SIAM J. MATH. ANAL.
Vol. 50, No. 3, pp. 2441–2465

THE MEAN FIELD EQUATION FOR THE KURAMOTO MODEL ON GRAPH SEQUENCES WITH NON-LIPSCHITZ LIMIT*

DMITRY KALIUZHNYI-VERBOVETSKYI† AND GEORGI S. MEDVEDEV†

Abstract. The Kuramoto model (KM) of coupled phase oscillators on graphs provides the most influential framework for studying collective dynamics and synchronization. It exhibits a rich repertoire of dynamical regimes. Since the work of Strogatz and Mirollo [*J. Stat. Phys.*, 63 (1991), pp. 613–635], the mean field equation derived in the limit as the number of oscillators in the KM goes to infinity has been the key to understanding a number of interesting effects, including the onset of synchronization and chimera states. In this work, we study the mathematical basis of the mean field equation as an approximation of the discrete KM. Specifically, we extend the Neunzert's method of rigorous justification of the mean field equation (cf. [H. Neunzert, *Fluid Dyn. Trans.*, 9 (1978), pp. 229–254]) to cover interacting dynamical systems on graphs. We then apply it to the KM on convergent graph sequences with non-Lipschitz limit. This family of graphs includes many graphs that are of interest in applications, e.g., nearest-neighbor and small-world graphs. The approaches for justifying the mean field limit for the KM proposed previously in [C. Lancellotti, *Transp. Theory Statist. Phys.*, 34 (2005), pp. 523–535; H. Chiba and G. S. Medvedev, arXiv:1612.06493, 2016] do not cover the non-Lipschitz case.

Key words. mean field limit, interacting dynamical systems, graph limit, small-world graph

AMS subject classifications. 34C15, 45J05, 45L05, 74A25, 92D25

DOI. 10.1137/17M1134007

1. Introduction.

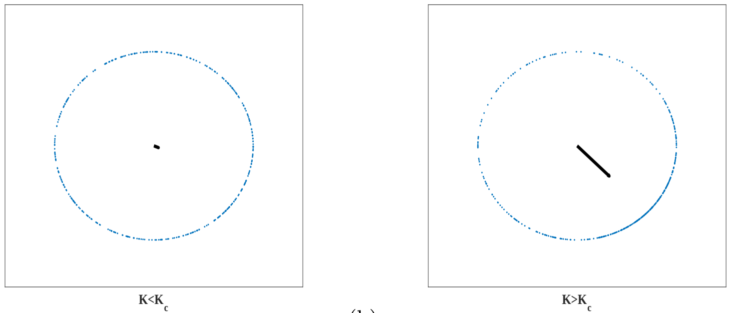
The Kuramoto model (KM) of coupled phase oscillators provides a useful framework for studying collective behavior in large ensembles of interacting dynamical systems. It is derived from a weakly coupled system of nonlinear oscillators, which are described by autonomous systems of ordinary differential equations possessing a stable limit cycle [8]. Originally, Kuramoto considered all-to-all coupled systems, in which each oscillator interacts with all other oscillators in exactly the same way. In this case, the KM has the following form:

$$\dot{u}_{n,i} = \omega_i + \frac{K}{n}\sum_{j=1}^{n} \sin\left(u_{n,j} - u_{n,i} + \alpha\right), \quad i \in [n] := \{1, 2, \ldots, n\}. \tag{1.1}$$

Here, $u_{n,i} : \mathbb{R}^+ \to \mathbb{S} := \mathbb{R}/2\pi\mathbb{Z}$ stands for the phase of oscillator i as a function of time, ω_i is its intrinsic frequency, K is the coupling strength, and α is a parameter defining the type of interactions.

Despite its simple form, the KM (1.1) features a rich repertoire of interesting dynamical effects. For the purpose of this review, we mention the onset of synchronization in (1.1) with randomly distributed intrinsic frequencies ω_i (Figure 1(a), (b)) (cf. [21]) and chimera states, interesting spatio-temporal patterns combining coherent and incoherent behaviors [10, 1]. The mathematical analysis of these and many other dynamical regimes uses the mean field equation, derived in the limit when the number of oscillators goes to infinity [22]. The mean field equation is a partial differential

*Received by the editors June 12, 2017; accepted for publication (in revised form) January 9, 2018; published electronically May 8, 2018.
http://www.siam.org/journals/sima/50-3/M113400.html
Funding: The work of the second author was funded in part by the NSF grants DMS 1412066 and 1715161.
†Department of Mathematics, Drexel University, Philadelphia, PA 19104 (dmitryk@math.drexel.edu, medvedev@drexel.edu).



(a) (b)

FIG. 1. *The distribution of the phase oscillators in the KM* (1.1) *for values of K below* (a) *and above* (b) *the critical value K_c. In the former plot, the distribution is approximately uniform, whereas the latter plot exhibits a pronounced cluster. The bold vectors depict the order parameter, whose length reflects the degree of coherence. The random distribution of the oscillators shown in these plots can be effectively analyzed with the mean field equation* (1.18)*. In particular, the mean field analysis determines the critical value K_c.*

equation for the probability density describing the distribution of the phases on $\mathbb{S}$. We discuss the mean field equation in more detail below.

Recently, there has been a growing interest in the dynamics of coupled dynamical systems on graphs [20]. In the KM on a graph, each oscillator is placed at a node of an undirected graph $\Gamma_n = \langle V(\Gamma_n), E(\Gamma_n)\rangle$. Here, $V(\Gamma_n) = [n]$ stands for the node set of Γ_n, and $E(\Gamma_n)$ denotes its edge set. The oscillator i interacts only with the oscillators at the adjacent nodes:

$$\dot{u}_{n,i} = \omega_i + \frac{K}{n} \sum_{j:j\sim i} \sin\left(u_{n,j} - u_{n,i} + \alpha\right), \quad i \in [n], \tag{1.2}$$

where $j \sim i$ is a shorthand for $\{i,j\} \in E(\Gamma_n)$.

Clearly, one can not expect limiting behavior of solutions of (1.2) as $n \to \infty$, unless the graph sequence $\{\Gamma_n\}$ is convergent in the appropriate sense. In the present paper, we use the following construction of the convergent sequence $\{\Gamma_n\}$. Let W be a symmetric measurable function on the unit square $I^2 := [0,1]^2$. W is called a graphon. It will be used to define the asymptotic behavior of $\{\Gamma_n\}$. Further, let

$$X_n = \{x_{n,1}, x_{n,2}, \ldots, x_{n,n}\}, \quad x_{n,i} = i/n,\ i \in [n], \tag{1.3}$$

and

$$W_{n,ij} := n^2 \int_{I_{n,i}\times I_{n,j}} W(x,y)dxdy, \quad I_{n,i} := [x_{n,(i-1)}, x_{n,i}),\ i,j \in [n]. \tag{1.4}$$

The weighted graph $\Gamma_n = G(W, X_n)$ on n nodes is defined as follows. The vertex set is $V(\Gamma_n) = [n]$ and the edge set is

$$E(\Gamma_n) = \left\{\{i,j\} :\ W_{n,ij} \neq 0,\ i,j \in [n]\right\}. \tag{1.5}$$

Each edge $\{i,j\} \in E(\Gamma_n)$ is supplied with the weight $W_{n,ij}$.[1]

[1]There are several possible ways of defining the weights $W_{n,ij}, i,j \in [n]$ (see Remark 3.5).

The KM on $\Gamma_n = G(W, X_n)$ has the following form:

$$\dot{u}_{n,i} = \omega_i + \frac{K}{n}\sum_{j=1}^{n} W_{n,ij} \sin\left(u_{n,j} - u_{n,i} + \alpha\right), \quad i \in [n]. \tag{1.6}$$

For different W (1.6) implements the KM on a variety of simple and weighted graphs. Moreover, it provides an effective approximation of the KM on random graphs. Indeed, let $\bar{\Gamma}_n = G_r(X_n, W)$ be a random graph on n nodes, whose edge set is defined as follows:

$$\mathbb{P}\left(\{i, j\} \in E(\Gamma_n)\right) = W_{n,ij}, \tag{1.7}$$

assuming the range of W is $[0, 1]$. The decision for each pair $\{i, j\}$ is made independently from the decisions on other pairs. $\bar{\Gamma}_n = G_r(X_n, W)$ is called a W-random graph [12].

The KM on the W-random graph $\bar{\Gamma}_n = G_r(X_n, W)$ has the form

$$\dot{\bar{u}}_{n,i} = \omega_i + Kn^{-1}\sum_{j=1}^{n} e_{n,ij} \sin(\bar{u}_{n,j} - \bar{u}_{n,i} + \alpha), \quad i \in [n], \tag{1.8}$$

where $e_{n,ij}, 1 \le i \le j \le n$, are independent Bernoulli RVs,

$$\mathbb{P}(e_{n,ij} = 1) = W_{n,ij},$$

and $e_{n,ij} = e_{n,ji}$.

The following lemma shows that the deterministic model (1.6) approximates the KM on the random graph $\bar{\Gamma}_n$ (1.8).

Lemma 1.1. *Let $u_n(t)$ and $\bar{u}_n(t)$ denote solutions of the IVP for* (1.6) *and* (1.8), *respectively. Suppose that the initial data for these problems coincide $u_n(0) = \bar{u}_n(0)$. Then for any $\delta \in (0, 1)$,*

$$\sup_{t\in[0,T]} \|u_n(t) - \bar{u}_n(t)\|_{1,n} \le C_1 e^{C_2 T} n^{\frac{-1+\delta}{4}} \quad \mathbb{P} - \textit{almost surely (a.s.)}, \tag{1.9}$$

where C_1 and C_2 are positive constants independent of n and T, $u_n = (u_1, u_2, \ldots, u_n)$, $\bar{u}_n = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$, and

$$\|u_n\|_{1,n} = \sqrt{n^{-1}\sum_{i=1}^{n} u_{ni}^2} \tag{1.10}$$

is a discrete L^2-norm.

Proof. See Appendix A. ∎

Example 1.2. A few examples are in order.

1. Let $W(x, y) \equiv p \in (0, 1)$. Then $\bar{\Gamma}_n = G_r(X_n, W)$ is an Erdős–Rényi graph (Figure 2(a)).
2. Let

$$W_{p,h}(x, y) = \begin{cases} 1 - p, & d_{\mathbb{S}}(2\pi x, 2\pi y) \le 2\pi h, \\ p & \text{otherwise,} \end{cases} \tag{1.11}$$

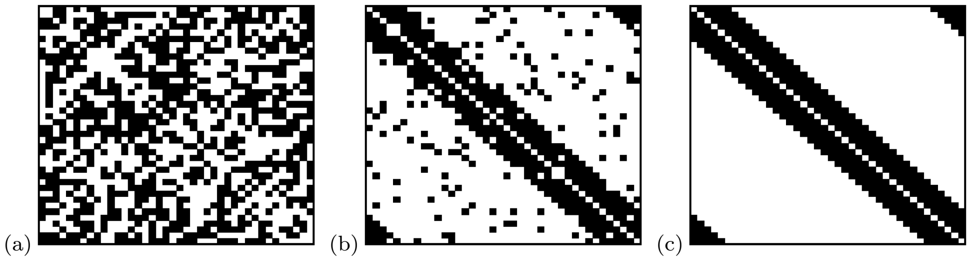


FIG. 2. *The pixel pictures representing adjacency matrices of the Erdős–Rényi* (a), *small-world* (b), *and nearest-neighbor* (c) *graphs.*

where $p, h \in (0, 1/2)$ are two parameters and

$$d_{\mathbb{S}}(\theta, \theta') = \min\{|\theta - \theta'|, 2\pi - |\theta - \theta'|\} \tag{1.12}$$

is the distance on $\mathbb{S}$. Then $\bar{\Gamma}_n = G_r(X_n, W_{p,h})$ is a W-small-world graph [17] (Figure 2(b)).

3. $\Gamma_n = G(X_n, W_{1,h})$ is a $\lfloor h^{-1} \rfloor$-nearest-neighbor graph (Figure 2(c)).

For more examples, we refer an interested reader to [3].

Remark 1.3. For simplicity, we restrict the presentation to the KM on dense graphs. The KM on W-random graphs (1.8) easily extends to sparse graphs like scale-free graphs. (See [9] for details.)

Below, we will focus on the deterministic model (1.6). All results for this model can be extended to the KM on random graphs via Lemma 1.1. Furthermore, from now on we will assume that all intrinsic frequencies in (1.6) are the same $\omega_i = \omega$, $i \in [n]$, and, thus, ω can be set to 0 by switching to the rotating frame of coordinates. Extending the analysis in the main part of this paper to models with distributed frequencies ω_i is straightforward (see, e.g., [3]), but it complicates the presentation. We will comment on the adjustments in the analysis that are necessary to cover the distributed intrinsic frequencies case in section 4. Until then we consider the following system of n coupled oscillators on $\Gamma_n = G(W, X_n)$:

$$\dot{u}_{n,i} = n^{-1} \sum_{j=1}^{n} W_{n,ij} D(u_{n,j} - u_{n,i}), \tag{1.13}$$

$$u_{n,i}(0) = u_{n,i}^0, \quad i \in [n], \tag{1.14}$$

where D is a Lipschitz continuous 2π-periodic function.

For the remainder of this section and in the next section, we assume $W \in L^\infty(I^2)$. Without loss of generality, we assume

$$\sup_{(x,y)\in I^2} |W(x,y)| \le 1, \quad \max_{u \in \mathbb{S}} |D(u)| \le 1, \tag{1.15}$$

and

$$|D(u) - D(v)| \le |u - v| \quad \forall u, v \in \mathbb{S}. \tag{1.16}$$

In addition, we assume that the graphon W satisfies the following condition:

$$\lim_{\delta \to 0} \int_I |W(x + \delta, y) - W(x, y)| dy = 0 \quad \forall x \in I. \tag{1.17}$$

Having defined the KMs on deterministic and random graphs (1.6) and (1.8), respectively, we will now turn to the mean field limit:

$$\frac{\partial}{\partial t}\rho(t,u,x) + \frac{\partial}{\partial u}\{V(t,u,x)\rho(t,u,x)\} = 0, \tag{1.18}$$

where

$$V(t,u,x) = \int_I \int_{\mathbb{S}} W(x,y)D(v-u)\rho(t,v,y)dvdy. \tag{1.19}$$

The initial condition

$$\rho(0,u,x) = \rho^0(u,x) \in L^1(G),\ G := \mathbb{S} \times I, \tag{1.20}$$

is a probability density function on $\mathbb{S}$ for $x \in I$ almost everywhere (a.e.), i.e., $\rho_0 \geq 0$ and

$$\int_{\mathbb{S}} \rho^0(u,x)du = 1 \quad x \in I \text{ a.e.} \tag{1.21}$$

In the continuum limit as $n \to \infty$, the nodes of Γ_n fill out I. Thus, heuristically, $\rho(t,u,x)$ in (1.18) stands for the density of the probability distribution of the phase of the oscillator at $x \in I$ on $\mathbb{S}$ at time $t \geq 0$. As we will see below, this probability distribution is indeed continuous for $t > 0$, provided that the initial conditions for the discrete problem (1.13), (1.14) converge weakly to the probability distribution with density (1.20). In fact, in [3] it is shown that in this case, the empirical measure on the Borel subsets of G,

$$\mu_t^n(A) = n^{-1}\sum_{i=1}^n \mathbf{1}_A((u_{ni}(t), x_{ni})), \quad A \in \mathcal{B}(G), \tag{1.22}$$

converges weakly to the absolutely continuous measure

$$\mu_t(A) = \int\int_A \rho(t,u,x)dudx, \quad A \in \mathcal{B}(G). \tag{1.23}$$

The analysis in [3], which extends the analysis of the all-to-all coupled KM (1.1) by Lancellotti [11], relies on the Lipschitz continuity of W. This is the essential assumption of the Neunzert's fixed point argument that lies at the heart of the method used in [11, 3]. This puts the KM on such common graphs as the small-world and k-nearest-neighbor ones out of the scope of applications of [3] (see Example 1.2). It is the goal of the present paper to fill this gap. Specifically, we extend the Neunzert's method to the KM on convergent families of graphs with non-Lipschitz limits. Our results apply to a general model of n interacting particles on a graph (cf. [7]). However, for concreteness and in view of the diverse applications of the KM, in this paper, we present our method in the context of the KM of coupled phase oscillators.

The organization of this paper is as follows. In the next section, we revise the Neunzert's fixed point theory to adapt it to the KM on convergent graph sequences. This includes a careful choice of the underlying metric space in subsection 2.1, setting up the fixed point equation in subsection 2.2, proving the existence and uniqueness of a solution of the fixed point equation in section 2, and showing continuous dependence on the initial data in subsection 2.4 and on the graphon W in subsection 2.5. In section 3, we apply the fixed point theory to the KM on graphs. To this end, we first apply it to an auxiliary problem and then show that this problem approximates the original KM on graphs. We conclude with a brief discussion of our results in section 4.

2. The fixed point equation. In this section, we extend Neunzert's fixed point method, so that it can be used for the analysis of the IVP for the mean field equation (1.18), (1.19) with non-Lipschitz kernel W.

2.1. The metric space. Let $\mathcal{M}$ denote the space of Borel probability measures on $\mathbb{S}$. The bounded Lipschitz distance on $\mathcal{M}$ is given by

$$d(\mu,\eta) = \sup_{f\in\mathcal{L}} \left| \int_{\mathbb{S}} f(v)\,(d\mu(v) - d\eta(v)) \right|, \tag{2.1}$$

where $\mathcal{L}$ stands for the class of Lipschitz continuous functions on $\mathbb{S}$ with Lipschitz constant at most 1 (cf. [4]). $\langle \mathcal{M}, d\rangle$ is a complete metric space.

Consider the set of measurable $\mathcal{M}$-valued functions[2] $\bar{\mu} : x \mapsto \mu^x \in \mathcal{M}$

$$\bar{\mathcal{M}} := \{\bar{\mu} : I \to \mathcal{M}\}.$$

Equip $\bar{\mathcal{M}}$ with the metric

$$\bar{d}(\bar{\mu},\bar{\eta}) = \int_I d(\mu^x,\eta^x)\,dx. \tag{2.2}$$

LEMMA 2.1. *$\langle\bar{\mathcal{M}},\bar{d}\rangle$ is a complete metric space.*

Proof. Since d is a metric, it is straightforward that $\bar{d}$ is a metric as well. In order to prove the completeness of $\langle\bar{\mathcal{M}},\bar{d}\rangle$, take a Cauchy sequence $\{\bar{\mu}_n\}$ in $\bar{\mathcal{M}}$. Then there is an increasing sequence of indices n_k such that

$$\bar{d}(\mu_{n_k},\mu_{n_{k+1}}) = \int_I d\left(\mu^x_{n_k},\mu^x_{n_{k+1}}\right)\,dx < \frac{1}{2^{k+1}}, \qquad k = 1,2,\ldots.$$

By Levi's theorem, the series

$$\sum_{k=1}^{\infty} d\left(\mu^x_{n_k},\mu^x_{n_{k+1}}\right)$$

converges for a.e. $x \in I$ to some measurable function $f(x)$, and

$$\sum_{k=1}^{\infty}\int_I d\left(\mu^x_{n_k},\mu^x_{n_{k+1}}\right)\,dx = \int_I f(x)\,dx.$$

Since, for every i,j with $j > i$,

$$d\left(\mu^x_{n_i},\mu^x_{n_j}\right) \le \sum_{k=i}^{j-1} d\left(\mu^x_{n_k},\mu^x_{n_{k+1}}\right),$$

the sequence $\{\mu^x_{n_k}\}$ is Cauchy for a.e. $x \in I$. Since the metric space $\langle\mathcal{M},d\rangle$ is complete, there exists the limit

$$\mu^x = \lim_{k\to\infty}\mu^x_{n_k}, \qquad \text{a.e. } x \in I.$$

Extending the definition of μ^x in an arbitrary way to all of I, we obtain a function

$$\bar{\mu} := \{\mu^x\}\colon I \to \mathcal{M}$$

[2] $\bar{\mu} : I \to \mathcal{M}$ is called measurable if the preimage of an open set in $\mathcal{M}$ is a Lebesgue measurable subset of I.

which is measurable as an a.e. pointwise limit of measurable functions. Thus $\bar{\mu} \in \bar{\mathcal{M}}$. Next, for every i, j with $j > i$, we have

$$\begin{aligned}\bar{d}(\bar{\mu}_{n_i}, \bar{\mu}_{n_j}) &= \int_I d\left(\mu^x_{n_i}, \mu^x_{n_j}\right) dx \le \int_I \sum_{k=i}^{j-1} d\left(\mu^x_{n_k}, \mu^x_{n_{k+1}}\right) dx = \sum_{k=i}^{j-1} \int_I d\left(\mu^x_{n_k}, \mu^x_{n_{k+1}}\right) dx \\ &\le \sum_{k=i}^{\infty} \int_I d\left(\mu^x_{n_k}, \mu^x_{n_{k+1}}\right) dx < \frac{1}{2^i}.\end{aligned}$$

We also have that, for all $j > i$, $d(\mu^x_{n_i}, \mu^x_{n_j}) \le f(x)$ a.e. $x \in I$ and the function f is Lebesque integrable on I. By the Lebesque dominated convergence theorem, letting $j \to \infty$, we obtain that

$$\begin{aligned}\bar{d}(\bar{\mu}_{n_i}, \bar{\mu}) &= \int_I d\left(\mu^x_{n_i}, \mu^x\right) dx = \int_I \lim_{j\to\infty} d\left(\mu^x_{n_i}, \mu^x_{n_j}\right) dx \\ &= \lim_{j\to\infty} \int_I d\left(\mu^x_{n_i}, \mu^x_{n_j}\right) dx \le \frac{1}{2^i} \to 0,\end{aligned}$$

as $i \to \infty$. Since the subsequence $\{\bar{\mu}_{n_i}\}$ of the Cauchy sequence $\{\bar{\mu}_n\}$ converges to $\bar{\mu}$ in $\bar{\mathcal{M}}$, we conclude that the sequence $\{\bar{\mu}_n\}$ converges to $\bar{\mu}$ as well. Thus $\langle \bar{\mathcal{M}}, \bar{d} \rangle$ is a complete metric space. ∎

Let $T > 0$ be arbitrary but fixed and denote $\mathcal{T} = [0, T]$. We define $\mathcal{M}_{\mathcal{T}} = C(\mathcal{T}, \bar{\mathcal{M}})$, the space of continuous $\bar{\mathcal{M}}$-valued functions.

For any $\alpha > 0$, the following is a metric on $\bar{\mathcal{M}}_{\mathcal{T}}$:

$$d_\alpha(\bar{\mu}_\cdot, \bar{\nu}_\cdot) = \sup_{t\in\mathcal{T}} e^{-\alpha t} \bar{d}(\bar{\mu}_t, \bar{\nu}_t) = \sup_{t\in\mathcal{T}} e^{-\alpha t} \int_I d(\mu^x_t, \nu^x_t) dx. \tag{2.3}$$

2.2. The equation of characteristics. Recall that $\mathcal{T} := [0, T]$, where $T > 0$ is arbitrary but fixed. For a given $\bar{\mu}_\cdot \in \mathcal{M}_{\mathcal{T}}$, consider the following equation of characteristics

$$\frac{d}{dt} u = V[W, \bar{\mu}_\cdot](u, x, t), \tag{2.4}$$

where

$$V[W, \bar{\mu}_\cdot](u, x, t) = \int_I W(x, y) \left\{ \int_{\mathbb{S}} D(v - u) d\mu^y_t(v) \right\} dy. \tag{2.5}$$

LEMMA 2.2. *For every $\bar{\mu}_\cdot \in \bar{\mathcal{M}}_{\mathcal{T}}$, $V[W, \bar{\mu}_\cdot](u, x, t)$ is Lipschitz continuous in u and continuous in x and t.*

Proof. The proof follows from the following estimates. First, using (1.15) and (1.16), we have

$$\begin{aligned}&|V[W, \bar{\mu}_\cdot](u, x, t) - V[W, \bar{\mu}_\cdot](v, x, t)| \\ &\quad= \left| \int_I W(x, y) \int_{\mathbb{S}} \left(D(w - u) - D(w - v)\right) d\mu^y_t(w) dy \right| \\ &\quad\le \int_I |W(x, y)| \int_{\mathbb{S}} |D(w - u) - D(w - v)| \, d\mu^y_t(w) dy \\ &\quad\le |u - v| \quad \forall u, v \in \mathbb{S},\ x \in I,\ t \in \mathcal{T}.\end{aligned}$$

Using the bound on D (cf. (1.15)), we obtain

$$
\begin{aligned}
&|V[W,\bar\mu_\cdot](u,x,t)-V[W,\bar\mu_\cdot](u,z,t)|\\
&\quad=\left|\int_I (W(x,y)-W(z,y))\int_{\mathbb{S}} D(v-u)d\mu_t^y(v)dy\right|\\
&\quad\le \int_I |W(x,y)-W(z,y)|\int_{\mathbb{S}} |D(v-u)|\,d\mu_t^y(v)dy\\
&\quad\le \int_I |W(x,y)-W(z,y)|\,dy \quad \forall u\in\mathbb{S},\ x,z\in I,\ t\in\mathcal{T}.
\end{aligned}
\tag{2.6}
$$

The continuity of $V[W,\bar\mu_\cdot]$ in x follows from (2.6) and (1.17).

Finally,

$$
\begin{aligned}
&|V[W,\bar\mu_\cdot](u,x,t)-V[W,\bar\mu_\cdot](u,x,s)|\\
&\quad=\left|\int_I W(x,y)\int_{\mathbb{S}} D(v-u)\,(d\mu_t^y(v)-d\mu_s^y(v))\,dy\right|\\
&\quad\le \int_I |W(x,y)|\left|\int_{\mathbb{S}} D(v-u)\,(d\mu_t^y(v)-d\mu_s^y(v))\right|dy\\
&\quad\le \int_I |W(x,y)|\,d(\mu_t^y,\mu_s^y)dy\\
&\quad\le \bar d(\bar\mu_t,\bar\mu_s).
\end{aligned}
$$

□

Similarly to the derivation of the last inequality, we prove the following lemma.

LEMMA 2.3.

$$
|V[W,\bar\mu_\cdot](u,x,t)-V[W,\bar\nu_\cdot](u,x,t)|\le \bar d(\bar\mu_t,\bar\nu_t)\quad \forall \bar\mu_\cdot,\bar\nu_\cdot\in\bar{\mathcal{M}}_{\mathcal{T}}.
\tag{2.7}
$$

Consider the initial value problem (IVP) for (2.4) subject to the initial condition at time $s\in\mathcal{T}$, $u(s)=u_s$. By Lemma 2.2, for every $x\in I$ and $u_s\in\mathbb{S}$, there exists a unique solution of the IVP for (2.4). Since $V[W,\bar\mu_\cdot](u,x,t)$ is uniformly Lipschitz in u, $u(t)$ can be extended to $t\in\mathcal{T}$. Thus, the equation of characteristics (2.4) generates the flow on $\mathbb{S}$:

$$
T^x_{t,s}[W,\bar\mu_\cdot]u_s=u(t),\ t,s\in\mathcal{T},u_s\in\mathbb{S}.
\tag{2.8}
$$

For every $x\in I$, $T^x_{t,s}[W,\bar\mu_\cdot], t,s\in\mathcal{T}$, is a two-parameter family of one-to-one transformations of $\mathbb{S}$ to itself depending continuously on x:

$$
T^x_{s,s}[W,\bar\mu_\cdot]=\mathrm{id},\quad (T^x_{t,s}[W,\bar\mu_\cdot])^{-1}=T^x_{s,t}[W,\bar\mu_\cdot].
$$

2.3. Existence of solution of the fixed point equation. In the remainder of this section, we will study the following fixed point equation. For a given $\bar\mu_0\in\bar{\mathcal{M}}$, consider the pushforward measure

$$
\bar\mu_t=\bar\mu_0\circ T_{0,t}[W,\bar\mu_\cdot]\quad \forall t\in\mathcal{T},
\tag{2.9}
$$

which is interpreted as

$$
\mu_t^x=\mu_0^x\circ T^x_{0,t}[W,\bar\mu_\cdot]\quad \text{a.e. } x\in I,\ \text{and } t\in\mathcal{T}.
\tag{2.10}
$$

First, we address existence and uniqueness of solution of (2.9).

THEOREM 2.4. *For every* $\bar{\mu}_0 \in \bar{\mathcal{M}}$, *the fixed point equation* (2.9) *has a unique solution* $\bar{\mu}_\cdot \in \bar{\mathcal{M}}_{\mathcal{T}}$.

For the proof of Theorem 2.4, we will need a variant of the Gronwall's inequality, which we formulate below for convenience.

LEMMA 2.5. *Let* $\phi(t)$ *and* $a(t)$ *be continuous functions on* $[0,T]$ *and*

$$\phi(t) \le A\int_0^t \phi(s)ds + B\int_0^t a(s)ds + C, \quad t\in[0,T], \tag{2.11}$$

where $A \ge 0$. *Then*

$$\phi(t) \le e^{At}\left(B\int_0^t a(s)e^{-As}ds + C\right). \tag{2.12}$$

Proof. The proof is standard (see, e.g., [5]). ∎

Proof of Theorem 2.4. Given $\bar{\mu}_0 \in \bar{\mathcal{M}}$, consider $A : \bar{\mathcal{M}}_{\mathcal{T}} \to \bar{\mathcal{M}}_{\mathcal{T}}$ defined by

$$A[W,\bar{\mu}_\cdot](t,x) = \mu_0^x \circ T^x_{0,t}[W,\bar{\mu}_\cdot], \quad \text{a.e. } x\in I. \tag{2.13}$$

Below we show that A is a contraction on $(\bar{\mathcal{M}}_{\mathcal{T}}, d_\alpha)$ with $\alpha > 2$. To this end,

$$\begin{aligned}
&\bar{d}\left(A[W,\bar{\mu}_\cdot](t,\cdot), A[W,\bar{\eta}_\cdot](t,\cdot)\right)\\
&\quad= \bar{d}(\bar{\mu}_0\circ T_{0,t}[W,\bar{\mu}_\cdot], \bar{\mu}_0\circ T_{0,t}[W,\bar{\eta}_\cdot])\\
&\quad= \int_I d(\mu_0^x\circ T^x_{0,t}[W,\bar{\mu}_\cdot], \mu_0^x\circ T^x_{0,t}[W,\bar{\eta}_\cdot])dx\\
&\quad= \int_I \sup_{f\in\mathcal{L}}\left|\int_{\mathbb{S}} f(v)\left(d\mu_0^x\circ T^x_{0,t}[W,\bar{\mu}_\cdot](v) - d\mu_0^x\circ T^x_{0,t}[W,\bar{\eta}_\cdot](v)\right)\right|dx\\
&\quad= \int_I \sup_{f\in\mathcal{L}}\left|\int_{\mathbb{S}} f\left(T^x_{t,0}[W,\bar{\mu}_\cdot]v\right)d\mu_0^x(v) - \int_{\mathbb{S}} f\left(T^x_{t,0}[W,\bar{\eta}_\cdot]v\right)d\mu_0^x(v)\right|dx\\
&\quad\le \int_I\int_{\mathbb{S}} \left|T^x_{t,0}[W,\bar{\mu}_\cdot]v - T^x_{t,0}[W,\bar{\eta}_\cdot]v\right| d\mu_0^x(v)dx =: \lambda(t).
\end{aligned} \tag{2.14}$$

The change of variables formula used in (2.14) is explained in [13, section 6.1]. Using (2.4) and (2.7), we obtain

$$\begin{aligned}
&\lambda(t)\\
&= \int_I\int_{\mathbb{S}} \left|T^x_{t,0}[W,\bar{\mu}_\cdot]v - T^x_{t,0}[W,\bar{\eta}_\cdot]v\right| d\mu_0^x(v)dx\\
&\le \int_0^t\int_I\int_{\mathbb{S}} \left|V[W,\bar{\mu}_\cdot]\left(T^x_{s,0}[W,\bar{\mu}_\cdot]v,x,s\right) - V[W,\bar{\eta}_\cdot]\left(T^x_{s,0}[W,\bar{\eta}_\cdot]v,x,s\right)\right| d\mu_0^x(v)dxds\\
&\le \int_0^t\int_I\int_{\mathbb{S}} \left|V[W,\bar{\mu}_\cdot]\left(T^x_{s,0}[W,\bar{\mu}_\cdot]v,x,s\right) - V[W,\bar{\eta}_\cdot]\left(T^x_{s,0}[W,\bar{\mu}_\cdot]v,x,s\right)\right| d\mu_0^x(v)dxds\\
&\quad+ \int_0^t\int_I\int_{\mathbb{S}} \left|V[W,\bar{\eta}_\cdot]\left(T^x_{s,0}[W,\bar{\mu}_\cdot]v,x,s\right) - V[W,\bar{\eta}_\cdot]\left(T^x_{s,0}[W,\bar{\eta}_\cdot]v,x,s\right)\right| d\mu_0^x(v)dxds\\
&\le \int_0^t \bar{d}(\bar{\mu}_s,\bar{\eta}_s)ds + \int_0^t\int_I\int_{\mathbb{S}} \left|T^x_{s,0}[W,\bar{\mu}_\cdot]v - T^x_{s,0}[W,\bar{\eta}_\cdot]v\right| d\mu_0^x(v)dxds\\
&\le \int_0^t \bar{d}(\bar{\mu}_s,\bar{\eta}_s)ds + \int_0^t \lambda(s)ds.
\end{aligned} \tag{2.15}$$

Using Gronwall's inequality (cf. Lemma 2.5), from (2.15) we obtain

$$\lambda(t) \le e^t \int_0^t \bar{d}(\bar{\mu}_s, \bar{\eta}_s) e^{-s} ds. \tag{2.16}$$

Combining (2.14), (2.15), and (2.16), we have

$$\bar{d}\left(A[W,\bar{\mu}](t,\cdot), A[W,\bar{\eta}](t,\cdot)\right) \le e^t \int_0^t \bar{d}(\bar{\mu}_s, \bar{\eta}_s) e^{-s} ds \tag{2.17}$$

and

$$\begin{aligned} d_\alpha\left(A[W,\bar{\mu}](t,\cdot), A[W,\bar{\eta}](t,\cdot)\right) &= \sup_{t\in\mathcal{T}} \left\{ e^{-\alpha t} \bar{d}\left(A[W,\bar{\mu}](t,\cdot), A[W,\bar{\eta}](t,\cdot)\right)\right\} \\ &\le \sup_{t\in\mathcal{T}} e^{-(\alpha-1)t} \int_0^t \bar{d}(\bar{\mu}_s, \bar{\eta}_s) e^{-s} ds \\ &\le d_\alpha\left(\bar{\mu}, \bar{\eta}\right) e^{-(\alpha-1)t} \int_0^t e^{(\alpha-1)s} ds \\ &\le (\alpha-1)^{-1} d_\alpha\left(\bar{\mu}_\cdot, \bar{\eta}_\cdot\right). \end{aligned} \tag{2.18}$$

We conclude the proof with using the contraction mapping principle to establish a unique solution of (2.9). ∎

2.4. Continuous dependence on initial data.

LEMMA 2.6. *Let $\bar{\mu}_\cdot, \bar{\eta}_\cdot \in \bar{\mathcal{M}}_\mathcal{T}$ be two solutions of* (2.9) *corresponding to initial conditions $\bar{\mu}_0, \bar{\eta}_0 \in \bar{\mathcal{M}}$, respectively. Then*

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\mu}_t, \bar{\eta}_t) \le e^{2T} \bar{d}(\bar{\mu}_0, \bar{\eta}_0). \tag{2.19}$$

Proof. For every $t \in \mathcal{T}$, by the triangle inequality, we have

$$\begin{aligned} \bar{d}(\bar{\mu}_t, \bar{\eta}_t) &= \bar{d}(\bar{\mu}_0 \circ T_{0,t}[W,\bar{\mu}_\cdot], \bar{\eta}_0 \circ T_{0,t}[W,\bar{\eta}_\cdot]) \\ &\le \bar{d}(\bar{\mu}_0 \circ T_{0,t}[W,\bar{\mu}_\cdot], \bar{\mu}_0 \circ T_{0,t}[W,\bar{\eta}_\cdot]) + \bar{d}(\bar{\mu}_0 \circ T_{0,t}[W,\bar{\eta}_\cdot], \bar{\eta}_0 \circ T_{0,t}[W,\bar{\eta}_\cdot]). \end{aligned} \tag{2.20}$$

Exactly in the same way as in (2.14), we estimate the first term on the right-hand side of (2.20) as follows:

$$\begin{aligned} &\bar{d}(\bar{\mu}_0 \circ T_{0,t}[W,\bar{\mu}_\cdot], \bar{\mu}_0 \circ T_{0,t}[W,\bar{\eta}_\cdot]) \\ &\quad = \int_I \sup_{f\in\mathcal{L}} \left| \int_\mathbb{S} \left(f(T^x_{t,0}[W,\bar{\mu}_\cdot]v) - f(T^x_{t,0}[W,\bar{\eta}_\cdot]v)\right) d\mu^x_0(v) \right| dx \\ &\quad \le \int_I \int_\mathbb{S} \left| T^x_{t,0}[W,\bar{\mu}_\cdot]v - T^x_{t,0}[W,\bar{\eta}_\cdot]v \right| d\mu^x_0(v) dx =: \lambda(t). \end{aligned} \tag{2.21}$$

Similarly, repeating the steps in (2.15)

$$\lambda(t) \le \int_0^t \bar{d}(\bar{\mu}_s, \bar{\eta}_s) ds + \int_0^t \lambda(s) ds. \tag{2.22}$$

Using Gronwall's inequality, from (2.22) we obtain

$$\lambda(t) \le e^t \int_0^t \bar{d}(\bar{\mu}_s, \bar{\eta}_s) e^{-s} ds. \tag{2.23}$$

Next, we turn to the second term on the right-hand side of (2.20):

$$
\begin{aligned}
&\bar d(\bar\mu_0\circ T_{0,t}[W,\bar\eta_\cdot],\bar\eta_0\circ T_{0,t}[W,\bar\eta_\cdot])\\
&= \int_I \sup_{f\in\mathcal{L}}\left|\int_{\mathbb{S}} f(v)d\mu_0^x\circ T^x_{0,t}[W,\bar\eta_\cdot](v)-\int_{\mathbb{S}} f(v)d\eta_0^x\circ T^x_{0,t}[W,\bar\eta_\cdot](v)\right|dx\\
&= \int_I \sup_{f\in\mathcal{L}}\left|\int_{\mathbb{S}} f(T^x_{t,0}[W,\bar\eta_\cdot]v)d\mu_0^x(v)-\int_{\mathbb{S}} f(T^x_{t,0}[W,\bar\eta_\cdot]v)d\eta_0^x(v)\right|dx\\
&\le \bar d(\bar\mu_0,\bar\eta_0).
\end{aligned}\tag{2.24}
$$

The combination of (2.20), (2.21), (2.23), and (2.24) yields

$$\bar d(\bar\mu_t,\bar\eta_t)\le e^t\int_0^t \bar d(\bar\mu_s,\bar\eta_s)e^{-s}ds+\bar d(\bar\mu_0,\bar\eta_0).\tag{2.25}$$

Denote $\phi(t):=\bar d(\bar\mu_t,\bar\eta_t)e^{-t}$ and rewrite (2.25) as

$$\phi(t)\le\int_0^t\phi(s)ds+e^{-t}\bar d(\bar\mu_0,\bar\eta_0)=:\psi(t).$$

Next,

$$
\begin{aligned}
\psi'(t)&=\phi(t)-e^{-t}\bar d(\bar\mu_0,\bar\eta_0)\\
&\le\psi(t)-e^{-t}\bar d(\bar\mu_0,\bar\eta_0)\\
&\le\psi(t).
\end{aligned}
$$

Further,

$$\frac{d}{ds}\left\{e^{-s}\psi(s)\right\}=e^{-s}\psi'(s)-e^{-s}\psi(s)\le 0,\tag{2.26}$$

and, thus,

$$\phi(t)\le\psi(t)\le e^t\psi(0)=e^t\bar d(\bar\mu_0,\bar\eta_0).$$

Recalling, the definition of $\phi(t)$, we arrive at

$$\bar d(\bar\mu_t,\bar\eta_t)\le e^{2t}\bar d(\bar\mu_0,\bar\eta_0),\quad t\in\mathcal{T},\tag{2.27}$$

from which (2.19) follows. ∎

2.5. Continuous dependence on the kernel. In this subsection, we study how the solution of the fixed point equation (2.9) changes under the perturbation of the kernel W. To this end, let W and U be two bounded measurable functions on I^2 satisfying (1.17). Then for a given $\mu_0\in\mathcal{M}$ each of the fixed point equations

$$\bar\mu_t=\bar\mu_0\circ T_{0,t}[W,\bar\mu_\cdot],\tag{2.28}$$

$$\bar\nu_t=\bar\mu_0\circ T_{0,t}[U,\bar\nu_\cdot]\tag{2.29}$$

has a unique solution in $\mathcal{M}_\mathcal{T}$, which we denote by $\bar\mu_t$ and $\bar\nu_t$, respectively.

LEMMA 2.7.

$$\sup_{t\in\mathcal{T}}\bar d(\bar\mu_t,\bar\nu_t)\le e^{2T}\|W-U\|_{L^1(I^2)}.\tag{2.30}$$

Proof. We proceed along the lines of the proof of Lemma 2.6. Replicating the steps in (2.14), we derive

$$\bar{d}(\bar{\mu}_t, \bar{\nu}_t) \leq \int_I \int_{\mathbb{S}} \left| T^x_{t,0}[W, \bar{\mu}_\cdot]v - T^x_{t,0}[U, \bar{\nu}_\cdot]v \right| d\mu^x_0 dx =: \lambda(t). \tag{2.31}$$

Further,

$$\begin{aligned} \lambda(t) &= \int_I\int_{\mathbb{S}} \left| \int_0^t \left\{ V[W, \bar{\mu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) - V[U, \bar{\nu}_\cdot] \left(T^x_{t,0}[U, \bar{\nu}_\cdot]v, x, s\right) \right\} ds \right| d\mu^x_0(v) dx \\ &\leq \int_0^t \int_I \int_{\mathbb{S}} \left| V[W, \bar{\mu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) - V[W, \bar{\nu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) \right| d\mu^x_0(v) dx ds \\ &\quad + \int_0^t \int_I \int_{\mathbb{S}} \left| V[W, \bar{\nu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) - V[U, \bar{\nu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) \right| d\mu^x_0(v) dx ds \\ &\quad + \int_0^t \int_I \int_{\mathbb{S}} \left| V[U, \bar{\nu}_\cdot] \left(T^x_{t,0}[W, \bar{\mu}_\cdot]v, x, s\right) - V[U, \bar{\nu}_\cdot] \left(T^x_{t,0}[U, \bar{\nu}_\cdot]v, x, s\right) \right| d\mu^x_0(v) dx ds \\ &=: \lambda_1(t) + \lambda_2(t) + \lambda_3(t). \end{aligned} \tag{2.32}$$

We estimate the first term on the right-hand side of (2.32), using Lemma 2.3:

$$\lambda_1(t) \leq \int_0^t \bar{d}(\bar{\mu}_s, \bar{\nu}_s) ds. \tag{2.33}$$

For the second term, we have

$$\begin{aligned} \lambda_2(t) &\leq \int_0^t \int_I \int_{\mathbb{S}} \left[\int_I |W(x,y) - U(x,y)| \right. \\ &\qquad \left. \left\{ \int_{\mathbb{S}} \left| D(w - T^x_{t,0}[W, \bar{\mu}_\cdot]v) \right| d\nu^y_s(w) \right\} dy \right] d\mu^x_0(v) dx ds \\ &\leq \int_0^t \int_{I^2} |W(x,y) - U(x,y)| dx dy ds, \end{aligned} \tag{2.34}$$

where we used $|D(u)| \leq 1$ to get the latter inequality. Finally, to estimate the third term, we use Lipschitz continuity of $V[\cdot,\cdot](u, x, t)$ in u:

$$\lambda_3(t) \leq \int_0^t \int_I \int_{\mathbb{S}} \left| T^x_{s,0}[W, \bar{\mu}_\cdot]v - T^x_{s,0}[U, \bar{\nu}_\cdot]v \right| d\mu^x_0(v) dx ds = \int_0^t \lambda(s) ds. \tag{2.35}$$

Plugging (2.33)–(2.35) into (2.32), we obtain

$$\lambda(t) \leq \int_0^t \lambda(s) ds + \int_0^t \left(\bar{d}(\bar{\mu}_s, \bar{\nu}_s) + \|W - U\|_{L^1(I^2)} \right) ds. \tag{2.36}$$

By Gronwall's inequality (cf. Lemma 2.5),

$$\lambda(t) \leq e^t \int_0^t e^{-s} \left(\bar{d}(\bar{\mu}_s, \bar{\nu}_s) + \|W - U\|_{L^1(I^2)} \right) ds. \tag{2.37}$$

The combination of (2.31) and (2.37) yields

$$\bar{d}(\bar{\mu}_t, \bar{\nu}_t) \leq e^t \int_0^t e^{-s} \left(\bar{d}(\bar{\mu}_s, \bar{\nu}_s) + \|W - U\|_{L^1(I^2)} \right) ds. \tag{2.38}$$

Denote $\phi(t) := e^{-t}\bar{d}(\bar{\mu}_t, \bar{\nu}_t)$ and rewrite (2.38) as

$$\begin{aligned}\phi(t) &\leq \int_0^t \phi(s)ds + \|W-U\|_{L^1(I^2)} \int_0^t e^{-s}ds \\ &\leq \int_0^t \phi(s)ds + \|W-U\|_{L^1(I^2)}.\end{aligned} \tag{2.39}$$

Using Lemma 2.5 again, we have

$$\phi(t) \leq e^t \|W-U\|_{L^1(I^2)}.$$

Recalling, the definition of $\phi(t)$, we finally get

$$\bar{d}(\bar{\mu}_t, \bar{\nu}_t) \leq e^{2t}\|W-U\|_{L^1(I^2)}.$$

∎

3. Application to coupled systems.

In this section, we apply the fixed point theory developed in the previous section to the proof of convergence of solutions of the KM on graphs (1.2).

3.1. The initial value problem.

We begin by addressing the well-posedness of the IVP for (1.18), i.e., review the notion of the weak solution of the mean field equation (1.18) and then prove the existence and uniqueness of the weak solution of the IVP (1.18)–(1.20). The following definition of the weak solution of (1.18) is adapted from [19].

DEFINITION 3.1. *A measurable function $\rho : \mathcal{T} \times G \to \mathbb{R}$ is called a weak solution of* (1.18)–(1.20) *if the following conditions hold a.e. $x \in I$.*

1. *$\rho(t,u,x)$ is weakly continuous in $t \in \mathcal{T}$, i.e., $t \mapsto \int_{\mathbb{S}} \rho(t,u,x) f(u)du$ is a continuous map for every $f \in C(\mathbb{S})$.*
2. *The following identity holds*

$$\begin{aligned}&\int_0^T \left\{ \int_{\mathbb{S}} \rho(t,u,x) \left(\frac{\partial}{\partial t} w(t,u) + V(t,u,x) \frac{\partial}{\partial u} w(t,u) \right) du \right\} dt \\ &\quad + \int_{\mathbb{S}} w(0,u)\rho_0(u,x)du = 0\end{aligned} \tag{3.1}$$

for every $w \in C^1(\mathcal{T} \times \mathbb{S})$ with support in $(0,T] \times \mathbb{S}$.

THEOREM 3.2. *Suppose $W \in L^1(I^2)$ satisfies* (1.17) *and $\rho_0 \in L^1(G)$. Then there is a unique weak solution to the IVP* (1.18)–(1.20).

Proof. Recall that $T^x_{0,t} := T^x_{0,t}[W, \bar{\mu}_\cdot]$ is the flow generated by the equation of characteristics (2.4) on $\mathbb{S}$ (see (2.8)). By Theorem 2.4, for μ^x_0, a family of measures on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ with densities $\rho_0(\cdot, x)$, there is a unique solution of the fixed point equation (2.9).

For a.e. $x \in I$ and $t \in \mathcal{T}$, $T^x_{0,t}$ is one-to-one and Lipschitz continuous. As an absolutely continuous function, $T^x_{0,t}$ is differentiable a.e. on $\mathbb{S}$ and has an essentially bounded weak derivative. Using (2.10) and the change of variables formula for Lipschitz maps (cf. [6, Theorem 2, section 3.3.3]), we have

$$\begin{aligned}\mu^x_t(A) = \mu^x_0 \circ T^x_{0,t}(A) &= \int_{T^x_{0,t}A} \rho_0(u,x)du \\ &= \int_A \rho_0(T^x_{0,t}u, x) \left| \frac{\partial}{\partial u} T^x_{0,t}u \right| du, \quad \text{a.e. } x \in I,\end{aligned} \tag{3.2}$$

for any Borel $A \subset \mathbb{S}$. In the last integral of (3.2), $\frac{\partial}{\partial u} T^x_{0,t} u$ is understood as a weak derivative. Thus, for $t \in \mathcal{T}$ and a.e. $x \in I$, μ^x_t is an absolutely continuous measure with density

$$\rho(t,u,x) = \rho_0(T^x_{0,t}u, x) \left| \frac{\partial}{\partial u} T^x_{0,t} u \right|. \tag{3.3}$$

To show that $\rho(t,u,x)$ is a weak solution of (1.18)–(1.20), as in the proof of [19, Theorem 1], we set

$$h := \frac{\partial}{\partial t} w + V(t,u,x) \frac{\partial}{\partial u} w(t,u) \tag{3.4}$$

and compute

$$\begin{aligned} \int_0^T \int_{\mathbb{S}} \rho(t,u,x) h(t,u,x) du dt &= \int_0^T \int_{\mathbb{S}} \left(\rho_0(T^x_{0,t}u, x) \left| \frac{\partial}{\partial u} T^x_{0,t} u \right| h(t,u,x) \right) du dt \\ &= \int_{\mathbb{S}} \rho_0(u,x) \left(\int_0^T h(t, T^x_{t,0} u) dt \right) du. \end{aligned} \tag{3.5}$$

Further, using the chain rule and (3.4), we have

$$\begin{aligned} \frac{d}{dt} w(t, T^x_{t,0} u) &= \partial_1 w(t, T^x_{t,0} u) + \partial_2 w(t, T^x_{t,0} u) \frac{d}{dt} T^x_{t,0} u \\ &= \partial_1 w(t, T^x_{t,0} u) + \partial_2 w(t, T^x_{t,0} u) V(t,u,x) = h(t, T^x_{t,0} u, x), \end{aligned} \tag{3.6}$$

where $\partial_{1,2}$ stand for the partial derivatives with respect to the first and second argument, respectively. The combination of (3.5) and (3.6) yields

$$\begin{aligned} \int_0^T \int_{\mathbb{S}} \rho(t,u,x) h(t,u,x) du dt &= \int_{\mathbb{S}} \rho_0(u,x) \left(\int_0^T h(t, T^x_{t,0} u) dt \right) du \\ &= \int_{\mathbb{S}} \rho_0(u,x) \left(\int_0^T \frac{\partial}{\partial t} w(t, T^x_{t,0} u) du \right) dt \\ &= - \int_{\mathbb{S}} \rho_0(u,x) w(0,u) du. \end{aligned}$$ □

3.2. Approximation. We continue by collecting several results on approximation, which will be used later in this section.

Let $W \in L^2(I^2)$ and consider a step function $W_n : I \to \mathbb{R}$, which on each cell $I_{n,i} \times I_{n,j}$ is equal to the average value $W_{n,ij}$, $i,j \in [n]$ (cf. (1.4)).

LEMMA 3.3. *$W_n \to W$ a.e. and in $L^2(I^2)$ as $n \to \infty$.*

Proof. The proof can be found in [9, Lemma 5.3]. □

Let $\Gamma_n = \langle [n], E(\Gamma_n), W_n \rangle$ be a weighted graph on n nodes. Here, $[n]$ is the set of vertices and

$$E(\Gamma_n) = \{\{i,j\} : \ i,j \in [n]\}$$

is the edge set. Each edge $\{i,j\}$ is equipped with a real weight $W_{n,ij}$, the ijth entry of the $n \times n$ weight matrix W_n.

On Γ_n consider a coupled system:

$$\dot{u}_{n,i} = n^{-1}\sum_{j=1}^{n} W_{n,ij} D(u_{n,j} - u_{n,i}), \quad i \in [n]. \tag{3.7}$$

By $u_n = (u_{n,1}, u_{n,2}, \dots, u_{n,n})^{\mathsf{T}} \in \mathbb{R}^n$ we denote a solution of the coupled system (3.7). Along with (3.7), consider the coupled system on a weighted graph $\tilde{\Gamma}_n = \langle [n], E(\Gamma_n), \tilde{W}_n \rangle$, whose solution is denoted by $\tilde{u}_n$.

Recall the discrete L^2-norm defined in (1.10). We will also need its analogue for discretizations of functions on I^2:

$$\|W_n\|_{2,n} = \sqrt{n^{-2}\sum_{i,j=1}^{n} W_{n,ij}^2}. \tag{3.8}$$

LEMMA 3.4 (see [3, Lemma 4.1]). *Let $u_n(t)$ and $\tilde{u}_n(t)$ denote solutions of the IVPs for the coupled system* (3.7) *on weighted graphs $\Gamma_n = \langle [n], E(\Gamma_n), W_n \rangle$ and $\tilde{\Gamma}_n = \langle [n], E(\Gamma_n), \tilde{W}_n \rangle$, respectively. Suppose that the initial data for these problems coincide*

$$u_n(0) = \tilde{u}_n(0). \tag{3.9}$$

Then for any $T > 0$,

$$\max_{t\in[0,T]} \|u_n(t) - \tilde{u}_n(t)\|_{1,n} \le C_1 \left\| W_n - \tilde{W}_n \right\|_{2,n}, \tag{3.10}$$

where $C_1 = \sqrt{Te^{5T}} > 0$.

Remark 3.5. Note that (3.10) implies that the solutions of any two discrete models (1.6) with weights given by W_n and $\tilde{W}_n$ will be close, provided that W_n and $\tilde{W}_n$ converge to W in L^2.

Let $\rho(t,u,x)$ and $\varrho_n(t,u,x)$ denote the solutions of the IVP (1.18), (1.19) subject to the initial conditions $\rho^0(u,x)$ and $\varrho_n^0(u,x)$, respectively. We assume that $\rho^0(u,x)$ and $\varrho_n^0(u,x)$ are nonnegative functions from $L^1(G)$ satisfying (1.21). Denote the measures generated by $\rho(t,u,x)$ and $\varrho_n(t,u,x)$ by

$$\nu_t^x(A) = \int_A \rho(t,u,x)du \quad \text{and} \quad \upsilon_{n,t}^x(A) = \int_A \varrho(t,u,x)du, \quad x \in I,\ A \in \mathcal{B}(\mathbb{S}). \tag{3.11}$$

LEMMA 3.6. *Suppose for every $u \in \mathbb{S}$, $\varrho^0(u,\cdot) \to \rho^0(u,\cdot)$ a.e. on I. Then*

$$\lim_{n\to\infty} \sup_{t\in\mathcal{T}} \bar{d}(\bar{\nu}_t, \bar{\upsilon}_{n,t}) = 0. \tag{3.12}$$

Proof. $\bar{\nu}_t$ and $\bar{\upsilon}_{n,t}$ solve the fixed point equation (2.9) subject to the initial conditions $\bar{\nu}_0$ and $\bar{\upsilon}_{n,0}$, respectively. By Lemma 2.6,

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\nu}_t, \bar{\upsilon}_{n,t}) \le e^{2T} \bar{d}(\bar{\nu}_{n,0}, \bar{\upsilon}_{n,0}). \tag{3.13}$$

Thus, it remains to estimate $\bar{d}(\bar{\nu}_0, \bar{\upsilon}_{n,0})$. To this end, we consider

$$\begin{aligned} \bar{d}(\bar{\nu}_0, \bar{\upsilon}_{n,0}) &= \int_I \sup_{f\in\mathcal{L}} \left| \int_{\mathbb{S}} f(v) \left[\rho^0(v,x) - \varrho_n^0(v,x)\right] dv \right| dx \\ &\le \int_I \int_{\mathbb{S}} \left|\rho^0(v,x) - \varrho_n^0(v,x)\right| dvdx \\ &= \int_{\mathbb{S}} \int_I \left|\rho^0(v,x) - \varrho_n^0(v,x)\right| dxdv =: \int_{\mathbb{S}} \phi_n(v)dv. \end{aligned} \tag{3.14}$$

Recall that for every $v \in \mathbb{S}$, $\varrho_n^0(v,\cdot) \to \rho^0(v,\cdot)$ a.e. Further, by (1.21) and the Fubini theorem,

$$\int_{\mathbb{S}}\int_I \varrho_n^0(v,x)dxdv = \int_{\mathbb{S}}\int_I \rho^0(v,x)dxdv = 1.$$

By Scheffé's lemma (cf. [24]), $\phi_n \to 0$ pointwise on $\mathbb{S}$. Since

$$0 \le \phi_n(u) \le \int_I \left(\varrho_n^0(u,x) + \rho^0(u,x)\right)dx,$$

the Lebesgue dominated convergence theorem yields $\int_{\mathbb{S}} \phi_n(u)du \to 0$. ◻

3.3. The auxiliary problems. Our next goal is to establish convergence to the mean field limit for an auxiliary discrete model. To this end, recall that $W_n : I^2 \to [-1,1]$ is a step function taking constant value $W_{n,ij}$ on $I_{n,i} \times I_{n,j},\ (i,j) \in [n]^2$ (cf. (1.4)).

Suppose $n \in \mathbb{N}$ is arbitrary but fixed. Let $N = nm$ for some $m \in \mathbb{N}$ and consider a coupled system

$$\dot v_{N,(k-1)m+l} = N^{-1}\sum_{i=1}^n\sum_{j=1}^m W_{n,ki} D(v_{N,(i-1)m+j} - v_{N,(k-1)m+l}), \tag{3.15}$$

$$v_{N,(k-1)m+l}(0) = u^0_{n,kl}, \quad k \in [n],\ l \in [m], \tag{3.16}$$

where $u^0 = (u^0_{n,pq}),\ (p,q) \in [n] \times \mathbb{N}$ is a random array of initial conditions

$$\begin{matrix} u^0_{n,11}, & u^0_{n,12}, & u^0_{n,13}, & \dots \\ u^0_{n,21}, & u^0_{n,22}, & u^0_{n,23}, & \dots \\ \dots & \dots & \dots & \dots \\ u^0_{n,n1}, & u^0_{n,n2}, & u^0_{n,n3}, & \dots \end{matrix} \quad . \tag{3.17}$$

Here, for every $k \in [n]$, $u^0_{n,ki}, i \in \mathbb{N}$, are independent identically distributed continuous random variables. The density of the probability distribution of $u^0_{n,ki}$ is given by

$$\rho^0_{n,k}(v) = n\int_{I_{n,k}} \rho^0(u,x)dx. \tag{3.18}$$

Thus, we have defined the probability measure on the measurable space $(\Omega^{(n)} = (\mathbb{R}^\infty)^n, \mathcal{F}^{(n)} = (\mathcal{B}(\mathbb{R}^\infty))^n)$. Denote the resultant probability space by $(\Omega^{(n)}, \mathcal{F}^{(n)}, \mathbb{P}_0^{(n)})$.

Remark 3.7. Note that $\rho^0_{n,k}(u)$ is indeed a probability density function on $\mathbb{S}$:

$$\int_{\mathbb{S}} \rho^0_{n,k}(u)du = n\int_{\mathbb{S}}\int_{I_{n,k}} \rho^0_n(u,x)dxdu = n\int_{I_{n,i}}\int_{\mathbb{S}} \rho^0_n(u,x)dudx = 1,$$

where we used the Fubini theorem and (1.21).

We are going to describe the solution of (3.15), (3.16) in terms of the local empirical measures

$$\mu^x_{n,m,t}(A) = m^{-1}\sum_{j=1}^m \mathbf{1}_A(v_{N,(i-1)m+j}(t)), \quad x \in I_{n,i},\ A \in \mathcal{B}(\mathbb{S}),\ i \in [n]. \tag{3.19}$$

We will show that in the large m limit, the behavior of solutions of the discrete model (3.15), (3.16) is effectively approximated by the IVP for the following integro-differential equation,

$$\frac{\partial}{\partial t}\rho_n(t,u,x) + \frac{\partial}{\partial u}\left\{V_n(t,u,x)\rho_n(t,u,x)\right\} = 0, \tag{3.20}$$

where

$$V_n(t,u,x) = \int_I \int_{\mathbb{S}} W_n(x,y) D(u-v)\rho_n(t,v,y)dvdy, \tag{3.21}$$

and subject to the initial condition

$$\rho_n(0,u,x) = \sum_{i=1}^n \rho^0_{n,i}(v)\mathbf{1}_{I_{n,i}}(x) =: \rho^0_n(u,x). \tag{3.22}$$

By Theorem 3.2, there is a unique weak solution of the IVP (3.20)–(3.22), which defines a family of absolutely continuous measures on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$:

$$\mu^x_{n,t}(A) = \int_A \rho_n(t,u,x)du, \quad x \in I,\ A \in \mathcal{B}(\mathbb{S}). \tag{3.23}$$

THEOREM 3.8. *There exists $\mathcal{U}^{(n)} \in \mathcal{F}^{(n)}$, $\mathbb{P}^{(n)}_0(\mathcal{U}^{(n)}) = 1$, such that*

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\mu}_{n,m,t}, \bar{\mu}_{n,t}) \to 0, \quad m \to \infty, \tag{3.24}$$

for every $u^0 \in \mathcal{U}^{(n)}$.

The proof of Theorem 3.8 follows from two lemmas, which we prove next.

LEMMA 3.9. *The empirical measure $\bar{\mu}_{n,m,t}$ and the absolutely continuous measure $\bar{\mu}_{n,t}$ satisfy the following estimate:*

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\mu}_{n,m,t}, \bar{\mu}_{n,t}) \le e^{2T}\bar{d}(\bar{\mu}_{n,m,0}, \bar{\mu}_{n,0}). \tag{3.25}$$

Proof. The proof follows from the fact that the empirical measure $\mu^x_{n,m,t}$ and the absolutely continuous measure $\mu^x_{n,t}$ satisfy the fixed point equation

$$\mu^x_t = \mu^x_0 \circ T^x_{0,t}[W_n, \bar{\mu}_\cdot] \quad \text{a.e. } x \in I \tag{3.26}$$

for $\bar{\mu}_0 := \bar{\mu}_{n,m,0}$ and $\bar{\mu}_0 := \bar{\mu}_{n,0}$, respectively.

For the continuous measures $\mu^x_{n,t}$ this follows from the proof of Theorem 3.2. Thus, it remains to verify (3.26) for the discrete measures $\mu^x_{n,m,t}$. Using (3.19), for every $x \in I_{n,k}$ we have

(3.27)

$$\begin{aligned} V[W_n, \bar{\mu}_{n,m,\cdot}](t,u,x) &= \int_I W_n(x,y)\left\{\int_{\mathbb{S}} D(v-u)d\mu^y_{n,m,t}(v)\right\}dy \\ &= \sum_{i=1}^n \int_{I_{n,i}} W_n(x,y)\left\{m^{-1}\sum_{j=1}^m D\left(v_{N,(i-1)m+j}(t)-u\right)\right\}dy \\ &= N^{-1}\sum_{i=1}^n\sum_{j=1}^m W_{n,ki} D\left(v_{N,(i-1)m+j}(t)-u\right). \end{aligned}$$

The right-hand side of (3.27) with $u := v_{N,(k-1)m+l}(t)$ yields the velocity field acting on the oscillator $((k-1)m+l)$ of the discrete system (3.15). Therefore, by construction of the empirical measure (3.19), for every $x \in I$ we have

$$\mu^x_{n,m,t} = \mu^x_{n,m,0} \circ T^x_{0,t}[W_n, \bar{\mu}_{n,m,.}].$$

Finally, (3.25) follows from Lemma 2.6. ∎

LEMMA 3.10.

$$\lim_{m\to\infty} \bar{d}(\bar{\mu}_{n,m,0}, \bar{\mu}_{n,0}) = 0 \quad \mathbb{P}^{(n)}_0\text{-}a.s. \tag{3.28}$$

Proof. Since $\mu^x_{n,m,0}$ and $\mu^x_{n,0}$ are constant on each $I_{n,i},\ i \in [n]$,

$$\begin{aligned}\bar{d}(\bar{\mu}_{n,m,0}, \bar{\mu}_{n,0}) &= \sum_{i=1}^{n} \int_{I_{n,i}} d\left(\mu^y_{n,m,0}, \mu^y_{n,0}\right) dy \\ &= n^{-1} \sum_{i=1}^{n} d\left(\mu^{x_{n,i}}_{n,m,0}, \mu^{x_{n,i}}_{n,0}\right).\end{aligned} \tag{3.29}$$

Let $i \in [n]$ be fixed. It is sufficient to show that

$$\lim_{m\to\infty} d\left(\mu^{x_{n,i}}_{n,m,0}, \mu^{x_{n,i}}_{n,0}\right) = 0. \tag{3.30}$$

Suppose $\epsilon > 0$ is given and denote

$$A_m = \left\{u^0 \in \Omega^{(n)} :\ \sup_{x\in\mathbb{R}} \left|\mu^{x_{n,i}}_{n,m,0}(R_x) - \mu^{x_{n,i}}_{n,0}(R_x)\right| > \epsilon\right\}, \tag{3.31}$$

where $R_x := (-\infty, x]$. By the Dvoretzky–Kiefer–Wolfowitz inequality [14],

$$\mathbb{P}^{(n)}_0(A_m) \le 2e^{-2m\epsilon^2}. \tag{3.32}$$

By the Borel–Cantelli lemma,

$$\mathbb{P}^{(n)}_0\left(\limsup_{m\to\infty} A_m\right) = 0, \tag{3.33}$$

i.e., there exist $m_1 = m_1(\epsilon)$ and $\mathcal{U}_\epsilon$, $\mathbb{P}^{(n)}_0(\mathcal{U}_\epsilon) = 1$, such that

$$\sup_{x\in\mathbb{R}} \left|\mu^{x_{n,i}}_{n,m,0}(R_x) - \mu^{x_{n,i}}_{n,0}(R_x)\right| \le \epsilon, \quad m \ge m_1, \tag{3.34}$$

provided $u^0 \in \mathcal{U}_\epsilon$. Define $\mathcal{U}_0 = \bigcap_{k=1}^{\infty} \mathcal{U}_{\frac{1}{k}}$. By continuity, $\mathbb{P}^{(n)}_0(\mathcal{U}_0) = 1$. For $u^0 \in \mathcal{U}_0$ and for every $\epsilon > 0$ there exists $m_1 \in \mathbb{N}$ such that (3.34) holds. Thus, for $u^0 \in \mathcal{U}_0$, $\mu^{x_{n,i}}_{n,m,0}$ converges to $\mu^{x_{n,i}}_{n,0}$ weakly as $m \to \infty$. By [4, Theorem 11.3.3], this implies that (3.30) holds for $u^0 \in \mathcal{U}_0$. ∎

For the proof of our main result, we also need to consider the IVP for (3.20), (3.21) with the initial condition

$$\rho_n(0, u, x) = \rho^0(u, x). \tag{3.35}$$

To distinguish the solution of the IVP with initial condition (3.35) from that with (3.22), we denote it by $\tilde{\rho}_n(t,u,x)$. The solution of (3.20), (3.21), (3.35) generates another family of absolutely continuous measures on $(\mathbb{S},\mathcal{B}(\mathbb{S}))$:

$$\nu^x_{n,t}(A)=\int_A \tilde{\rho}_n(t,u,x)du, \quad x\in I,\ A\in\mathcal{B}(\mathbb{S}). \tag{3.36}$$

The well-posedness of the initial value problems (3.20), (3.21), (3.35) and (1.18), (1.20) on $\mathcal{T}$ is established via the same argument used for the auxiliary problem in the proof of Theorem 3.8.

3.4. The main result. We now turn to the original model (1.13), (1.14). In analogy to how it was done for the auxiliary problem in the previous subsection, for given $n,m\in\mathbb{N}$ and $N=nm$, we define the empirical measure

$$\nu^x_{n,m,t}(A)=m^{-1}\sum_{j=1}^m \mathbf{1}_A\left(u_{N,(i-1)m+j}(t)\right), \quad A\in\mathcal{B}(\mathbb{S}),\ x\in I_{n,i}, i\in[n], \tag{3.37}$$

where u_N is the solution of the IVP (1.13), (1.14) subject to the initial condition $u_{N,(i-1)m+j}=u^0_{n,ij}$, $(i,j)\in[n]\times[m]$ (cf. (3.17)). Likewise, $\bar{\nu}_t$ stands for the $\mathcal{M}$-valued function defined using the solution of the IVP for (1.18), (1.20):

$$\nu^x_t(A)=\int_A \rho(t,u,x)du, \quad A\in\mathcal{B}(\mathbb{S}),\ x\in I. \tag{3.38}$$

Our goal is to show that for large $n,m\in\mathbb{N}$, the continuous measure (3.38) approximates the empirical measure (3.37), and, thus, describes the behavior of solutions of the discrete model (1.13) with $N=nm$. This is achieved in the following theorem.

THEOREM 3.11. *For a given $\epsilon>0$ there exist $n_1=n_1(\epsilon)\in\mathbb{N}$ such that the following holds. For any $n\ge n_1$ there exist $\mathcal{U}^{(n)}\in\mathcal{F}^{(n)}$, $\mathbb{P}^{(n)}_0(\mathcal{U}^{(n)})=1$, and $m_1=m_1(\epsilon,n_1)$ such that*

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\nu}_{n,m,t},\bar{\nu}_t)\le\epsilon \tag{3.39}$$

for every $m\ge m_1$ and $u^0\in\mathcal{U}^{(n)}$.

Proof.

1. Recall that W_n stands for the step function taking on each cell $I_{n,i}\times I_{n,j}$, $i,j\in[n]$, the average value of W on this cell (cf. (1.4)). By Lemma 3.3,
$$\exists n_2\in\mathbb{N}: \quad \|W_n-W\|_{L^2(I^2)}\le\frac{\epsilon}{8\left(\sqrt{Te^{5T}}+e^{2T}\right)}, \quad n\ge n_2. \tag{3.40}$$
2. By Lemma 3.6,
$$\exists n_3: \quad \sup_{t\in\mathcal{T}}\bar{d}(\bar{\mu}_{n,t},\bar{\nu}_{n,t})\le\frac{\epsilon}{4}, \quad n\ge n_3.$$
3. For the remainder of the proof, let $n\ge n_1:=\max\{n_2,n_3\}$ be arbitrary but fixed.
4. By Theorem 3.8, there exist $\mathcal{U}^{(n)}\in\mathcal{F}^{(n)}$, $\mathbb{P}^{(n)}_0(\mathcal{U}^{(n)})=1$, and $m_1=m_1(\epsilon,n_1)$ such that
$$\sup_{t\in\mathcal{T}}\bar{d}(\bar{\mu}_{n,m,t},\bar{\mu}_{n,t})\le\frac{\epsilon}{4} \quad \forall m\ge m_1. \tag{3.41}$$

5. By Lemma 2.7 and (3.40),

$$\sup_{t\in\mathcal{T}} \bar{d}(\bar{\nu}_t, \bar{\nu}_{n,t}) < e^{2T}\|W - W_n\|_{L^2(I^2)} \le \frac{\epsilon}{8}. \tag{3.42}$$

6. We estimate

$$\begin{aligned}
&\bar{d}(\bar{\mu}_{n,m,t}, \bar{\nu}_{n,m,t}) \\
&\quad = \int_I d\left(\mu^x_{n,m,t}, \nu^x_{n,m,t}\right) dx \\
&\quad = n^{-1}\sum_{i=1}^n d(\mu^{x_{n,i}}_{n,m,t}, \nu^{x_{n,i}}_{n,m,t}) \\
&\quad = n^{-1}\sum_{i=1}^n \sup_{f\in\mathcal{L}} \left| \int_{\mathbb{S}} f(v)\left(d\mu^{x_i}_{n,m,t}(v) - d\nu^{x_i}_{n,m,t}(v)\right)\right| \\
&\quad = n^{-1}\sum_{i=1}^n \sup_{f\in\mathcal{L}} \left| m^{-1}\sum_{j=1}^m \left(f(v_{N,(i-1)m+j}(t)) - f(u_{N,(i-1)m+j}(t))\right)\right| \\
&\quad \le n^{-1}\sum_{i=1}^n \sup_{f\in\mathcal{L}} m^{-1}\sum_{j=1}^m \left|f(v_{N,(i-1)m+j}(t)) - f(u_{N,(i-1)m+j}(t))\right| \\
&\quad \le (nm)^{-1}\sum_{i=1}^n\sum_{j=1}^m \left|v_{N,(i-1)m+j}(t) - u_{N,(i-1)m+j}(t)\right|.
\end{aligned} \tag{3.43}$$

Further, by the Schwarz inequality followed by the application of Lemma 3.4 and (3.40), we continue the string of estimates in (3.43) as follows:

$$\begin{aligned}
\bar{d}(\bar{\mu}_{n,m,t}, \bar{\nu}_{n,m,t}) &\le N^{-1}\sum_{j=1}^N |v_{N,j}(t) - u_{N,j}(t)| \\
&\le \|v_N - u_N\|_{1,N} \le \sqrt{Te^{5T}}\|W_n - W_N\|_{L^2(I^2)} \\
&\le \sqrt{Te^{5T}}\left(\|W_n - W\|_{L^2(I^2)} + \|W - W_N\|_{L^2(I^2)}\right) \\
&\le \frac{\epsilon}{4},
\end{aligned} \tag{3.44}$$

where we used (3.40) to derive the last inequality.

7. By combining the estimates in 1–6, we have

$$\bar{d}(\bar{\nu}_{n,m,t}, \nu_t) \le \bar{d}(\bar{\nu}_{n,m,t}, \bar{\mu}_{n,m,t}) + \bar{d}(\bar{\mu}_{n,m,t}, \bar{\mu}_{n,t}) + \bar{d}(\bar{\mu}_{n,t}, \bar{\nu}_{n,t}) + \bar{d}(\bar{\nu}_{n,t}, \bar{\nu}_t) \le \epsilon \tag{3.45}$$

uniformly in $t \in \mathcal{T}$. ∎

4. Discussion. The choice of the KM in this paper was motivated by its role in the theory of synchronization [21] and analytical convenience. The analysis in the previous sections can be naturally extended to other models. First, by extending the phase space to include ω it applies easily to the KM with distributed intrinsic frequencies (1.6). Specifically, let $G = \mathbb{S} \times \mathbb{R}$ be an extended phase space, $\mathcal{M}_G$ be the space of Borel probability measures on G, $\bar{\mu} : x \in I \mapsto \mu^x \in \mathcal{M}_G$ be an $\mathcal{M}_G$-measurable function, and $\bar{\mu}_t : t \in \mathcal{T} \mapsto \bar{\mu}_t^x$ be a weakly continuous function. Then we rewrite the equation of characteristics in the following form:

(4.1)
$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ \phi \end{pmatrix} = \begin{pmatrix} \phi + \int_I W(x,y)\left\{\int_{\mathbb{S}}\int_{\mathbb{R}} D(v-u)d\mu_t^y(v,\lambda)\right\}dy \\ 0 \end{pmatrix} =: V[W,\bar{\mu}_\cdot](u,\phi,x,t).$$

By assigning initial condition $(u(0), \phi(0)) = (u_i^0, \omega_i)$, the right-hand side of the top equation in (4.1) yields the velocity field acting on oscillator i with the intrinsic frequency ω_i. Since the newly added second component of the vector field in (4.1) is trivial, all necessary estimates for $V[W,\bar{\mu}_\cdot](u,\phi,x,t)$ are as before. With the equation of characteristics (4.1) in hand, one can set up the fixed point equation and then analyze it exactly in the same way as it was done in sections 2 and 3. We refer an interested reader to [3], where the KM with distributed frequencies was analyzed albeit for Lipschitz graph limits. More generally, our formalism allows us to deal with coupled models like (1.6), for which the distributions of the initial data and parameters may depend on a given oscillator. In conventional mean field models, initial data (and parameters) are assumed to be independent and identically distributed (see, e.g., [7]).

Likewise, there are no principal difficulties in applying our analysis to other models of interacting dynamical systems on graphs, such as the Cucker–Smale model of flocking [23, 18], consensus protocols [16], as well as neuronal networks [2]. The multi-dimensionality of the phase space can be handled in the same way as explained above for the KM with distributed frequencies. The treatment of the coupling term, the key ingredient in the analysis, remains the same as in the present paper. On the other hand, the weighted graph model, which we adopted in this paper (see (1.3)–(1.5)), provides a simple unified treatment of interacting dynamical systems on a variety of deterministic and random graphs, including Erdős–Rényi and small-world, and certain approximations of power law graphs [15, 9].

Appendix A. Proof of Lemma 1.1. We will prove the proximity between the solutions of the coupled systems on random and averaged deterministic graphs for the following generalized KM:

$$\dot{u}_{n,i} = f(u_{n,i}, \omega_i, t) + \frac{1}{n}\sum_{j=1}^{n} W_{n,ij} D(u_{n,j} - u_{n,i}), \quad i \in [n], \tag{A.1}$$

and

$$\dot{\bar{u}}_{n,i} = f(\bar{u}_{n,i}, \omega_i, t) + \frac{1}{n}\sum_{j=1}^{n} e_{n,ij} D(\bar{u}_{n,j} - \bar{u}_{n,i}), \quad i \in [n]. \tag{A.2}$$

Here, $f(u, \omega, t)$ is a Lipschitz continuous function in u and $\omega \in \mathbb{R}^r$ and continuous in t; D is a 2π-periodic Lipschitz continuous function. Recall that $e_{n,ij}$, $1 \leq i < j \leq n$, are Bernoulli random variables

$$\mathbb{P}(e_{n,ij} = 1) = W_{n,ij} \tag{A.3}$$

and $e_{n,ji} = e_{n,ij}$. Denote the Lipschitz constants of f and D by L_f and L_D, respectively, and let $L = \max\{L_f, L_D, 1\}$. The proof below is a modification of the proof of [3, Lemma 4.3]. (See also [9, Lemma 4.1].

Denote $\phi_{n,i} := u_{n,i} - \bar{u}_{n,i}$. By subtracting (A.2) from (A.1), multiplying the result by $n^{-1}\phi_{n,i}$, and summing over $i \in [n]$, we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\phi_n\|^2_{1,n} \\
&= \underbrace{n^{-1}\sum_{i=1}^{n}(f(u_{n,i},\omega_i,t) - f(u_{n,i},\omega_i,t))\phi_{n,i} + n^{-2}\sum_{i,j=1}^{n}(W_{n,ij} - e_{n,ij})D(u_{n,j} - u_{n,i})\phi_{n,i}}_{I_1} \\
&\quad + \underbrace{n^{-2}\sum_{i,j=1}^{n} e_{n,ij}\left[D(u_{n,j} - u_{n,i}) - D(\bar{u}_{n,j} - \bar{u}_{n,i})\right]\phi_{n,i}}_{I_2} =: I_1 + I_2,
\end{aligned} \tag{A.4}$$

where $\|\cdot\|^2_{1,n}$ is defined in (1.10).

Using Lipschitz continuity of D and the triangle inequality, we have

$$\begin{aligned}
|I_2| &\le L_D n^{-2}\sum_{i,j=1}^{n}\left(|\phi_{n,i}| + |\phi_{n,j}|\right)|\phi_{n,i}| \\
&\le L_D n^{-1}\sum_{i=1}^{n}\phi^2_{n,i} + \frac{L_D}{2n^2}\sum_{i,j=1}^{n}\left(\phi^2_{n,i} + \phi^2_{n,j}\right) \le 2L_D\|\phi_n\|^2_{1,n}.
\end{aligned} \tag{A.5}$$

To estimate I_1, we will need the following definitions:

$$\begin{aligned}
Z_{n,i}(t) &= n^{-1}\sum_{j=1}^{n} a_{n,ij}(t)\eta_{n,ij}, \\
a_{n,ij}(t) &= D\left(u_{n,j}(t) - \bar{u}_{n,i}(t)\right), \\
\eta_{n,ij} &= W_{n,ij} - e_{n,ij},
\end{aligned}$$

and $Z_n = (Z_{n,1}, Z_{n,2}, \dots, Z_{n,n})$. With these definitions in hand, we estimate I_1 as follows:

$$|I_1| = L_f\|\phi_n\|^2_{1,n} + \left|n^{-1}\sum_{i=1}^{n} Z_{n,i}\phi_{n,i}\right| \le L_f\|\phi_n\|^2_{1,n} + 2^{-1}\left(\|Z_n\|^2_{1,n} + \|\phi_n\|^2_{1,n}\right). \tag{A.6}$$

The combination of (A.4), (A.5), and (A.6) yields

$$\frac{d}{dt}\|\phi_n(t)\|^2_{1,n} \le 7L\|\phi_n(t)\|^2_{1,n} + \|Z_n(t)\|^2_{1,n}. \tag{A.7}$$

Using the Gronwall's inequality, we have

$$\|\phi_n(t)\|^2_{1,n} \le e^{7Lt}\left(\|\phi_n(0)\|^2_{n,1} + \int_0^t e^{-7Ls}\|Z_n(s)\|^2_{1,n}ds\right)$$

and

$$\sup_{t\in[0,T]}\|\phi_n(t)\|^2_{1,n} \le e^{7LT}\left(\|\phi_n(0)\|^2_{n,1} + \int_0^\infty e^{-7Ls}\|Z_n(s)\|^2_{1,n}ds\right). \tag{A.8}$$

Our next goal is to estimate $\int_0^\infty e^{-7Ls}\|Z_n(s)\|_{1,n}^2 ds$. To this end, we will use the following observations. Note that $\eta_{n,ik}$ and $\eta_{n,il}$ are independent for $k \neq l$ and

$$\mathbb{E}\eta_{n,ij} = \mathbb{E}\left(W_{n,ij} - e_{n,ij}\right) = 0, \tag{A.9}$$

by (A.3).

By straightforward estimation (cf. [3, Lemma 4.3]), we have

$$\mathbb{E}\eta_{n,ij}^2 \leq 2^{-2} \quad \text{and} \quad \mathbb{E}\left(\eta_{n,ij}^4\right) \leq 2^{-4}, \quad (i,j) \in [n]^2. \tag{A.10}$$

Next,

$$\int_0^\infty e^{-7Lt} Z_{n,i}(t)^2 dt = n^{-2} \sum_{k,l=1}^n c_{n,ikl}\eta_{n,ik}\eta_{n,il}, \tag{A.11}$$

where

$$c_{n,ikl} = \int_0^\infty e^{-7Lt} a_{n,ik}(t) a_{ni,l}(t) dt \quad \text{and} \quad |c_{n,ikl}| \leq (7L)^{-1} =: C_1. \tag{A.12}$$

Further, from (A.11) and (A.12), we have

$$\int_0^\infty e^{-7Lt}\|Z_n(t)\|_{1,n}^2 dt = n^{-3} \sum_{i,k,l=1}^n c_{n,ikl}\eta_{n,ik}\eta_{n,il} \tag{A.13}$$

and, finally,

$$\mathbb{E}\left(\int_0^\infty e^{-7Lt}\|Z_n(t)\|_{1,n}^2 dt\right)^2 = n^{-6} \sum_{i,k,l,j,p,q=1}^n c_{n,ikl}c_{n,jpq}\mathbb{E}\left(\eta_{n,ik}\eta_{n,il}\eta_{n,jp}\eta_{n,jq}\right). \tag{A.14}$$

We have six summation indices i,k,l,j,p,q ranging from 1 to n. Since $\mathbb{E}\eta_{n,ik} = 0$ for $i,k \in [n]$, and RVs $\eta_{n,ik}$ and $\eta_{n,jp}$ are independent whenever $\{i,k\} \neq \{j,p\}$, the nonzero terms on the right-hand side of (A.14) fall into two groups:

I : $c_{n,ikk}^2\eta_{n,ik}^4$,

II : $c_{n,ikk}c_{n,jpp}\eta_{n,ik}^2\eta_{n,jp}^2$ $(i \neq j)$ or $c_{n,ikl}^2\eta_{n,ik}^2\eta_{n,il}^2$ $(k \neq l)$.

There are n^2 terms of type I and $3n^3(n-1)$ terms of type II. Thus,

$$\mathbb{E}\left(\int_0^\infty e^{-7Lt}\|Z_n(t)\|_{1,n}^2 dt\right)^2 \leq C_1^2 n^{-6}\left(n^2 + 3n^3(n-1)\right) = O\left(n^{-2}\right). \tag{A.15}$$

For a given $\epsilon > 0$ and arbitrary $0 < \delta < 1$, define

$$A_n = \left\{\left|\int_0^\infty e^{-7Ls}\|Z_n(s)\|_{1,n}^2 ds\right| > \epsilon n^{\frac{-(1-\delta)}{2}}\right\}.$$

By the Markov inequality and (A.15), we obtain

$$\sum_{n=1}^\infty \mathbb{P}(A_n) \leq \epsilon^{-2}\sum_{n=1}^\infty \mathbb{E}\left(\int_0^\infty \|Z_n(t)\|_{1,n}^2 dt\right)^2 = \epsilon^{-2}\sum_{n=1}^\infty n^{-(1+\delta)} < \infty. \tag{A.16}$$

By the Borel–Cantelli lemma,

$$\lim_{n\to\infty} n^{\frac{1-\delta}{2}} \mathbb{E}\left(\int_0^\infty e^{-7Lt}\|Z_n(t)\|_{1,n}^2 dt\right)^2 = 0, \quad \mathbb{P}\text{-a.s.} \tag{A.17}$$

The combination of (3.33) and (A.8) yields

$$\sup_{t\in[0,T]} \|u_n(t) - \bar{u}_n(t)\|_{1,n} \le e^{4LT}\left(\|u_n(0) - \bar{u}_n(0)\|_{1,n} + O\left(n^{\frac{-(1-\delta)}{4}}\right)\right), \quad \mathbb{P}\text{-a.s.} \tag{A.18}$$

Acknowledgments. The authors thank Carlo Lancellotti for sharing his insights on the mean field limit, an anonymous referee for pointing out Dvoretzky–Kiefer–Wolfowitz inequality, and Mokshay Madiman for a useful conversation.

REFERENCES

[1] D.M. Abrams and S.H. Strogatz, *Chimera states in a ring of nonlocally coupled oscillators*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 16 (2006), pp. 21–37.

[2] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul, *Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons*, J. Math. Neurosci., 2 (2012), 10.

[3] H. Chiba and G.S. Medvedev, *The Mean Field Analysis for the Kuramoto Model on Graphs* I. *The Mean Field Equation and Transition Point Formulas*, arXiv:1612.06493, 2016.

[4] R.M. Dudley, *Real Analysis and Probability*, Cambridge Stud. Adv. Math. 74, Cambridge University Press, Cambridge, 2002.

[5] L.C. Evans, *Partial Differential Equations*, 2nd ed., Grad. Stud. Math. 19, AMS, Providence, RI, 2010.

[6] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1992.

[7] F. Golse, *On the dynamics of large particle systems in the mean field limit*, Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity, Lect. Notes Appl. Math. Mech. 3, Springer, New York, 2016, pp. 1–144.

[8] F.C. Hoppensteadt and E.M. Izhikevich, *Weakly Connected Neural Networks*, Appl. Math. Sci. 126, Springer, New York, 1997.

[9] D. Kaliuzhnyi-Verbovetskyi and G.S. Medvedev, *The semilinear heat equation on sparse random graphs*, SIAM J. Math. Anal., 49 (2017), pp. 1333–1355.

[10] Y. Kuramoto and D. Battogtokh, *Coexistence of coherence and incoherence in nonlocally coupled phase oscillators*, Nonlinear Phenom. Complex Syst., 5 (2002), pp. 380–385.

[11] C. Lancellotti, *On the Vlasov limit for systems of nonlinearly coupled oscillators without noise*, Transp. Theory Statist. Phys., 34 (2005), pp. 523–535.

[12] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B, 96 (2006), pp. 933–957.

[13] B. Makarov and A. Podkorytov, *Real Analysis: Measures, Integrals and Applications*, Universitext, Springer, New York, 2013.

[14] P. Massart, *The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality*, Ann. Probab., 18 (1990), pp. 1269–1283.

[15] G.S. Medvedev and X. Tang, *The Kuramoto Model on Power Law Graphs*, arXiv:1705.04905, 2017.

[16] G.S. Medvedev, *Stochastic stability of continuous time consensus protocols*, SIAM J. Control Optim., 50 (2012), pp. 1859–1885.

[17] G.S. Medvedev, *The nonlinear heat equation on W-random graphs*, Arch. Ration. Mech. Anal., 212 (2014), pp. 781–803.

[18] S. Motsch and E. Tadmor, *Heterophilious dynamics enhances consensus*, SIAM Rev., 56 (2014), pp. 577–621.

[19] H. Neunzert, *Mathematical investigations on particle-in-cell methods*, Fluid Dyn. Trans., 9 (1978), pp. 229–254.

[20] M.A. Porter and J.P. Gleeson, *Dynamical systems on networks*, Front. Appl. Dyn. Syst. Rev. Tutor. 4, Springer, New York, 2016.

[21] S.H. Strogatz, *From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators*, Phys. D, 143 (2000), pp. 1–20.

[22] S.H. Strogatz and R.E. Mirollo, *Stability of incoherence in a population of coupled oscillators*, J. Stat. Phys., 63 (1991), pp. 613–635.

[23] J.H. von Brecht, B. Sudakov, and A.L. Bertozzi, *Swarming on random graphs* II, J. Stat. Phys., 158 (2015), pp. 699–734.

[24] D. Williams, *Probability with Martingales*, Cambridge Math. Textb., Cambridge University Press, Cambridge, 1991.