SIAM J. MATH. ANAL.
Vol. 49, No. 2, pp. 1333–1355

THE SEMILINEAR HEAT EQUATION ON SPARSE RANDOM GRAPHS*

DMITRY KALIUZHNYI-VERBOVETSKYI† AND GEORGI S. MEDVEDEV†

Abstract. Using the theory of L^p-graphons [C. Borgs et al., preprint, arXiv:1401.2906, 2014; C. Borgs et al., preprint, arXiv:1408.0744, 2014], we derive and rigorously justify the continuum limit for systems of differential equations on sparse random graphs. Specifically, we show that the solutions of the initial value problems for the discrete models can be approximated by those of an appropriate nonlocal diffusion equation. Our results apply to a range of spatially extended dynamical models of different physical, biological, social, and economic networks. Importantly, our assumptions cover network topologies featured in many important real-world networks. In particular, we derive the continuum limit for coupled dynamical systems on power law graphs. The latter is the main motivation for this work.

Key words. coupled dynamical systems, continuum limit, power law graph, graph limit, Kuramoto model, opinion dynamics

AMS subject classifications. 34C15, 45J05, 45L05, 74A25, 05C90, 92D25

DOI. 10.1137/16M1075831

1. Introduction.

Reaction-diffusion equations describe the change of concentration of chemical, biological, or other species as a result of local reaction and spatial diffusion:

$$\frac{\partial}{\partial t}u = \Delta u + f(u). \tag{1.1}$$

Here, $u : Q \times \mathbb{R}^+$ is an unknown function, whose interpretation depends on the model at hand, defined on spatial domain $Q \subset \mathbb{R}^n$ and evolving in time. Reaction-diffusion systems have been successfully used to study pattern formation and propagation phenomena in such diverse areas of science as ecology, molecular biology, morphogenesis, neuroscience, and material science, to name a few [6].

In many models of collective behavior of discrete agents, one is led to replace the spatial domain Q by a graph and the Laplace operator Δ by the graph Laplacian [2]. Specifically, let $\Gamma_n = \langle V(\Gamma_n), E(\Gamma_n)\rangle$ denote a graph on n nodes. Here, $V(\Gamma_n)$ and $E(\Gamma_n)$ stand for the sets of nodes and edges, respectively. Without loss of generality, let $V(\Gamma_n) = \{1, 2, \ldots, n\} =: [n]$ and consider the following nonlinear evolution equation on Γ_n:

$$\dot{u}_{ni} = \frac{1}{\deg_{\Gamma_n}(i)} \sum_{j:\{i,j\}\in E(\Gamma_n)} D(u_{nj} - u_{ni}) + f(u_{ni}),\ i \in [n], \tag{1.2}$$

where D and f are Lipschitz continuous functions and $\deg_{\Gamma_n}(i)$ stands for the degree of node $i \in [n]$. The sum on the right-hand side of (1.2) models nonlinear diffusion on Γ_n. Discrete diffusion operators of this form have been used in various models of collective behavior. For instance, with $D(u) = \sin u$ it appears in the Kuramoto model [11] and

*Received by the editors May 18, 2016; accepted for publication (in revised form) February 16, 2017; published electronically April 13, 2017.
http://www.siam.org/journals/sima/49-2/M107583.html
Funding: The work of the second author was supported in part by the NSF DMS 1412066.
†Department of Mathematics, Drexel University, Philadelphia, PA 19104 (dmitryk@math.drexel.edu, medvedev@drexel.edu).

in the power network models [8], with $D(u) = \phi(|u|)u$ for an appropriate function ϕ, it is used in models of flocking [7] and opinion dynamics [17], and with $D(u) = u$—in consensus protocols [13]. In the latter case, (1.2) becomes a semilinear heat equation on Γ_n:

$$\dot{u}_{ni} = \frac{1}{\deg_{\Gamma_n}(i)} \sum_{j:\{j,i\}\in E(\Gamma_n)} (u_{nj} - u_{ni}) + f(u_{ni}),\ i \in [n]. \tag{1.3}$$

Understanding the dynamics of coupled systems (1.2) and (1.3) on graphs modeling connectivity in real-life systems like neuronal networks, power grids, or the internet, can be quite challenging. Recently, new powerful techniques for describing and analyzing the structure of large graphs, based on the appropriate notions of convergence, have been developed in graph theory [12]. Many nontrivial graph sequences that are of interest in applications, such as Erdős–Rényi, small-world, and preferential attachment graphs, have relatively simple limits, expressed by symmetric measurable functions on a unit square, called graphons [12]. These graph limits can be used for developing continuum models approximating the dynamics of (1.2) for large n:

$$\frac{\partial}{\partial t} u(x,t) = \int_I W(x,y) D(u(y,t) - u(x,t))dy + f(u(x,t)), \tag{1.4}$$

where W is the graphon describing the limiting behavior of $\{\Gamma_n\}$.

In [14, 15], the continuum limit (1.4) was derived and rigorously justified for coupled dynamical systems on convergent families of dense graphs.[1] The analysis in [14, 15] covers systems on many interesting graphs including small-world and Erdős–Rényi graphs. However, many real-world networks feature sparse connectivity. Thus, in this paper, we focus on coupled systems on convergent families of sparse graphs.

Our work is inspired by the recent progress made by Borgs et al., who extended the theory of graph limits originally developed for dense graphs to sparse graphs of unbounded degree [4, 5]. The new theory covers many interesting examples of graphs. Notably, it applies to graphs with power law degree distribution, which was identified in many different systems [1]. A distinctive feature of the convergence theory for sparse graphs is that the graphons are no longer bounded, as in the dense case, but in general are functions from $L^p(I^2)$, $p > 1$. This leads to continuum model (1.4) with $W \in L^p(I^2)$. The analysis of (1.4) with an L^p kernel presents new challenges that were not present in the L^∞-case, analyzed in [14, 15]. Overcoming these problems is the goal of this paper.

In the next section, we adapt the notion of W-random sparse graph from [4] to define a sequence of random graphs $\Gamma_n = G(W, \rho_n, X_n)$ with edge density $\rho_n \to 0$ as $n \to \infty$ and with the graph limit $W \in L^p(I^2), p > 1$. This random graph model will be used throughout this paper. It covers power law graphs, our main motivating example, as well as sparse stochastic block and sparse Erdős–Rényi graphs (cf. Examples 2.1 and 2.4) among many other sparse graphs. In section 2.1, we compute the expected degree and edge density of $\Gamma_n = G(W, \rho_n, X_n)$. We then formulate the dynamical model on Γ_n and formally derive the continuum limit (1.4). The derivation includes two steps. First, we average the right-hand side of the coupled model (which depends on the random realization of Γ_n) to obtain a deterministic equation. We then send $n \to \infty$ in the averaged system to derive the continuum limit. This derivation is done for the semilinear heat equation (1.3), which will be studied in the main part of the

[1] If $|E| = O(|V|^2)$, $\Gamma = \langle V, E\rangle$ is called dense, otherwise it is called sparse.

paper. However, the same derivation easily translates to the nonlinear equation (1.2), which results in (1.4). In section 3, we establish well-posedness of the initial value problem (IVP) for (1.4) and derive certain a priori estimates for solutions of the IVPs. For technical reasons, we restrict ourselves to studying the semilinear heat equation for the remainder of this paper. In the last section, we comment on how this analysis extends to cover certain nonlinear models arising in applications. In particular, we discuss the Kuramoto model on power law graphs.

The main result of this work is formulated in section 2.3. Under the appropriate assumptions on the graphon $W \in L^2(I^2)$ and the nonlinearity f, we prove that the solutions of the IVP for the semilinear heat equation (1.3) on Γ_n converge in $L^2(0,T;L^2(I))$ (for any $T>0$) in probability to the solution of the continuum limit (1.4) subject to the appropriate initial condition as $n\to\infty$. This is the subject of Theorem 2.6, which is proved in sections 4 and 5. In the former section, the justification for the averaging is provided. In the latter, we show that the solutions of the IVP for the averaged equation on Γ_n converge to those for the continuum limit as $n\to\infty$. To this end, we show that the solutions of the averaged equation can be approximated by the solutions of certain Galerkin problems, which, in turn, converge to the solution of the continuum limit. The final section discusses extensions of our work to certain nonlinear models that are important in applications.

2. The model.

2.1. The random graph model. We start with the description of the sparse random graphs that will be used in this paper. Our random graph model is motivated by the construction of sparse W-random graphs in [4, 5]. Specifically, let W be a symmetric nonnegative function on a unit square I^2, X_n be a discretization of I,

$$X_n = \{x_{n0}, x_{n1}, x_{n2}, \dots, x_{nn}\},\ x_{ni} = i/n,\ i = 0,1,\dots,n, \tag{2.1}$$

and $\{\rho_n\}$ be a sequence of positive numbers such that $\rho_n \to 0$ and $n\rho_n \to \infty$ as $n\to\infty$.

$\Gamma_n = G(W,\rho_n,X_n)$ stands for a random graph with the node set $V(\Gamma_n)=[n]$ and the edge set $E(\Gamma_n)$ such that the probability that $\{i,j\}$ forms an edge is

$$\mathbb{P}\left(\{i,j\}\in E(\Gamma_n)\right) = \rho_n \bar{W}_n(x_{ni},x_{nj}),\ i,j\in[n],^2 \tag{2.2}$$

where

$$\bar{W}_n(x,y) = \rho_n^{-1}\wedge W(x,y).^3 \tag{2.3}$$

The decision on whether a given pair of nodes is included in the edge set is made independently from other pairs. In other words, $G(W,\rho_n,X_n)$ is a product probability space

$$(\Omega_n = \{0,1\}^{n(n+1)/2}, 2^{\Omega_n}, \mathbb{P}\ ). \tag{2.4}$$

By $\Gamma_n(\omega), \omega\in\Omega_n$, we will denote a random graph drawn from the probability distribution $G(W,\rho_n,X_n)$.

Throughout this paper, we use Bernoulli random variables

$$\xi_{nij}(\omega) = \mathbf{1}_{\{i,j\}\in E(\Gamma_n)}(\omega),\ i,j\in[n], \tag{2.5}$$

to describe the edge set of Γ_n. Random variable ξ_{nij} takes value 1 if $\{i,j\}$ forms an

[2] To keep notation simple, we allow for loops, i.e., edges connecting a node to itself, in our random graph model. Excluding loops would not lead to any changes in the analysis.

[3] Throughout this paper, we use $a\wedge b$ and $a\vee b$ to denote $\min\{a,b\}$ and $\max\{a,b\}$, respectively.

edge and 0 otherwise. In particular,

$$\mathbb{E}\,\xi_{nij} = \mathbb{P}\left(\{i,j\} \in E(\Gamma_n)\right) = \rho_n \bar{W}_n(x_{ni}, x_{nj}),\ i,j \in [n], \tag{2.6}$$

and the expected degree of node i of Γ_n

$$\mathbb{E}\ \deg_{\Gamma_n}(i) = \mathbb{E}\left(\sum_{j=1}^n \xi_{nij}\right) = \rho_n \sum_{j=1}^n \bar{W}_n(x_{ni}, x_{nj}). \tag{2.7}$$

Next, we formulate our assumptions on the graphon W.

(W-1) $W \in L^2(I^2)$ is a nonnegative symmetric function on the unit square I^2 that is continuous on its interior.

(W-2) $\int_{I^2} W(x,y)dxdy > 0$ and

$$n^{-2} \sum_{i,j=1}^n \bar{W}_n(x_{ni}, x_{nj}) = \int_{I^2} W(x,y)dxdy + o(1). \tag{2.8}$$

(W-3) For every $x \in (0,1]$, $W(x,\cdot) \in L^1(I)$, and

$$\inf_{x\in(0,1)} \int_I W(x,z)dz =: \nu > 0. \tag{2.9}$$

Moreover,

$$n^{-1} \sum_{j=1}^n \bar{W}_n(x, x_j) = \int_I W(x,y)dy\,(1 + \delta_n(x)), \tag{2.10}$$

where $\delta_n(x) \to 0$ as $n \to \infty$ uniformly in $x \in (0,1)$.

Conditions in (W-2) and (W-3) guarantee that the expected edge density and expected degrees of nodes of Γ_n for $n \gg 1$ are well-defined and are well-approximated by the corresponding integrals of W. The above assumptions on graphon W are dictated by the random graph model and are practically minimal.

We will now introduce two more technical assumptions that are needed for the proof of our main result:

(W-4) $W \in L^4(I^2)$ and

$$\limsup_n n^{-2} \sum_{i,j=1}^n \bar{W}_n(x, x_j)^2 < \infty. \tag{2.11}$$

Assumptions in (W-4) will not be used until section 5.3.

The main example motivating our random graph model is the following configuration model of a power law graph.

Example 2.1. Let $0 < \alpha < \gamma < 1$ and consider $G(W, \rho_n, X_n)$, where $\rho_n = n^{-\gamma}$ and

$$W(x,y) = (1-\alpha)^2 (xy)^{-\alpha}. \tag{2.12}$$

LEMMA 2.2. *$\Gamma_n = G(W, \rho_n, X_n)$ of Example 2.1 is a sparse graph with a power law degree distribution. In particular, we have the following:*

(A) *The expected degree of node $i \in [n]$ of Γ_n is*

$$\mathbb{E}\ \deg_{\Gamma_n}(i) = (1-\alpha) n^{1+\alpha-\gamma} i^{-\alpha} (1 + o(1)). \tag{2.13}$$

(B) *The expected edge density of Γ_n is $n^{-\gamma}(1 + o(1))$.*

Proof. By (2.2),

$$\mathbb{E}\ \deg_{\Gamma_n}(i) = \sum_{j=1}^{n} \mathbb{P}\left(\{i,j\} \in E(\Gamma_n)\right) = \rho_n n \sum_{j=1}^{n} \bar{W}_n(x_{ni}, x_{nj}) n^{-1}. \tag{2.14}$$

Plugging $\rho_n = n^{-\gamma}$ and (2.3) into (2.14), we have

$$\mathbb{E}\ \deg_{\Gamma_n}(i) = (1-\alpha) n^{1+\alpha-\gamma} i^{-\alpha} \sum_{j=1}^{n} \left[\delta_n \wedge W^{(1)}(x_{nj})\right] n^{-1}, \tag{2.15}$$

where $W^{(1)}(x) = (1-\alpha)x^{-\alpha}$ and $\delta_n = i^{\alpha} n^{\gamma-\alpha}(1-\alpha)^{-1}$.

Denote

$$I_{ni} := (x_{n(i-1)}, x_{ni}], \quad i \in [n]. \tag{2.16}$$

Next, let

$$W_n^{(1)}(x) = \sum_{j=1}^{n} (\delta_n \wedge W^{(1)}(x_{nj})) \mathbf{1}_{I_{nj}}(x) \tag{2.17}$$

and note that

$$n^{-1} \sum_{j=1}^{n} \delta_n \wedge W_n^{(1)}(x_{nj}) = \int_I W_n^{(1)} dx. \tag{2.18}$$

Furthermore, $W_n^{(1)} \le W^{(1)}$ and $W_n^{(1)} \to W^{(1)}$ pointwise on $(0,1]$ as $n \to \infty$. By the dominated convergence theorem [3],

$$\lim_{n\to\infty} \int_I W_n^{(1)} dx = \int_I W^{(1)} dx = 1. \tag{2.19}$$

The combination of (2.15), (2.18), and (2.19) yields (2.13). This shows (A).

A similar argument is used to estimate the expected number of edges in Γ_n,

$$\mathbb{E}\ |E(\Gamma_n)| = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_n \bar{W}_n(x_{ni}, x_{nj}) = \frac{1}{2} n^{2-\gamma} \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{W}_n(x_{ni}, x_{nj}) n^{-2}.$$

Define

$$W_n(x,y) = \sum_{i,j=1}^{n} (\rho_n^{-1} \wedge W(x_{ni}, x_{nj})) \mathbf{1}_{I_{ni} \times I_{nj}}(x,y). \tag{2.20}$$

Then

$$\mathbb{E}\ |E(\Gamma_n)| = \frac{n^{2-\gamma}}{2} \int_{I^2} W_n dxdy. \tag{2.21}$$

By construction, $W_n \le W$ and $W_n \to W$ as $n \to \infty$ on $(0,1] \times (0,1]$. By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_{I^2} W_n dxdy = \int_{I^2} W dxdy = 1. \tag{2.22}$$

Equations (2.21) and (2.22) imply

$$\mathbb{E}\ |E(\Gamma_n)| = \frac{n^{2-\gamma}}{2}(1 + o(1)). \tag{2.23}$$

By dividing both sides of (2.23) by $n(n+1)/2$, the total number of possible edges, we obtain the statement in (B). ∎

Remark 2.3. The power law graphs defined above are sparse, because the expected edge density is $O(n^{-\gamma})$, $\gamma > 0$. On the other hand, the expected number of edges grows superlinearly as $n^{2-\gamma}$, because $\gamma < 1$. To preserve these features, in the general random graph model $G(W, \rho_n, X_n), n \in \mathbb{N}$, above it is assumed that $\rho_n \to 0$ and $n\rho_n \to \infty$ as $n \to \infty$.

We conclude the discussion of the graph model with two more examples of sparse graphs covered by our assumptions. Both examples are taken from [5].

Example 2.4. Consider $\Gamma_n = G(W, n^{-\beta}, X_n)$, $\beta \in (0, 1)$ for the following choices of W.

(1) Let $W \equiv 1$. Then Γ_n is a generalization of an Erdős–Rényi random graph $G_{n,p}$ with $p = n^{-\beta}$. Note that the edge density in this case is $n^{-\beta}$. For the classical Erdős–Rényi graph $G_{n,p}$ with constant $p \in (0, 1)$, the edge density is equal to p. The latter graph is dense, whereas the former is sparse.
(2) Let $W(x, y) = b_{ij} \geq 0$, $(x, y) \in V_i \times V_j$, $(i, j) \in [k]^2$, where $\sum_{i,j=1}^{k} b_{ij} > 0$ and $(V_1, V_2, \ldots, V_k)$ is a partition of I into k disjoint intervals. In this case, Γ_n is a sparse stochastic block graph with edge density $n^{-\beta}$.

2.2. The dynamical model. Having defined the structure of the network, we next turn to its dynamics. Let $\Gamma_n = \Gamma_n(\omega)$, $\omega \in \Omega_n$ (cf. (2.4)), be a random graph taken from the probability distribution $G(W, \rho_n, X_n)$ and consider the following system of differential equations,

$$\dot{u}_{ni} = \frac{1}{d_{ni}} \sum_{j=1}^{n} \xi_{nij}(\omega) \left(u_{nj} - u_{ni}\right) + f(u_{ni}),\ i \in [n], \tag{2.24}$$

where $u_n(t) = (u_{n1}(t), u_{n2}(t), \ldots, u_{nn}(t))$, ξ_{nij}, $i, j \in [n]$, are Bernoulli random variables defined in (2.5), $d_{ni} = \mathbb{E}\ \deg_{\Gamma_n}(i)$ (cf. (2.7)), and $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous function

$$\exists L_f > 0 : \quad |f(x) - f(y)| \leq L_f |x - y|\ \forall x, y \in \mathbb{R}. \tag{2.25}$$

The sum on the right-hand side of (2.24) defines a discrete diffusion operator. For simplicity, we scale the sum on the right-hand side of (2.24) by the expected degree rather than by the actual degree. Thus, (2.24) is an evolution equation on a random graph Γ_n. Specifically, it is a semilinear heat equation on Γ_n, since the sum on the right-hand side of (2.24) is a discrete graph Laplacian.

We are interested in describing the dynamics of (2.24) for $n \gg 1$. However, the right-hand side of (2.24) depends on the random graph $\Gamma_n(\omega)$, i.e., on the random event $\omega \in \Omega_n$:

$$F_{ni}(u_n, \omega) = \frac{1}{d_{ni}} \sum_{j=1}^{n} \xi_{nij}(\omega) \left(u_{nj} - u_{ni}\right) + f(u_{ni}), \quad i \in [n].$$

As the first step in the analysis of (2.24), we approximate it by the deterministic problem by averaging the right-hand side of (2.24) over all realizations of Γ_n:

$$\dot{v}_{ni}(t) = \bar{F}_{ni}(v_n),\ v_n(t) = (v_{n1}(t), v_{n2}(t), \ldots, v_{nn}(t)), \tag{2.26}$$

where

$$\bar{F}_{ni}(v) = \mathbb{E}\, F_{ni}(v_n, \cdot) = \frac{\rho_n}{d_{ni}} \sum_{j=1}^{n} \bar{W}_n(x_{ni}, x_{nj})(v_{nj} - v_{ni}) + f(v_{ni}),$$

$$= \frac{1}{n} \sum_{j=1}^{n} V_{nij}(v_{nj} - v_{ni}) + f(v_{ni}), \quad i \in [n], \tag{2.27}$$

where

$$V_{nij} = \frac{\bar{W}_n(x_{ni}, x_{nj})}{n^{-1} \sum_{k=1}^{n} \bar{W}_n(x_{ni}, x_{nk})}. \tag{2.28}$$

Next, we take the limit in the averaged equation (2.27) as $n \to \infty$. To this end, we represent the solution of (2.27) as a step function,

$$v_n(x,t) = \sum_{i=1}^{n} v_{ni}(t) \mathbf{1}_{I_{ni}}(x), \tag{2.29}$$

and rewrite (2.27) as

$$\frac{\partial}{\partial t} v_n(x,t) = \int_I V_n(x,y)(v_n(y,t) - v_n(x,t))dy + f(v_n(x,t)), \tag{2.30}$$

where

$$V_n(x,y) = \sum_{i,j=1}^{n} V_{nij} \mathbf{1}_{I_{ni} \times I_{nj}}(x,y). \tag{2.31}$$

Assuming that $v_n(x,t) \to u(x,t)$ in the appropriate sense, and using the integrability assumptions (W-2) and (W-3), in the limit as $n \to \infty$ we formally obtain the following continuum limit of (2.26):

$$\frac{\partial}{\partial t} u(x,t) = \int_I U(x,y)\,(u(y,t) - u(x,t))\,dy + f(u(x,t)), \tag{2.32}$$

$$U(x,y) = W(x,y) \left(\int_I W(x,z)dz \right)^{-1}. \tag{2.33}$$

Note that $U \in L^2(I^2)$ (cf. (W-1) and (2.9)).

Example 2.5. For the power law graphs defined in Example 2.1 with square integrable graphons, the continuum limit takes the following form

$$\frac{\partial}{\partial t} u(x,t) = \int_I y^{-\alpha}\,(u(y,t) - u(x,t))\,dy + f(u(x,t)),\ 0 < \alpha < 1/2.$$

The goal of this paper is to describe the relation between the solutions of the IVPs for the discrete model (2.24) on sparse graph $\Gamma_n, n \gg 1$, the averaged model (2.26), and the continuum limit (2.32).

2.3. The main result. Let $g \in L^\infty(I)$ and consider the IVP for (2.32) subject to the initial condition

$$u(x,0) = g(x), \quad x \in I. \tag{2.34}$$

Likewise, we supply the discrete problem (2.24) with the initial condition

$$u_{ni}(0) = n \int_{I_{ni}} g(x)dx, \quad i \in [n]. \tag{2.35}$$

To compare solutions of the IVPs for the discrete and continuous models, we define the step function

$$u_n(x,t) = \sum_{i=1}^{n} u_{ni}(t)\mathbf{1}_{I_{ni}}(x). \tag{2.36}$$

The main result of this paper concerns the L^2-proximity of $u_n(\cdot, t)$ and $u(\cdot, t)$ on finite time intervals for large n.

THEOREM 2.6. *Let* $\Gamma_n = G(W, \rho_n, X_n)$ *be a sequence of random graphs, where* W *satisfies conditions* (W-1)–(W-4), $X_n, n \in \mathbb{N}$, *are defined in* (2.1), *and the positive sequence* $\{\rho_n\}$ *is such that* $\rho_n \to 0$ *and* $n\rho_n \to \infty$ *as* $n \to \infty$. *Suppose* $f : \mathbb{R} \to \mathbb{R}$ *is Lipschitz continuous function (cf.* (2.25)*),* $g \in L^\infty(I)$*, and* $T > 0$ *is arbitrary.*

Then for solutions of the IVPs (2.32), (2.34) *and* (2.24), (2.35)*, we have*

$$\int_0^T \|u_n(\cdot, t) - u(\cdot, t)\|^2_{L^2(I)} dt \to 0 \quad \textit{as } n \to \infty \quad \textit{in probability.}$$

3. The IVP for the nonlocal equation.

3.1. Existence and uniqueness of solutions.

In this section, we show that the IVP for (2.32), (2.33) has a unique solution. The contraction mapping principle used below applies to the nonlinear heat equation

$$\frac{\partial}{\partial t} u(x,t) = \int_I U(x,y) D\left(u(y,t) - u(x,t)\right) dy + f(u(x,t)), \tag{3.1}$$

where $D\colon \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous:

$$|D(u) - D(v)| \le L_D |u - v| \; \forall u, v \in \mathbb{R}. \tag{3.2}$$

Below, we study the well-posedness of the IVP for (3.1). The results of this section will obviously hold for (2.32) as well.

With the definition (2.33) in mind, in this section, we assume that $U \in L^p(I^2), p \ge 2$, is a nonnegative function, satisfying

$$\int_I U(x,y)dy = 1. \tag{3.3}$$

We interpret the solution of the IVP for (3.1), $u(x,t)$, as a vector-valued map $\mathbf{u} : \mathbb{R} \to L^q(I)$, i.e., $[\mathbf{u}(t)](x) = u(x,t)$.

THEOREM 3.1. *Suppose* $U \in L^p(I^2)$, $p \ge 2$, *is a nonnegative function satisfying* (3.3) *and functions* f *and* D *satisfy* (2.25) *and* (3.2)*, respectively. Then the IVP for* (3.1) *with initial data* $\mathbf{u}(0) = g \in L^q(I)$, $q = p/(p-1)$ *has a unique solution* $\mathbf{u} \in C^1(\mathbb{R}; L^q(I))$*, which depends continuously on* g.

Proof. Denote

$$\tau = (2L(\|U\|_{L^p(I^2)} + 2))^{-1}, \tag{3.4}$$

where $L = L_f \vee L_D$ is the largest of the two Lipschitz constants of D and f (cf. (2.25), (3.2)). Denote $\mathcal{M} = C(0, \tau; L^q(I))$ and define $K\colon \mathcal{M} \to \mathcal{M}$ as follows:

$$[K\mathbf{u}](t) = g + \int_0^t \left(\int_I U(\cdot, y) D\left(u(y,s) - u(\cdot, s)\right) dy + f(u(\cdot, s)) \right) ds. \tag{3.5}$$

(The correctness of this definition will be shown later.) We rewrite the IVP for (3.1) as a fixed point equation for the mapping K,

$$\mathbf{u} = K\mathbf{u}, \tag{3.6}$$

and show that K is a contraction on $\mathcal{M}$.

The following inequalities hold for any $u \in L^q(I)$ and $W \in L^p(I^2)$, $p > 1$, $q = p/(p-1)$,

$$\|u\|_{L^q(I)} \le \|u\|_{L^{p\vee q}(I)}, \ \|W\|_{L^p(I^2)} \le \|W\|_{L^{p\vee q}(I^2)}. \tag{3.7}$$

They follow from the Hölder inequality applied to functions defined on the unit interval I and the unit square I^2, respectively. In particular, for $q \le 2 \le p$, we have

$$\|u\|_{L^q(I)} \le \|u\|_{L^p(I)}, \ \|U\|_{L^q(I^2)} \le \|U\|_{L^p(I^2)}. \tag{3.8}$$

For any $\mathbf{u}, \mathbf{v} \in \mathcal{M}$, we have

$$\begin{aligned}
&\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} \\
&\quad = \max_{t\in[0,\tau]} \|K\mathbf{u}(t) - K\mathbf{v}(t)\|_{L^q(I)} \\
&\quad \le \max_{t\in[0,\tau]} \left\| \int_0^t \left(\int_I U(\cdot, y) \left| D\left(u(y,s) - u(\cdot,s)\right) - D\left(v(y,s) - v(\cdot, s)\right)\right| dy \right.\right. \\
&\qquad\qquad \left.\left. + L|u(\cdot, s) - v(\cdot, s)| \right) ds \right\|_{L^q(I)} \\
&\quad \le L \max_{t\in[0,\tau]} \left\| \int_0^t \left(\int_I U(\cdot, y) \left| u(y,s) - u(\cdot, s) - v(y,s) + v(\cdot, s)\right| dy \right.\right. \\
&\qquad\qquad \left.\left. + |u(\cdot, s) - v(\cdot, s)| \right) ds \right\|_{L^q(I)} \\
&\quad \le \tau L \max_{t\in[0,\tau]} \left\{ \left\| \int_I U(\cdot, y) \left| u(y,t) - v(y,t)\right| dy \right\|_{L^q(I)} \right. \\
&\qquad\qquad \left. + \left\| \int_I U(\cdot, y) \left| u(\cdot,t) - v(\cdot,t)\right| dy \right\|_{L^q(I)} + \|u(\cdot, t) - v(\cdot, t)\|_{L^q(I)} \right\} \\
&\quad = \tau L \max_{t\in[0,\tau]} \left\{ \left\| \int_I U(\cdot, y) \left| u(y,t) - v(y,t)\right| dy \right\|_{L^q(I)} + 2\left\|u(\cdot, t) - v(\cdot, t)\right\|_{L^q(I)} \right\},
\end{aligned} \tag{3.9}$$

where we used Lipschitz continuity of D and f, and (3.3). Using the Hölder inequality and the second inequality in (3.8), we have

$$\begin{aligned}
\left\| \int_I U(\cdot, y) \left| u(y,t) - v(y,t)\right| dy \right\|_{L^q(I)} &\le \left\| \|U(x, \cdot)\|_{L^p(I)} \|u(\cdot, t) - v(\cdot, t)\|_{L^q(I)} \right\|_{L^q(I_x)} \\
&\le \|U\|_{L^p(I^2)} \|\mathbf{u}(t) - \mathbf{v}(t)\|_{L^q(I)},
\end{aligned} \tag{3.10}$$

where $I_x = [0, 1]$ refers to the domain of a function of x.

The combination of (3.9) and (3.10) yields

$$\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} \le L\tau \left(\|U\|_{L^{p\vee q}(I^2)} + 2\right) \|\mathbf{u} - \mathbf{v}\|_{\mathcal{M}}. \tag{3.11}$$

Thus, using (3.4), we have

$$\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} \le \frac{1}{2} \|\mathbf{u} - \mathbf{v}\|_{\mathcal{M}}.$$

It follows that K is a correctly defined contraction on $\mathcal{M}$.

Next, we show $K(\mathcal{M}) \subset \mathcal{M}$. To this end, for $\mathbf{z} \equiv 0$ on $I \times [0, \tau]$, we have

$$\begin{aligned}\|K\mathbf{u}\|_{\mathcal{M}} &\le \|K\mathbf{u} - K\mathbf{z}\|_{\mathcal{M}} + \|K\mathbf{z}\|_{\mathcal{M}} \\ &\le \frac{1}{2}\|\mathbf{u}\|_{\mathcal{M}} + \|K\mathbf{z}\|_{\mathcal{M}}.\end{aligned} \tag{3.12}$$

Further,

$$[K\mathbf{z}](t) = g + t\,(D(0) + f(0)),$$

so that $Kz \in \mathcal{M}$, and then (3.12) implies that $K\mathbf{u} \in \mathcal{M}$.

From (3.11), by the Banach contraction mapping principle, there exists a unique solution of the IVP for (3.1), $\bar{\mathbf{u}} \in \mathcal{M} \subset C(0, \tau; L^q(I))$. Using $\bar{\mathbf{u}}(\tau)$ as the initial condition, the local solution can be extended to $[0, 2\tau]$, and by repeating this argument to $[0, T]$ for any $T > 0$. In a similar fashion, we can prove the existence and uniqueness of the solution of (3.6) on $[-T, 0]$ for any $T > 0$. Thus, we have a unique solution of (3.6) on the whole real axis, i.e., $\mathbf{u} \in C(0, \mathbb{R}; L^q(I))$. The integrand in (3.5) is continuous as a map $L^q(I) \to L^q(I)$. Thus, (3.5) and (3.6) imply that $\mathbf{u}$ is continuously differentiable and we obtain a classical solution of the IVP for (3.1) on the whole real axis. Finally, since $K : \mathcal{M} \to \mathcal{M}$ is a uniform contraction (cf. (3.11)), which depends on g continuously (cf.(3.5)), the fixed point is a continuous function of g as well (cf. [10, section 1.2.6, Exercise 3]). ∎

3.2. A priori estimates.

THEOREM 3.2. *Let* $\mathbf{u}(t)$ *denote the solution of the IVP for* (3.1) *with* $U \in L^1(I^2)$ *and initial condition* $\mathbf{u}(0) = g \in L^\infty(I)$. *Then* $\mathbf{u} \in C(\mathbb{R}; L^\infty(I))$ *and for any* $T > 0$, *there exists* $C > 0$ *depending on* T *but not on* U *such that*

$$\|\mathbf{u}\|_{C(0,T;L^\infty(I))} \le C\|\mathbf{u}(0)\|_{L^\infty(I)}. \tag{3.13}$$

Proof. If $U \in L^1(I^2)$ and $\mathbf{u}(0) = g \in L^\infty(I)$ then the contraction mapping argument used in the proof of Theorem 3.1 yields the existence of the unique solution $\mathbf{u} \in C^1(\mathbb{R}; L^\infty(I))$. Indeed, let

$$\mathcal{M} := C(0, \tau; L^\infty(I)) \quad \text{for} \quad \tau := (6L)^{-1}$$

and consider the operator K defined by (3.5). As before, we show that K is a well-defined contraction on $\mathcal{M}$.

Indeed, for any $\mathbf{u}, \mathbf{v} \in \mathcal{M}$, we have

$$\begin{aligned}\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} &= \max_{t\in[0,\tau]} \|K\mathbf{u}(t) - K\mathbf{v}(t)\|_{L^\infty(I)} \\ &\le \max_{t\in[0,\tau]} \left\| \int_0^t \left(\int_I U(\cdot, y)\,|D\,(u(y,s) - u(\cdot,s)) - D\,(v(y,s) - v(\cdot,s))|\,dy \right.\right. \\ &\qquad \left.\left. + |f(u(\cdot,s)) - f(v(\cdot,s))| \right) ds \right\|_{L^\infty(I)}.\end{aligned}$$

Using the Lipschitz continuity of D and f and the triangle inequality, we further obtain

$$\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} \leq \max_{t\in[0,\tau]} L \left\| \int_0^t \left(\int_I U(\cdot, y)\, |u(y,s) - u(\cdot, s) - v(y,s) + v(\cdot, s)|\, dy \right.\right.$$
$$\left.\left. + |u(\cdot, s) - (v(\cdot, s)| \right) ds \right\|_{L^\infty(I)}$$
$$\leq \max_{t\in[0,\tau]} \left(2L \int_I U(\cdot, y) dy + L \right) \int_0^t \|u(\cdot, s) - v(\cdot, s)\|_{L^\infty(I)} ds$$
$$\leq 3L\tau \|\mathbf{u} - \mathbf{v}\|_{\mathcal{M}}.$$

Recalling, the definition of τ, we arrive at

$$\|K\mathbf{u} - K\mathbf{v}\|_{\mathcal{M}} \leq \frac{1}{2} \|\mathbf{u} - \mathbf{v}\|_{\mathcal{M}}.$$

Following the lines of the proof of Theorem 3.1, it is straightforward to show that the fixed point of (3.6) is the unique solution of the IVP for (3.1), $\mathbf{u} \in C^1(\mathbb{R}, L^\infty(I))$, which depends continuously on the initial data $g \in L^\infty(I)$.

Denote

$$m(t) := \|u(\cdot, t)\|_{L^\infty(I)}.$$

From (3.1), using the Lipschitz continuity of D and f, we have

$$|u(x,t)| = |g(x)| + L \int_0^t \left(\int_I U(x,y) \left(|u(y,s) - v(y,s)| + |u(x,s) - v(x,s)|\right) dy \right.$$
$$\left. + |u(x,s) - v(x,s)| \right) ds$$
$$\leq m(0) + L \left(2 \int_I U(x,y) dy + 1 \right) \int_0^t m(s) ds.$$

Thus,

$$m(t) \leq m(0) + 3L \int_0^t m(s) ds.$$

Since $\mathbf{u} \in C(0,T; L^\infty(I))$, by Gronwall's inequality (cf. [9, Appendix B]), for any $t \in [0,T]$

$$m(t) \leq m(0) \left(1 + 3Lte^{3Lt}\right) \leq Cm(0), \quad C := 1 + 3LTe^{3LT},$$

and (3.13) follows. ∎

We will also use the following observation.

LEMMA 3.3. *Let $W \in L^2(I^2)$ be a symmetric function and $u \in L^\infty(I)$. Then*

$$\int_{I^2} W(x,y) \left(u(y) - u(x)\right) u(x) dx dy = \frac{-1}{2} \int_{I^2} W(x,y) \left(u(y) - u(x)\right)^2 dx dy. \tag{3.14}$$

Proof. Rewrite the left-hand side of (3.14) as

$$\int_{I^2} W(x,y) \left(u(y) - u(x)\right) u(x) dx dy = - \int_{I^2} W(x,y) \left(u(y) - u(x)\right)^2 dx dy$$
$$+ \int_{I^2} W(x,y) \left(u(y) - u(x)\right) u(y) dx dy. \tag{3.15}$$

Using the symmetry of $W(x,y)$, for the second term on the right-hand side of (3.15) we have

$$\int_{I^2} W(x,y)\,(u(y)-u(x))\,u(y)dxdy = -\int_{I^2} W(x,y)\,(u(y)-u(x))\,u(x)dxdy. \tag{3.16}$$

After plugging (3.16) into (3.15), we obtain (3.14). □

Next, we formulate the discrete counterparts of Theorem 3.2 and Lemma 3.3. To this end, consider an IVP for the semilinear discrete heat equation

$$\dot{u}_{ni} = \frac{1}{n}\sum_{j=1}^{n} V_{nij} D(u_{nj}-u_{ni}) + f(u_{ni}),\ i \in [n], \tag{3.17}$$

where (V_{nij}) is a nonnegative matrix with entries derived from the graphon W (see (2.28)).

Let $u_n(t) = (u_{n1}(t), u_{n2}(t), \ldots, u_{nn}(t))$ be a solution of (3.17). Denote

$$\|u_n\|_{2,n} = \sqrt{n^{-1}\sum_{i=1}^{n} u_{ni}^2} \text{ and } \|u_n\|_{\infty,n} = \max_{i\in[n]} |u_{ni}|. \tag{3.18}$$

Recall that the discrete problem (3.17) can be rewritten as the nonlocal equation (2.30). By applying Theorem 3.2 to (2.30), we obtain the following theorem.

THEOREM 3.4. *For the solution of the IVP for* (3.17)*, we have*

$$\max_{t\in[0,T]} \|u_n(t)\|_{\infty,n} \le C\|u_n(0)\|_{\infty,n}\ \forall n, \tag{3.19}$$

where $C > 0$ depends on T only.

Finally, we state a discrete version of Lemma 3.3. It can be derived from Lemma 3.3, or proved directly.

LEMMA 3.5. *Let (W_{ij}) be an $n\times n$ symmetric matrix. Then for any $(\theta_1, \theta_2, \ldots, \theta_n) \in \mathbb{R}^n$,*

$$\sum_{i,j=1}^{n} W_{ij}(\theta_j-\theta_i)\theta_i = \frac{-1}{2}\sum_{i,j=1}^{n} W_{ij}(\theta_j-\theta_i)^2. \tag{3.20}$$

4. Averaging. In this section, we show that for large n the solutions of the heat equation (2.24) on Γ_n can be approximated by the solutions of the averaged equation (2.26), (2.27).

For convenience, we rewrite the original and the averaged models. For the former model, we plug in the expression for the mean degree d_{ni} (2.7) into (2.24) to obtain

$$\dot{u}_{ni} = n^{-1}\sum_{j=1}^{n} \eta_{nij}(u_{nj}-u_{ni}) + f(u_{ni}), \quad i \in [n], \tag{4.1}$$

where

$$\eta_{nij} = \xi_{nij}\left(\rho_n n^{-1}\sum_{j=1}^{n} \bar{W}_{nij}\right)^{-1}. \tag{4.2}$$

Recall the averaged model (2.26)

$$\dot{v}_{ni} = n^{-1} \sum_{j=1}^{n} V_{nij}(v_{nj} - v_{ni}) + f(v_{ni}), \quad i \in [n], \tag{4.3}$$

where

$$V_{nij} = \bar{G}_{ni}^{-1} \bar{W}_{nij}, \quad G_{ni} := n^{-1} \sum_{j=1}^{n} \bar{W}_{nij}. \tag{4.4}$$

Note that for fixed $i \in [n]$, $\{\eta_{nij},\ j \in [n]\}$ are independent random variables and

$$\mathbb{E}\ \eta_{nij} = V_{nij}, \quad i, j \in [n]. \tag{4.5}$$

Below, we use the following weighted norm in $\mathbb{R}^n$:

$$\|\psi_n\|_{G_n} := \sqrt{n^{-1} \sum_{i=1}^{n} G_{ni} \psi_{ni}^2}. \tag{4.6}$$

Here, we implicitly assume that n is large enough, so that $\min_{i \in [n]} G_{ni} > 0$ (cf. (W-3)).

We now formulate the main result of this section.

THEOREM 4.1. *Let $u_n(t)$ and $v_n(t)$ denote solutions of the IVP for* (4.1) *and* (4.3)*, respectively. Suppose that the initial data for these problems satisfy*

$$\max\{|u_n(0)|, |v_n(0)|\} \leq C_1 \ \text{ uniformly in } n \text{ and} \tag{4.7}$$

$$\lim_{n \to \infty} \|v_n(0) - u_n(0)\|_{G_n} = 0. \tag{4.8}$$

Then

$$\lim_{n \to \infty} \max_{t \in [0,T]} \|v_n(t) - u_n(t)\|_{G_n} = 0 \quad \text{in probability.} \tag{4.9}$$

For the proof of Theorem 4.1, we will need the following lemma.

LEMMA 4.2. *Let $T > 0$ and (a_{nij}) be an $n \times n$ matrix, whose entries depend on $t \in [0, T]$. Suppose*

$$\sup_{t \in [0,T]} \max_{(i,j) \in [n]^2} |a_{nij}(t)| \leq C_2 \ \ \forall n \in \mathbb{N}. \tag{4.10}$$

Denote $Z_n(t) = (Z_{n1}(t), Z_{n2}(t), \ldots, Z_{nn}(t))$, where

$$Z_{ni}(t) = n^{-1} \sum_{j=1}^{n} a_{nij}(t)(\eta_{nij} - V_{nij}),\ i \in [n],\ t \in [0, T], \tag{4.11}$$

η_{nij} are defined in (4.2) *(see* (2.5) *for the definition of ξ_{nij}.)*

Then

$$\lim_{n \to \infty} \int_0^T \|Z_n(t)\|_{G_n}^2 dt = 0 \quad \text{in probability.} \tag{4.12}$$

Proof. Let

$$c_{nij} := \int_0^T a_{nij}(t)a_{nik}(t)dt, \quad i,j \in [n], \tag{4.13}$$

and note that by (4.10),

$$|c_{nij}| \le C_2^2 T \quad i,j \in [n], \tag{4.14}$$

uniformly in n. Using (4.2), (4.13), and (4.14) we have

$$\begin{aligned}
\mathbb{E} \int_0^T Z_{ni}^2(t)dt &= n^{-2}\mathbb{E} \left(\sum_{k,j=1}^n \int_0^T a_{nij}(t)a_{nik}(t)dt \, (\eta_{nij} - V_{nij})\,(\eta_{nik} - V_{nik}) \right) \\
&= n^{-2}\mathbb{E} \sum_{j=1}^n c_{nij}^2 \mathbb{E}\, (\eta_{nij} - V_{nij})^2 = n^{-2}\mathbb{E} \sum_{j=1}^n c_{nij}^2 (\mathbb{E}\, \eta_{nij}^2 - V_{nij}^2) \\
&\le C_2^4 T^2 \rho_n^{-1} n^{-2} G_{ni}^{-1} \sum_{j=1}^n V_{nij} \le C_2^4 T^2 \rho_n^{-1} n^{-2} G_{ni}^{-1} \sum_{j=1}^n V_{nij} \\
&= C_2^4 T^2 (n\rho_n)^{-1} G_{ni}^{-1}.
\end{aligned} \tag{4.15}$$

By Markov's inequality, for arbitrary $\epsilon > 0$,

$$\mathbb{P}\left(\int_0^T \|Z_n(t)\|_{G_n}^2 dt \ge \epsilon \right) = \mathbb{P}\left(n^{-1} \sum_{i=1}^n G_{ni} \int_0^T Z_{ni}^2 dt \ge \epsilon \right) \le C_2^2 T (\epsilon \rho_n n)^{-1} \to 0 \tag{4.16}$$

as $n \to \infty$, because $\rho_n n \to \infty$. This proves the lemma. ∎

Proof of Theorem 4.1. Denote $\phi_{ni} = u_{ni} - v_{ni}$. By subtracting (2.26) from (2.24), multiplying the result by $n^{-1}G_{ni}\phi_{ni}$, and summing over $i \in [n]$, we obtain

$$\begin{aligned}
2^{-1}\frac{d}{dt}\|\phi_n\|_{G_n}^2 = n^{-2} &\sum_{i,j=1}^n \bar{W}_{nij}(\phi_{nj} - \phi_{ni})\phi_{ni} \\
&+ n^{-2} \sum_{i,j=1}^n G_{ni}(\eta_{nij} - V_{nij})(u_{nj} - u_{ni})\phi_{ni} \\
&+ n^{-1} \sum_{i=1}^n G_{ni}\left[f(u_{ni}) - f(v_{ni})\right]\phi_{ni}.
\end{aligned} \tag{4.17}$$

By Lemma 3.3, the first term on the right-hand side of (4.17) is nonpositive:

$$\sum_{i,j=1}^n \bar{W}_{nij}(\phi_{nj} - \phi_{ni})\phi_{ni} = -2^{-1} \sum_{j=1}^n \bar{W}_{nij}(\phi_{nj} - \phi_{ni})^2 \le 0. \tag{4.18}$$

Thus, using (4.18) and (2.25), from (4.17) we have

$$2^{-1}\frac{d}{dt}\|\phi_n\|_{G_n}^2 \le n^{-2} \sum_{i,j=1}^n G_{ni}(\eta_{nij} - V_{nij})(u_{nj} - u_{ni})\phi_{ni} + L_f\|\phi_n\|_{G_n}^2. \tag{4.19}$$

Further, let

(4.20)

$$a_{nij}(t) := u_{nj}(t) - u_{ni}(t),\ Z_{ni}(t) := n^{-1}\sum_{j=1}^{n} a_{nij}(t)(\eta_{nij} - V_{nij}),\ (i,j) \in [n]^2,\ t \in [0,T].$$

From Theorem 3.4 and (4.7), we have

$$\max_{t\in[0,T]} \max_{(i,j)\in[n]^2} |a_{nij}(t)| \leq C_3 \quad \forall n.$$

Thus, $Z_{ni}(t),\ i \in [n]$, in (4.20) satisfy the assumptions of Lemma 4.2.

Using $ab \leq 2^{-1}(a^2+b^2)$, we rewrite the first term on the right-hand side of (4.19):

$$\left| n^{-2} \sum_{i,j=1}^{n} G_{ni}(\eta_{nij} - V_{nij})(u_{nj} - u_{ni})\phi_{ni} \right| \leq 2^{-1}(\|Z_n\|^2_{G_n} + \|\phi_n\|^2_{G_n}). \tag{4.21}$$

Using (4.21), from (4.19) we obtain

$$\frac{d}{dt}\|\phi_n\|^2_{G_n} \leq (2L_f + 1)\|\phi_n\|^2_{G_n} + \|Z_n\|^2_{G_n}. \tag{4.22}$$

By Gronwall's inequality,

$$\sup_{t\in[0,T]} \|\phi_n(t)\|^2_{G_n} \leq e^{(2L_f+1)T}\left(\|\phi_n(0)\|^2_{G_n} + \int_0^T \|Z_n(t)\|^2_{G_n} dt \right). \tag{4.23}$$

Thus, by Lemma 4.2 and (4.8), the right-hand side of (4.23) tends to 0 as $n \to \infty$ in probability. □

5. The continuum limit. Having justified averaging in (2.26), our next goal is to show that the IVP for the averaged equation (2.26) can be approximated by that for the continuum limit (2.32), (2.33), subject to the initial condition

$$u(x,0) = g(x), \tag{5.1}$$

where $g \in L^\infty(I)$. To compare the solutions of the discrete problem (2.26) and continuous equation (2.32) we supply the former problem with the initial condition that is consistent with (5.1):

$$v_{ni}(0) = n\int_{I_{ni}} g(x)dx,\ i \in [n]. \tag{5.2}$$

Below, we construct a finite-dimensional Galerkin approximation of (2.32) and (5.1) and prove its convergence. In the next section, we compare solutions obtained by the Galerkin scheme with the solutions of the IVP for (2.26).

Throughout this section, we assume that conditions (W-1)–(W-3) and (2.25) hold.

5.1. The Galerkin problem. Let $X = L^2(I)$, define $K : X \to X$ by

$$[K(u)](x) = \int_I U(x,y)(u(y) - u(x))dy, \tag{5.3}$$

and rewrite (2.32) as follows:

$$\mathbf{u}' = K(\mathbf{u}) + f(\mathbf{u}), \tag{5.4}$$

$$\mathbf{u}(0) = g. \tag{5.5}$$

Recall that $\mathbf{u} : \mathbb{R} \to X$ stands for the vector-valued function defined by $[\mathbf{u}(t)](x) = u(x,t)$ for each $t \in \mathbb{R}$.

DEFINITION 5.1. *Function* $\mathbf{u} \in H^1(0,T;X)$ *is called a weak solution of the IVP* (5.4), (5.5) *on* $[0,T]$ *if*

$$(\mathbf{u}'(t) - K(\mathbf{u}(t)) - f(\mathbf{u}(t)), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in X \tag{5.6}$$

almost everywhere (a.e.) on $[0,T]$ *and* $\mathbf{u}(0) = g$.

To construct a finite-dimensional problem approximating (5.6), we introduce $X_n = \operatorname{span}\{\phi_{ni} : i \in [n]\}$, a linear subspace of X. Here,

$$\phi_{ni}(x) = \mathbf{1}_{I_{ni}}(x) = \begin{cases} 1, & x \in I_{ni}, \\ 0, & x \notin I_{ni}, \end{cases} \quad i \in [n]. \tag{5.7}$$

Next, we construct the Galerkin approximation of the solution of (5.4), (5.5). To this end, we fix $n \in \mathbb{N}$ and look for the approximate solution in the form

$$\mathbf{u}_n(t) = \sum_{i=1}^{n} u_{ni}(t)\phi_{ni}. \tag{5.8}$$

The differentiable coefficients $u_{ni}(t),\ i \in [n]$, are determined by projecting the original equation and the initial condition onto X_n:

$$(\mathbf{u}_n'(t) - K(\mathbf{u}_n(t)) - f(\mathbf{u}_n(t)), \phi) = 0 \quad \forall \phi \in X_n, \tag{5.9}$$

$$\mathbf{u}_n(0) = P_{X_n} g = \sum_{i=1}^{n} \frac{(g, \phi_{ni})}{(\phi_{ni}, \phi_{ni})}\phi_{ni}, \tag{5.10}$$

where $P_{X_n} : X \to X_n$ stands for the orthogonal projector onto X_n. After plugging (5.8) into (5.9) and setting $\mathbf{v} = \phi_{ni},\ i \in [n]$, we arrive at the following IVP for the unknown coefficients $u_{ni}(t),\ i \in [n]$:

$$\dot{u}_{ni}(t) = n^{-1} \sum_{j=1}^{n} U_{nij} \left(u_{nj}(t) - u_{ni}(t)\right) + f(u_{ni}), \tag{5.11}$$

$$u_{ni}(0) = \frac{(g, \phi_{ni})}{(\phi_{ni}, \phi_{ni})} = n \int_{I_{ni}} g(x)dx. \tag{5.12}$$

Here,

$$U_{nij} = n^2 \int_{I_{ni} \times I_{nj}} U(x,y)dxdy = n^2 \int_{I_{ni} \times I_{nj}} \frac{W(x,y)}{\int_I W(x,z)dz} dxdy \le n. \tag{5.13}$$

Note that the right-hand side of (5.11) is uniformly Lipschitz continuous, which guarantees the existence of a unique solution of the IVP (5.11), (5.12) on $\mathbb{R}$.

It will be convenient to have the Galerkin equation (5.11) rewritten as the integral equation

$$\frac{\partial}{\partial t} u_n(x,t) = \int_I U_n(x,y) \left(u_n(y,t) - u_n(x,t)\right) dy + f(u_n(x,t)), \tag{5.14}$$

where U_n and u_n are step functions

$$U_n(x,y) = \sum_{i,j=1}^{n} U_{nij} \mathbf{1}_{I_{ni} \times I_{nj}}(x,y), \tag{5.15}$$

$$u_n(x,t) = \sum_{i=1}^{n} u_{ni}(t) \mathbf{1}_{I_{ni}}(x).$$

5.2. Convergence of the Galerkin scheme. In this section, we show that the solutions of the Galerkin problems (5.9), (5.10), $\mathbf{u}_n$, converge to $\mathbf{u}$, a unique weak solution of (5.4), (5.5), in the $L^2(0,T;X)$ norm as $n \to \infty$.

THEOREM 5.2. *For any $T > 0$, there is a unique weak solution of* (5.4), (5.5), $\mathbf{u} \in H^1(0,T;X)$. *The solutions of the Galerkin problems* (5.9), (5.10), $\mathbf{u}_n$, *converge to $\mathbf{u}$ in the $L^2(0,T;X)$ norm as $n \to \infty$.*

Proof.

1. We shall first establish the following bounds for the solutions $\mathbf{u}_n$ of the Galerkin problem (5.9), (5.10) that hold uniformly in n:

$$\exists C_4 = C_4(T, \|\mathbf{u}(0)\|_{L^\infty(I)}) : \quad \max\{\|\mathbf{u}_n\|_{C(0,T;L^\infty(I))}, \|\mathbf{u}_n\|_{C(0,T;X)}, \|\mathbf{u}_n'\|_{C(0,T;X)}\} \le C_4. \tag{5.16}$$

The L^∞-bound and, therefore, the X-bound follow from Theorem 3.2. These bounds are uniform in n, because

$$\|\mathbf{u}_n(0)\|_{L^\infty(I)} = \|P_{X_n} g\|_{L^\infty(I)} \le \|g\|_{L^\infty(I)}.$$

To bound $\|\mathbf{u_n'}\|_{C(0,T;X)}$ we proceed as follows:

$$|(\mathbf{u_n'}(t), \mathbf{v})| \le \int_I \bar{U}_n(x,y)|u_n(x,t) - u_n(y,t)||v(x,t)|dxdy + \int_I |f(u_n(x))||v(x)|dx.$$

Using the L^∞-bound for $\mathbf{u}_n$ (5.16), the continuity of f, $\|\bar{U}_n\|_{L^2(I^2)} \le \|U\|_{L^2(I^2)}$, and the triangle and Cauchy–Schwarz inequalities, we obtain

$$|(\mathbf{u_n'}(t), \mathbf{v})| \le C_5(\|U\|_{L^2(I^2)} + C_6)\|\mathbf{v}\| \;\forall \mathbf{v} \in X.$$

Thus,

$$\|\mathbf{u_n'}(t)\| \le C_6, t \ge 0 \tag{5.17}$$

uniformly in n.

2. Estimates in (5.16) imply

$$\|\mathbf{u}_n\|_{L^2(0,T;X)} \le C_4, \tag{5.18}$$

$$\|\mathbf{u}_n(t+h) - \mathbf{u}_n(t)\|_X \le C_4|h|, \tag{5.19}$$

respectively. From (5.19), we further have

$$\int_0^T \|\mathbf{u}_n(t+h) - \mathbf{u}_n(t)\|_X^2 dt \le C_4^2 T h^2. \tag{5.20}$$

From (5.18) and (5.20), using the Frechet–Kolmogorov theorem (cf. [19]), we see that $(\mathbf{u}_n)$ is precompact in $L^2(0,T;X)$. Thus, one can select a subsequence $(\mathbf{u}_{n_k})$ that converges to $\mathbf{u} \in L^2(0,T;X)$.

3. Likewise, integrating both sides of (5.17) from 0 to T, we obtain

$$\|\mathbf{u_n'}\|_{L^2(0,T;X)} \le C_6\sqrt{T}$$

uniformly in n. Thus, $(\mathbf{u}'_{n_k})$ is weakly precompact in $L^2(0,T;X)$, and one can select a subsequence $(\mathbf{u}'_{n_{k'}})$ that weakly converges to $\mathbf{w} \in L^2(0,T;X)$ and strongly converges to $\mathbf{u}' \in L^2(0,T;X)$. Clearly, $\mathbf{w} = \mathbf{u}'$. Indeed, taking $\phi \in C^1(0,T;X)$ with compact support in $(0,T)$ and using integration by parts, we obtain

$$\int_0^T \mathbf{u}'_{n_{k'}}(t)\phi(t)dt = -\int_0^T \mathbf{u}_{n_{k'}}(t)\phi'(t)dt. \tag{5.21}$$

By sending $k' \to \infty$ in (5.21), we see that $\mathbf{u}' \in H^1(0,T;X)$ and $\mathbf{u}' = \mathbf{w}$.

4. Next, we show that $\mathbf{u}$ is a unique weak solution of (5.6) satisfying $\mathbf{u}(0) = \mathbf{g}$. This follows from a standard argument (see, e.g., [9, Theorem 7.1.3]).

   Fix $N \in \mathbb{N}$ and choose a function of the form

$$\mathbf{v}(t) = \sum_{j=1}^{N} d_j(t)\phi_{\mathbf{nj}}, \tag{5.22}$$

where $d_j(t)$ are continuously differentiable functions and the ϕ_{nj} are defined in (5.7). Next, we multiply (5.9) with $n > N$ and $\phi := \phi_{nj}$ by $d_j(t)$, sum over j, and integrate the result from 0 to T to obtain

$$\int_0^T (\mathbf{u}'_n(t) - K(\mathbf{u}_n(t)) - f(\mathbf{u}_n(t)), \mathbf{v}(t))dt = 0.$$

Passing to the limit along $n = n_k$, we have

$$\int_0^T (\mathbf{u}'(t) - K(\mathbf{u}(t)) - f(\mathbf{u}(t)), \mathbf{v}(t))dt = 0. \tag{5.23}$$

This equality holds for an arbitrary $\mathbf{v}$ as in (5.22). Since functions of this form for $N \in \mathbb{N}$ are dense in $L^2(0,T;X)$, we conclude that (5.23) holds for all $\mathbf{v} \in L^2(0,T;X)$. Therefore,

$$(\mathbf{u}' - K(\mathbf{u}) - f(\mathbf{u}), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in L^2(0,T;X) \tag{5.24}$$

a.e. on $[0,T]$.

5. To show that $\mathbf{u}$ is a weak solution of (2.32), (5.1), it remains to verify $\mathbf{u}(0) = \mathbf{g}$. To this end, we choose $\mathbf{v} \in C^1(0,T;X)$ vanishing at $t = T$ as a test function in (5.6) and integrate by parts to obtain

$$-\int_0^T (\mathbf{u}(t), \mathbf{v}'(t))\, dt = \int_0^T (K(\mathbf{u}(t)) + f(\mathbf{u}(t)), \mathbf{v}(t))\, dt + (\mathbf{u}(0), \mathbf{v}(0)). \tag{5.25}$$

Using the same test functions in (5.9), we have

$$-\int_0^T (\mathbf{u}_{n_k}(t), \mathbf{v}'(t))\, dt = (K(\mathbf{u}_{n_k}(t)) + f(\mathbf{u}_{n_k}(t)), \mathbf{v}(t))\, dt + (\mathbf{u}_{n_k}(0), \mathbf{v}(0)). \tag{5.26}$$

Passing to the limit in (5.26) yields

$$-\int_0^T (\mathbf{u}(t), \mathbf{v}'(t))\, dt = \int_0^T (K(\mathbf{u}(t)) + f(\mathbf{u}(t)), \mathbf{v}(t))\, dt + (\mathbf{g}, \mathbf{v}(0)). \tag{5.27}$$

Comparing the limiting equation (5.27) with (5.25) we conclude that $\mathbf{u}(0) = \mathbf{g}$. Thus, $\mathbf{u}$ is a weak solution of (5.4).

6. To show that the just constructed weak solution is unique, suppose that there is another solution

$$\mathbf{w}' = K(\mathbf{w}) + f(\mathbf{w}) \tag{5.28}$$

satisfying the same initial condition $\mathbf{w}(0) = \mathbf{g}$. Denote $\xi = \mathbf{u} - \mathbf{w}$. By subtracting (5.28) from (5.4), multiplying both sides by $G(x)\xi$, and integrating over I, we have

$$\frac{1}{2}\frac{d}{dt}\|\sqrt{G}\xi(\cdot,t)\|_X^2 = \int_{I^2} W(x,y)\left(\xi(y,t) - \xi(x,t)\right)\xi(x,t)dydx + \int_I G(x)\left(f(u(x,t)) - f(w(x,t)\right)\xi(x,t)dx.$$

Using Lemma 3.3 and Lipschitz continuity of f, we obtain

$$\frac{d}{dt}\|\sqrt{G}\xi(\cdot,t)\|_X^2 \le L\|\sqrt{G}\xi(\cdot,t)\|_X^2.$$

Since G is strictly positive on I (cf. (2.9)), from the last inequality and $\xi(0) = 0$, we conclude that $\mathbf{u}(t) = \mathbf{w}(t)$ for all $t \in [0,T]$. This proves uniqueness.

7. The uniqueness of the weak solution entails $\mathbf{u}_n \to \mathbf{u}$ as $n \to \infty$. Indeed, suppose on the contrary that there exists a subsequence $\mathbf{u}_{n_l}$, which is not converging to $\mathbf{u}$. Then for a given $\epsilon > 0$ one can select a subsequence $\mathbf{u}_{n_{l_i}}$ such that

$$\|\mathbf{u}_{n_{l_i}} - \mathbf{u}\|_{L^2(0,T;X)} > \epsilon \;\forall i \in \mathbb{N}.$$

However, $(\mathbf{u}_{n_{l_i}})$ is precompact in $L^2(0,T,X)$ and contains a subsequence converging to a weak solution of (5.4), which must be $\mathbf{u}$ by uniqueness. This is a contradiction. ∎

5.3. Approximation. It remains to estimate the difference between the solutions of the averaged equation (2.26) and that of the Galerkin problem (5.11). The key is the estimate for the L^4-norm of the difference between V_n and U_n (see (2.31) and (5.15)), the kernels used in the averaged and the Galerkin problems, respectively.

LEMMA 5.3.

$$\|U_n - V_n\|_{L^4(I^2)} \to 0 \quad \text{as } n \to \infty. \tag{5.29}$$

Proof. First, we show that $U_n, n \in \mathbb{N}$, form a sequence of L^4-bounded martingales [18]. To this end, we consider a probability space $(I^2, \mathcal{B}(I^2), \lambda)$ with I^2 as a sample space equipped with the σ–algebra of Borel sets, and the Lebesgue measure as probability. Let $\mathcal{A}_n$ denote the algebra of subsets of I^2 generated by the sets $I_{ni} \times I_{nj}$, $(i,j) \in [n]^2$. Then U_n can be represented as the conditional expectation

$$U_n = \mathbb{E}\,(U|\mathcal{A}_n), \quad n \in \mathbb{N}.$$

Since $U \in L^4(I^2)$ (cf. (W-4) and (2.9)), the L^p–martingale convergence theorem yields

$$U_n \to U \quad \text{a.e. and in } L^4(I^2) \text{ as } n \to \infty. \tag{5.30}$$

Next, we turn to functions $V_n, n \in \mathbb{N}$ (cf. (2.31)):

$$
\begin{aligned}
V_n(x,y) &= \sum_{i,j=1}^n \frac{\bar W_n(x_{ni},x_{nj})}{n^{-1}\sum_{k=1}^n \bar W_n(x_{ni},x_{nk})} \mathbf{1}_{I_{ni}\times I_{nj}}(x,y) \\
&= \frac{\sum_{i,j=1}^n \bar W_n(x_{ni},x_{nj}) \mathbf{1}_{I_{ni}\times I_{nj}}(x,y)}{\sum_{i=1}^n n^{-1}\sum_{k=1}^n \bar W_n(x_{ni},x_{nk}) \mathbf{1}_{I_{ni}}(x)} \\
&=: \frac{P_n(x,y)}{Q_n(x)}. \qquad (5.31)
\end{aligned}
$$

From (2.3) and (W-1), we have $P_n \to W$ a.e. on I^2. Likewise, by (2.10),

$$Q_n = \int_I W(\cdot, z) dz (1+\delta_n) \geq \nu > 0 \text{ as } n \to \infty$$

uniformly on any closed interval lying in $(0,1)$. Thus, $\frac{P_n}{Q_n} \to U$ a.e. on I^2 as $n \to \infty$. Furthermore, by (2.3) and (W-3),

$$\left| \frac{P_n}{Q_n} \right| \leq \frac{W}{\nu}.$$

Since $V_n, U \geq 0$,

$$(V_n - U)^4 \leq 4\left(V_n^4 + U^4\right) \leq \frac{4W^2}{\nu^4}.$$

Thus, $V_n - U \to 0$ a.e. on I^2 and

$$|V_n - U| \leq \sqrt{2}\nu^{-1} W \in L^4(I^2).$$

By the dominated convergence theorem $V_n \to U$ in $L^4(I^2)$. From this and (5.30), we conclude

$$\|U_n - V_n\|_{L^4(I^2)} \to 0, \quad n \to \infty. \qquad \square$$

LEMMA 5.4. *For any $T > 0$, solutions of the IVPs for* (2.26) *and* (5.11) *satisfy*

$$\lim_{n\to\infty} \max_{t\in[0,T]} \|u_n(t) - v_n(t)\|_{G_n} = 0, \tag{5.32}$$

provided

$$\lim_{n\to\infty} \|u_n(0) - v_n(0)\|_{G_n} = 0. \tag{5.33}$$

Proof. Denote $\phi_{ni} := u_{ni} - v_{ni},\ i \in [n]$. By subtracting (2.26) from (5.11), multiplying the result by $n^{-1} G_{ni}\phi_{ni}$ (see (4.2) for the definition of G_{ni}), and summing over $i \in [n]$, we obtain

$$
\begin{aligned}
&2^{-1}\frac{d}{dt}\|\phi_n\|_{G_n}^2 \\
&\quad = n^{-2}\sum_{i,j=1}^n \bar W_{nij}(\phi_{nj} - \phi_{ni})\phi_{ni} + n^{-2}\sum_{i,j=1}^n G_{ni}(\bar U_{nij} - V_{nij})(u_{nj} - u_{ni})\phi_{ni} \\
&\quad + n^{-1}\sum_{i=1}^n G_{ni}\left[f(u_{ni}) - f(v_{ni})\right]\phi_{ni}. \qquad (5.34)
\end{aligned}
$$

As before, we use Lemma 3.5 and (2.25) to obtain

$$n^{-2}\sum_{i,j=1}^{n}\bar{W}_{nij}(\phi_{nj}-\phi_{ni})\phi_{ni}\le 0, \tag{5.35}$$

$$\left|n^{-1}\sum_{i=1}^{n}G_{ni}\left[f(u_{ni})-f(v_{ni})\right]\phi_{ni}\right|\le L_f\|\phi_n\|^2_{G_n}. \tag{5.36}$$

Using $\max_{t\in[0,T]}\|u_n(t)\|_{\infty,n}\le C_7$ (cf. Theorem 3.4), we estimate

$$\left|n^{-2}\sum_{i,j=1}^{n}G_{ni}(U_{nij}-V_{nij})(u_{nj}-u_{ni})\phi_{ni}\right|\le C_8\left(\Delta_n(W)+\|\phi_n\|^2_{G_n}\right),$$

where

$$\Delta_n(W):=\frac{1}{n^2}\sum_{i,j=1}^{n}G_{ni}(U_{nij}-V_{nij})^2.$$

Further, using the Cauchy–Schwarz inequality, we have

$$\Delta_n(W)\le\left(n^{-1}\sum_{i=1}^{n}G_{ni}^2\right)^{1/2}\|U_n-V_n\|^2_{L^4(I^2)}. \tag{5.37}$$

Recalling the definition of G_{ni} and using the Cauchy–Schwarz inequality again, we obtain

$$n^{-1}\sum_{i=1}^{n}G_{ni}^2=n^{-1}\sum_{i=1}^{n}\left(n^{-1}\sum_{j=1}^{n}\bar{W}_{nij}\right)^2\le n^{-2}\sum_{i,j=1}^{n}\bar{W}_{nij}^2. \tag{5.38}$$

Using (2.11), (5.38), and Lemma 5.4, we obtain

$$\lim_{n\to\infty}\Delta_n(W)=0. \tag{5.39}$$

The combination of (5.34)–(5.37) yields

$$\frac{d}{dt}\|\phi_n\|^2_{G_n}\le 2(C_8+L_f)\|\phi_n\|^2_{G_n}+2C_7\Delta_n(W).$$

By Gronwall's inequality,

$$\max_{t\in[0,T]}\|\phi_n(t)\|^2_{G_n}\le\left(\|\phi_n(0)\|^2_{2,n}+\frac{C_8}{C_8+L_f}\Delta_n(W)\right)\exp\{(C_8+L_f)T\}. \tag{5.40}$$

The right-hand side in (5.40) tends to 0 as $n\to\infty$, as follows from (5.33) and (5.39). This proves the lemma. ∎

Theorem 2.6 now follows from Theorems 4.1 and 5.2 and Lemma 5.4.

6. Discussion. The analysis in the preceding sections justifies the continuum limit (2.32) for the semilinear heat equation (2.24) on sparse W–random graphs. In conclusion, we outline several extensions of this work to certain nonlinear models, which are of interest in applications.

6.1. The nonlinear model. The analysis in sections 4 and 5 can be extended to cover the following nonlinear heat equation on $\Gamma_n = \langle V(\Gamma_n), E(\Gamma_n)\rangle$:

$$\dot{u}_{ni} = \frac{1}{\deg_{\Gamma_n}(i)} \sum_{j:\{i,j\}\in E(\Gamma_n)} D(u_{nj} - u_{ni}) + f(u_{ni}), \tag{6.1}$$

where D and f are Lipschitz continuous functions (cf. (3.2), (2.25)). In addition, we assume that D is an odd function satisfying the sign condition

$$uD(u) \geq 0. \tag{6.2}$$

Both conditions hold for the original Kuramoto model with $D(u) = \sin u$ [11] as well as for the models of opinion dynamics where $D(u) = \phi(|u|)u$ for an appropriate influence function ϕ [17].

Under the above assumptions on f and D, we can justify the continuum limit for (6.1).

THEOREM 6.1. *Let $g \in L^\infty(I)$ and $T > 0$ be arbitrary. Denote the solutions of* (6.1) *and* (1.4) *subject to the initial conditions* (2.35) *and* (2.34) *by $u_{ni}(t)$, $i \in [n]$, and $u(x,t)$, respectively.*

Then

$$\lim_{n\to 0} \int_0^T \|u_n(\cdot,t) - u(\cdot,t)\|^2_{L^2(I)} dt = 0 \quad \text{in probability},$$

where

$$u_n(x,t) := \sum_{i=1}^n u_{ni}(t)\mathbf{1}_{I_{ni}}(x).$$

For the proof of Theorem 6.1, one needs the following modification of Lemma 3.3.

LEMMA 6.2. *Let $W \in L^2(I)$ be a symmetric function and D be an odd symmetric continuous function. Then for any $u \in L^\infty(I)$,*

$$\int_{I^2} W(x,y)D(u(y)-u(x))u(x)dxdy$$
$$= -2^{-1} \int_{I^2} W(x,y)D(u(y)-u(x))(u(y)-u(x))dxdy.$$

If, in addition, $W \geq 0$ and D satisfies (6.2)*, then*

$$\int_{I^2} W(x,y)D(u(y)-u(x))u(x)dxdy \leq 0.$$

With Lemma 6.2 in hand, the proofs of the statements in sections 4 and 5 can be translated to the nonlinear equation (6.1) with minor changes.

6.2. An alternative scaling and other graph models. If the diffusion term is scaled by $n\rho_n$ instead of $d_{ni} = O(n\rho_n)$ as in (2.24), the formal derivation of the continuum limit yields

$$\frac{\partial}{\partial t} u(x,t) = \int_I W(x,y) D\left(u(y,t) - u(x,t)\right) dy + f(u(x,t)). \tag{6.3}$$

Here, the kernel is W instead of U (cf. (2.33)). In particular, for the Kuramoto model on the power law family of graphs, the alternative scaling yields

$$\frac{\partial}{\partial t} u(x,t) = x^{-\alpha} \int_I y^{-\alpha} \sin\left(u(y,t) - u(x,t)\right) dy. \tag{6.4}$$

The presence of the x-dependent factor on the right-hand side of (6.4) has interesting implications for the spatial patterns generated by the Kuramoto model. In particular, it is responsible for the existence of the chimera-like patterns in the Kuramoto model with repulsive coupling on power law graphs (cf. [16]).

The proof of existence of the strong solution of the IVP in section 3 does not cover (6.3), because it relies on condition (3.3), which does not hold for W in general (see (3.9)). However, one can show the existence of the weak solution for the IVP for (6.3) (cf. Definition 5.1) by constructing it as the limit of solutions of the Galerkin problems following the lines of the analysis in section 5.2.

Likewise, there are many different ways to define a convergent family of sparse random graphs. Instead of (2.2) one could, for example, define the probability for a given pair of nodes to belong to the edge set using averaging:

$$\mathbb{P}\ (\{i,j\} \in E(\Gamma_n)) = \rho_n n \int_{I_{ni}\times I_{nj}} W(x,y)dxdy. \tag{6.5}$$

The analysis of this paper can be used to justify the continuum limit for coupled systems on $\{\Gamma_n\}$ defined by (6.5).

REFERENCES

[1] A.-L. Barabási and A. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–512.
[2] N. Biggs, *Algebraic Graph Theory*, 2nd ed., Cambridge University Press, Cambridge, 1993.
[3] V. I. Bogachev, *Measure theory*, Vol. I, Springer, Berlin, 2007.
[4] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao, *An L^p Theory of Sparse Graph Convergence* I: *Limits, Sparse Random Graph Models, and Power Law Distributions*, preprint, arXiv:1401.2906, 2014.
[5] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao, *An L^p Theory of Sparse Graph Convergence* II: *LD Convergence, Quotients, and Right Convergence*, preprint, arXiv:1408.0744, 2014.
[6] M. C. Cross and P. C. Hohenberg, *Pattern formation out of equilibrium*, Rev. Modern Phys., 65 (1993), pp. 851–1112.
[7] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52 (2007), pp. 852–862.
[8] F. Dörfler and F. Bullo, *Synchronization and transient stability in power networks and nonuniform Kuramoto oscillators*, SIAM J. Control Optim., 50 (2012), pp. 1616–1642.
[9] L. C. Evans, *Partial Differential Equations*, AMS, Providence, RI, 2010.
[10] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math. 840, Springer, Berlin, 1981.
[11] Y. Kuramoto, *Cooperative dynamics of oscillator community*, Progr. Theor. Phys. Suppl. 79, (1984), pp. 223–240.
[12] L. Lovász, *Large Networks and Graph Limits*, AMS, Providence, RI, 2012.
[13] G. S. Medvedev, *Stochastic stability of continuous time consensus protocols*, SIAM J. Control Optim., 50 (2012), pp. 1859–1885.
[14] G. S. Medvedev, *The nonlinear heat equation on dense graphs and graph limits*, SIAM J. Math. Anal., 46 (2014), pp. 2743–2766.
[15] G. S. Medvedev, *The nonlinear heat equation on W-random graphs*, Arch. Ration. Mech. Anal., 212 (2014), pp. 781–803.
[16] G. S. Medvedev and X. Tang, *The Kuramoto Model on Power Law Graphs*, manuscript.
[17] S. Motsch and E. Tadmor, *Heterophilious dynamics enhances consensus*, SIAM Rev., 56 (2014), pp. 577–621.
[18] D. Williams, *Probability with martingales*, Cambridge Math. Textb. Cambridge University Press, Cambridge, 1991.
[19] K. Yosida, *Functional analysis*, Classics Math., Springer, Berlin, 1995.