of $\nu^x_0$ along $\Phi^x_{0,t}[\eta,\nu.,h]$. In particular, if $\nu. \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_{+,\mathsf{abs}}(Y)))$, then $\mathcal{A}[\eta,h]\nu. \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_{+,\mathsf{abs}}(Y)))$ by the positive invariance of $Y$. Hence the *Vlasov operator* can be represented in terms of the density $\rho(t,y,\phi) := \frac{\mathrm{d}\nu^y_t(\phi)}{\mathrm{d}\mu_X(y)\mathrm{d}\phi}$ for every $t \in \mathcal{I}$:

(4.2)

$$\widehat{V}[\eta,\rho(\cdot),h](t,x,\phi) = \sum_{\ell=1}^{r} \int_{X^{k_\ell-1}} \underbrace{\int_Y \cdots \int_Y}_{k_\ell-1} g_\ell(t,\phi,\psi_1,\psi_2,\dots,\psi_{k_\ell-1}) \prod_{\ell=1}^{k_\ell-1} \rho(t,y_\ell,\psi_\ell)\mathrm{d}\psi_1\cdots\mathrm{d}\psi_{k_\ell-1}$$
$$\mathrm{d}\eta^x_\ell(y_1,\dots,y_{k_\ell-1}) + h(t,x,\phi).$$

Let $\mathsf{L}^1(X\times Y; \mu_X\otimes\mathfrak{m})$ be the space of all integrable functions w.r.t. the reference measure $\mu_X \otimes \mathfrak{m}$. Let

$$\mathsf{L}^1_+(X\times Y;\mu_X\otimes\mathfrak{m}) = \Big\{ f \in \mathsf{L}^1(X\times Y;\mu_X\otimes\mathfrak{m}) \colon \int_{X\times Y} f \mathrm{d}\mu_X\mathrm{d}\mathfrak{m} = 1,$$
$$f \ge 0,\ \mu_X\otimes\mathfrak{m} \text{ a.e. on } X\times Y \Big\},$$

be the space of densities of probabilities on $X \times Y$. Conversely, for every function $\rho\colon \mathcal{I} \to \mathsf{L}^1_+(X\times Y;\mu_X\otimes\mathfrak{m})$, for $(t,y) \in \mathcal{I}\times X$,

$$\mathrm{d}\nu^y_t(\phi) = \rho(t,y,\phi)\mathrm{d}\phi$$

defines $\nu. \in \mathcal{B}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_+(Y)))$. Hence (4.2) can be transformed to the Vlasov operator (3.1) in terms of $\nu.$.

Let $\rho_0 \colon X \times Y \to \mathbb{R}_+$ be continuous in $x$ for $\mathfrak{m}$-a.e. $\phi \in Y$, and integrable in $\phi$ for every $x \in X$ such that

$$\int_X \int_Y \rho_0(x,\phi)\mathrm{d}\phi\mathrm{d}\mu_X(x) = 1.$$

Consider the VE

$$\begin{aligned}&\frac{\partial\rho(t,x,\phi)}{\partial t} + \mathrm{div}_\phi\left(\rho(t,x,\phi)\widehat{V}[\eta,\rho(\cdot),h](t,x,\phi)\right) = 0, t\in(0,T],\ x\in X,\ \mathfrak{m}\text{-a.e. } \phi\in Y,\\ &\rho(0,\cdot) = \rho_0(\cdot).\end{aligned} \tag{4.3}$$

First, let us define the weak solution to (4.3). This definition is analogous to [31, Definition 4.6].

**Definition 4.1.** Let $Y$ be a compact positively invariant subset of (3.2) given in Theorem 3.2. We say $\rho\colon \mathcal{I}\times X\times Y\to\mathbb{R}^+$ is a *uniformly weak solution* to the IVP (4.3) if for every $x\in X$, the following three conditions are satisfied:

(i) Normalization. $\int_X\int_Y \rho(t,x,\phi)\mathrm{d}\phi\mathrm{d}x=1$, for all $t\in\mathcal{I}$.

(ii) Uniform weak continuity. The map $t\mapsto\int_Y f(\phi)\rho(t,x,\phi)\mathrm{d}\phi$ is continuous uniformly in $x\in X$, for every $f\in\mathcal{C}(Y)$.

(iii) Integral identity: For all test functions $w\in\mathcal{C}^1(\mathcal{I}\times Y)$ with $\operatorname{supp} w\subseteq[0,T[\times U$ and $U\subset\subset Y$, the equation below holds:

$$\int_0^T\int_Y \rho(t,x,\phi)\left(\frac{\partial w(t,\phi)}{\partial t}+\widehat{V}[\eta,\rho(\cdot),h](t,x,\phi)\cdot\nabla_\phi w(t,\phi)\right)\mathrm{d}\phi\mathrm{d}t \\ +\int_Y w(0,\phi)\rho_0(x,\phi)\mathrm{d}\phi=0, \tag{4.4}$$

where $\operatorname{supp} w=\overline{\{(t,u)\in\mathcal{I}\times Y\colon w(t,u)\neq 0\}}$ is the support of $w$, and $\widehat{V}[\eta,\rho(\cdot),h]$ is given in (4.2).

The well-posedness of the VE associated with the generalized IPS depends on continuity properties of the operator $\mathcal{A}$. We comment that Definition 4.1 is well-posed, c.f. [31, Remark 4.7].

We first provide the continuity property of $\mathcal{A}$.

**Proposition 4.2.** *Assume* (**A1**)-(**A6**). *Denote* $\eta=(\eta_\ell)_{\ell=1}^r$.

*(i) Continuity in $t$.*

$$t\mapsto(\mathcal{A}[\eta,h]\nu_t\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(Y))).$$

*In particular, if* $\nu_\cdot\in\mathcal{C}(\mathcal{I},\,\mathcal{C}_*(X,\mathcal{M}_+(Y)))$, *then* $\mathcal{A}[\eta,h]\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\,\mathcal{M}_+(Y)))$. *Moreover, the mass conservation law holds:*

$$\mathcal{A}[\eta,h]\nu_t^x(Y)=\nu_0^x(Y),\quad\forall x\in X.$$

*(ii) Lipschitz continuity in $\nu_\cdot$. For $\nu^1,\nu^2\in\mathcal{C}_*(X,\mathcal{M}_+(Y)))$, we have*

$$d_\infty(\mathcal{A}[\eta,h]\nu_t^1,\mathcal{A}[\eta,h]\nu_t^2)\le \mathrm{e}^{L_1t}d_\infty(\nu_0^1,\nu_0^2)+L_2\mathrm{e}^{L_1t}\int_0^t d_\infty(\nu_\tau^1,\nu_\tau^2)\mathrm{e}^{-L_1\tau}\mathrm{d}\tau,$$

*where $L_1\coloneqq L_1(\nu_\cdot^2)$ and $L_2\coloneqq L_2(\nu_\cdot^1,\nu_\cdot^2)$ are constants.*

*(iii) Lipschitz continuity of $\mathcal{A}[\eta,h]$ in $h$. For $h_1$, $h_2$ satisfying* (**A3**) *and* (**A7**) *with $h$ replaced by $h_i$ for $i=1,2$,*

$$d_\infty(\mathcal{A}[\eta,h_1]\nu_t,\mathcal{A}[\eta,h_2]\nu_t)\le L_3\|h_1-h_2\|_{\infty,\mathcal{I}},$$

*where $L_3\coloneqq L_3(\nu_\cdot)$ is a constant.*

*(iv) Absolute continuity. If $\nu_0\in\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y))$, then*

$$\mathcal{A}[\eta,h]\nu_t\in\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y)),\quad\forall t\in\mathcal{I}.$$

The proof of Proposition 4.2 is provided in Appendix A.

**Proposition 4.3.** *Assume* (**A1**)-(**A4**) *and* (**A6**)-(**A7**). *Let $\nu_0\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$, and $L_1$, $L_2$, and $L_3$ be defined as in Proposition 4.2. Then there exists a unique solution $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(Y)))$ to the fixed point equation (4.1). In addition, if* (**A4**)$'$ *holds and $\nu_0\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$, then $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\,\mathcal{M}_+(Y)))$; if $\nu_0\in\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y))$, then*

$$\nu_t\in\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y)),\quad\forall t\in\mathcal{I}.$$

*Moreover, the solutions have continuous dependence on*

(i) *the initial conditions:*

$$d_\infty(\nu_t^1,\nu_t^2)\le \mathrm{e}^{(L_1(\nu_\cdot^2)+L_2\|\nu_\cdot^1\|)t}d_\infty(\nu_0^1,\nu_0^2),\quad t\in\mathcal{I},$$

*where $\nu_\cdot^i$ is the solution to (4.1) with initial condition $\nu_0^i$ for $i=1,2$.*

(ii) *$h$: Assume $\nu_0^1=\nu_0^2$. Then*

$$d_\infty(\nu_t^1,\nu_t^2)\le \frac{1}{L_3(\nu_\cdot^2)}\|\nu_\cdot^1\|\mathrm{e}^{\mathcal{BL}(h_2)\mathrm{e}^{L_1(\nu_\cdot^2)T}T}\mathrm{e}^{(L_1(\nu_\cdot^2)+L_2\|\nu_\cdot^1\|)t}\|h_1-h_2\|_\infty,$$

*where $\nu_\cdot^i$ is the solution to (4.1) with $h$ replaced by $h_i$ for $i=1,2$.*

(iii) *$\eta$: Let $\{\eta_\ell^K\}_{K\in\mathbb{N}}\subseteq\mathcal{B}(X,\mathcal{M}_+(X^{k_\ell-1}))$ such that $\lim_{K\to\infty}d_\infty(\eta_\ell,\eta_\ell^K)=0$, for $\ell=1,\ldots,r$. Assume $\nu_0,\ \nu_0^K\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$ with*

$$\lim_{K\to\infty}d_\infty(\nu_0,\nu_0^K)=0.$$

*Then*

$$\lim_{K\to\infty}\sup_{t\in\mathcal{I}}d_\infty(\nu_t,\nu_t^K)=0,$$

*where $\nu_\cdot^K$ is the solution to (4.1) with $\eta_\ell$ replaced by $\eta_\ell^K$ for $\ell=1,\ldots,r$ and $K\in\mathbb{N}$.*

The proof of Proposition 4.3 is provided in Appendix B.

With the above assumptions and under appropriate metrics, one can show that the operator defined in (4.1)

$$\mathcal{A}=(\mathcal{A}^x)_{x\in X}\colon\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(Y)))\to\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(Y)))$$

is a contraction. Now we obtain the well-posedness of the VE (4.3).

**Theorem 4.4.** *Assume* (**A1**)*-*(**A4**) *and* (**A6**)*. Let $\rho_0\in\mathsf{L}_+^1(X\times Y;\mu_X\otimes\mathfrak{m})$. Assume additionally that $\rho_0(x,\phi)$ is continuous in $x\in X$ for $\mathfrak{m}$-a.e. $\phi\in Y$. Then there exists a unique uniform weak solution to the IVP of (4.3) with initial condition $\rho(0,x,\phi)=\rho_0(x,\phi)$, $x\in X$, $\phi\in Y$.*

*Proof.* The proof is the same as that of [31, Theorem 4.8], which is independent of the specific form of $V$ and $\widehat{V}$, but based on the continuous dependence properties given in Proposition 4.2 and Proposition 4.3. $\square$

## 5. Approximation of time-dependent solutions to VE

In this section, we study approximation of the solution to the VE (4.4).

Based on the continuous dependence of solutions to the fixed point equation on the underlying DHGMs $\{\eta_\ell\}_{\ell=1}^r$, on the initial measure $\nu_0$, as well as on function $h$ established in Proposition 4.3, together with the recently established results on *deterministic empirical approximation of positive measures* [50, 14, 3] (see Propositions 5.2 and 5.3 below), we will establish the discretization of solutions of VE over finite time interval $\mathcal{I}$ by a sequence of discrete ODE systems coupled on finite directed hypergraphs converging *weakly* to the DHGMs $\{\eta_\ell\}_{\ell=1}^r$ (Theorem 5.6 below).

Beforehand, let us recall some approximation results from [31].

**Proposition 5.1** (Partition of $X$)**.** [31, Lemma 5.4] *Assume* (**A1**)*. Then there exists a sequence of pairwise disjoint partitions $\{A_i^m\colon i=1,\ldots,m\}_{m\in\mathbb{N}}$ of $X$ such that $X=\cup_{i=1}^mA_i^m$ for every $m\in\mathbb{N}$ and*

$$\lim_{m\to\infty}\max_{1\le i\le m}\mathit{Diam}\,A_i^m=0.$$

**Proposition 5.2** (Approximation of the initial distribution)**.** [31, Lemma 5.5] *Assume* (**A1**) *and $\nu_0\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$. Let $\{A_i^m\}_{1\le i\le m}$ be a partition of $X$ for $m\in\mathbb{N}$ satisfying*

$$\lim_{m\to\infty}\max_{1\le i\le m}\mathit{Diam}\,A_i^m=0.$$

*Let $x_i^m \in A_i^m$, for $i=1,\dots,m$, $m\in\mathbb{N}$. Then there exists a sequence $\{\varphi^{m,n}_{(i-1)n+j}\colon i=1,\dots,m, j=1,\dots,n\}_{n,m\in\mathbb{N}}\subseteq Y$ such that*

$$\lim_{m\to\infty}\lim_{n\to\infty} d_\infty(\nu_0^{m,n},\nu_0)=0,$$

*where $\nu_0^{m,n}\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$ with*

$$\nu_0^{m,n,x}\coloneqq\sum_{i=1}^m \mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n\delta_{\varphi^{m,n}_{(i-1)n+j}},\quad x\in X,$$

$$a_{m,i}=\begin{cases}\dfrac{\int_{A_i^m}\nu_0^x(Y)\mathrm{d}\mu_X(x)}{\mu_X(A_i^m)}, & \text{if}\quad \mu_X(A_i^m)>0,\\ \nu_0^{x_i^m}(Y), & \text{if}\quad \mu_X(A_i^m)=0.\end{cases}$$

**Proposition 5.3** (Approximation of the DHGM)**.** *Assume* (**A1**) *and* (**A4**)$'$. *For every $m\in\mathbb{N}$, let $A_i^m$ and $x_i^m$ be defined in Proposition 5.2 for $i=1,\dots,m$, $m\in\mathbb{N}$. Then for every $\ell=1,\dots,r$, there exists a sequence $\{y^{\ell,m,n}_{(i-1)n+j}\colon i=1,\dots,m, j=1,\dots,n\}_{m,n\in\mathbb{N}}\subseteq X^{k_\ell-1}$ such that*

$$\lim_{m\to\infty}\lim_{n\to\infty} d_\infty(\eta_\ell^{m,n},\eta_\ell)=0,$$

*where $\eta_\ell^{m,n}\in\mathcal{B}(X,\mathcal{M}_+(X^{k_\ell-1}))$ with*

$$\eta_\ell^{m,n,x}\coloneqq\sum_{i=1}^m \mathbb{1}_{A_i^m}(x)\frac{b_{\ell,m,i}}{n}\sum_{j=1}^n\delta_{y^{\ell,m,n}_{(i-1)n+j}},\quad x\in X,$$

$$b_{\ell,m,i}=\begin{cases}\dfrac{\int_{A_i^m}\eta_\ell^x(X)\mathrm{d}\mu_X(x)}{\mu_X(A_i^m)}, & \text{if}\quad \mu_X(A_i^m)>0,\\ \eta_\ell^{x_i^m}(X), & \text{if}\quad \mu_X(A_i^m)=0.\end{cases}$$

The proof of Proposition 5.3 is analogous to that of [31, Lemma 5.6] and thus is omitted.

**Proposition 5.4** (Approximation of $h$)**.** [31, Lemma 5.9] *Assume* (**A3**) *and* (**A7**).

*For every $m\in\mathbb{N}$, let $x_i^m$ be defined in Proposition 5.2 and*

$$h^m(t,z,\phi)=\sum_{i=1}^m\mathbb{1}_{A_i^m}(z)h(t,x_i^m,\phi),\quad t\in\mathcal{I},\ z\in X,\ \phi\in Y.$$

*Then*

$$\lim_{m\to\infty}\int_0^T\int_Y\sup_{x\in X}|h^m(t,x,\phi)-h(t,x,\phi)|\,\mathrm{d}\phi\mathrm{d}t=0.$$

Now we are ready to provide a discretization of the VE on the DHGM by a sequence of ODEs. To summarize, there exists

- a partition $\{A_i^m\}_{1\le i\le m}$ of $X$ and points $x_i^m\in A_i^m$ for $i=1,\dots,m$, for every $m\in\mathbb{N}$,
- a sequence $\{\varphi^{m,n}_{(i-1)n+j}\colon i=1,\dots,m, j=1,\dots,n\}_{n,m\in\mathbb{N}}\subseteq Y^{k_\ell-1}$ and $\{a_{m,i}\colon i=1,\dots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$, for $\ell=1,\dots,r$, and
- a sequence $\{y^{\ell,m,n}_{(i-1)n+j}\colon i=1,\dots,m, j=1,\dots,n\}_{m,n\in\mathbb{N}}\subseteq X^{k_\ell-1}$ and $\{b_{\ell,m,i}\colon i=1,\dots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$, for $\ell=1,\dots,r$,

such that

$$\lim_{m\to\infty}\lim_{n\to\infty} d_\infty(\nu_0^{m,n},\nu_0)=0,$$

$$\lim_{m\to\infty}\lim_{n\to\infty} d_\infty(\eta_\ell^{m,n},\eta_\ell)=0,\quad \ell=1,\dots,r,$$

$$\lim_{m\to\infty}\int_0^T\int_Y\sup_{x\in X}|h^m(t,x,\phi)-h(t,x,\phi)|\,\mathrm{d}\phi\mathrm{d}t=0,$$

where

$$
\begin{aligned}
\nu_0^{m,n,x} &:= \sum_{i=1}^m \mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n \delta_{\varphi^{m,n}_{(i-1)n+j}}, \quad x\in X,\\
\eta_\ell^{m,n,x} &:= \sum_{i=1}^m \mathbb{1}_{A_i^m}(x)\frac{b_{\ell,m,i}}{n}\sum_{j=1}^n \delta_{y^{\ell,m,n}_{(i-1)n+j}}, \quad x\in X,\\
h^m(t,z,\phi) &:= \sum_{i=1}^m \mathbb{1}_{A_i^m}(z)h(t,x_i^m,\phi), \quad t\in\mathcal{I},\ z\in X,\ \phi\in Y.
\end{aligned}
\tag{5.1}
$$

Consider the following IVP of a coupled ODE system:

$$
\dot\phi_{(i-1)n+j} = F_i^{m,n}(t,\phi_{(i-1)n+j},\Phi),\quad 0<t\le T,\quad \phi_{(i-1)n+j}(0)=\varphi^{m,n}_{(i-1)n+j},\\
i=1,\dots,m,\ j=1,\dots,n,
\tag{5.2}
$$

where $\Phi=(\phi_{(i-1)n+j})_{1\le i\le m,1\le j\le n}$ and

$$
\begin{aligned}
F_i^{m,n}(t,\psi,\Phi) = &\sum_{\ell=1}^r \frac{b_{\ell,m,i}}{n}\sum_{j=1}^n\sum_{p_1=1}^m \frac{a_{m,p_1}}{n}\mathbb{1}_{A^m_{p_1}}(y^{\ell,m,n}_{(i-1)n+j,1})\cdots\sum_{p_{k_\ell-1}=1}^m\frac{a_{m,p_{k_\ell-1}}}{n}\\
&\mathbb{1}_{A^m_{p_{k_\ell-1}}}(y^{\ell,m,n}_{(i-1)n+j,k_\ell-1})\sum_{q_1=1}^n\cdots\sum_{q_{k_\ell-1}=1}^n\\
&g_\ell(t,\psi,\phi^{m,n}_{(p_1-1)n+q_1},\dots,\phi^{m,n}_{(p_{k_\ell-1}-1)n+q_{k_\ell-1}})+h^m(t,x_i^m,\psi)
\end{aligned}
$$

The following well-posedness result is akin to Proposition 3.1 and hence the proof is omitted.

**Proposition 5.5.** *Then there exists a unique solution $\phi^{m,n}(t)=(\phi^{m,n}_{(i-1)n+j}(t))$ to (5.2), for $m,n\in\mathbb{N}$.*

Based on Proposition 5.5, let

$$
\nu_t^{m,n,x} := \sum_{i=1}^m \mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n \delta_{\phi^{m,n}_{(i-1)n+j}(t)}, \quad x\in X.
\tag{5.3}
$$

Now we present the approximation of solutions to the VE (4.3).

**Theorem 5.6.** *Assume (**A1**)-(**A3**), (**A4**)$'$, (**A6**)-(**A7**). Assume $\rho_0(x,\phi)$ is continuous in $x\in X$ for $\mathfrak{m}$-a.e. $\phi\in Y$ such that $\rho_0\in\mathsf{L}^1_+(X\times Y;\mu_X\otimes\mathfrak{m})$ and*

$$
\sup_{x\in X}\|\rho_0(x,\cdot)\|_{\mathsf{L}^1(Y;\mathfrak{m})}<\infty.
$$

*Let $\rho(t,x,\phi)$ be the uniformly weak solution to the VE (4.3) with initial condition $\rho_0$. Let $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y)))$ be the measure-valued function defined in terms of the uniform weak solution to (4.3):*

$$
\mathrm{d}\nu_t^x(\phi)=\rho(t,x,\phi)\mathrm{d}\phi,\quad \textit{for every}\quad t\in\mathcal{I}\quad\textit{and}\quad x\in X,\quad \mathfrak{m}\ \textit{a.e.}\ \phi\in Y.
$$

*Then $\nu_t\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$, for all $t\in\mathcal{I}$, provided $\nu_0\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$. Moreover, let $\nu_0^{m,n}\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$, $\eta^{\ell,m,n}\in\mathcal{B}(X,\mathcal{M}_+(X^{k_\ell-1}))$, and $h^m\in\mathcal{C}(\mathcal{I}\times X\times Y,\mathbb{R}^{r_2})$ be defined in (5.1), and $\nu_t^{m,n}$ be defined in (5.3). Then*

$$
\lim_{n\to\infty} d_0(\nu_\cdot^{m,n},\nu_\cdot)=0.
$$

The proof of Theorem 5.6 is provided in Section 8.

## 6. Applications

In this section, we apply our main results to investigate the MFL of three networks with higher order interactions emerging from physics, epidemiology, and ecology.

6.1. **A Kuramoto-Sakaguchi model with heterogeneous higher order interactions.** Consider the following Kuramoto-Sakaguchi phase reduction network proposed in [5]:

$$\begin{aligned}\dot{\phi}_i^N =& h_i^N + \frac{1}{N}\sum_{j=1}^N W_{i,j}^{(1,N)} g_1(\phi_i^N,\phi_j^N) + \frac{1}{N^2}\sum_{j=1}^N\sum_{k=1}^N W_{i,j,k}^{(2,N)} g_2(\phi_i^N,\phi_j^N,\phi_k^N)\\ &+ \frac{1}{N^3}\sum_{j=1}^N\sum_{k=1}^N\sum_{p=1}^N W_{i,j,k,p}^{(3,N)} g_3(\phi_i^N,\phi_j^N,\phi_k^N,\phi_p^N),\end{aligned}$$

where $\phi_i^N\in\mathbb{T}$ is the phase and $h_i$ the natural frequency of the $i$-th oscillator, and $g_\ell$ ($\ell=1,2,3$) are the coupling functions of different higher-order interactions.

Let $(I_j^N)_{1\le j\le N}$ be an equi-partition of $X=[0,1]$ defined as in(2.6), and let $h^N\coloneqq\sum_{j=1}^N\mathbb{1}_{I_j^N}h_j^N$.

Assume

(**H1.1**) $g_\ell$ ($\ell=1,2,3$) are Lipschitz continuous.
(**H1.2**) $h_j^N$ ($j=1,\ldots,N$) fulfill that there exists $h\in\mathcal{C}(X)$ such that

$$\lim_{N\to\infty}\sup_{x\in X}|h(x)-h^N(x)|=0.$$

(**H1.3**) For $\ell=1,2,3$, $W^{(\ell,N)}$ converges in the uniform bounded Lipschitz metric to $\eta_\ell\in\mathcal{C}(X,\mathcal{M}_+(X^\ell))$.

Note that (**H1.2**) implies that

$$\lim_{N\to\infty}\|h-h^N\|_1=0.$$

Now we consider the VE

$$\begin{aligned}&\frac{\partial\rho(t,x,\phi)}{\partial t}+\mathrm{div}_\phi\Big(\rho(t,x,\phi)\widehat{V}[\eta,\rho,h](t,x,\phi)\Big)=0,\ t\in(0,T],\ x\in X,\ \mathfrak{m}\text{-a.e. }\phi\in\mathbb{T},\\ &\rho(0,\cdot)=\rho_0(\cdot),\end{aligned}\tag{6.1}$$

where

$$\begin{aligned}\widehat{V}[\eta,\rho,h](t,x,\phi)=&h(x)+\int_X\int_{\mathbb{T}} g_1(\phi(t,x),\psi)\rho(t,y,\psi)\mathrm{d}\psi\mathrm{d}\eta_1^x(y)\\ &+\int_{X^2}\int_{\mathbb{T}}\int_{\mathbb{T}} g_2(\phi(t,x),\psi_1,\psi_2)\rho(t,y_1,\psi_1)\rho(t,y_2,\psi_2)\mathrm{d}\psi_1\mathrm{d}\psi_2\mathrm{d}\eta_2^x(y_1,y_2)\\ &+\int_{X^3}\int_{\mathbb{T}}\int_{\mathbb{T}}\int_{\mathbb{T}} g_3(\phi(t,x),\psi_1,\psi_2,\psi_3)\rho(t,y_1,\psi_1)\rho(t,y_2,\psi_2)\rho(t,y_3,\psi_3)\\ &\mathrm{d}\psi_1\mathrm{d}\psi_2\mathrm{d}\psi_3\mathrm{d}\eta_3^x(y_1,y_2,y_3).\end{aligned}$$

**Theorem 6.1.** *Assume* (**H1.1**)-(**H1.3**). *Let $T>0$. Assume $\rho_0(x,\phi)$ is continuous in $x\in X$ for $\mathfrak{m}$-a.e. $\phi\in Y$ such that $\rho_0\in\mathsf{L}^1_+(X\times Y;\mu_X\otimes\mathfrak{m})$ and*

$$\sup_{x\in X}\|\rho_0(x,\cdot)\|_{\mathsf{L}^1(Y;\mathfrak{m})}<\infty.$$

*Define $\nu_0\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$ by*

$$\rho_0(x,\phi)=\frac{\mathrm{d}\nu_0(x,\phi)}{\mathrm{d}\phi}\quad\text{for}\quad x\in X\quad\text{and}\quad\mathfrak{m}\text{-a.e.}\quad\phi\in\mathbb{T}.$$

*Then there exists a unique uniformly weak solution $\rho(t,\cdot)$ to the VE (6.1). Moreover, if $\nu_0 \in \mathcal{C}_*(X,\mathcal{M}_+(Y))$ and $\lim_{N\to\infty} d_\infty(\nu_{N,0},\nu_0)=0$, then*

$$\lim_{N\to\infty} d_0(\nu_{N,\cdot},\nu_\cdot)=0.$$

*In particular,*

$$\lim_{N\to\infty} d_{\mathsf{BL}}\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\phi_i^N(t)}, \int_X \nu_t^x(\bullet)\mathrm{d}\mu_X(x)\Big)=0, \quad \textit{for} \quad t\in\mathcal{I}.$$

*Proof.* The proof is similar to those of Theorem 4.4 and Theorem 5.6. Here we still have the invariance under the flow of fiberized equation of characteristics of a compact set which is $\mathbb{T}$ (in lieu of a compact set $Y$ in a Euclidean space). □

6.2. **An epidemic model with higher-order interactions.** Assume

(**H2.1**) $(X,\mathfrak{B}(X),\mu_X)$ is a compact probability space.

(**H2.2**) For $(u_1,u_2)\in\mathbb{R}_+^2$, let $\beta(t,u_1,u_2)\ge 0$ be the disease transmission function satisfying $\beta(t,u_1,u_2)=0$ provided $u_1u_2=0$. Moreover, $\beta$ is continuous in $t$, and locally Lipschitz continuous in $u_1,\ u_2$ uniformly in $t$.

(**H2.3**) For $u\in\mathbb{R}_+$, let $\gamma(t,x,u)\ge 0$ be the recovery rate function, and for every $x\in X$, $\gamma(t,x,0)=0$. Moreover, $\gamma$ is continuous in $t$, and continuous in $x$ uniformly in $u$, and Lipschitz continuous in $u\in\mathbb{R}_+$ uniformly in $t$.

For any fixed $N\in\mathbb{N}$, let

$$Y=\{u\in\mathbb{R}_+^2\colon u_1+u_2=N\}. \tag{6.2}$$

(**H2.4**) $\eta\in\mathcal{C}(X,\mathcal{M}_+(X^2))$.

(**H2.5**) $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(\mathbb{R}^2)))$ is uniformly compactly supported within $Y\subseteq\mathbb{R}_+^2$.

Under (**H2.1**)-(**H2.5**), motivated by [9], we propose the following generalized non-local multi-group SIS epidemic model on a DHGM $\eta$ incorporating the higher-order interactions due to the nonlinear dependence of both the infection pressure and the community structure (home and workplace):

$$\begin{aligned}\frac{\partial S_x}{\partial t}=&-\int_{X^2}\int_{\mathbb{R}_+^2}\int_{\mathbb{R}_+^2}\left(\beta(t,\psi_{1,2},S_x)+\beta(t,\psi_{2,2},S_x)\right)\mathrm{d}\nu_t^{y_2}(\psi_2)\mathrm{d}\nu_t^{y_1}(\psi_1)\mathrm{d}\eta^x(y_1,y_2)\\&+\gamma(t,x,I_x),\\ \frac{\partial I_x}{\partial t}=&\int_{X^2}\int_{\mathbb{R}_+^2}\int_{\mathbb{R}_+^2}\left(\beta(t,\psi_{1,2},S_x)+\beta(t,\psi_{2,2},S_x)\right)\mathrm{d}\nu_t^{y_2}(\psi_2)\mathrm{d}\nu_t^{y_1}(\psi_1)\mathrm{d}\eta^x(y_1,y_2)\\&-\gamma(t,x,I_x),\end{aligned}$$

which further generalizes the epidemic network on a digraph measure proposed in [31]. Here $S_x$ and $I_x$ stand for the number of susceptible and infected individuals at location $x\in X$ (or interpreted as in the group with label $x$), $\beta(t,\psi_2^1,S_x)$ stands for the infection caused by family members of $S_x$ at home while $\beta(t,\psi_2^2,S_x)/S_x$ the infection rate caused by colleagues of $S_x$ in the workplace, and $\eta\colon X\to\mathcal{M}_+(X^2)$ is the generalized hypergraph.

By (**H2.3**), let

$$g(t,\phi,\psi^1,\psi^2)=\beta(t,\psi_{1,2},\phi_1)+\beta(t,\psi_{2,2},\phi_1)\begin{pmatrix}-1\\1\end{pmatrix},\quad h(t,x,\phi)=\gamma(t,x,\phi_1)\begin{pmatrix}1\\-1\end{pmatrix},$$

$$V[\eta,\nu_\cdot,h](t,x,\phi)=\int_{X^2}\int_Y\int_Y g(t,\phi,\psi_1,\psi_2)\mathrm{d}\nu_t^{y_2}(\psi_2)\mathrm{d}\nu_t^{y_1}(\psi_1)\mathrm{d}\eta^x(y_1,y_2)+h(t,x,\phi),$$

and

$$\widehat{V}[\eta,\rho.,h](t,x,\phi)=\int_{X^2}\int_Y\int_Y g(t,\phi,\psi_1,\psi_2)\rho(t,y_1,\psi_1)\rho(t,y_2,\psi_2)\mathrm{d}\psi_2\mathrm{d}\psi_1\mathrm{d}\eta^x(y_1,y_2)+h(t,x,\phi).$$

Consider the VE

(6.3)

$$\frac{\partial\rho(t,x,\phi)}{\partial t}+\mathrm{div}_\phi\left(\rho(t,x,\phi)\widehat{V}[\eta,\rho(\cdot),h](t,x,\phi)\right)=0,\quad t\in(0,T],\ x\in X,\ \mathfrak{m}\text{-a.e. }\phi\in Y,$$
$$\rho(0,\cdot)=\rho_0(\cdot).$$

According to Propositions 5.1-5.4, there exists

- a partition $\{A_i^m\}_{1\le i\le m}$ of $X$ and points $x_i^m\in A_i^m$ for $i=1,\ldots,m$, for every $m\in\mathbb{N}$,
- a sequence $\{\varphi_{(i-1)n+j}^{m,n}\coloneqq i=1,\ldots,m,j=1,\ldots,n\}_{n,m\in\mathbb{N}}\subseteq Y$ and $\{a_{m,i}\colon i=1,\ldots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$, and
- a sequence $\{y_{(i-1)n+j}^{m,n}\colon i=1,\ldots,m,j=1,\ldots,n\}_{m,n\in\mathbb{N}}\subseteq X^2$ and $\{b_{m,i}\colon i=1,\ldots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$,

such that

$$\lim_{m\to\infty}\lim_{n\to\infty}d_\infty(\nu_0^{m,n},\nu_0)=0,$$
$$\lim_{m\to\infty}\lim_{n\to\infty}d_\infty(\eta^{m,n},\eta)=0,$$
$$\lim_{m\to\infty}\int_0^T\int_Y\sup_{x\in X}|h^m(t,x,\phi)-h(t,x,\phi)|\,\mathrm{d}\phi\mathrm{d}t=0,$$

where

$$\nu_0^{m,n,x}\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n\delta_{\varphi_{(i-1)n+j}^{m,n}},\quad x\in X,$$
$$\eta^{m,n,x}\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(x)\frac{b_{m,i}}{n}\sum_{j=1}^n\delta_{y_{(i-1)n+j}^{m,n}},\quad x\in X,\tag{6.4}$$
$$h^m(t,z,\phi)\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(z)h(t,x_i^m,\phi),\quad t\in\mathcal{I},\ z\in X,\ \phi\in Y.$$

Consider the following IVP of a coupled ODE system:

$$\dot{\phi}_{(i-1)n+j}=F_i^{m,n}(t,\phi_{(i-1)n+j},\Phi),\quad 0<t\le T,\quad \phi_{(i-1)n+j}(0)=\varphi_{(i-1)n+j}^{m,n},\quad i=1,\ldots,m,\ j=1,\ldots,n,\tag{6.5}$$

where $\Phi=(\phi_{(i-1)n+j})_{1\le i\le m,1\le j\le n}$ and

$$F_i^{m,n}(t,\psi,\Phi)=\frac{b_{m,i}}{n}\sum_{j=1}^n\sum_{p_1=1}^m\frac{a_{m,p_1}}{n}\mathbb{1}_{A_{p_1}^m}(y_{(i-1)n+j,1}^{m,n})\sum_{p_2=1}^m\frac{a_{m,p_2}}{n}\mathbb{1}_{A_{p_2}^m}(y_{(i-1)n+j,2}^{m,n})$$
$$\sum_{q_1=1}^n\sum_{q_2=1}^n g(t,\psi,\phi_{(p_1-1)n+q_1},\phi_{(p_2-1)n+q_2})+h^m(t,x_i^m,\psi).$$

Then by Proposition 5.5, there exists a unique solution $\phi^{m,n}(t)=(\phi_{(i-1)n+j}^{m,n}(t))_{1\le i\le m,1\le j\le n}$ to (6.5), for $m,n\in\mathbb{N}$.

For $t\in\mathcal{I}$, define

$$\nu_t^{m,n,x}\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n\delta_{\phi_{(i-1)n+j}^{m,n}(t)},\quad x\in X.\tag{6.6}$$

**Theorem 6.2.** *Assume* (**H3.1**)*-*(**H3.4**)*. Then there exists a unique uniformly weak solution* $\rho(t,x,\phi)$ *to* (6.3)*. Assume additionally* $\rho_0(x,\phi)$ *is continuous in* $x\in X$ *for* $\mathfrak{m}$*-a.e.* $\phi\in Y$ *such that* $\rho_0\in \mathsf{L}^1_+(X\times Y;\mu_X\otimes\mathfrak{m})$ *and*

$$\sup_{x\in X}\|\rho_0(x,\cdot)\|_{\mathsf{L}^1(Y;\mathfrak{m})}<\infty.$$

*Let* $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_{+,\mathsf{abs}}(Y)))$ *be the measure-valued function defined in terms of the uniformly weak solution to* (6.3)*:*

$$\mathrm{d}\nu_t^x=\rho(t,x,\phi)\mathrm{d}\phi,\quad \textit{for every}\quad t\in\mathcal{I},\quad \textit{and}\quad x\in X.$$

*Then* $\nu_t\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$*, for all* $t\in\mathcal{I}$*, provided* $\nu_0\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$*. Moreover, let* $\nu_0^{m,n}\in\mathcal{B}_*(X,\mathcal{M}_+(Y))$*,* $\eta^{m,n}\in\mathcal{B}(X,\,\mathcal{M}_+(Y))$*, and* $h^m\in\mathcal{C}(\mathcal{I}\times X\times Y,\mathbb{R}^2)$ *be defined in* (6.4)*, and* $\nu_t^{m,n}$ *be defined in* (6.6)*. Then*

$$\lim_{n\to\infty}d_0(\nu_\cdot^{m,n},\nu_\cdot)=0.$$

*Proof.* It is straightforward to verify that (**H3.1**) implies (**A1**), (**H3.2**) implies (**A2**), and (**H3.3**) implies (**A3**) and (**A7**). It remains to show (**A6**) is fulfilled with $Y$ defined in (6.2). This is a simple consequence of the fact that this SIS model is conservative:

$$\frac{\partial}{\partial t}(S_x(t)+I_x(t))=0.$$

□

6.3. **Lotka-Volterra model with dispersal on a hypergraph.** Assume that

(**H3.1**) $(X,\mathfrak{B}(X),\mu_X)$ be a compact probability space.

(**H3.2**) $W_{i,j}$ are odd functions and locally Lipschitz continuous satisfying $0\le W_{i,j}(u)\le u$ for all $u\in\mathbb{R}^+$ for all $i,j=1,2$.

(**H3.3**) $\eta_1,\ \eta_2\in\mathcal{B}(X,\mathcal{M}_+(X^2))$.

Let $\Lambda_1,\ \Lambda_2>0$ satisfy

$$\Lambda_1\ge\frac{\alpha}{\beta},\quad \Lambda_2\ge-\frac{\iota}{\theta}+\frac{\sigma}{\theta}\Lambda_1. \tag{6.7}$$

Let $Y=\{\phi\in\mathbb{R}^2_+\colon\phi_1\le\Lambda_1,\ \phi_2\le\Lambda_2\}$ be the rectangle in the positive cone, which is a convex compact set.

(**H3.4**) $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\mathcal{M}_+(\mathbb{R}^2)))$ is uniformly compactly supported within $Y\subseteq\mathbb{R}^2_+$.

Under (**H3.1**)-(**H3.4**), consider the general Lotka-Volterra with the species of two types moving on generalized directed hypergraphs:

$$\begin{aligned}\frac{\partial\phi_1(t,x)}{\partial t}=&\phi_1(t,x)(\alpha-\beta\phi_1(t,x)-\gamma\phi_2(t,x))+\int_X\int_X\int_Y\int_Y\\&(W_{1,1}(\psi_1-\phi_1(t,x))+W_{1,2}(\psi_2-\phi_1(t,x)))\,\mathrm{d}\nu_t^{y_1}(\psi_1)\mathrm{d}\nu_t^{y_2}(\psi_2)\mathrm{d}\eta_1^x(y_1,y_2)\\\frac{\partial\phi_2(t,x)}{\partial t}=&\phi_2(t,x)(-\iota+\sigma\phi_1(t,x)-\theta\phi_2(t,x))+\int_X\int_X\int_Y\int_Y\\&(W_{2,1}(\psi_1-\phi_2(t,x))+W_{2,2}(\psi_2-\phi_2(t,x)))\,\mathrm{d}\nu_t^{y_1}(\psi_1)\mathrm{d}\nu_t^{y_2}(\psi_2)\mathrm{d}\eta_2^x(y_1,y_2),\end{aligned}$$

where $\phi_1(t)$ and $\phi_2(t)$ stand for population densities of two competing species at time $t$, respectively, and all given functions and parameters are non-negative. The model can be regarded as a generalization of the Lotka-Volterra model on the graph proposed in [45].

Let

$$g_1(t,\phi,\psi_1,\psi_2)=\begin{pmatrix}W_{1,1}(\psi_1-\phi_1)+W_{1,2}(\psi_2-\phi_1)\\0\end{pmatrix},$$

$$g_2(t,\phi,\psi_1,\psi_2)=\begin{pmatrix}0\\W_{2,1}(\psi_1-\phi_2)+W_{2,2}(\psi_2-\phi_2)\end{pmatrix},$$

$$h(\phi)=\begin{pmatrix}\phi_1(\alpha-\beta\phi_1-\gamma\phi_2)\\ \phi_2(-\iota+\sigma\phi_1-\theta\phi_2)\end{pmatrix},$$

and

$$\begin{aligned}\widehat{V}[\eta,\rho.,h](t,x,\phi)=&\sum_{\ell=1}^2\int_X\int_X\int_Y\int_Y g_\ell(t,\phi,\psi_1,\psi_2)\rho(t,y_1,\psi_1)\rho(t,y_2,\psi_2)\mathrm{d}\psi_1\mathrm{d}\psi_2\mathrm{d}\eta_\ell^x(y_1,y_2)\\&+h(\phi).\end{aligned}$$

Consider the VE

$$\begin{aligned}&\frac{\partial\rho(t,x,\phi)}{\partial t}+\mathrm{div}_\phi\Big(\rho(t,x,\phi)\widehat{V}[\eta,\rho(\cdot),h](\phi)\Big)=0,\quad t\in(0,T],\ x\in X,\ \mathfrak{m}\text{-a.e. }\phi\in Y,\\&\rho(0,\cdot)=\rho_0(\cdot).\end{aligned}\tag{6.8}$$

Note that there exists

- a partition $\{A_i^m\}_{1\le i\le m}$ of $X$ and points $x_i^m\in A_i^m$ for $i=1,\ldots,m$, for every $m\in\mathbb{N}$,
- a sequence $\{\varphi^{m,n}_{(i-1)n+j}\colon i=1,\ldots,m,j=1,\ldots,n\}_{n,m\in\mathbb{N}}\subseteq Y$ and $\{a_{m,i}\colon i=1,\ldots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$, and
- a sequence $\{y^{\ell,m,n}_{(i-1)n+j}\colon i=1,\ldots,m,j=1,\ldots,n\}_{m,n\in\mathbb{N}}\subseteq X^2$ and $\{b_{\ell,m,i}\colon i=1,\ldots,m\}_{m\in\mathbb{N}}\subseteq\mathbb{R}_+$, for $\ell=1,2$,

such that

$$\lim_{m\to\infty}\lim_{n\to\infty}d_\infty(\nu_0^{m,n},\nu_0)=0,$$

$$\lim_{m\to\infty}\lim_{n\to\infty}d_\infty(\eta_\ell^{m,n},\eta_\ell)=0,\quad \ell=1,2,$$

$$\lim_{m\to\infty}\int_0^T\int_Y\sup_{x\in X}|h^m(t,x,\phi)-h(t,x,\phi)|\,\mathrm{d}\phi\mathrm{d}t=0,$$

where

$$\begin{aligned}\nu_0^{m,n,x}&\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(x)\frac{a_{m,i}}{n}\sum_{j=1}^n\delta_{\varphi^{m,n}_{(i-1)n+j}},\quad x\in X,\\ \eta_\ell^{m,n,x}&\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(x)\frac{b_{\ell,m,i}}{n}\sum_{j=1}^n\delta_{y^{\ell,m,n}_{(i-1)n+j}},\quad x\in X,\quad \ell=1,2,\\ h^m(t,z,\phi)&\coloneqq\sum_{i=1}^m\mathbb{1}_{A_i^m}(z)h(t,x_i^m,\phi),\quad t\in\mathcal{I},\ z\in X,\ \phi\in Y.\end{aligned}\tag{6.9}$$

Consider the following IVP of a coupled ODE system:

$$\begin{aligned}\dot\phi_{(i-1)n+j}=F_i^{m,n}(t,\phi_{(i-1)n+j},\Phi),\quad 0<t\le T,\quad \phi_{(i-1)n+j}(0)=\varphi^{m,n}_{(i-1)n+j},\\ i=1,\ldots,m,\ j=1,\ldots,n,\end{aligned}\tag{6.10}$$

where $\Phi=(\phi_{(i-1)n+j})_{1\le i\le m,1\le j\le n}$ and $\Phi=(\phi_{(i-1)n+j})_{1\le i\le m,1\le j\le n}$ and

$$\begin{aligned}F_i^{m,n}(t,\psi,\Phi)=&\sum_{\ell=1}^2\frac{b_{\ell,m,i}}{n}\sum_{j=1}^n\sum_{p_1=1}^m\frac{a_{m,p_1}}{n}\mathbb{1}_{A_{p_1}^m}(y^{\ell,m,n}_{(i-1)n+j,1})\sum_{p_2=1}^m\frac{a_{m,p_2}}{n}\mathbb{1}_{A_{p_2}^m}(y^{\ell,m,n}_{(i-1)n+j,2})\\&\sum_{q_1=1}^n\sum_{q_2=1}^n g_\ell(t,\psi,\phi_{(p_1-1)n+q_1},\phi_{(p_2-1)n+q_2})+h^m(t,x_i^m,\psi).\end{aligned}$$

Then by Proposition 5.5, there exists a unique solution $\phi^{m,n}(t)=(\phi^{m,n}_{(i-1)n+j}(t))_{1\le i\le n,1\le j\le m}$ to (6.10), for $m,n\in\mathbb{N}$.

For $t \in \mathcal{I}$, define

$$\nu_t^{m,n,x} := \sum_{i=1}^m \mathbb{1}_{A_i^m}(x) \frac{a_{m,i}}{n} \sum_{j=1}^n \delta_{\phi^{m,n}_{(i-1)n+j}(t)}, \quad x \in X. \tag{6.11}$$

**Theorem 6.3.** *Assume* (**H3.1**)*-*(**H3.4**)*. Additionally assume $\Lambda_1, \Lambda_2$ satisfy* (6.7)*. Then there exists a unique uniformly weak solution $\rho(t,x,\phi)$ to* (6.8)*. Assume additionally $\rho_0(x,\phi)$ is continuous in $x \in X$ for $\mathfrak{m}$-a.e. $\phi \in Y$ such that $\rho_0 \in \mathsf{L}^1_+(X \times Y; \mu_X \otimes \mathfrak{m})$ and*

$$\sup_{x \in X} \|\rho_0(x,\cdot)\|_{\mathsf{L}^1(Y;\mathfrak{m})} < \infty.$$

*Let $\nu_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_{+,\mathsf{abs}}(Y)))$ be the measure-valued function defined in terms of the uniformly weak solution to* (6.8)*:*

$$\mathrm{d}\nu_t^x = \rho(t,x,\phi)\mathrm{d}\phi, \quad \textit{for every} \quad t \in \mathcal{I}, \quad \textit{and} \quad x \in X.$$

*Then $\nu_t \in \mathcal{C}_*(X, \mathcal{M}_+(Y))$, for all $t \in \mathcal{I}$, provided $\nu_0 \in \mathcal{C}_*(X, \mathcal{M}_+(Y))$. Moreover, let $\nu_0^{m,n}$, $\eta^{\ell,m,n}$ for $\ell = 1,2$, and $h^m$ be defined in* (6.9)*, and $\nu_t^{m,n}$ be defined in* (6.11)*. Then*

$$\lim_{n\to\infty} d_0(\nu_\cdot^{m,n}, \nu_\cdot) = 0.$$

*Proof.* First note that (**H3.1**) implies (**A1**) and (**H3.2**) implies (**A2**). It is readily verified that (**A3**) and (**A7**) are fulfilled since $Y$ is compact. In addition, (**A4**)$'$ follows from (**H3.3**). Hence it suffices to show that (**A6**) holds with $Y$ for some $c, \Lambda > 0$.

Note that $\partial Y = \{\phi_1 = 0\} \cup \{\phi_2 = 0\} \cup \{\phi_1 = \Lambda_1\} \cup \{\phi_2 = \Lambda_2\}$. In the following, we will show that

$$V[\eta, \nu_\cdot, h](t,x,\phi) \cdot \upsilon(\phi) \le 0, \quad \text{for all } t \in \mathcal{T},\ x \in X, \quad \phi \in \partial Y,$$

where $\upsilon(\phi)$ is the outer normal vector at $\phi$. We prove it case by case.

(i) For $\phi \in \{\varphi \in Y \colon \varphi_1 = 0\}$, $\upsilon(\phi) = (-1,0)$, and

$$\begin{aligned}&V[\eta, \nu_\cdot, h](t,x,\phi) \cdot \upsilon(\phi)\\ =& -\int_X\int_X\int_Y\int_Y (W_{1,1}(\psi_1 - \phi_1) + W_{1,2}(\psi_2 - \phi_1))\, \mathrm{d}\nu_t^{y_1}(\phi_1)\mathrm{d}\nu_t^{y_2}(\phi_2)\mathrm{d}\eta_1^x(y_1,y_2) \le 0.\end{aligned}$$

(ii) For $\phi \in \{\varphi \colon \varphi_2 = 0\}$, $\upsilon(\phi) = (0,-1)$, and

$$\begin{aligned}&V[\eta, \nu_\cdot, h](t,x,\phi) \cdot \upsilon(\phi)\\ =& -\int_X\int_X\int_Y\int_Y (W_{2,1}(\psi_1 - \phi_2) + W_{2,2}(\psi_2 - \phi_2))\, \mathrm{d}\nu_t^{y_1}(\phi_1)\mathrm{d}\nu_t^{y_2}(\phi_2)\mathrm{d}\eta_2^x(y_1,y_2) \le 0.\end{aligned}$$

(iii) For $\phi \in \{\varphi \colon \varphi_1 = \Lambda_1\}$, $\upsilon(\phi) = (1,0)$. By (6.7),

$$\begin{aligned}&V[\eta, \nu_\cdot, h](t,x,\phi) \cdot \upsilon(\phi)\\ =&\Lambda_1(\alpha - \beta\Lambda_1 - \gamma\phi_2)\\ &+ \int_X\int_X\int_Y\int_Y (W_{1,1}(\psi_1 - \Lambda_1) + W_{1,2}(\psi_2 - \Lambda_1))\, \mathrm{d}\nu_t^{y_1}(\phi_1)\mathrm{d}\nu_t^{y_2}(\phi_2)\mathrm{d}\eta_1^x(y_1,y_2)\\ \le&\Lambda_1(\alpha - \beta\Lambda_1) \le 0,\end{aligned}$$

since $\Lambda_1 \ge \frac{\alpha}{\beta}$, due to (6.7).

(iv) For $\phi \in \{\varphi \colon \varphi_2 = \Lambda_2\}$, $\upsilon(\phi) = (1,0)$. By (6.7), and $\psi_2 \le \Lambda_2$ for $\psi \in Y$, we have

$$\begin{aligned}&V[\eta, \nu_\cdot, h](t,x,\phi) \cdot \upsilon(\phi)\\ =&\Lambda_2(-\iota + \sigma\phi_1 - \theta\Lambda_2)\\ &+ \int_X\int_X\int_Y\int_Y (W_{2,1}(\psi_1 - \Lambda_2) + W_{2,2}(\psi_2 - \Lambda_2))\, \mathrm{d}\nu_t^{y_1}(\psi_1)\nu_t^{y_2}(\psi_2)\mathrm{d}\eta_2^x(y_1,y_2)\\ \le&\Lambda_2(-\iota + \sigma\Lambda_1 - \theta\Lambda_2) \le 0,\end{aligned}$$

since by (6.7), we have $\Lambda_2 \ge -\frac{\iota}{\theta} + \frac{\sigma}{\theta}\Lambda_1$.

□

## 7. Discussion

In this paper we regard directed hypergraph limits as elements in $\mathcal{B}(X,\mathcal{M}_+(X^{k-1}))$ for some $k\in\mathbb{N}\setminus\{1\}$, the space of bounded measure-valued functions. The motivation of doing such a work comes from the emergent demanding applications from networks models of higher-order interactions [8]. We extend the idea proposed in [31] in a plain way from *directed graph measures* to *directed hyper-graph measures*. We apply our main results to Kuramoto networks of higher-order interactions in physics as well as models in epidemiology and ecology. There are also models of higher-order interactions which do *not* admit a *compact* positively invariant set [2], and hence our results cannot directly apply. It would be desirable to derive similar results on MFLs of IPS with an unbounded invariant set. In addition, our results only apply to the case where the hypergraph measure is a limit of the sequence of hypergraphs of *uniformly bounded* cardinalities. It will be challenging while exciting to study the case e.g., where the hypergraph measure is a limit of the sequence of hypergraphs of *unbounded* cardinalities , e.g., a sequence of simpicial complexes of expanding cardinalities.

We also point out that there is a generic drawback of our definition of DHGM as a generalization of hypergraphs. In the theory of large limits of graphs [35], the topology is induced by the cut distance which is sufficient to gurantee that a Cauchy sequence of finite graphs has a graph limit object. However, the extension to hypergraphs with the same property seems rather non- trivial and involved [19, 35], which makes it impossible to define hypergraphon in a trivial way as direct translation of graphons from graphs to hypergraphs, while preserving the above property ensuring the existence of a limit object. Such topology induced by cut distance is weaker than the weak topology induced by bounded Lipschitz functions. For this reason, our definition of DHGM can be narrower than a more desriable and natural one to be defined for limits of hypergraphs. In other words, potentially there may exist a limit of a sequence of hypergraphs with a uniform cardinality which may not be defined as a DHGM; and there can be discretizations as ODE networks by construction or obstruction which are different from the concrete constructions we provide in this paper. Nevertheless, our definition makes a decent trade-off in the applications, in the light of the rather rare literature [**?**] which uses the weaker topology aforementioned in the analysis of MFL of networks of dynamical systms, due to the obvious difficulty in obtaining estimates of distances of the empirical distribution and the MFL, caused by the combinatorial nature of the definition of cut distance. We leave all these worth questions for our future work.

## 8. Proof of Theorem 5.6

*Proof.* First, the proof of $\nu_\cdot\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\mathcal{M}_+(Y)))$ follows directly from that of [31, Theorem 5.15]. The strategy based on four steps to prove the approximation result is identical to that given in the proof of [31, Theorem 5.15]. Steps II-IV among the four steps are independent of $\widehat{V}$. We essentially need to verify that $\nu_\cdot^{m,n}$ is the unique solution to the fixed point equation associated with $\eta^{m,n}$ and $h^m$:

$$\nu_\cdot^{m,n}=\mathcal{A}[\eta^{m,n},h^m]\nu_\cdot^{m,n}.$$

In the light of Step I in the proof of [31, Theorem 5.15], it suffices to express the Vlasov operator in the discrete context and show it is consistent with the definition of $F_i^{m,n}$.

Denote $\eta^{m,n}=(\eta_\ell^{m,n})_{1\le\ell\le r}$. For $x\in I_i^N$, $t\in\mathcal{I}$, by (5.1) and (5.3), we have

$$\begin{aligned}&V^{m,n}[\eta^{m,n},\nu_\cdot^{m,n},h^m](t,x,\phi)\\ =&\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\underbrace{\int_{\mathbb{R}^{r_2}}\cdots\int_{\mathbb{R}^{r_2}}}_{k_\ell-1}g_\ell(t,\phi,\psi_1,\ldots,\psi_{k_\ell-1})\mathrm{d}\nu_t^{m,n,y_{k_\ell-1}}(\psi_{k_\ell-1})\end{aligned}$$

$$\begin{aligned}
&\cdots \mathrm{d}\nu_t^{m,n,y_1}(\psi_1)\cdot \mathrm{d}\eta_\ell^{m,n,x}(y_1,\dots,y_{k_\ell-1}) + h^m(t,x,\phi),\\
=&\sum_{\ell=1}^r \frac{b_{\ell,m,i}}{n}\sum_{j=1}^n \underbrace{\int_{\mathbb{R}^{r_2}}\cdots\int_{\mathbb{R}^{r_2}}}_{k_\ell-1} g_\ell(t,\phi,\psi_1,\dots,\psi_{k_\ell-1})\\
&\cdot \mathrm{d}\nu_t^{m,n,y^{\ell,m,n}_{(i-1)n+j,k_\ell-1}}(\psi_{k_\ell-1})\cdots \mathrm{d}\nu_t^{m,n,y^{\ell,m,n}_{(i-1)n+j,1}}(\psi_1) + h^m(t,x_i^m,\phi)\\
=&\sum_{\ell=1}^r \frac{b_{\ell,m,i}}{n}\sum_{j=1}^n\sum_{p_1=1}^m \frac{a_{m,p_1}}{n}\mathbb{1}_{A^m_{p_1}}(y^{\ell,m,n}_{(i-1)n+j,1})\cdots\sum_{p_{k_\ell-1}=1}^m \frac{a_{m,p_{k_\ell-1}}}{n}\\
&\mathbb{1}_{A^m_{p_{k_\ell-1}}}(y^{\ell,m,n}_{(i-1)n+j,k_\ell-1})\sum_{q_1=1}^n\cdots\sum_{q_{k_\ell-1}=1}^n\\
&g_\ell(t,\phi,\phi^{m,n}_{(p_1-1)n+q_1},\dots,\phi^{m,n}_{(p_{k_\ell-1}-1)n+q_{k_\ell-1}}) + h^m(t,x_i^m,\phi)\\
=&F_i^{m,n}(t,\phi,\Phi^{m,n}(t)).
\end{aligned}$$

□

## Acknowledgements

Both authors thank Giulio Zucal for bringing [11] to their attention. We also thank Christian Bick, Tobias Böhle, Raffaella Mulas, Davide Sclosa, and Giulio Zucal (i.e., the other members of the "H-Team") for helpful discussions on hypergraph limits and networks of higher-order interactions. CK also acknowledges support via a Lichtenberg Professorship of the VolkswagenStiftung. CX acknowledges the support from TUM University Foundation and the Alexander von Humboldt Foundation and an internal start-up funding from the University of Hawai'i at Mānoa.

## References

[1] Backhausz, A. and Szegedy, B. Action convergence of operators and graphs. *Canad. J. Math.*, 74:72–121, 2022.
[2] Battiston, F. et al. Networks beyond pairwise interactions: Structure and dynamics. *Phys. Rep.*, 874:1–92, 2020.
[3] Bencheikh, O. and Jourdain, B. Approximation rate in Wasserstein distance of probability measures on the real line by deterministic empirical measures. *J. Approx. Theory*, 274:105684, 2022.
[4] Benjamini, I. and Schramm, O. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:1–13, 2001.
[5] Bick, C., Ashwin, P., and Rodrigues, A. Chaos in generically coupled phase oscillator networks with nonpairwise interactions. *Chaos*, 26:094814, 2016.
[6] Bick, C., Böhle, T., and Kuehn, C. Multi-population phase oscillator networks with higher-order interactions. *NoDEA*, 29(64), 2022.
[7] Bick, C., Böhle, T., and Kuehn, C. Phase oscillator networks with nonlocal higher-order interactions: twisted states, stability and bifurcations. *arXiv:2206.01951*, 2022.
[8] Bick, C. and et al. What are higher-order networks? *SIAM Review*, 2022.
[9] Bodó, Á., Katona, G.Y., and Simon, P.L. SIS epidemic propagation on hypergraphs. *Bull. Math. Biol.*, 78:713–735, 2016.
[10] Bogachev, V.I. *Measure Theory: Volume I.* Springer-Verlag, Berlin Heidelberg, 2007.
[11] Bombal, F. Continuous multilinear operators on $c(k)$ spaces and polymeasures. *Extracta Mathematicae*, 22:127–146, 2007.
[12] Borgs., C. et al. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219:1801–1851, 2008.
[13] Bretto, A. *Hypergraph Theory: An Introduction*. Mathematical Engineering. Springer International Publishing, Cham, Switzerland, 2013.
[14] Chevallier, J. Uniform decomposition of probability measures: quantization, clustering and rate of convergence. *J. Appl. Probab.*, 55:1037–1045, 2018.
[15] Chiba, H. and Medvedev, G.S. The mean field analysis of the Kuramoto model on graphs I. the mean field equation and transition point formulas. *Discrete Contin. Dyn. Syst. Ser. A*, 39:131–155, 2019.

[16] Cobzaş, S., Miculescu, R., and Nicolae, A. *Lipschitz Functions*, volume 2241 of *Lect. Notes Math.* Springer-Verlag, Cham, Switzerland, 2019.
[17] Cucker, F. and Smale, S. Emergent behavior in flocks. *IEEE Trans. Automatic Control*, 52:852–862, 2006.
[18] Cui, K., KhudaBukhsh, W.R., and Koeppl, H. Hypergraphon mean field games. *Chaos*, 32:113129, 2022.
[19] Elek, G. and Szegedy, B. A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.*, 231:1731–1772, 2012.
[20] Feinberg, M. *Foundations of Chemical Reaction Network Theory.* Applied Mathematical Sciences. Springer International Publishing, Cham, 2019.
[21] Gambuzza, L. et al. Stability of synchronization in simplicial complexes. *Nature Commun.*, 12:1255, 2021.
[22] Gkogkas, M.A. and Kuehn, C. Graphop mean-field limits for Kuramoto-type models. *SIAM J. Appl. Dyn. Syst.*, 21:248–283, 2022.
[23] Golse, F. On the dynamics of large particle systems in the mean field limit. *arXiv:1301.5494*, 2013.
[24] Gowers, W.T. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. Math.*, 166:897–946, 2007.
[25] Hegselmann, R. and Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.*, 5(3), 2002.
[26] Jabin, P.-E., Poyato, D., and Soler, J. Mean field limit of non-exchangable systems. *arXiv:2112.15406*, 2021.
[27] Jonsson, J. *Simplicial Complexes of Graphs*, volume 1928 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin Heidelberg, 2008.
[28] Kaliuzhnyi-Verbovetskyi, D. and Medvedev, G.S. The mean field equation for the Kuramoto model on graph sequences with non-Lipschitz limit. *SIAM J. Math. Anal.*, 50:2441–2465, 2018.
[29] Keliger, D. and Horváth, I. Accuracy criterion for mean field approximations of Markov processes on hypergraphs. *Phys. A: Stat. Mech.*, 609:128370, 2023.
[30] Kiss, I.Z., Miller, J.C., and Simon, P.L. *Mathematics of Epidemics on Networks: From Exact to Approimate Models*, volume 46 of *Interdisciplinary Applied Mathematics*. Springer International Publishing, Cham, 2019.
[31] Kuehn, C. and Xu, C. Vlasov equations on digraph measures. *J. Diff. Equ.*, 339:261–349, 2022.
[32] Kunszenti-Kovács, D., Szegedy, B., and Lovász, L. Measures on the square as sparse graph limits. *J. Combin. Theory Ser. B*, 138:1–40, 2019.
[33] Lacker, D., Ramanan, K., and Wu, R. Large sparse networks of interacting diffusions. *arXiv:1904.02585*, 2019.
[34] Lacker, D., Ramanan, K., and Wu, R. Local weak convergence for sparse networks of interacting processes. *Ann. Appl. Probab.*, 33:643–688, 2023.
[35] Lovász, L. *Large Networks and Graph Limits*, volume 60 of *AMS*. Colloquium Publications, Providence,UAS, 2012.
[36] Lovász, L. and Szegedy, B. Limits of dense graph sequences. *J. Comb. Theory, Ser. B.*, 96:933–957, 2006.
[37] Oelschläger, K. Morale, D., Capasso, V. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.*, 50:49–66, 2005.
[38] Mulas, R., Kuehn, C., and Jost, J. Coupled dynamics on hypergraphs: Master stability of steady states and synchronization. *Phys. Rev. E*, 101:062313, 2020.
[39] Neunzert, H. An introduction to the nonlinear Boltzmann-Vlasov equation. In Cercignani, C., editor, *Kinetic Theories and the Boltzmann Equation*, volume 1048 of *Lect. Notes Math.*, pages 60–110. Springer-Verlag, Berlin, Heidelberg, 1984.
[40] Newman, M.E.J. *Networks.* Oxford Univ. Press, Oxford, UK, 2nd edition, 2018.
[41] Nijholt, E. and DeVille, L. Dynamical systems defined on simplicial complexes: Symmetries, conjugacies, and invariant subspaces. *Chaos*, 32:093131, 2022.
[42] Oliveira, R.I., Reis, G.H., and Stolerman, L.M. Interacting diffusions on sparse graphs: hydrodynamics from local weak limits. *Electron. J. Probab.*, 25:1–35, 2020.
[43] Pielou, E.C. *An Introduction to Mathematical Ecology*, volume 46. Wiley-Interscience, New York, USA, 1969.
[44] Roddenberry, T.M. and Segarra, S. Limits of dense simplicial complexes. arXiv:2207.07980, 2022.
[45] Slavík, A. Lotka–Volterra competition model on graphs. *SIAM J. Appl. Dyn. Syst.*, 19:725–762, 2020.
[46] Spohn, H. *Large Scale Dynamics of Interacting Particles.* Texts and Monographs in Physics. Springer-Verlag, Berlin, Heidelberg, 1st edition, 1991.
[47] Tao, T. A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A*, 113:1257–1280, 2006.
[48] Teschl, G. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Grad. Stud. Math.* AMS, Providence, Rhode Island, 2012.

[49] Torres, L. et al. The why, how, and when of representations for complex systems. *SIAM Rev*, 63:435–485, 2021.
[50] Xu, C. and Berger, A. Best finite constrained approximations of one-dimensional probabilities. *J. Approx. Theory*, 244:1–36, 2019.
[51] Zhao, Y.F. Hypergraph limits: A regularity approach. *Random Struct. and Alg.*, 47:205–226, 2015.

## Appendix A. Proof of Proposition 4.2

*Proof.* Denote $\eta=(\eta_\ell)_{\ell=1}^r$ We will suppress the variables in $V[\eta,\nu_\cdot,h](t,x,\psi)$ and $\Phi_{s,t}^x[\eta,\nu_\cdot,h]$ whenever they are clear and not the emphasis from the context.

The proof of the absolute continuity property is analogous to that of [31, Proposition 4.4(iv)]. In the following, we prove the rest three continuous dependence properties item by item.

The properties of $\mathcal{A}$ follows from that of $\Phi_{0,t}^x[\eta,\nu_\cdot,h]$. Hence in the following, we will first establish corresponding continuity and Lipschitz continuity for $\Phi_{0,t}^x[\eta,\nu_\cdot,h]$ and then apply the results to derive respective properties for $\mathcal{A}$.

(i) Continuity in $t$. Indeed,

$$\begin{aligned}
&d_\infty(\mathcal{A}[\eta,h]\nu_t,\mathcal{A}[\eta,h]\nu_s)\\
=&\sup_{x\in X} d_{\mathsf{BL}}(\Phi_{t,0}^x[\eta,\nu_\cdot,h]\#\nu_0^x,\Phi_{s,0}^x[\eta,\nu_\cdot,h]\#\nu_0^x)\\
=&\sup_{x\in X}\sup_{f\in\mathcal{BL}_1(Y)}\left|\int_Y\left(f\circ\Phi_{t,0}^x[\eta,\nu_\cdot,h]\phi-f\circ\Phi_{s,0}^x[\eta,\nu_\cdot,h]\phi\right)\mathrm{d}\nu_0^x(\phi)\right|\\
\le&\sup_{x\in X}\int_Y\left|\Phi_{t,0}^x[\eta,\nu_\cdot,h]\phi-\Phi_{s,0}^x[\eta,\nu_\cdot,h]\phi\right|\mathrm{d}\nu_0^x(\phi)\\
=&\sup_{x\in X}\int_Y\left|\int_s^t\Big(h(\tau,x,\Phi_{\tau,0}^x[\eta,\nu_\cdot,h]\phi)\right.\\
&+\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\underbrace{\int_Y\cdots\int_Y}_{k_\ell-1}g_\ell(\tau,\Phi_{\tau,0}^x[\eta,\nu_\cdot,h]\phi,\psi_1,\psi_2,\ldots,\psi_{k_\ell-1})\mathrm{d}\nu_\tau^{y_{k_\ell-1}}(\psi_{k_\ell-1})\cdots\mathrm{d}\nu_t^{y_1}(\psi_1)\\
&\left.\mathrm{d}\eta_\ell^x(y_1,\ldots,y_{k_\ell-1})\Big)\mathrm{d}\tau\right|\mathrm{d}\nu_0^x(\phi)\\
\le&\sup_{x\in X}\int_Y\int_s^t\Big(\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\int_Y\cdots\int_Y\|g_\ell\|_\infty\mathrm{d}\nu_\tau^{y_{k_\ell-1}}\cdots\mathrm{d}\nu_\tau^{y_1}\mathrm{d}\eta_\ell^x(y_1,\ldots,y_{k_\ell-1})+\|h\|_{\infty,\mathcal{I}}\Big)\mathrm{d}\tau\mathrm{d}\nu_0^x(\phi)\\
\le&\|\nu_\cdot\|\Big(\sum_{\ell=1}^r\|\nu_\cdot\|^{k_\ell-1}\|g_\ell\|_\infty\|\eta_\ell\|+\|h\|_{\infty,\mathcal{I}}\Big)|t-s|\\
\le&L_1\|\nu_\cdot\||t-s|\to0,\quad\text{as}\quad|s-t|\to0,
\end{aligned}$$

where

$$L_1\coloneqq L_1(\nu_\cdot)=\mathcal{BL}(h)+\sum_{\ell=1}^r\mathcal{BL}(g_\ell)\|\eta_\ell\|\|\nu_\cdot\|^{k_\ell-1},$$

$\mathcal{BL}(h)=\sup_{t\in\mathcal{I}}\sup_{x\in X}\mathcal{BL}(h(t,x,\cdot))$ and $\mathcal{BL}(g_\ell)=\sup_{t\in\mathcal{I}}\sup_{x\in X}\mathcal{BL}(g_\ell(t,x,\cdot))$ for $\ell=1,\ldots,r$.

This shows that

$$t\mapsto\mathcal{A}[\eta,h]\nu_t\in\mathcal{C}(\mathcal{I},\mathcal{B}_*(X,\mathcal{M}_+(Y))).$$

The proof of the mass conservation law is the same as that of [31, Proposition 4.4].

(ii) Next, we show $\mathcal{A}^x[\eta,h]\nu_\cdot$ is Lipschitz continuous in $\nu_\cdot$. We first prove Lipschitz continuity of $\Phi_{t,0}^x\phi$ in the initial condition $\phi$. Note that

$$|V[\eta,\nu_\cdot,h](t,x,\phi_1)-V[\eta,\nu_\cdot,h](t,x,\phi_2)|$$

$$\le\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\underbrace{\int_Y\cdots\int_Y}_{k_\ell-1}|g_\ell(t,\phi_1,\psi_1,\ldots,\psi_{k_\ell-1})-g_\ell(t,\phi_2,\psi_1,\ldots,\psi_{k_\ell-1})|\mathrm{d}\nu_t^{y_{k_\ell-1}}(\psi_{k_\ell-1})$$

$$\cdots\mathrm{d}\nu_t^{y_1}(\psi_1)\cdot\mathrm{d}\eta_\ell^x(y_1,\ldots,y_{k_\ell-1})+|h(t,x,\phi_1)-h(t,x,\phi_2)|$$

$$\le L_1(\nu_\cdot)|\phi_1-\phi_2|,\quad t\in\mathcal{I},\tag{A.1}$$

This yields

$$|\Phi^x_{t,0}\phi_1(x)-\Phi^x_{t,0}\phi_2(x)|$$

$$\le|\phi_1(x)-\phi_2(x)|+\int_0^t|V[\eta,\nu_\cdot,h](\tau,x,\Phi^x_{\tau,0}\phi_1(x))-V[\eta,\nu_\cdot,h](\tau,x,\Phi^x_{\tau,0}\phi_2(x))|\mathrm{d}\tau$$

$$\le|\phi_1(x)-\phi_2(x)|+L_1(\nu_\cdot)\int_0^t|\Phi^x_{\tau,0}\phi_1(x)-\Phi^x_{\tau,0}\phi_2(x)|\mathrm{d}\tau.$$

By Gronwall's inequality,

$$|\Phi^x_{t,0}\phi_1(x)-\Phi^x_{t,0}\phi_2(x)|\le\mathrm{e}^{L_1(\nu_\cdot)t}|\phi_1(x)-\phi_2(x)|,\tag{A.2}$$

Similarly, one can also show that

$$|\Phi^x_{0,t}\phi_1(x)-\Phi^x_{0,t}\phi_2(x)|\le\mathrm{e}^{L_1(\nu_\cdot)t}|\phi_1(x)-\phi_2(x)|.$$

Now we are ready to show $\mathcal{A}^x[\eta,h]\nu_\cdot$ is Lipschitz continuous in $\nu_\cdot$. Observe that

$$d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x},\Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{2,x})$$

$$\le d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x},\Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{1,x})+d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{1,x},\Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{2,x}).\tag{A.3}$$

Note that

$$|V[\nu^1_\cdot](t,x,\phi)-V[\nu^2_\cdot](t,x,\phi)|$$

$$\le\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\Big|\int_{Y^{k_\ell-1}}g_\ell(t,\phi,\psi_1,\ldots,\psi_{k_\ell-1})\mathrm{d}(\otimes_{j=1}^{k_\ell-1}\nu_t^{1,y_j}(\psi_j)-\otimes_{j=1}^{k_\ell-1}\nu_t^{2,y_j}(\psi_j))\Big|$$

$$\cdot\,\mathrm{d}\eta_\ell^x(y_1,\ldots,y_{k_\ell-1})$$

$$=\sum_{\ell=1}^r\int_{X^{k_\ell-1}}\Big|\int_{Y^{k_\ell-1}}g_\ell(t,\phi,\psi_1,\ldots,\psi_{k_\ell-1})\mathrm{d}(\nu_t^{1,y_1}(\psi_1)\otimes\nu_t^{1,y_2}(\psi_2)\cdots\otimes\nu_t^{1,y_{k_\ell-1}}(\psi_{k_\ell-1})$$

$$-\nu_t^{2,y_1}(\psi_1)\otimes\nu_t^{1,y_2}(\psi_2)\cdots\otimes\nu_t^{1,y_{k_\ell-1}}(\psi_{k_\ell-1})+\nu_t^{2,y_1}(\psi_1)\otimes\nu_t^{1,y_2}(\psi_2)\cdots\otimes\nu_t^{1,y_{k_\ell-1}}(\psi_{k_\ell-1})$$

$$-\nu_t^{2,y_1}(\psi_1)\otimes\nu_t^{2,y_2}(\psi_2)\otimes\nu_t^{1,y_3}(\psi_3)\cdots\otimes\nu_t^{1,y_{k_\ell-1}}(\psi_{k_\ell-1})+\cdots$$

$$+\nu_t^{2,y_1}(\psi_1)\otimes\nu_t^{2,y_2}(\psi_2)\cdots\otimes\nu_t^{2,y_{k_\ell-2}}(\psi_{k_\ell-2})\otimes\nu_t^{1,y_{k_\ell-1}}(\psi_{k_\ell-1})-\otimes_{j=1}^{k_\ell-1}\nu_t^{2,y_j}(\psi_j))\Big|$$

$$\cdot\,\mathrm{d}\eta_\ell^x(y_1,\ldots,y_{k_\ell-1})$$

$$\le L_2d_\infty(\nu^1_t,\nu^2_t),\quad t\in\mathcal{I},$$

where $L_2=L_2(\eta,\nu^1_\cdot,\nu^2_\cdot)\coloneqq\begin{cases}\sum_{\ell=1}^r\|\eta_\ell\|\mathcal{BL}(g_\ell)\sum_{i=1}^{k_\ell-2}\|\nu^1_\cdot\|^i\|\nu^2_\cdot\|^{k_\ell-2-i}, & \text{if}\quad k_\ell>2,\\ \sum_{\ell=1}^r\|\eta_\ell\|\mathcal{BL}(g_\ell), & \text{if}\quad k_\ell=2.\end{cases}$

We now estimate the first term.

$$d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x},\Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{1,x})$$

$$=\sup_{f\in\mathcal{BL}_1(Y)}\int_Y f(\phi)\mathrm{d}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x}-\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{2,x})$$

$$=\sup_{f\in\mathcal{BL}_1(Y)}\int_Y((f\circ\Phi^x_{t,0}[\nu^1_\cdot])(\phi)-(f\circ\Phi^x_{t,0}[\nu^2_\cdot])(\phi))\mathrm{d}\nu_0^{1,x}(\phi)$$

$$\leq \int_Y \left|\Phi^x_{t,0}[\nu^1_\cdot](\phi) - \Phi^x_{t,0}[\nu^2_\cdot](\phi)\right| \mathrm{d}\nu_0^{1,x}(\phi) =: \lambda_x(t)$$

$$\leq \int_Y \int_0^t \left|V[\nu^1_\cdot](\tau,x,\phi) - V[\nu^2_\cdot](\tau,x,\phi)\right| \mathrm{d}\tau \mathrm{d}\nu_0^{1,x}(\phi)$$

$$\leq L_2(\nu^1_\cdot,\nu^2_\cdot)\nu_0^{1,x}(Y)\int_0^t d_\infty(\nu^1_\tau,\nu^2_\tau)\mathrm{d}\tau. \tag{A.4}$$

Next, we estimate the second term. For $f \in \mathcal{BL}_1(Y)$, from (A.2) it follows that

$$\mathcal{L}(f\circ\Phi^x_{t,0}[\nu^2_\cdot]) \leq \mathcal{L}(f)\mathcal{L}(\Phi^x_{t,0}[\nu^2_\cdot]) \leq \mathcal{L}(f)\mathrm{e}^{L_1(\nu^2_\cdot)t}, \quad \|f\circ\Phi^x_{t,0}[\nu^2_\cdot]\|_\infty \leq \|f\|_\infty.$$

Hence

$$\mathcal{BL}(f\circ\Phi^x_{t,0}) \leq \mathrm{e}^{L_1(\nu^2_\cdot)t}. \tag{A.5}$$

For every $x \in X$,

$$d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x}, \Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{1,x})$$

$$= \sup_{f\in\mathcal{BL}_1(Y)} \int_Y (f\circ\Phi^x_{t,0}[\nu^2_\cdot])(\phi)\mathrm{d}(\nu_0^{1,x}(\phi) - \nu_0^{2,x}(\phi))$$

$$\leq \mathrm{e}^{L_1(\nu^2_\cdot)t} d_{\mathsf{BL}}(\nu_0^{1,x},\nu_0^{2,x}) \leq \mathrm{e}^{L_1(\nu^2_\cdot)t} d_\infty(\nu^1_0,\nu^2_0). \tag{A.6}$$

Combining (A.4) and (A.6), it follows from (A.3) that

$$d_\infty(\mathcal{A}[\eta,h]\nu^1_t, \mathcal{A}[\eta,h]\nu^2_t) = \sup_{x\in X} d_{\mathsf{BL}}(\Phi^x_{t,0}[\nu^1_\cdot]_\#\nu_0^{1,x}, \Phi^x_{t,0}[\nu^2_\cdot]_\#\nu_0^{1,x})$$

$$\leq \mathrm{e}^{L_1(\nu^2_\cdot)t} d_\infty(\nu^1_0,\nu^2_0) + \|\nu^1_\cdot\| L_2(\eta,\nu^1_\cdot,\nu^2_\cdot)\mathrm{e}^{L_1(\nu^2_\cdot)t}\int_0^t d_\infty(\nu^1_\tau,\nu^2_\tau)\mathrm{e}^{-L_1(\nu^2_\cdot)\tau}\mathrm{d}\tau.$$

(iii) Lipschitz continuity of $\mathcal{A}[\eta,h]$ in $h$.

We first need to establish the Lipschitz continuity for $\Phi^x_{s,t}[h]$ on $h$. Note that

$$\begin{aligned}
&|\Phi^x_{t,0}[h_1]\phi - \Phi^x_{t,0}[h_2]\phi| \\
\leq& \int_0^t |h_1(\tau,x,\Phi^x_{\tau,0}[h_1]\phi) - h_2(\tau,x,\Phi^x_{\tau,0}[h_2]\phi)|\mathrm{d}\tau \\
&+ \int_0^t \sum_{\ell=1}^r \int_{X^{k_\ell-1}}\int_Y\cdots\int_Y \Big|g_\ell(\tau,\Phi^x_{0,\tau}[h_1]\phi,\psi_1,\psi_2,\ldots,\psi_{k_\ell-1}) \\
&- g_\ell(\tau,\Phi^x_{0,\tau}[h_2]\phi,\psi_1,\psi_2,\ldots,\psi_{k_\ell-1})\Big|\mathrm{d}\nu_\tau^{y_{k_\ell-1}}(\psi_{k_\ell-1})\cdots\mathrm{d}\nu_\tau^{y_1}(\psi_1)\mathrm{d}\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})\mathrm{d}\tau \\
\leq& \int_0^t |h_1(\tau,x,\Phi^x_{\tau,0}[h_1]\phi) - h_2(\tau,x,\Phi^x_{\tau,0}[h_1]\phi)|\mathrm{d}\tau + \int_0^t |h_2(\tau,x,\Phi^x_{\tau,0}[h_1]\phi) - h_2(\tau,x,\Phi^x_{\tau,0}[h_2]\phi)|\mathrm{d}\tau \\
&+ \int_0^t \sum_{\ell=1}^r \int_{X^{k_\ell-1}}\int_Y\cdots\int_Y \Big|g_\ell(\tau,\Phi^x_{\tau,0}[h_1]\phi,\psi_1,\psi_2,\ldots,\psi_{k_\ell-1}) - g_\ell(\tau,\Phi^x_{\tau,0}[h_2]\phi,\psi_1,\psi_2,\ldots,\psi_{k_\ell-1})\Big| \\
&\mathrm{d}\nu_\tau^{y_{k_\ell-1}}(\psi_{k_\ell-1})\cdots\mathrm{d}\nu_\tau^{y_1}(\psi_1)\mathrm{d}\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})\mathrm{d}\tau \\
\leq& T\|h_1-h_2\|_{\infty,\mathcal{I}} + \Big(\mathcal{L}(h_2) + \sum_{\ell=1}^r \mathcal{L}(g_\ell)\|\eta_\ell\|\|\nu_\cdot\|^{k_\ell-1}\Big)\int_0^t |\Phi^x_{\tau,0}[h_1]\phi - \Phi^x_{\tau,0}[h_2]\phi|\mathrm{d}\tau \\
\leq& T\|h_1-h_2\|_{\infty,\mathcal{I}} + L_1(\nu_\cdot)\int_0^t |\Phi^x_{\tau,0}[h_1]\phi - \Phi^x_{\tau,0}[h_2]\phi|\mathrm{d}\tau.
\end{aligned}$$

By Gronwall's inequality,

$$|\Phi^x_{t,0}[h_1]\phi - \Phi^x_{t,0}[h_2]\phi| \leq T\mathrm{e}^{L_1(\nu_\cdot)t}\|h_1-h_2\|_{\infty,\mathcal{I}}.$$

This shows that

$$\begin{aligned}
&d_\infty(\mathcal{A}[\eta,h_1]\nu_t,\mathcal{A}[\eta,h_2]\nu_t)\\
=&\sup_{x\in X} d_{\mathsf{BL}}(\Phi^x_{t,0}[\eta,\nu.,h_1]\#\nu^x_0,\Phi^x_{t,0}[\eta,\nu.,h_2]\#\nu^x_0)\\
=&\sup_{x\in X}\sup_{f\in\mathcal{BL}_1(Y)}\left|\int_Y\left(f\circ\Phi^x_{t,0}[\eta,\nu.,h_1]\phi-f\circ\Phi^x_{t,0}[\eta,\nu.,h_2]\phi\right)\mathrm{d}\nu^x_0(\phi)\right|\\
\le&\int_Y|\Phi^x_{t,0}[\eta,\nu.,h_1]\phi-\Phi^x_{t,0}[\eta,\nu.,h_2]\phi|\mathrm{d}\nu^x_0(\phi)\\
\le& L_3\|h_1-h_2\|_{\infty,\mathcal{I}},
\end{aligned}$$

where

$$L_3=L_3(\nu.)\coloneqq T\mathrm{e}^{L_1(\nu.)T}\|\nu.\|. \tag{A.7}$$

□

## Appendix B. Proof of Proposition 4.3

*Proof.* The unique existence of solutions to the fixed point equation (4.1) is proved by the Banach fixed point theorem, which is analogous to that of [31, Proposition 4.5].

- Continuity in $t$. It follows directly from Proposition 4.2(i).
- Continuity in $x$. Assume $\nu.\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\mathcal{M}_+(Y)))$. We will show $\mathcal{A}[\eta,h]\nu.\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\mathcal{M}_+(Y)))$. It suffices to show that the continuity of measures in $x$ is preserved: $x\mapsto\nu^x_0\circ\Phi^x_{t,0}[\eta,\nu.,h]$ is continuous. Indeed,

$$\begin{aligned}
&d_{\mathsf{BL}}(\Phi^x_{t,0}[\eta,\nu.,h]\#\nu^x_0,\Phi^{x'}_{t,0}[\eta,\nu.,h]\#\nu^{x'}_0)\\
=&\sup_{f\in\mathcal{BL}_1(Y)}\left|\int_Y f\circ\Phi^x_{t,0}[\eta,\nu.,h]\phi\mathrm{d}\nu^x_0(\phi)-f\circ\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi\mathrm{d}\nu^{x'}_0(\phi)\right|\\
\le&\int_Y\left|\Phi^x_{t,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi\right|\mathrm{d}\nu^x_0(\phi)\\
&+\sup_{f\in\mathcal{BL}_1(Y)}\left|\int_Y f\circ\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi\mathrm{d}(\nu^x_0(\phi)-\nu^{x'}_0(\phi))\right|.
\end{aligned}$$

It follows from (A.2) that $\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi$ is Lipschitz continuous in $\phi$ with constant $\mathrm{e}^{L_1(\nu.)T}$, and from (A.5) it follows that

$$\frac{f\circ\Phi^{x'}_{t,0}[\eta,\nu.,h]}{\mathrm{e}^{L_1(\nu.)T}}\in\mathcal{BL}_1(Y).$$

In addition, from (A.1), we have

$$\begin{aligned}
&|V[\eta,\nu.,h](\tau,x,\Phi^x_{\tau,0}[\eta,\nu.,h]\phi)-V[\eta,\nu.,h](\tau,x',\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)|\\
\le&|V[\eta,\nu.,h](\tau,x,\Phi^x_{\tau,0}[\eta,\nu.,h]\phi)-V[\eta,\nu.,h](\tau,x,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)|\\
&+|V[\eta,\nu.,h](\tau,x,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)-V[\eta,\nu.,h](\tau,x',\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)|\\
\le&L_1(\nu.)|\Phi^x_{\tau,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi|+\sum_{\ell=1}^r\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}}g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1}))\\
&\otimes_{j=1}^{k_\ell-1}\mathrm{d}\nu^{y_j}_\tau(\psi_j)\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|\\
&+|h(\tau,x,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)-h(\tau,x',\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)|\\
\le&L_1(\nu.)|\Phi^x_{\tau,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi|+\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|
\end{aligned}$$

$$+\sum_{\ell=1}^{r}\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1})\otimes_{j=1}^{k_\ell-1}\mathrm{d}\nu^{y_j}_\tau(\psi_j)$$
$$\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|$$
$$\leq L_1(\nu.)|\Phi^x_{\tau,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi|+\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|$$
$$+\sum_{\ell=1}^{r}\Big|\int_{X^{k_\ell-1}}\Big(\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1})\otimes_{j=1}^{k_\ell-1}\mathrm{d}\nu^{y_j}_\tau(\psi_j)\Big)$$
$$\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|.$$

This implies that

$$|\Phi^x_{t,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi|$$
$$=\int_0^t|V[\eta,\nu.,h](\tau,x,\Phi^x_{\tau,0}[\eta,\nu.,h]\phi)-V[\eta,\nu.,h](\tau,x',\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi)|\mathrm{d}\tau$$
$$\leq L_1(\nu.)\int_0^t|\Phi^x_{\tau,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi|\mathrm{d}\tau+\int_0^t\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|\mathrm{d}\tau$$
$$+\sum_{\ell=1}^{r}\int_0^t\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\cdots,\psi_{k_\ell-1})\otimes_{j=1}^{k_\ell-1}\mathrm{d}\nu^{y_j}_\tau(\psi_j)$$
$$\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|\mathrm{d}\tau.$$

By Gronwall's inequality,

$$|\Phi^x_{t,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi|$$
$$\leq\Big(\sum_{\ell=1}^{r}\int_0^t\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\cdots,\psi_{\ell_k-1})\otimes_{j=1}^{k_\ell-1}\mathrm{d}\nu^{y_j}_\tau(\psi_j)\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})$$
$$-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|\mathrm{d}\tau+\int_0^t\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|\mathrm{d}\tau\Big)\mathrm{e}^{L_1(\nu.)t}.$$

By (A.5), this further shows that

$$d_{\mathsf{BL}}(\Phi^x_{t,0}[\eta,\nu.,h]\#\nu^x_0,\Phi^{x'}_{t,0}[\eta,\nu.,h]\#\nu^{x'}_0)$$
$$\leq\int_Y|\Phi^x_{t,0}[\eta,\nu.,h]\phi-\Phi^{x'}_{t,0}[\eta,\nu.,h]\phi|\mathrm{d}\nu^x_0(\phi)+\mathrm{e}^{L_1(\nu.)T}d_{\mathsf{BL}}(\nu^x_0,\nu^{x'}_0)$$
$$\leq\mathrm{e}^{L_1(\nu.)T}\Big(\int_0^t\int_Y\sum_{\ell=1}^{r}\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu^{y_j}_\tau(\psi_j)\tag{B.1}$$
$$\mathrm{d}(\eta^x_\ell(y_1,\ldots,y_{k_\ell-1})-\eta^{x'}_\ell(y_1,\ldots,y_{k_\ell-1}))\Big|\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau$$
$$+\|\nu_0\|\int_0^t\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|\mathrm{d}\tau+d_{\mathsf{BL}}(\nu^x_0,\nu^{x'}_0)\Big).$$

Since $\nu.\in\mathcal{C}(\mathcal{I},\mathcal{C}_*(X,\mathcal{M}_+(Y)))$, by Proposition 2.3(iv), $\otimes_{j=1}^{k_\ell-1}\nu^{y_j}_t$ is weakly continuous in $(y_1,\ldots,y_{k_\ell-1})$. By (**A2**), $g_\ell$ is bounded Lipschitz continuous in $\phi,\psi_1,\ldots,\psi_{k_\ell-1}$; and $\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi$ is Lipschitz continuous in $\phi$ by (A.2), we have $g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1})$ is bounded continuous in $\phi$ and $\psi_1\ldots,\psi_{k_\ell-1}$. Hence for $\ell=1,\ldots,r$, $\int_{Y^{k_\ell-1}}g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\ldots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu^{y_j}_\tau(\psi_j)$ is continuous in $(y_1,\ldots,y_{k_\ell-1})$.

Moreover,

$$\Big|\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\dots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu_\tau^{y_j}(\psi_j)\Big|\le \mathcal{BL}(g_\ell)\|\nu.\|^{k_\ell-1}<\infty$$

is also uniformly bounded for $x\in X$. Hence by Proposition 2.3(iii), we know

$$\begin{aligned}\lim_{|x-x'|\to 0}&\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\dots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu_\tau^{y_j}(\psi_j)\\ &\mathrm{d}(\eta_\ell^x(y_1,\dots,y_{k_\ell-1})-\eta_\ell^{x'}(y_1,\dots,y_{k_\ell-1}))\Big|=0,\end{aligned}\tag{B.2}$$

since $\eta\in\mathcal{C}(X,\mathcal{M}_+(X^{k_\ell-1}))$. Notice that

$$\begin{aligned}&\int_0^t\Big|\int_{X^{k_\ell-1}}\int_{Y^{k_\ell-1}} g_\ell(\tau,\Phi^{y}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\dots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu_\tau^{y_j}(\psi_j)\mathrm{d}\eta_\ell^x(y_1,\dots,y_{k_\ell-1})\Big|\mathrm{d}\tau\\ \le&\mathcal{BL}(g_\ell)\|\nu.\|^{k_\ell-1}\|\eta_\ell\|T<\infty,\end{aligned}$$

by the Dominated Convergence Theorem, it follows from (B.2) that

$$\begin{aligned}&\lim_{|x-x'|\to 0}\int_0^t\Big|\int_{X^{k_\ell-1}}\int_Y\cdots\int_Y g_\ell(\tau,\Phi^{x'}_{\tau,0}[\eta,\nu.,h]\phi,\psi_1,\dots,\psi_{k_\ell-1})\mathrm{d}\otimes_{j=1}^{k_\ell-1}\nu_\tau^{y_j}(\psi_j)\\ &\mathrm{d}(\eta_\ell^x(y_1,\dots,y_{k_\ell-1})-\eta_\ell^{x'}(y_1,\dots,y_{k_\ell-1}))\Big|\mathrm{d}\tau=0.\end{aligned}$$

Moreover, by (**A7**) as well as the Dominated Convergence Theorem again,

$$\lim_{|x-x'|\to 0}\int_0^t\sup_{\varphi\in Y}|h(\tau,x,\varphi)-h(\tau,x',\varphi)|\mathrm{d}\tau=0.$$

Since $\nu_0\in\mathcal{C}_*(X,\mathcal{M}_+(Y))$, we have

$$\lim_{|x-x'|\to 0} d_{\mathsf{BL}}(\nu_0^x,\nu_0^{x'})=0.$$

Hence from (B.1) it follows that

$$\lim_{|x-x'|\to 0} d_{\mathsf{BL}}(\Phi^x_{t,0}[\eta,\nu.,h]\#\nu_0^x,\Phi^{x'}_{t,0}[\eta,\nu.,h]\#\nu_0^{x'})=0.$$

The absolute continuity of solutions follow from Proposition 4.2(iv). In the following, we prove properties (i)-(iii) item by item.

(i) Lipschitz continuity in $\nu_0$. It follows from Proposition 4.2(ii) via Gronwall inequality.

(ii) Lipschitz continuity of $\nu.$ in $h$. Assume $\nu_0^1=\nu_0^2$.

We first need to establish the Lipschitz continuity for $\Phi^x_{s,t}[h]$. Note that

$$\begin{aligned}&|\Phi^x_{0,t}[\nu^1,h_1]\phi-\Phi^x_{0,t}[\nu^2,h_2]\phi|\\ \le&|\Phi^x_{0,t}[\nu^1,h_1]\phi-\Phi^x_{0,t}[\nu^1,h_2]\phi|+|\Phi^x_{0,t}[\nu^1,h_2]\phi-\Phi^x_{0,t}[\nu^2,h_2]\phi|.\end{aligned}$$

The second term follows from (A.6). The estimate for the first term follows from Proposition 4.2 (iii). It follows from (A.1) that

$$\begin{aligned}&\le\Big|\int_0^t\big(V[\nu^1,h_1](\tau,x,\Phi^x_{0,\tau}[h_1]\phi)-V[\nu^1,h_2](\tau,x,\Phi^x_{0,\tau}[h_2]\phi)\big)\,\mathrm{d}\tau\Big|\\ &\le\int_0^t\big|V[\nu^1,h_1](\tau,x,\Phi^x_{0,\tau}[h_1]\phi)-V[\nu^1,h_2](\tau,x,\Phi^x_{0,\tau}[h_1]\phi(x))\big|\,\mathrm{d}\tau\\ &\quad+\int_0^t\big|V[\nu^1,h_2](\tau,x,\Phi^x_{0,\tau}[h_1]\phi)-V[\nu^1,h_2](\tau,x,\Phi^x_{0,\tau}[h_2]\phi)\big|\,\mathrm{d}\tau\\ &\le\int_0^t\big|h_1(\tau,x,\Phi^x_{0,\tau}[h_1]\phi(x))-h_2(\tau,x,\Phi^x_{0,\tau}[h_1]\phi(x))\big|\,\mathrm{d}\tau\end{aligned}$$

$$+ L_1(\nu_\cdot)\int_0^t \left|\Phi_{0,\tau}^x[h_1]\phi - \Phi_{0,\tau}^x[h_2]\phi\right| \mathrm{d}\tau.$$

By Gronwall's inequality, we have

$$\left|\Phi_{0,t}^x[\nu^1, h_1]\phi - \Phi_{0,t}^x[\nu^2, h_2]\phi\right| \le \mathrm{e}^{L_1(\nu_\cdot)t}\int_0^t |h_1(\tau, x, \Phi_{0,\tau}^x[h_1]\phi) - h_2(\tau, x, \Phi_{0,\tau}^x[h_1]\phi)|\mathrm{d}\tau.$$

Hence

$$\begin{aligned}
&d_{\mathsf{BL}}(\Phi_{0,t}^x[h_1]\#\nu_0^x, \Phi_{0,t}^x[h_2]\#\nu_0^x)\\
=&\sup_{f\in\mathcal{BL}_1(Y)}\int_Y f(\phi)\mathrm{d}(\Phi_{0,t}^x[h_1]\#\nu_0^x - \Phi_{0,t}^x[h_2]\#\nu_0^x)\\
=&\sup_{f\in\mathcal{BL}_1(Y)}\int_Y \big((f\circ\Phi_{t,0}^x[h_1])(\phi) - (f\circ\Phi_{t,0}^x[h_2])(\phi)\big)\,\mathrm{d}\nu_0^x(\phi)\\
\le&\int_Y \left|\Phi_{t,0}^x[h_1]\phi - \Phi_{t,0}^x[h_2]\phi)\right| \mathrm{d}\nu_0^x(\phi)\\
\le&\mathrm{e}^{L_1(\nu_\cdot)t}\int_0^t\int_Y |h_1(\tau, x, \Phi_{0,\tau}^x[h_1]\phi) - h_2(\tau, x, \Phi_{0,\tau}^x[h_1]\phi)|\mathrm{d}\nu_0^x(\phi)\mathrm{d}\tau\\
\le&\mathrm{e}^{L_1(\nu_\cdot)t}\int_0^t\int_Y |h_1(\tau, x, \phi) - h_2(\tau, x, \phi)|\mathrm{d}\nu_\tau^x(\phi)\mathrm{d}\tau\\
\le&L_3\|h_1 - h_2\|_{\infty,\mathcal{I}},
\end{aligned}$$

where $L_3$ is defined in (A.7), which further implies that

$$\begin{aligned}
&d_\infty(\mathcal{A}[\eta, h_1](\nu_t), \mathcal{A}[\eta, h_2](\nu_t))\\
=&\sup_{x\in X} d_{\mathsf{BL}}(\Phi_{0,t}^x[h_1]\#\nu_0^x, \Phi_{0,t}^x[h_2]\#\nu_0^x) \le L_3(\nu_\cdot)\|h_1 - h_2\|_{\infty,\mathcal{I}}.
\end{aligned}$$

(iii) Continuous dependence on $\eta$. Since $\nu_0 \in \mathcal{C}_*(X, \mathcal{M}_+(Y))$, we have $\nu_\cdot \in \mathcal{C}(\mathcal{I}, \mathcal{C}_*(X, \mathcal{M}_+(Y)))$. Based on the continuous dependence on the initial distributions proved in (i), as well as a triangle inequality, it suffices to prove the case assuming

$$\nu_0^K = \nu_0$$

Let $\nu_\cdot^K \in \mathcal{C}(\mathcal{I}, \mathcal{B}_*(X, \mathcal{M}_+(Y)))$ with $\nu_0^K = \nu_0$ be the solutions to the fixed point equations

$$\nu_t = \mathcal{A}[\eta, h]\nu_t, \quad \nu_t^K = \mathcal{A}[\eta^K, h]\nu_t^K, \quad t \in \mathcal{I},$$

where $\eta^K = (\eta_\ell^K)_{\ell=1}^r$. Assume

$$\lim_{K\to\infty} d_\infty(\eta^\ell, \eta^{K,\ell}) = 0, \quad \ell = 1, \ldots, r.$$

In the following, we show

$$\lim_{K\to\infty} d_\infty(\mathcal{A}[\eta, h]\nu_t, \mathcal{A}[\eta^K, h]\nu_t^K) = 0, \quad t \in \mathcal{I}.$$

By the triangle inequality,

$$\begin{aligned}
d_{\mathsf{BL}}(\nu_t^x, \nu_t^{K,x}) =&d_{\mathsf{BL}}(\Phi_{t,0}^x[\eta, \nu_\cdot]_\#\nu_0^x, \Phi_{t,0}^x[\eta^K, \nu_\cdot^K]_\#\nu_0^x)\\
\le&d_{\mathsf{BL}}(\Phi_{t,0}^x[\eta^K, \nu_\cdot]_\#\nu_0^x, \Phi_{t,0}^x[\eta^K, \nu_\cdot^K]_\#\nu_0^x)\\
&+ d_{\mathsf{BL}}(\Phi_{t,0}^x[\eta, \nu_\cdot]_\#\nu_0^x, \Phi_{t,0}^x[\eta^K, \nu_\cdot]_\#\nu_0^x).
\end{aligned} \tag{B.3}$$

From (A.4) it follows that

$$\begin{aligned}
&d_{\mathsf{BL}}(\Phi_{t,0}^x[\eta^K, \nu_\cdot]_\#\nu_0^x, \Phi_{t,0}^x[\eta^K, \nu_\cdot^K]_\#\nu_0^x)\\
\le&\int_Y |\Phi_{t,0}^x[\eta^K, \nu_\cdot]\phi - \Phi_{t,0}^x[\eta^K, \nu_\cdot^K]\phi|\mathrm{d}\nu_0^x(\phi) =: \beta_x^K(t),
\end{aligned}$$

$$\leq L_{2,K}(\eta^K)\|\nu_\cdot\|\mathrm{e}^{L_{1,K}(\nu^K_\cdot)t}\int_0^t d_\infty(\nu_\tau,\nu^K_\tau)\mathrm{e}^{-L_{1,K}(\nu^K_\cdot)\tau}\mathrm{d}\tau, \tag{B.4}$$

where the index $K$ in the constants indicates the dependence on $K$.

We now estimate the second term. By (A.1),

$$\begin{aligned}
&d_{\mathsf{BL}}(\Phi^x_{t,0}[\eta,\nu_\cdot]_\#\nu^x_0,\Phi^x_{t,0}[\eta^K,\nu_\cdot]_\#\nu^x_0)\\
=&\sup_{f\in\mathcal{BL}_1(Y)}\int_Y f\mathrm{d}(\Phi^x_{t,0}[\eta,\nu_\cdot]_\#\nu^x_0-\Phi^x_{t,0}[\eta^K,\nu_\cdot]_\#\nu^x_0)\\
=&\sup_{f\in\mathcal{BL}_1(Y)}\int_Y \big((f\circ\Phi^x_{t,0}[\eta,\nu_\cdot])(\phi)-(f\circ\Phi^x_{t,0}[\eta^K,\nu_\cdot])(\phi)\big)\,\mathrm{d}\nu^x_0(\phi)\\
\leq&\int_Y|\Phi^x_{t,0}[\eta,\nu_\cdot]\phi-\Phi^x_{t,0}[\eta^K,\nu_\cdot]\phi)|\mathrm{d}\nu^x_0(\phi)=:\gamma^K_x(t)\\
=&\int_Y\left|\int_0^t(V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta,\nu_\cdot]\phi)-V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi))\mathrm{d}\tau\right|\mathrm{d}\nu^x_0(\phi)\\
\leq&\int_Y\left|\int_0^t(V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta,\nu_\cdot]\phi)-V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi))\mathrm{d}\tau\right|\mathrm{d}\nu^x_0(\phi)\\
&+\int_Y\left|\int_0^t(V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)-V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi))\mathrm{d}\tau\right|\mathrm{d}\nu^x_0(\phi)\\
\leq&\int_Y\int_0^t\big|V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta,\nu_\cdot]\phi)-V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)\big|\,\mathrm{d}\tau\mathrm{d}\nu^x_0(\phi)\\
&+\int_Y\int_0^t\big|V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)-V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)\big|\,\mathrm{d}\tau\mathrm{d}\nu^x_0(\phi)\\
\leq&L_1(\nu_\cdot)\int_0^t\int_Y\big|\Phi^x_{\tau,0}[\eta,\nu_\cdot]\phi-\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi\big|\,\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau\\
&+\int_0^t\int_Y\big|V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)-V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi)\big|\,\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau\\
+&\int_0^t\int_Y\big|V[\eta,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi)-V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi)\big|\,\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau\\
+&\int_0^t\int_Y\big|V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi)-V[\eta^K,\nu_\cdot](\tau,x,\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi)\big|\,\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau\\
\leq&L_1(\nu_\cdot)\int_0^t\gamma^K_x(\tau)\mathrm{d}\tau+L_1(\nu_\cdot)\int_0^t\int_Y\big|\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi-\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi\big|\,\mathrm{d}\nu^x_0(\phi)\mathrm{d}\tau\\
&+\int_0^t\int_Y\big|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\big|\,\mathrm{d}\nu^{K,x}_\tau(\phi)\mathrm{d}\tau\\
&+L_{1,K}(\nu_\cdot)\int_0^t\int_Y\big|\Phi^x_{\tau,0}[\eta^K,\nu^K_\cdot]\phi-\Phi^x_{\tau,0}[\eta^K,\nu_\cdot]\phi\big|\,\mathrm{d}\nu^x_0(\psi)\mathrm{d}\tau\\
=&L_1(\nu_\cdot)\int_0^t\gamma^K_x(\tau)\mathrm{d}\tau+(L_1(\nu_\cdot)+L_{1,K}(\nu_\cdot))\int_0^t\beta^K_x(\tau)\mathrm{d}\tau\\
&+\int_0^t\int_Y\big|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\big|\,\mathrm{d}\nu^{K,x}_\tau(\phi)\mathrm{d}\tau.
\end{aligned}$$

To obtain further estimates, let

$$\zeta^K_x(\tau)\coloneqq\int_Y\big|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\big|\,\mathrm{d}\nu^x_\tau(\phi).$$

By the triangle inequality,

$$\begin{aligned}
&\int_0^t \int_Y \left|V[\eta,\nu.](\tau,x,\phi)-V[\eta^K,\nu.](\tau,x,\phi)\right| \mathrm{d}\nu_\tau^{K,x}(\phi)\mathrm{d}\tau\\
\le& \int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\\
&+\int_0^t \left|\int_Y \left|V[\eta,\nu.](\tau,x,\phi)-V[\eta^K,\nu.](\tau,x,\phi)\right| \mathrm{d}(\nu_\tau^{K,x}(\phi)-\nu_\tau^x(\phi))\right|\mathrm{d}\tau.
\end{aligned}$$

Recall that

$$\sup_{x\in X}\eta_\ell^{K,x}(Y)\le \sup_{x\in X}\Big(\eta_\ell^x(Y)+d_\infty(\eta_\ell^x,\eta_\ell^{K,x})\Big),\quad \ell=1,\dots,r.$$

Moreover, since

$$\lim_{k\to\infty} d_\infty(\nu_0^K,\nu_0)=0$$

and

$$\sum_{\ell=1}^r \left|\|\eta_\ell\|-\|\eta_\ell^K\|\right|\le \sum_{\ell=1}^r d_\infty(\eta_\ell,\eta_\ell^K)\to 0,\quad \text{as}\quad K\to\infty,$$

we have there exists some $b>0$ independent of $K$ such that

(B.5) $$\sup_{K\in\mathbb{N}}(L_1(\nu.)+L_{1,K}(\nu.))\le b,\quad \sup_{K\in\mathbb{N}}(L_2(\eta,\nu.,\nu.^K)+L_2(\eta^K,\nu.,\nu.^K))\le b.$$

Let

$$f_K(\tau,x,\varphi):=\left|V[\eta,\nu.](\tau,x,\varphi)-V[\eta^K,\nu.](\tau,x,\varphi)\right|.$$

Using analogous arguments as in the proof for the limit [31, (A.2)], we have

$$\lim_{K\to\infty}\sup_{x\in X}|f_K(\tau,x,\varphi)|=0,$$

which further implies that $f_K$ is bounded. Moreover, it follows from (A.1) again that

$$\begin{aligned}
&|f_K(\tau,x,\varphi)-f_K(\tau,x,\phi)|\\
\le&|V[\eta,\nu.](\tau,x,\varphi)-V[\eta,\nu.](\tau,x,\phi)|+\left|V[\eta^K,\nu.](\tau,x,\varphi)-V[\eta^K,\nu.](\tau,x,\phi)\right|\\
\le&(L_1+L_{1,K})|\varphi-\phi|\le b|\varphi-\phi|.
\end{aligned}$$

Further, by (A.1), one can show that $f_K(\tau,x,\varphi)$ is bounded Lipschitz continuous in $\varphi$ with some constant $\widehat{b}>0$ such that

$$\sup_{K\in\mathbb{N}}\sup_{\tau\in\mathcal{I}}\sup_{x\in X}\mathcal{BL}(f_K(\tau,x,\cdot))\le\widehat{b}.$$

Hence

$$\begin{aligned}
&\left|\int_0^t\int_Y \left|V[\eta,\nu.](\tau,x,\phi)-V[\eta^K,\nu.](\tau,x,\phi)\right| \mathrm{d}(\nu_\tau^{K,x}(\phi)-\nu_\tau^x(\phi))\mathrm{d}\tau\right|\\
\le&\widehat{b}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau.
\end{aligned}$$

This further implies that

$$\gamma_x^K(t)\le L_1\int_0^t\gamma_x^K(\tau)\mathrm{d}\tau+b\int_0^t\beta_x^K(\tau)\mathrm{d}\tau+\widehat{b}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau+\int_0^t\zeta_x^K(\tau)\mathrm{d}\tau.$$

By Gronwall's inequality, we have

$$\gamma_x^K(t)\le \mathrm{e}^{L_1t}\left(b\int_0^t\beta_x(\tau)\mathrm{d}\tau+\widehat{b}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau+\int_0^t\zeta_x^K(\tau)\mathrm{d}\tau\right)$$

Hence by (A.4), (B.3), (B.4), and (B.5), we have for $t\in\mathcal{I}$,

$$d_{\mathsf{BL}}(\nu_t^x,\nu_t^{K,x})\le\beta_x^K(t)+\gamma_x^K(t)$$

$$
\begin{aligned}
&\leq \beta_x^K(t) + \mathrm{e}^{L_1 t}\left(b\int_0^t \beta_x(\tau)\mathrm{d}\tau + \widehat{b}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau + \int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\right)\\
&\leq L_{2,K}\|\nu_\cdot\|\int_0^t \mathrm{e}^{L_{1,K}(t-\tau)}d_\infty(\nu_\tau^K,\nu_\tau)\mathrm{d}\tau\\
&\quad + \mathrm{e}^{L_1 t}b\int_0^t \beta_x(\tau)\mathrm{d}\tau + \widehat{b}\mathrm{e}^{L_1 t}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau + \mathrm{e}^{L_1 t}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\\
&\leq (b-L_2)\|\nu_\cdot\|\int_0^t \mathrm{e}^{(b-L_1)(t-\tau)}d_\infty(\nu_\tau^K,\nu_\tau)\mathrm{d}\tau\\
&\quad + \mathrm{e}^{L_1 t}b\int_0^t (b-L_2)\|\nu_\cdot\|\int_0^\tau \mathrm{e}^{(b-L_1)(\tau-s)}d_\infty(\nu_s^K,\nu_s)\mathrm{d}s\mathrm{d}\tau\\
&\quad + \widehat{b}\mathrm{e}^{L_1 t}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau + \mathrm{e}^{L_1 t}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\\
&\leq (b-L_2)\|\nu_\cdot\|(1+\mathrm{e}^{L_1 t}bt)\int_0^t \mathrm{e}^{(b-L_1)(t-\tau)}d_\infty(\nu_\tau^K,\nu_\tau)\mathrm{d}\tau\\
&\quad + \widehat{b}\mathrm{e}^{L_1 t}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau + \mathrm{e}^{L_1 t}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\\
&\leq (b-L_2)\|\nu_\cdot\|(1+bT)\mathrm{e}^{L_1 t}\int_0^t \mathrm{e}^{(b-L_1)(t-\tau)}d_\infty(\nu_\tau^K,\nu_\tau)\mathrm{d}\tau\\
&\quad + \widehat{b}\mathrm{e}^{L_1 t}\int_0^t d_{\mathsf{BL}}(\nu_\tau^{K,x},\nu_\tau^x)\mathrm{d}\tau + \mathrm{e}^{L_1 t}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau\\
&\leq L_4\int_0^t d_\infty(\nu_\tau^K,\nu_\tau)\mathrm{d}\tau + \mathrm{e}^{L_1 T}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau,
\end{aligned}
$$

where $L_4 \coloneqq \mathrm{e}^{bT}(b-L_2)\|\nu_\cdot\|(1+bT)+\mathrm{e}^{L_1 T}\widehat{b}$. By Gronwall's inequality,

$$d_\infty(\nu_t^K,\nu_t) \leq \mathrm{e}^{L_4+L_1 T}\sup_{x\in X}\int_0^t \zeta_x^K(\tau)\mathrm{d}\tau.$$

To show $\lim_{K\to\infty} d_0(\nu_\cdot^K,\nu_\cdot)=0$, it suffices to show

$$\lim_{K\to\infty}\sup_{x\in X}\int_0^T \zeta_x^K(\tau)\mathrm{d}\tau = 0.$$

For every $t\in\mathcal{I}$, define $\widehat{\nu}_t \equiv \sup_{x\in X}\nu_t^x$:

$$\widehat{\nu}_t(E) = \sup_{x\in X}\nu_t^x(E),\quad \forall E\in\mathcal{B}(Y).$$

Since $\nu_t\in\mathcal{B}(X,\mathcal{M}_+(Y))$, it is easy to show that $\widehat{\nu}_t\in\mathcal{M}_+(Y)$. By Fatou's lemma,

$$
\begin{aligned}
\sup_{x\in X}\zeta_x^K(\tau) &= \sup_{x\in X}\int_Y \left|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\right|\mathrm{d}\nu_\tau^x(\phi)\\
&\leq \sup_{x\in X}\int_Y \left|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\right|\mathrm{d}\widehat{\nu}_\tau(\phi)\\
&\leq \int_Y \sup_{x\in X}\left|V[\eta,\nu_\cdot](\tau,x,\phi)-V[\eta^K,\nu_\cdot](\tau,x,\phi)\right|\mathrm{d}\widehat{\nu}_\tau(\phi).
\end{aligned}
$$

Since $\nu_\cdot\in\mathcal{B}(\mathcal{I},\mathcal{M}_+(Y))$, we have $\widehat{\nu}_\cdot\in\mathcal{B}(\mathcal{I},\mathcal{M}_+(Y))$. Using analogous arguments as those for proving [31, Proposition 3.2], we have

$$\sup_{x\in X}\int_0^T \zeta_x^K(\tau)\mathrm{d}\tau \leq \int_0^T \sup_{x\in X}\zeta_x^K(\tau)\mathrm{d}\tau$$

$$\leq \int_0^T \int_Y \sup_{x\in X} \left|V[\eta,\nu.](\tau,x,\phi) - V[\eta^K,\nu.](\tau,x,\phi)\right| \mathrm{d}\widehat{\nu}_\tau(\phi)\mathrm{d}\tau \to 0, \quad \text{as } K\to\infty.$$

$\square$