

SPARSE MONTE CARLO METHOD FOR NONLOCAL DIFFUSION PROBLEMS*

DMITRY KALIUZHNYI-VERBOVETSKYI[†] AND GEORGI S. MEDVEDEV[‡]

Abstract. A class of evolution equations with nonlocal diffusion is considered in this work. These are integro-differential equations arising as models of propagation phenomena in continuum media with nonlocal interactions including neural tissue, porous media, peridynamics, and models with fractional diffusion, as well as continuum limits of interacting dynamical systems. The principal challenge of numerical integration of nonlocal systems stems from the lack of spatial regularity of the data and solutions intrinsic to nonlocal models. To overcome this problem we propose a semidiscrete numerical scheme based on the combination of sparse Monte Carlo and discontinuous Galerkin methods. Our method requires minimal assumptions on the regularity of the data. In particular, the kernel of the nonlocal diffusivity is assumed to be a square integrable function and may be singular or discontinuous. An important feature of our method is the use of sparsity. Sparse sampling of points in the Monte Carlo approximation of the nonlocal term allows us to use fewer discretization points without compromising the accuracy. For kernels with singularities, more points are selected automatically in the regions near the singularities. We prove convergence of the numerical method and estimate the rate of convergence. There are two principal ingredients in the error of the numerical method related to the use of Monte Carlo and Galerkin approximations, respectively. We analyze both errors. Two representative examples of discontinuous kernels are presented. The first example features a kernel with a singularity, while the kernel in the second example experiences jump discontinuity. We show how the information about the singularity in the former case and the geometry of the discontinuity set in the latter translate into the rate of convergence of the numerical procedure. In addition, we illustrate the rate of convergence estimate with a numerical example of an initial value problem, for which an explicit analytic solution is available. Numerical results are consistent with analytical estimates.

Key words. nonlocal diffusion equation, initial value problem, Monte Carlo method, Galerkin method, convergence analysis

MSC codes. 65M75, 65M60, 65M15, 60H15, 34C15

DOI. 10.1137/19M1308657

1. Introduction. We propose a numerical method for the initial value problem (IVP) for a nonlinear heat equation with nonlocal diffusion,

$$(1.1) \quad \partial_t u(t, x) = f(u, x, t) + \int W(x, y) D(u(t, y) - u(t, x)) dy, \quad x \in Q \subset \mathbb{R}^d,$$

$$(1.2) \quad u(0, x) = g(x).$$

For analytical convenience, we take $Q = [0, 1]^d$ as a spatial domain. Throughout this paper, when the domain of integration is not specified, it is assumed to be Q . Further,

*Received by the editors December 23, 2019; accepted for publication (in revised form) July 18, 2022; published electronically November 1, 2022.

<https://doi.org/10.1137/19M1308657>

Funding: The work of the second author was partially supported by National Science Foundation grants DMS-1715161 and DMS-2009233.

[†]Department of Higher Mathematics and Statistics, South-Ukrainian National Pedagogical University named after K. D. Ushinsky, Staroportofrankivska St., 26, Odesa, Ukraine, 65020 (kaliuzhnyi.DS@pdu.edu.ua).

[‡]Department of Mathematics, Drexel University, Philadelphia, PA 19104 USA (medvedev@drexel.edu).

$g \in L^2(Q)$, $W \in L^2(Q^2)$, f is a bounded measurable function on $\mathbb{R} \times Q \times \mathbb{R}^+$, which is Lipschitz continuous in u , continuous in t , and integrable in x , and D is a Lipschitz continuous function on $\mathbb{R}$:

$$(1.3) \quad |D(u_1) - D(u_2)| \leq L_D |u_1 - u_2| \quad \text{and} \quad |f(u_1, x, t) - f(u_2, x, t)| \leq L_f |u_1 - u_2|$$

for all $(x, t) \in Q \times \mathbb{R}$. Throughout this paper we assume

$$(1.4) \quad \sup_{u \in \mathbb{R}} |D(u)| \leq 1.$$

This assumption may be dropped if an a priori estimate on $\|u\|_{C(0,T;L^\infty(Q))}$ for $T > 0$ is available. Here and below, $T > 0$ is an arbitrary but fixed time horizon. Furthermore, the analysis below applies to models with the interaction function of a more general form $D(u_1, u_2)$ provided

$$(1.5) \quad |D(u_1, v_1) - D(u_2, v_2)| \leq L_D (|u_1 - u_2| + |v_1 - v_2|) \quad \forall u_1, u_2, v_1, v_2 \in \mathbb{R}.$$

However, we keep $D(u_1, u_2) := D(u_2 - u_1)$ to emphasize the connection to diffusion problems.

Equation (1.1) is a nonlocal diffusion problem. It arises as a continuum limit of interacting particle systems [26, 20, 18]. Equations of this form are used for modeling population dynamics [33, 8, 32, 2, 7], neural tissue [9], porous media flows [11, 12], and various other biological and physicochemical processes involving anomalous diffusion [34, 5]. The key distinction of the evolution equations with nonlocal diffusion from their classical counterparts is the lack of smoothening property. A priori the solution of (1.1), (1.2) is a square integrable function in x for all $t > 0$ [26] and it may have no more regularity beyond that, unless the initial data and kernel W are smooth [23, Theorem 3.3]. The lack of smoothness is a serious challenge for constructing numerical schemes for (1.1), (1.2) and for analyzing their convergence. All deterministic quadrature formulas require at least piecewise differentiability for a guaranteed convergence rate. The problem is even more challenging in high dimensional spatial domains. The main idea underlying our approach is to use the Monte Carlo approximation of the nonlocal term in (1.1). We take advantage of the essential feature of the Monte Carlo method: the independence of the convergence rate on the regularity of the integrand. The second key idea is the use of sparsity, which is twofold. First, sparse sampling of points in the Monte Carlo method is used to minimize computation without compromising the accuracy. For W with jump discontinuity across Lipschitz hypersurfaces, sparse sampling of points for evaluating integrals is computationally beneficial for $d \geq 2$. If the discontinuity set has nontrivial fractal dimension, the sparse Monte Carlo method performs better than its dense counterpart for $d = 1$ (cf. Lemma 4.3). Furthermore, sparsity is the key for extending the Monte Carlo method for models with singular kernels (see section 4.1). Not only does it allow us to apply the Monte Carlo method for unbounded functions, it also makes it adaptive: more sample points are selected automatically near the singularities. The combination of these ideas together with the discontinuous Galerkin method yields a numerical scheme for the IVP (1.1), (1.2) that performs well under minimal assumptions on the regularity of W and initial data.

This paper is based on our previous work on the continuum limit of interacting particle systems on convergent graph sequences [23, 24, 19, 26]. The continuum limit is a powerful tool for studying various aspects of network dynamics including existence, stability, and bifurcations of spatiotemporal patterns [35, 25, 27, 28]. Very often the derivation of the continuum limit is based on heuristic considerations and its rigorous

mathematical justification is a nontrivial problem. Recently, motivated by the theory of graph limits [22, 21, 6], we proved convergence to the continuum limit for a broad class of dynamical systems on graphs [23]. Importantly, our proof applies to models on random graphs [24] including sparse random graphs [19, 26]. These results prepared the ground for the numerical method proposed in this paper. There is an intimate relation between the problem of the continuum limit for interacting particle systems and numerical integration of nonlocal diffusion models. Given a continuum model (1.1), one can construct the corresponding particle system, approximating (1.1). This idea was proposed already in [23] but has not been detailed. Further, recent results for the continuum limit of coupled systems on sparse graphs indicate a strong potential of sparse discretization for numerical integration of nonlocal problems. It is the goal of this paper to present these ideas in detail.

In the next section, we present a discretization scheme for (1.1), which can be viewed as an interacting dynamical system on a *sparse random* graph. The structure of the graph is determined by the kernel W , which defines the asymptotic connectivity of the graph sequence parametrized by the size of the graph. In the theory of graph limits, such functions are called graphons [21]. In section 3, we prove convergence of the semidiscrete (discrete in space and continuous in time) approximation of (1.1) and turn to estimating the rate of convergence in sections 4 and 5. There are two main factors contributing to the error of approximation. The first is due to approximating the nonlocal term in (1.1) by a random sum (Monte Carlo method), while the second is due to approximating the kernel and the initial data by piecewise constant functions (discontinuous Galerkin method). The rate of convergence of the sparse Monte Carlo approximation follows from our previous results [26, Theorem 4.1]. Convergence of piecewise constant approximation in the L^2 -norm follows from classical theorems of analysis (cf. the Lebesgue–Besicovitch theorem [16] or the Martingale convergence theorem [36]). However, neither of these theorems provides the rate of convergence. In fact, the example in section 4.2 shows that without additional hypotheses the algebraic convergence may be arbitrarily slow. To this end, we study what determines the rate of convergence of piecewise constant approximations for a square integrable function. For Hölder continuous functions the answer is simple (cf. Lemma 4.1). For discontinuous functions, on the other hand, the answer naturally depends on the type of discontinuity. In section 4, we consider two examples elucidating this issue. The first example is based on a singular (unbounded) graphon. It shows how the information about the singularity translates into the rate of convergence estimate. Here, we also see how to use sparsity to optimize computation. The second example is adapted from [23]. This is a bounded graphon with jump discontinuity (section 4.2). In this case, the convergence rate depends on the geometry of the set of the discontinuity (cf. Lemma 4.3). In light of these examples, in section 5, we perform convergence analysis under general assumptions on W . In section 6, we extend the rate of convergence analysis to the fully discrete scheme. In section 7, we illustrate rate of convergence estimates with a numerical example. Here, we choose a nonlinear problem, which has an explicit solution. This allows us to verify the rate of convergence of the L^2 -error as the discretization step tends to zero. Special attention is paid to the dependence of the convergence rate on sparsity. Finally, in section 8 we present a proof of a technical lemma, Lemma 3.3, which extends the corresponding result in [26] to models in multidimensional domains and affords a wider range of sparsification.

Numerical methods for nonlocal diffusion problems have been the subject of intense research recently due to their increased use in modeling [30, 15, 14, 5, 4, 31]. Compared to the existing results, the contribution of the present work is that our

method applies to problems with nonlinear diffusivity as well as to problems with a more general form of the interaction function (cf. (1.5)). The main focus of this paper is how to deal with models with low regularity of the data. We believe that the combination of the Monte Carlo and discontinuous Galerkin methods provides an effective tool for numerical integration of nonlocal problems under minimal regularity assumptions.

2. The model and its discretization. In this section, we formulate the technical assumptions on the kernel W and describe the numerical scheme for solving the IVP (1.1)–(1.2).

We assume that $W \in L^2(Q^2)$ is subject to the following assumptions:

$$(W-1) \quad \max \left\{ \operatorname{ess\,sup}_{x \in Q} \int |W(x, y)| dy, \operatorname{ess\,sup}_{y \in Q} \int |W(x, y)| dx \right\} \leq W_1.$$

THEOREM 2.1. *Let $W \in L^2(Q^2)$ satisfy (W-1). Then for any $g \in L^2(Q)$ and $T > 0$ there is a unique solution of the IVP (1.1), (1.2) $u \in C^1(0, T; L^2(Q))$.*

Proof. The proof is as in [19, Theorem 3.1] with minor adjustments. $\square$

Next, we note that the kernel in the nonlocal term may be assumed nonnegative. Indeed, by writing $W = W^+ - W^-$ as the difference of its positive and negative parts, one can rewrite (1.1) as

$$(2.1) \quad \begin{aligned} \partial_t u(t, x) = & f(u, x, t) + \int W^+(x, y) D(u(t, y) - u(t, x)) dy \\ & - \int W^-(x, y) D(u(t, y) - u(t, x)) dy, \end{aligned}$$

where the nonlocal term splits into the difference of two terms with nonnegative kernels. Thus, without loss of generality, in the remainder of this paper we will assume

$$(2.2) \quad W \geq 0.$$

We approximate the IVP (1.1), (1.2) by the following system of ordinary differential equations:

$$(2.3) \quad \dot{u}_{n,\bar{i}} = f_{n,\bar{i}}(u_{n,\bar{i}}, t) + (\alpha_n n^d)^{-1} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} D(u_{n,\bar{j}} - u_{n,\bar{i}}),$$

$$(2.4) \quad u_{n,\bar{i}}(0) = g_{n,\bar{i}},$$

where $\bar{i} := (i_1, i_2, \dots, i_d) \in [n]^d$, $[n] := \{1, 2, \dots, n\}$, and

$$(2.5) \quad g_{n,\bar{i}} = n^d \int_{Q_{n,\bar{i}}} g(x) dx, \quad f_{n,\bar{i}}(u, t) = n^d \int_{Q_{n,\bar{i}}} f(u, x, t) dx,$$

$$(2.6) \quad Q_{n,\bar{i}} = \left[\frac{i_1 - 1}{n}, \frac{i_1}{n} \right) \times \left[\frac{i_2 - 1}{n}, \frac{i_2}{n} \right) \times \dots \times \left[\frac{i_d - 1}{n}, \frac{i_d}{n} \right).$$

The approximate solution of (1.1), (1.2) is then defined as follows:

$$(2.7) \quad u_n(t, x) = \sum_{\bar{i} \in [n]^d} u_{n,\bar{i}}(t) \mathbf{1}_{Q_{n,\bar{i}}}(x).$$

Next, we define random $\{0, 1\}$ -valued matrix $(a_{n,\bar{i}\bar{j}})$. To this end, let

$$(2.8) \quad \alpha_n = n^{-d\gamma}, \quad 0 \leq \gamma < 1.$$

The case $W \in L^\infty(Q^2)$ is slightly different and so we treat it separately. Thus, there are two cases to consider.

- (i) Suppose $W \in L^\infty(Q^2)$. Without loss of generality, we further assume that $0 \leq W \leq 1$. Then let

$$(2.9) \quad W_{n,\bar{i}\bar{j}} = n^{2d} \int_{Q_{n,\bar{i}} \times Q_{n,\bar{j}}} W(x, y) dx dy$$

and

$$(2.10) \quad \mathbb{P}(a_{n,\bar{i}\bar{j}} = 1) = \alpha_n W_{n,\bar{i}\bar{j}}, \quad \bar{i}, \bar{j} \in [n]^d.$$

- (ii) Alternatively, if W is in $L^2(Q^2)$ but not in $L^\infty(Q^2)$ then let

$$(2.11) \quad \tilde{W}_n(x, y) := \alpha_n^{-1} \wedge W(x, y) \quad \text{and} \quad W_{n,\bar{i}\bar{j}} = n^{2d} \int_{Q_{n,\bar{i}} \times Q_{n,\bar{j}}} \tilde{W}(x, y) dx dy,$$

where α_n is defined in (2.8) with $\gamma \in (0, 1)$. Then

$$(2.12) \quad \mathbb{P}(a_{n,\bar{i}\bar{j}} = 1) = \alpha_n W_{n,\bar{i}\bar{j}}, \quad \bar{i}, \bar{j} \in [n]^d.$$

Matrix $(a_{n,\bar{i}\bar{j}})$ defines a random graph $\Gamma_n = \langle V(\Gamma_n), E(\Gamma_n) \rangle$ with the set of nodes $V(\Gamma_n) = [n]^d$ and the edge set $E(\Gamma_n) \subset V(\Gamma_n) \times V(\Gamma_n)$. Two nodes $\bar{i} \in [n]^d$ and $\bar{j} \in [n]^d$ are connected by an edge iff $a_{n,\bar{i}\bar{j}} = 1$. The set of all pairs of connected nodes forms the edge set $E(\Gamma_n)$. Graphs constructed using the algorithm given above are called W-random graphs [22, 6]. The asymptotic behavior of such graphs is determined by a nonnegative integrable function W , called a graphon. The construction based on (2.9) yields a dense graph sequence (Γ_n) , i.e., the cardinality of the edge set $|E(\Gamma_n)|$ is $O(|V(\Gamma_n)|^2)$. If we use (2.10) or (2.12) instead, the resulting graph sequence (Γ_n) is sparse, i.e., $|E(\Gamma_n)| = o(|V(\Gamma_n)|^2)$. W-random graphs play an important role in the theory of graph limits [21]. The discretization scheme (2.3) can be interpreted as an interacting particle system on a random graph, whose continuum limit is (1.1) [26].

3. Convergence of the semidiscrete scheme. In this section, we study convergence of the discrete scheme (2.3), (2.4). We first deal with the more general case of unbounded graphon W and then specialize the result for $W \in L^\infty(Q^2)$.

The following technical assumption on W is needed to justify a wider range of sparse sampling in the Monte Carlo method. Let nonnegative $W \in L^4(Q^2)$ satisfy

(W-1s)

$$\max \left\{ \text{ess sup}_{x \in Q} \int W^k(x, y) dy, \text{ess sup}_{y \in Q} \int W^k(x, y) dx \right\} \leq \bar{W}_k, \quad k \in \{1, 2, 3, 4\}.$$

THEOREM 3.1. *Suppose nonnegative $W \in L^4(Q^2)$ is subject to (W-1s), and D , f , and g are as in (1.1), (1.2). Further, $\alpha_n = n^{-d\gamma}$ for some $\gamma \in (0, 1)$. Then for arbitrary $0 < \delta < 1 - \gamma$, we have*

$$\begin{aligned}
& \sup_{t \in [0, T]} \|u(t, \cdot) - u_n(t, \cdot)\|_{L^2(Q)} \\
& \leq C \left(\|g - g_n\|_{L^2(Q)} + \sup_{u \in \mathbb{R}, t \in [0, T]} \|f(u, \cdot, t) - f_n(u, \cdot, t)\|_{L^2(Q)} \right. \\
(3.1) \quad & \left. + \|\tilde{W}_n - W\|_{L^2(Q^2)} + \|\tilde{W}_n - P_n \tilde{W}_n\|_{L^2(Q^2)} + n^{-d(1-\gamma-\delta)/2} \right),
\end{aligned}$$

where C is a positive constant independent of n , and $P_n \tilde{W}_n =: W_n$ stands for the L^2 -projection of $\tilde{W}_n$ onto the finite-dimensional subspace

$$\begin{aligned}
(3.2) \quad X_n &= \text{span}\{\mathbf{1}_{Q_{n,\bar{i}} \times Q_{n,\bar{j}}}, (\bar{i}, \bar{j}) \in [n]^{2d}\}, \\
W_n(x, y) &= \sum_{(\bar{i}, \bar{j}) \in [n]^{d^2}} W_{n,\bar{i}\bar{j}} \mathbf{1}_{Q_{n,\bar{i}} \times Q_{n,\bar{j}}}(x, y),
\end{aligned}$$

and

$$f_n(u, x, t) = \sum_{\bar{i} \in [n]^d} f_{n,\bar{i}}(u, t) \mathbf{1}_{Q_{n,\bar{i}}}(x).$$

Estimate (3.1) holds almost surely (a.s.) with respect to the random graph model.

Remark 3.1. The theorem still holds without (W-1s), i.e., for square integrable W subject to (W-1). In this case, the last term on the right-hand side of (3.1) is replaced by $n^{-d(1/2-\gamma-\delta)}$, $\gamma \in (0, 1/2)$, and $\delta < 1/2 - \gamma$. Unlike (W-1), (W-1s) is not needed for the wellposedness of the IVP (1.1), (1.2).

The first two terms on the right-hand side of (3.1) correspond to the error of approximation of the initial data $g \in L^2(Q)$ and $f(u, x, t)$ by the step functions in x . Further, $\|\tilde{W}_n - W\|_{L^2(Q^2)}^2$ and $\|\tilde{W}_n - P_n \tilde{W}_n\|_{L^2(Q^2)}^2$ bound the error of approximation of W by a bounded step function W_n . Here, the first term $\|\tilde{W}_n - W\|_{L^2(Q^2)}^2$ is the error of truncating W and the second term $\|\tilde{W}_n - P_n \tilde{W}_n\|_{L^2(Q^2)}^2$ is the error of approximation of the truncated function $\tilde{W}_n$ by projecting it onto a finite-dimensional subspace. Finally, the last term on the right-hand side of (3.1) is the error of the approximation of the nonlocal term by the random sum in (2.3).

For bounded graphons W , Theorem 3.1 implies the following result.

COROLLARY 3.2. *Let $W \in L^\infty(Q^2)$. Then under the assumptions of Theorem 3.1 we have*

$$\begin{aligned}
(3.3) \quad & \sup_{t \in [0, T]} \|u(t, \cdot) - u_n(t, \cdot)\|_{L^2(Q)} \\
& \leq C \left(\|g - g_n\|_{L^2(Q)} + \sup_{u \in \mathbb{R}, t \in [0, T]} \|f(u, \cdot, t) - f_n(u, \cdot, t)\|_{L^2(Q)} \right. \\
& \quad \left. + \|W_n - P_n W_n\|_{L^2(Q^2)} + n^{-d(1-\gamma-\delta)/2} \right) \quad \text{a.s.},
\end{aligned}$$

where C is a positive constant independent of n .

Remark 3.2. From (3.3) one can see how to use sparse sampling of points in the Monte Carlo method to optimize computation. If the larger of the two errors of

approximation of g and W by step functions is $O(n^{-\kappa})$ with $0 < \kappa < d/2$ (cf. section 4.2) and the nonlinearity $f(\cdot)$ does not depend on x , then taking $\gamma = 1 - 2\kappa$ one can use sparse discretization without compromising the accuracy. This has the following computational advantages over dense random and, moreover, deterministic spatial discretization schemes. The evaluation of an integral over $[0, 1]^d$ using the rectangular rule requires $O(n^d)$ operations, whereas using a given set of random sample of points requires on average $O(n^{d(1-\gamma)})$.

The proof of Theorem 3.1 modulo a few minor details proceeds as the proof of convergence to the continuum limit in [24, 26]. First, the solution of the IVP (2.3), (2.4) is compared to that of the IVP for the averaged equation:

$$(3.4) \quad \dot{v}_{n,\bar{i}} = f_{n,\bar{i}}(v_{n,\bar{i}}, t) + n^{-d} \sum_{\bar{j} \in [n]^d} W_{n,\bar{i}\bar{j}} D(u_{n,\bar{j}} - u_{n,\bar{i}}), \quad \bar{i} \in [n]^d,$$

$$(3.5) \quad v_{n,\bar{i}}(0) = g_{n,\bar{i}}.$$

Then the solution of the averaged problem is compared to the solution of the IVP (1.1), (1.2). It is convenient to view the solution of the averaged problem as a function on $\mathbb{R}^+ \times Q$:

$$(3.6) \quad v_n(t, x) = \sum_{\bar{i} \in [n]^d} v_{n,\bar{i}}(t) \mathbf{1}_{Q_{n,\bar{i}}}(x).$$

Likewise, the solution of the discrete problem (2.3), (2.4) $u_n(t, x)$ is defined in (2.7).

We recast the IVP (3.4), (3.5) as follows:

$$(3.7) \quad \partial_t v_n(t, x) = f_n(v_n(t, x), x, t) + \int W_n(x, y) D(v_n(t, y) - v_n(t, x)) dy,$$

$$(3.8) \quad v_n(0, x) = g_n(x).$$

The first step of the proof of convergence of the numerical scheme (2.3), (2.4) is accomplished in the following lemma.

LEMMA 3.3. *Let nonnegative $W \in L^4(Q^2)$ subject to (W-1s), and $\alpha_n = n^{-d\gamma}$, $\gamma \in (0, 1)$ (cf. (2.11)). Then for any $T > 0$ for solutions of (2.3) and (3.4) subject to the same initial conditions, we have*

$$(3.9) \quad \sup_{t \in [0, T]} \|u_n(t, \cdot) - v_n(t, \cdot)\|_{L^2(Q)} \leq C n^{-d(1-\gamma-\delta)/2} \quad \text{a.s.,}$$

where $0 < \delta < 1 - \gamma$, and positive constant C independent of n .

The proof of the lemma is technical and is relegated to section 8. The result still holds for square integrable functions without the additional assumption (W-1s), albeit for a narrower range of $\gamma \in (0, 0.5)$ (cf. [26, Theorem 4.1]).

Proof of Theorem 3.1. Denote the difference between the solutions of the original IVP (1.1), (1.2) and the averaged IVP (3.7), (3.8):

$$(3.10) \quad w_n(t, x) = u(t, x) - v_n(t, x).$$

By subtracting (3.7) from (1.1), multiplying the resultant equation by w_n , and integrating over Q , we obtain

$$\begin{aligned}
 (3.11) \quad & \int \partial_t w_n(t, x) w_n(t, x) dx \\
 &= \int (f(u(t, x), x, t) - f(v_n(t, x), x, t)) w_n(t, x) dx \\
 &+ \int (f(v_n(t, x), x, t) - f_n(v_n(t, x), x, t)) w_n(t, x) dx \\
 &+ \int \int W(x, y) [D(u(t, y) - u(t, x)) - D(v_n(t, y) - v_n(t, x))] w_n(t, x) dy dx \\
 &+ \int \int (W(x, y) - W_n(x, y)) D(v_n(t, y) - v_n(t, x)) w_n(t, x) dy dx.
 \end{aligned}$$

Using Lipschitz continuity of $f(u, x, t)$ in u and an elementary case of the Young's inequality, we obtain

$$(3.12) \quad \left| \int (f(u(t, x), x, t) - f(v_n(t, x), x, t)) w_n(t, x) dx \right| \leq L_f \int w_n(t, x)^2 dx,$$

$$\begin{aligned}
 (3.13) \quad & \left| \int (f(v_n(t, x), x, t) - f_n(v_n(t, x), x, t)) w_n(t, x) dx \right| \\
 & \leq \frac{1}{2} \int (f(v_n(t, x), x, t) - f_n(v_n(t, x), x, t))^2 dx \\
 & + \frac{1}{2} \|w_n(t, \cdot)\|^2,
 \end{aligned}$$

where $\|\cdot\|$ stands for the $L^2(Q)$ -norm. Recall that D is bounded by 1 (cf. (1.4)). Using this bound and the Young's inequality, we obtain

$$\begin{aligned}
 (3.14) \quad & \left| \int \int (W(x, y) - W_n(x, y)) D(v_n(t, y) - v_n(t, x)) w_n(t, x) dy dx \right| \leq \frac{1}{2} \|W - W_n\|_{L^2(Q^2)} \\
 & + \frac{1}{2} \|w_n\|^2.
 \end{aligned}$$

Finally, using Lipschitz continuity of D and Young's inequality, we estimate

$$\begin{aligned}
 (3.15) \quad & \left| \int \int W(x, y) [D(u(t, y) - u(t, x)) - D(v_n(t, y) - v_n(t, x))] w_n(t, x) dy dx \right| \\
 & \leq L_D \int \int W(x, y) (|w_n(t, y)| + |w_n(t, x)|) |w_n(t, x)| dy dx \\
 & \leq L_D \int \int W(x, y) \left(\frac{1}{2} |w_n(t, y)|^2 + \frac{3}{2} |w_n(t, x)|^2 \right) dy dx \\
 & \leq \frac{3L_D}{2} \int \int W(x, y) |w_n(t, x)|^2 dy dx + \frac{L_D}{2} \int \int W(x, y) |w_n(t, y)|^2 dy dx \\
 & \leq 2W_1 L_D \|w_n\|^2,
 \end{aligned}$$

where we used the Fubini theorem and (W-1) in the last line.

By combining (3.11)–(3.15), we arrive at

(3.16)

$$\frac{d}{dt} \|w_n(t, \cdot)\|^2 \leq L \|w_n(t, \cdot)\|^2 + \sup_{u \in \mathbb{R}, t \in [0, T]} \|f(u, \cdot, t) - f_n(u, \cdot, t)\|^2 + \|W_n - W\|^2,$$

where $L = 1 + 2L_f + L_D(3W_1 + W_2)$.

By Gronwall's inequality, we have

$$\begin{aligned} & \sup_{t \in [0, T]} \|w_n(t, \cdot)\| \\ & \leq e^{LT/2} \sqrt{\|w_n(0, \cdot)\|^2 + \sup_{u \in \mathbb{R}, t \in [0, T]} \|f(u, \cdot, t) - f_n(u, \cdot, t)\|^2 + \|W_n - W\|_{L^2(Q^2)}^2} \\ & \leq e^{LT/2} \left(\|g - g_n\|_{L^2(Q)} + \sup_{u \in \mathbb{R}, t \in [0, T]} \|f(u, \cdot, t) - f_n(u, \cdot, t)\| + \|W_n - W\|_{L^2(Q^2)} \right). \end{aligned}$$

□

4. Two examples. The error of approximation of the nonlocal term by a random sum, the last term on the right-hand side of (3.1), is known explicitly. Next in importance is the error of approximation of the square integrable graphon W by the step function W_n . This error depends on the regularity of the graphon W . In this section, we consider two representative examples of W : a singular graphon (section 4.1) and a bounded graphon with jump discontinuities (section 4.2). Motivated by these examples in the next section, we will analyze the rate of convergence estimates under general assumptions on graphon W .

We will begin with the following estimate for Hölder continuous functions. To this end, $\phi \in L^p(Q)$, $p \geq 1$, and

$$\phi_n(x) = \sum_{\bar{i} \in [n]^d} \phi_{n,\bar{i}} \mathbf{1}_{Q_{n,\bar{i}}}(x), \quad \phi_{n,\bar{i}} = n^d \int_{Q_{n,\bar{i}}} \phi(x) dx.$$

LEMMA 4.1. Suppose $\phi \in L^p(Q)$, $p \geq 1$, is a Hölder continuous function

$$(4.1) \quad |\phi(x) - \phi(y)| \leq C|x - y|^\beta, \quad x, y \in Q, \quad \beta \in (0, 1].$$

Then

$$(4.2) \quad \|\phi - \phi_n\|_{L^p(Q)} \leq Ch^\beta.$$

Here and below, $h := n^{-1}$.

Proof. Using Jensen's inequality and (4.2), we have

$$\begin{aligned} \|\phi - \phi_n\|_{L^p(Q)}^p &= \int_Q \left| \sum_{\bar{i} \in [n]^d} \left(\phi(x) - n^d \int_{Q_{n,\bar{i}}} \phi(y) dy \right) \mathbf{1}_{Q_{n,\bar{i}}}(x) \right|^p dx \\ (4.3) \quad &= \sum_{\bar{i} \in [n]^d} \int_{Q_{n,\bar{i}}} \left| n^d \int_{Q_{n,\bar{i}}} (\phi(x) - \phi(y)) dy \right|^p dx \\ &\leq n^d \sum_{\bar{i} \in [n]^d} \int_{Q_{n,\bar{i}}} \int_{Q_{n,\bar{i}}} |\phi(x) - \phi(y)|^p dx dy \\ &\leq C^p h^{p\beta}. \end{aligned}$$

□

Remark 4.1. Below, we will freely apply Lemma 4.1 to functions on Q and on Q^2 . The latter are clearly covered by the lemma by taking $Q := Q^2$.

4.1. A singular graphon. Consider the problem of approximation by step functions of the singular kernel graphon

$$(4.4) \quad W(x, y) = \frac{1}{|x - y|^\lambda}, \quad x, y \in Q = [0, 1]^d,$$

where $0 < \lambda < d/2$.

LEMMA 4.2. For $\gamma \in (0, 1/2)$ and $0 < \lambda < d/2$ we have

$$(4.5) \quad \|W - W_n\|_{L^2(Q^2)} \leq \max \left\{ O \left(h^{d\gamma(\frac{d}{2\lambda}-1)} \right), O \left(h^{1-d\gamma(1+\frac{1}{\lambda})} \right) \right\}.$$

Proof.

1. Below, we will use the change of variables $(x, y) = T(u, v)$ for (x, y) and (u, v) from $\mathbb{R}^{2d}$, defined by

$$(4.6) \quad u_i = x_i - y_i \quad \text{and} \quad v_i = x_i + y_i, \quad i \in [d].$$

2. Let $\alpha_n = n^{-d\gamma}$ and recall that $\tilde{W}_n = h^{-d\gamma} \wedge W$. Denote $\tilde{Q} = \left\{ (x, y) \in Q^2 : |x - y|^{-\lambda} \geq h^{-d\gamma} \right\}$. Further,

$$(4.7) \quad \begin{aligned} \|W - \tilde{W}_n\|_{L^2(Q^2)}^2 &= \int_{\tilde{Q}} \left(\frac{1}{|x - y|^\lambda} - n^{d\gamma} \right)^2 dx dy \\ &\leq C_1 \int_{\{|u| \leq h^{\frac{d\gamma}{\lambda}}\}} \left(\frac{1}{|u|^\lambda} - n^{d\gamma} \right)^2 du \\ &\leq C_2 \int_0^{h^{\frac{d\gamma}{\lambda}}} \left(\frac{1}{r^\lambda} - n^{d\gamma} \right)^2 r^{d-1} dr \\ &= O \left((d - 2\lambda)^{-1} h^{2d\gamma(\frac{d}{2\lambda}-1)} \right), \end{aligned}$$

where we used (4.6) followed by the change to polar coordinates. Thus,

$$(4.8) \quad \|W - \tilde{W}_n\|_{L^2(Q^2)} = O \left(h^{d\gamma(\frac{d}{2\lambda}-1)} \right), \quad 0 < \lambda < d/2, \quad \gamma > 0.$$

3. Next we turn to estimating $\|\tilde{W}_n - W_n\|_{L^2(Q^2)}$. Since the truncated function $\tilde{W}_n$ is Lipschitz continuous on Q^2 , by Lemma 4.1,

$$\|\tilde{W}_n - W_n\|_{L^2(Q^2)} \leq L(\tilde{W}_n)h.$$

It remains to estimate the Lipschitz constant $L(\tilde{W}_n) \leq \text{ess sup}_{Q^2} |\nabla \tilde{W}_n|$. On $Q^2 - \tilde{Q}$,

$$(4.9) \quad |\nabla \tilde{W}_n| = \lambda |x - y|^{-1-\lambda} |\nabla_{(x,y)} |x - y|| = \sqrt{2}\lambda |x - y|^{-1-\lambda}.$$

The gradient approaches its greatest value as $|x - y| \searrow h^{\frac{d\gamma}{\lambda}}$. Thus,

$$\text{ess sup}_{Q^2} |\nabla \tilde{W}_n| = \sqrt{2}\lambda h^{-d\gamma(\frac{1}{\lambda}+1)}$$

and

$$(4.10) \quad \|\tilde{W}_n - W_n\|_{L^2(Q^2)} = O\left(h^{1-d\gamma(1+\frac{1}{\lambda})}\right).$$

4. The statement of the lemma follows (2.11) and (4.10) and the triangle inequality. $\square$

Next we choose γ to optimize the rate of convergence in (4.5). By setting the two exponents of h on the right-hand side of (4.5) equal, we see that the rate is optimal for

$$\gamma = \frac{2\lambda}{d(d+2)}.$$

With this choice of γ ,

$$\|W - W_n\|_{L^2(Q^2)} = O(h^{\frac{2\lambda}{d+2}(\frac{d}{2\lambda}-1)}).$$

To optimize the rate of convergence of the numerical scheme (2.3), (2.4), one has to choose $\gamma \in (0, 1)$ to maximize the smallest of the following three exponents:

$$d\gamma\left(\frac{d}{2\lambda}-1\right), \quad 1-d\gamma\left(1+\frac{1}{\lambda}\right), \quad \frac{d}{2}(1-\gamma),$$

where the last exponent comes from the error of the Monte Carlo approximation (cf. (3.1)).

4.2. $\{0,1\}$ -valued functions. The following example is adapted from [23]. It shows how jump discontinuities affect the rate of convergence of approximation by piecewise constant functions. The accuracy of approximation depends on the geometry of the hypersurface of discontinuity, more precisely, on its fractal dimension.

Let Q^+ be a closed subset of Q and consider

$$(4.11) \quad \phi(x) = \begin{cases} 1, & x \in Q^+, \\ 0 & \text{otherwise.} \end{cases}$$

Denote by ∂Q^+ the boundary of Q and recall the upper box-counting dimension of ∂Q^+

$$(4.12) \quad \beta := \overline{\lim}_{h \rightarrow 0} \frac{\log N_h(\partial Q^+)}{-\log h},$$

where $N_h(\partial Q^+)$ stands for the number of $Q_{n,\bar{i}}$, $\bar{i}, \bar{j} \in [n]^d$, having nonempty intersection with ∂Q^+ (cf. [17]). For instance, the upper box-counting dimension of the von Koch curve is $\log 4 / \log 3$ (see [17, Problem 2.5]).

LEMMA 4.3. *Let ϕ be defined by (4.11) and let ϕ_n stand for the L^2 -projection of ϕ on X_n (cf. (3.2)). Then*

$$(4.13) \quad \|\phi - \phi_n\|_{L^p(Q)} \leq Ch^{\frac{d-\beta}{p}}$$

for some positive C independent on n .

Proof. As in (4.3), we have

$$(4.14) \quad \|\phi - \phi_n\|_{L^p(Q^2)}^p \leq h^{-d} \sum_{\bar{i} \in [n]^d} \int_{Q_{n,\bar{i}}} \int_{Q_{n,\bar{i}}} |f(x) - f(z)|^p dz dx.$$

Note that the only nonzero terms in the sum on the right-hand side of (4.14) are the integrals over $Q_{n,\bar{i}} \times Q_{n,\bar{j}}$'s having nonempty intersection with ∂Q^+ . Thus,

$$(4.15) \quad \|\phi - \phi_n\|_{L^p(Q)}^p = h^d N_h(\partial Q^+) \leq Ch^{d-\beta},$$

where we used (4.12). $\square$

Remark 4.2. Note that as $\beta \rightarrow d - 0$ the rate of convergence in (4.13) can be made arbitrarily low.

5. The rate of convergence of the semidiscrete scheme.

5.1. Approximation by step functions. In this section, we address the rate of convergence of the Galerkin component of the numerical scheme (2.3), (2.4). Specifically, we study the error of the approximation of the graphon $W \in L^2(Q^2)$ and the initial data $g \in L^2(Q)$ by step functions.

We will need an L^p -modulus of continuity of function on a unit d -cube $Q = [0, 1]^d$. In fact, we only need the L^2 -modulus of continuity but present the analysis in a more general L^p -setting, since this does not require any extra effort. For functions on the real line, the definition of the L^p -modulus of continuity can be found in [1, 13]. Here, we present a suitable adaptation of this definition for the problem at hand.

DEFINITION 5.1. For $\phi \in L^p(Q)$ we define the L^p -modulus of continuity

$$(5.1) \quad \omega_p(\phi, \delta) = \sup_{|\xi|_\infty \leq \delta} \|\phi(\cdot + \xi) - \phi(\cdot)\|_{L^p(Q_\xi)}, \quad \delta > 0,$$

where $|\xi|_\infty := \max_{i \in [d]} |\xi_i|$ and $Q_\xi = \{x \in Q : x + \xi \in Q\}$.

For $\alpha \in (0, 1]$, we define a generalized Lipschitz space¹

$$(5.2) \quad \text{Lip}(\alpha, L^p(Q)) = \{\phi \in L^p(Q) : \exists C > 0 : \omega_p(\phi, \delta) \leq C\delta^\alpha\}.$$

Clearly $\text{Lip}(\alpha, L^p(Q))$ contains α -Hölder continuous functions. However, Lipschitz spaces are much larger than Hölder spaces. For instance, $\text{Lip}(1/p, L^p(Q))$ contains discontinuous functions.

Below, we express the error of approximation of $\phi \in L^p(Q)$ by a step function through $\omega_p(\phi, h)$. The analysis works out a little cleaner for dyadic discretization of Q , which will be assumed for the remainder of this section. Thus, we approximate $\phi \in L^p(Q)$ by a piecewise constant function

$$(5.3) \quad \phi_{2^m}(x) = \sum_{\bar{i} \in [2^m]^d} \phi_{Q_{2^m,\bar{i}}} \mathbf{1}_{Q_{2^m,\bar{i}}}(x),$$

where $\phi_{Q_{2^m,\bar{i}}}$ stands for the mean value of ϕ on $Q_{2^m,\bar{i}}$,

$$\phi_{Q_{2^m,\bar{i}}}(x) = 2^{md} \int_{Q_{2^m,\bar{i}}} \phi(x) dx, \quad i \in [2^m].$$

¹Below, we will freely apply the definitions and various estimates established for functions on Q to functions on Q^2 , for which they are trivially valid by setting $d := 2d$. In particular, the definitions of the modulus of continuity and the corresponding Lipschitz spaces obviously translate to functions on Q^2 .

LEMMA 5.2. For $\phi \in \text{Lip}(\alpha, L^p(Q))$, we have

$$(5.4) \quad \|\phi - \phi_{2^m}\|_{L^p(Q)} \leq C 2^{-\alpha m},$$

where C is independent of m .

Proof. Fix $m \in \mathbb{N}$ and denote $h := 2^{-m}$. To simplify notation, throughout the proof we drop 2^m in the subscript of $x_{2^m, i}$ and $Q_{2^m, \bar{i}}$.

Denote

$$Q_{\bar{i}}^{\bar{j}} = \left[x_{i_1-1} + j_1 \frac{h}{2}, x_{i_1-1} + (j_1+1) \frac{h}{2} \right) \times \cdots \times \left[x_{i_d-1} + j_d \frac{h}{2}, x_{i_d-1} + (j_d+1) \frac{h}{2} \right)$$

and

$$Q_{\bar{i}}' = \left[x_{i_1-1}, x_{i_1-1} + \frac{h}{2} \right) \times \cdots \times \left[x_{i_d-1}, x_{i_d-1} + \frac{h}{2} \right).$$

Note

$$(5.5) \quad \int_{Q_{\bar{i}}^{\bar{j}}} \phi(s) ds = \int_{Q_{\bar{i}}'} \phi\left(s + \bar{j} \frac{h}{2}\right) ds,$$

where

$$s + \bar{j} \frac{h}{2} = \left(s_1 + j_1 \frac{h}{2}, s_2 + j_2 \frac{h}{2}, \dots, s_d + j_d \frac{h}{2} \right).$$

Using (5.5), we write

$$(5.6) \quad \phi_{2^{m+1}}(x) = \sum_{\bar{i} \in [2^m]^d} \sum_{\bar{j} \in \{0,1\}^d} \left(\frac{2}{h}\right)^d \int_{Q_{\bar{i}}'} \phi\left(s + \bar{j} \frac{h}{2}\right) ds \mathbf{1}_{Q_{\bar{i}}^{\bar{j}}}(x).$$

Further, using

$$\begin{aligned} \mathbf{1}_{Q_{\bar{i}}}(x) &= \sum_{\bar{j} \in \{0,1\}^d} \mathbf{1}_{Q_{\bar{i}}^{\bar{j}}}(x), \\ \int_{Q_{\bar{i}}} \phi(s) ds &= \sum_{\bar{j} \in \{0,1\}^d} \int_{Q_{\bar{i}}'} \phi\left(s + \bar{k} \frac{h}{2}\right) ds, \end{aligned}$$

we rewrite (5.3) as follows:

$$(5.7) \quad \phi_{2^m}(x) = \sum_{\bar{i} \in [2^m]^d} \sum_{\bar{j} \in \{0,1\}^d} \left(\frac{1}{h}\right)^d \left\{ \sum_{\bar{k} \in \{0,1\}^d} \int_{Q_{\bar{i}}'} \phi\left(s + \bar{k} \frac{h}{2}\right) ds \right\} \mathbf{1}_{Q_{\bar{i}}^{\bar{j}}}(x).$$

By subtracting (5.6) from (5.7) we have

$$\phi_{2^m} - \phi_{2^{m+1}} = \sum_{\bar{i} \in [2^m]^d} \sum_{\bar{j} \in \{0,1\}^d} \frac{1}{h^d} \left\{ \sum_{\substack{\bar{k} \in \{0,1\}^d \\ \bar{k} \neq \bar{j}}} \int_{Q_{\bar{i}}'} \left[\phi\left(s + \bar{k} \frac{h}{2}\right) - \phi\left(s + \bar{j} \frac{h}{2}\right) \right] ds \right\} \mathbf{1}_{Q_{\bar{i}}^{\bar{j}}}.$$

Further,

$$(5.8) \quad |\phi_{2^m} - \phi_{2^{m+1}}|^p \leq \sum_{\bar{i} \in [2^m]^d} \sum_{\bar{j} \in \{0,1\}^d} \sum_{\substack{\bar{k} \in \{0,1\}^d \\ \bar{k} \neq \bar{j}}} \left| \int_{Q_{\bar{i}}'} h^{-d} \left[\phi \left(s + \bar{k} \frac{h}{2} \right) - \phi \left(s + \bar{j} \frac{h}{2} \right) \right] ds \right|^p \mathbf{1}_{Q_{\bar{i}}^{\bar{j}}}.$$

Integrating both sides of (5.8) over Q and using Jensen's inequality, we continue

$$(5.9) \quad \begin{aligned} \|\phi_{2^m} - \phi_{2^{m+1}}\|_{L^p(Q)}^p &\leq \sum_{\bar{j} \in \{0,1\}^d} \sum_{\substack{\bar{k} \in \{0,1\}^d \\ \bar{k} \neq \bar{j}}} \sum_{\bar{i} \in [2^m]^d} \int_{Q_{\bar{i}}'} \left| \phi \left(s + \bar{k} \frac{h}{2} \right) - \phi \left(s + \bar{j} \frac{h}{2} \right) \right|^p ds \\ &\leq 2^d (2^d - 1) \omega_p(\phi, h)^p. \end{aligned}$$

Thus,

$$\|\phi_{2^m} - \phi_{2^{m+1}}\|_{L^p(Q)} \leq [2^d (2^d - 1)]^{1/p} \omega_p(\phi, 2^{-(m+1)}) =: C_{d,p} \omega_p(\phi, 2^{-(m+1)}).$$

Since $\phi \in \text{Lip}(\alpha, L^p(Q))$, we have

$$(5.10) \quad \|\phi_{2^m} - \phi_{2^{m+1}}\|_{L^p(Q)} \leq C 2^{-\alpha m},$$

where C depends on ϕ, d , and p but not m .

Let $m \in \mathbb{N}$ be arbitrary but fixed. For any integer $M > m$ we have

$$(5.11) \quad \begin{aligned} \|\phi_{2^M} - \phi_{2^m}\|_{L^p(Q)} &= \left\| \sum_{k=m}^{M-1} (\phi_{2^{k+1}} - \phi_{2^k}) \right\|_{L^p(Q)} \leq \sum_{k=m}^{\infty} \|\phi_{2^{k+1}} - \phi_{2^k}\|_{L^p(Q)} \\ &\leq \sum_{k=m}^{\infty} \omega_p(\phi, 2^{-(k+1)}) \leq 2^{-p+1} \sum_{k=m}^{\infty} C 2^{\alpha(k+1)} \leq C 2^{-\alpha m}. \end{aligned}$$

By passing M to infinity in (5.11), we get (5.4). $\square$

5.2. The rate of convergence. We now can combine Theorem 3.1 and Lemma 5.4 to estimate the convergence rate for (2.3), (2.4). For the model with a bounded graphon W (cf. (i) in section 2) we have the following theorem.

THEOREM 5.3. *Suppose that in addition to the assumptions of Theorem 3.1, for some $\alpha_i \in (0, 1]$, $i \in [3]$, $g \in \text{Lip}(\alpha_1, L^2(Q))$, $W \in \text{Lip}(\alpha_2, L^2(Q^2)) \cap L^\infty(Q^2)$, and $f(u, \cdot, t) \in \text{Lip}(\alpha_3, L^2(Q))$ uniformly for $(u, t) \in \mathbb{R} \times [0, T]$, i.e.,*

$$(5.12) \quad \omega_2(f(u, \cdot, t), \delta) \leq C \delta^{\alpha_3},$$

where $C > 0$ is independent of (u, t) .

Then

$$(5.13) \quad \max_{t \in [0, T]} \|u(t, \cdot) - u_n(t, \cdot)\| \leq C n^{-\alpha}, \quad \alpha = \min \left\{ \alpha_1, \alpha_2, \alpha_3, \frac{d}{2} (1 - \gamma) \right\}, \quad \text{a.s.}$$

Here, C is independent of n .

Remark 5.1. Note the error estimate of the semidiscrete scheme holds with probability 1 with respect to the realization of the random graph sequence used to approximate the nonlocal term.

Remark 5.2. As an example of f satisfying condition (5.12), consider $f(u, x, t) = \chi(x)F(u, t)$, where $F(u, t)$ is a bounded function and $\chi(x)$ is a piecewise constant function. Such nonlinearities arise naturally in modeling piecewise constant media.

If $W \in L^2(Q^2)$ has singularities, then the convergence rate may also depend on the accuracy of approximation of W by the truncated function $\tilde{W}_n$. We do not estimate the truncation error for a general $W \in L^2(Q^2)$. For an example of how this error can be estimated for a given graphon in practice, we refer to the example in section 4.1.

6. Fully discrete scheme. Our main focus so far has been the accuracy of the semidiscrete scheme (2.3), i.e., the accuracy of the spatial approximation. To find a numerical solution of (2.3), one also needs to discretize time. To this end, one can use the Euler method or a higher order scheme, e.g., one of the methods of the Runge–Kutta family. In this section, we show that the standard error estimates available for these methods (cf. [3]) apply to the semidiscrete scheme (2.3). We show that the right-hand side of (2.3) satisfies the assumptions used in the error analysis of the Euler’s method in [3] and comment on the conditions used in the analysis of the Runge–Kutta methods.

Given $0 < \tau \ll 1$, consider the Euler discretization of (2.3)

$$(6.1) \quad u_{n,\bar{i}}^\tau(t_{i+1}) = u_{n,\bar{i}}^\tau(t_i) + \tau F_{n,\bar{i}}(u_n^\tau, t_i), \quad \bar{i} \in [n]^d, \quad t_i = i\tau, \quad i = 0, 1, \dots, M,$$

where

$$(6.2) \quad F_{n,\bar{i}}(v, t) = f_{n,\bar{i}}(v_{\bar{i}}, t) + \frac{1}{\alpha_n n^d} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} D(v_{\bar{j}} - v_{\bar{i}})$$

and $M = \lceil \frac{T}{\tau} \rceil$.

Further, by using linear interpolation, we extend the time domain of u_n^τ from the discrete set of points $\{t_i, i = 0, 1, \dots, M\}$ to $[0, T]$,

$$(6.3) \quad u_{n,\bar{i}}^\tau(t) := \sum_{i=0}^M \left[u_{n,\bar{i}}^\tau(t_i) + \frac{t - t_i}{\tau} (u_{n,\bar{i}}^\tau(t_{i+1}) - u_{n,\bar{i}}^\tau(t_i)) \right] \mathbf{1}_{[t_i, t_{i+1})}(t), \quad \bar{i} \in [n]^d, \quad t \in [0, T],$$

and define

$$(6.4) \quad u_n^\tau(t, x) = \sum_{\bar{i} \in [n]^d} u_{n,\bar{i}}^\tau(t) \mathbf{1}_{Q_{n,\bar{i}}}(x).$$

THEOREM 6.1. *In addition to the assumptions of Theorem 5.3, suppose that D and f are continuously differentiable functions on $\mathbb{R}$ and $\mathbb{R} \times [0, T_1]$, $T_1 := T + 1$, respectively, and*

$$(6.5) \quad \max \left\{ \sup_{v \in \mathbb{R}} |D'(v)|, \sup_{(v,t) \in \mathbb{R} \times [0, T_1]} \{ |\partial_t f(v, t)| + |\partial_v f(v, t)| \} \right\} \leq C.$$

Then

$$(6.6) \quad \max_{t \in [0, T]} \|u(t, \cdot) - u_n^\tau(t, \cdot)\| \leq C_1 n^{-\alpha} + C_2 \tau, \quad \text{a.s.,}$$

where C_1 and C_2 are independent of n and τ , and α is the same as in (5.13).

Remark 6.1. If f and D have continuous partial derivatives up to order $k \geq 2$, the analysis below can be extended to the k th order Runge–Kutta method. Verifying the Lipschitz continuity of the right-hand side of the numerical scheme (6.1), the key step in the analysis, remains the same as in the proof of Lemma 6.2 for the Euler method.

The accuracy of the fully discrete scheme relies on the the well-known error estimates for the Euler’s method, which we review next. To this end, we consider

$$(6.7) \quad \begin{aligned} \dot{Y} &= F(Y, t), \\ Y(0) &= Y_0 \in \mathbb{R}^{n^d}, \end{aligned}$$

where F is a continuously differentiable mapping from $U := \mathbb{R}^{n^d} \times [0, T_1]$ to $\mathbb{R}^{n^d}$.

The Euler scheme for (6.7) has the following form:

$$(6.8) \quad \begin{aligned} Y^\tau(t_{i+1}) &= Y^\tau(t_i) + \tau F(Y^\tau(t_i), t_i), \\ Y^\tau(0) &= Y_0, \end{aligned}$$

where $t_i = i\tau$, $i = 0, 1, \dots, M$, $M = \lceil \frac{T}{\tau} \rceil$.

In order to apply standard error estimates for the Euler’s method to (6.8), we need to impose the following assumptions:

$$(6.9) \quad \sup_{(Y,t) \in U} \|\partial_t F(Y, t)\| \leq L_t,$$

$$(6.10) \quad \sup_{(Y,t) \in U} \|\partial_Y F(Y, t)\| \leq L_Y,$$

where

$$\|B\| := \max_{|\xi|_\infty=1} |B\xi|_\infty, \quad \xi = (\xi_1, \xi_2, \dots, \xi_{n^d}),$$

and $|\xi|_\infty = \max_{i \in [n^d]} |\xi_i|$. Conditions (6.9) and (6.10) guarantee that the second derivative $\ddot{Y}(t)$ remains bounded. In addition, (6.10) implies Lipschitz continuity of $F(t, \cdot)$:

$$|F(Y, t) - F(Z, t)|_\infty \leq L_Y |Y - Z|_\infty$$

uniformly in t . Under these conditions, [3, Theorem 6.3] suitably adjusted to the multidimensional setting, we have

$$(6.11) \quad \max_{0 \leq t_i \leq T} |Y(t_i) - Y^\tau(t_i)|_\infty \leq C\tau,$$

where C is independent of τ .

Since $\dot{Y}(t)$ is bounded on $[0, T_1]$, from (6.11) one can further show

$$(6.12) \quad \max_{0 \leq t \leq T} |Y(t) - Y^\tau(t)|_\infty \leq C\tau,$$

where

$$(6.13) \quad Y^\tau(t) := \sum_{i=0}^M \left[Y^\tau(t_i) + \frac{t - t_i}{\tau} (Y^\tau(t_{i+1}) - Y^\tau(t_i)) \right] \mathbf{1}_{[t_i, t_{i+1})}(t).$$

To apply the error bound (6.12) to (6.1) we need to verify the bounds in (6.9) and (6.10) for F_n defined by (6.2). This is done in the following lemma.

LEMMA 6.2. *Suppose that D and f are continuously differentiable functions on $\mathbb{R}$ and $\mathbb{R} \times [0, T_1]$, respectively, and*

$$(6.14) \quad \max \left\{ \sup_{v \in \mathbb{R}} |D'(v)|, \sup_{(v,t) \in \mathbb{R} \times [0, T_1]} \{ |\partial_t f(v, t)| + |\partial_v f(v, t)| \} \right\} \leq C.$$

Then (6.9) and (6.10) hold for F_n defined by (6.2) a.s.

Proof. Below, we use the following vectorization of various arrays on $[n]^d$. Let ℓ be a fixed bijection from $[n]^d$ to $[n^d]$. Given $(\sigma_{\bar{i}}, \bar{i} \in [n]^d)$ define

$$\text{vec}(\sigma_{\bar{i}}, \bar{i} \in [n]^d) = (\sigma_1, \sigma_2, \dots, \sigma_{n^d}), \quad \text{where } \sigma_{\ell(\bar{i})} = \sigma_{\bar{i}}, \quad \forall \bar{i} \in [n]^d.$$

We need to verify conditions (6.9) and (6.10) for F_n . The first condition follows immediately from (6.5). To verify (6.10), we need to show

$$(6.15) \quad \sup_{(u,t) \in U} \max_{\bar{i} \in [n]^d} |\langle \text{vec}(\partial_{u_{\bar{k}}} F_{n,\bar{i}}(u, t); \bar{k} \in [n]^d), \text{vec}(\xi_{\bar{k}}; \bar{k} \in [n]^d) \rangle| \leq C \quad \text{a.s.},$$

where $|\text{vec}(\xi_{\bar{k}}; \bar{k} \in [n]^d)|_\infty = 1$ and $\langle \cdot, \cdot \rangle$ stands for the inner product in $\mathbb{R}^{n^d}$.

By direct computation,

$$(6.16) \quad \partial_{u_{\bar{k}}} F_{n,\bar{i}}(u, t) = \begin{cases} (\alpha_n n^d)^{-1} a_{n,\bar{i}\bar{k}} D'(u_{\bar{k}} - u_{\bar{i}}), & \bar{k} \neq \bar{i}, \\ \partial_u f(u_{\bar{i}}, t) - (\alpha_n n^d)^{-1} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} D'(u_{\bar{j}} - u_{\bar{i}}), & \bar{k} = \bar{i}, \end{cases}$$

and

$$(6.17) \quad \begin{aligned} S_{n,\bar{i}}(u, t) &:= \langle \text{vec}(\partial_{u_{\bar{k}}} F_{n,\bar{i}}(u, t); \bar{k} \in [n]^d), \text{vec}(\xi_{\bar{k}}; \bar{k} \in [n]^d) \rangle \\ &= \partial_u f(u_{\bar{i}}, t) \xi_{\bar{i}} - (\alpha_n n^d)^{-1} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \eta_{\bar{i}\bar{j}} \xi_{\bar{i}} + (\alpha_n n^d)^{-1} \sum_{\bar{k} \in [n]^d} a_{n,\bar{i}\bar{k}} \eta_{\bar{i}\bar{k}} \xi_{\bar{k}}, \end{aligned}$$

where $|\text{vec}(\xi_{\bar{k}}; \bar{k} \in [n]^d)|_\infty = 1$ and

$$\eta_{\bar{i}\bar{j}} = D'(u_{\bar{j}} - u_{\bar{i}}).$$

We need to bound $|S_{n,\bar{i}}(u, t)|$ uniformly in $\bar{i} \in [n]^d$ and $(u, t) \in U$. The bound for the first term on the right-hand side of (6.17) follows from (6.5). It remains to estimate the two random sums in (6.17). To this end, first note that

$$\sup_{(u,t) \in U} \max_{\bar{i}, \bar{j} \in [n]^d} |\eta_{\bar{i}\bar{j}} \xi_{\bar{j}}| \leq C,$$

because of the bound on $D'(u)$ in (6.5) and $|\xi_{\bar{k}}| \leq 1$. Since $a_{n,\bar{i}\bar{j}} \geq 0$, we have

$$\left| (\alpha_n n^d)^{-1} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \eta_{\bar{i}\bar{j}} \xi_{\bar{i}} \right| \leq C \left| (\alpha_n n^d)^{-1} \sum_{\bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \right|.$$

The last sum is bounded a.s. as follows from Lemmas 8.3 and 8.2. The bound for the other sum in (6.17) is established similarly. $\square$

Lemma 6.2 allows us to apply the error bound (6.12) to the time discretization of the semidiscrete problem:

$$\max_{t \in [0, T]} \|u_n^\tau(t, \cdot) - u_n(t, \cdot)\|_{L^\infty(Q)} \leq C\tau.$$

This estimate combined with (5.13) proves Theorem 6.1.

We conclude this section with the analysis of the computational cost versus accuracy. Denote the error of approximation

$$\Delta := \max_{t \in [0, T]} \|u(t, \cdot) - u_n^\tau(t, \cdot)\|$$

and suppose $\Delta = \epsilon$ for a given $\epsilon > 0$. By (6.6), we have

$$\tau = O(\epsilon) \quad \text{and} \quad n = O\left(\log \frac{1}{\epsilon}\right).$$

From this, we have that the number of time steps needed to compute the approximate solution on $O(1)$ time interval is $O(\epsilon^{-1})$. At each time step, the evaluation of the sum on the right-hand side of (2.3) requires $O(n^{d(1-\gamma)})$ operations. Thus, the total number of operations is

$$O\left(\frac{1}{\epsilon} n^{d(1-\gamma)}\right) = O\left(\frac{1}{\epsilon} \left(\log \frac{1}{\epsilon}\right)^{d(1-\gamma)}\right).$$

7. Numerical example. In this section, we illustrate convergence analysis in the previous sections with a numerical example. To this end, we consider an IVP for the continuum Kuramoto model with nonlocal nearest-neighbor coupling [29]:

$$(7.1) \quad \partial_t u(t, x) = \omega + \int_{[0, 1]} K(y - x) \sin(u(t, x) - u(t, y)),$$

$$(7.2) \quad u(0, x) = u^{(q)}(x),$$

where $u(t, x) \in \mathbb{T}$, $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, stands for the phase of the oscillator at $x \in [0, 1]$, and $\omega \in \mathbb{R}$ is its intrinsic frequency. Function K , describing the connectivity of the network, is first defined on $[0, 1/2]$ by

$$(7.3) \quad K(x) = \mathbf{1}_{\{y: |y| \leq r\}}(x), \quad r \in (-1/2, 1/2),$$

and then extended as a 1-periodic function on $\mathbb{R}$. The initial condition

$$(7.4) \quad u^{(q)}(x) = 2\pi(qx \bmod 1), \quad q \in \mathbb{Z},$$

is called a q -twisted state (Figure 1(a)). For $\omega = 0$, $u^{(q)}$ is a stationary solution of (7.1). Thus,

$$(7.5) \quad u(t, x) = (2\pi qx + \omega t) \bmod 2\pi$$

solves the IVP (7.1), (7.2). We use the explicit solution (7.5) to compute the error of the numerical integration of (7.1), (7.2).

To estimate the rate of convergence of the numerical scheme (2.3), (2.4) we use the following values of parameters: $r = 0.2$, $\omega = 0.5$, and $q = 3$. For these parameter

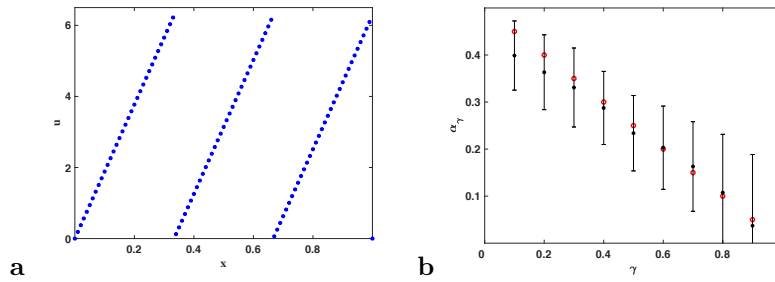


FIG. 1. (a) The 3-twisted state used to initialize the Kuramoto model (7.1). (b) The numerically estimated exponent characterizing convergence of the numerical scheme (2.3), (2.4), α_γ , is plotted as a function of γ (see section 7). The numerical estimates and the theoretical predictions are plotted using the black stars and red circles, respectively. Vertical bars indicate one standard deviation of the 200 random samples.

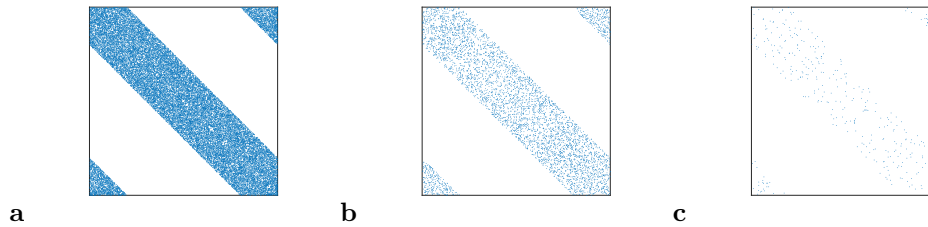


FIG. 2. Pixel pictures of the adjacency matrices of sparse graphs generated with the following values of γ : (a) 0.2, (b) 0.5, and (c) 0.9.

values, the traveling wave solution (7.5) is unstable. We integrated (7.1) numerically for $t \in [0, 0.5]$, using the fourth order Runge–Kutta method with the time step $\tau = 0.01$. The error of the Runge–Kutta method, i.e., of the discretization in time, is expected to be significantly smaller than that of the discretization in space (cf. (2.3), (2.4)). We verified this by repeating our numerical experiments for different values of τ (see Figure 2). This did not affect the results.

We integrated (7.1) numerically for different values of $\gamma \in (0, 1)$. For each value of γ we compute approximate solutions of (7.1), (7.2) for $n = 256$ and $n = 512$. Then we computed the error of approximation for each n : $e_{\gamma,256}$ and $e_{\gamma,512}$ and the rate of convergence

$$(7.6) \quad \alpha_\gamma = \frac{\ln(e_{\gamma,256}/e_{\gamma,512})}{\ln 2}.$$

We repeated this experiment 200 times. The means and standard deviations of α_γ are plotted in Figure 1(b). The numerical estimates of the convergence rate show very good agreement with the theoretical estimate $(1 - \gamma)/2$ for most values of γ . For the two smallest values of γ the deviation from the theoretical estimates is bigger, but it is still within one standard deviation. We attribute this to the decrease of accuracy of the algorithm for generating sparse graphs for small γ for values of n used in these experiments.

8. Proof of Lemma 3.3. In this section, we prove Lemma 3.3. The proof follows the lines of the proof of Theorem 4.1 in [26], which covers $\gamma \in (0, 1/2)$ for $d = 1$. Extension to the multidimensional case $d > 1$ is straightforward. Lemmas 8.3

and 8.4 adapted from [10] allow us to extend the range of γ to $(0, 1)$. The reader not interested in the extended range of γ may find a simpler proof in [26] easier to follow. For those interested in the full range of γ , below we present the proof of Lemma 3.3.

THEOREM 8.1. *Let nonnegative $W \in L^4(Q^2)$ satisfy*

(W-1s)

$$\max \left\{ \operatorname{ess\,sup}_{x \in Q} \int W^k(x, y) dy, \operatorname{ess\,sup}_{y \in Q} \int W^k(x, y) dx \right\} \leq \bar{W}_k, \quad k \in [4],$$

and

$$(8.1) \quad \liminf_{n \rightarrow \infty} \frac{\alpha_n n^d}{\ln n} > 0.$$

Then for solutions of (2.3) and (3.4) subject to the same initial conditions and arbitrary $0 < \epsilon < 1/2$, we have

$$(8.2) \quad \sup_{t \in [0, T]} \|u_n(t, \cdot) - v_n(t, \cdot)\|_{L^2(Q)} \leq C(\alpha_n n^d)^{1/2-\epsilon} \quad \text{a.s.}$$

for arbitrary $T > 0$ and positive constant C independent of n . In particular, for $\alpha_n = n^{-d\gamma}$, $\gamma \in (0, 1)$, we have

$$(8.3) \quad \sup_{t \in [0, T]} \|u_n(t, \cdot) - v_n(t, \cdot)\|_{L^2(Q)} \leq C n^{-d(1-\gamma-\delta)/2} \quad \text{a.s.},$$

where $0 < \delta < 1 - \gamma$ can be taken arbitrarily small.

We precede the proof of Theorem 8.1 with several auxiliary estimates.

LEMMA 8.2. *From (W-1s) it follows that*

$$(8.4) \quad \max \left\{ \sup_{n \in \mathbb{N}} \max_{\bar{i} \in [n]^d} n^{-d} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i}\bar{j}}^k, \sup_{n \in \mathbb{N}} \max_{\bar{j} \in [n]^d} n^{-d} \sum_{\bar{i} \in [n]^d} W_{n, \bar{i}\bar{j}}^k \right\} \leq \bar{W}_k, \quad k \in [4].$$

Proof. We prove (8.4) assuming that nonnegative W is in $L^2(Q^2)$, but not in $L^\infty(Q^2)$. In this case, $W_{n, \bar{i}\bar{j}}$ are defined by (2.11). For arbitrary $\bar{k} \in [n]^d$ and $n \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{j \in [n]^d} W_{n, \bar{i}, \bar{j}}^k &= \sum_{\bar{j} \in [n]^d} \left(n^{2d} \int_{Q_{n, \bar{i}} \times Q_{n, \bar{j}}} \alpha_n^{-1} \wedge W(x, y) dx dy \right)^k \\ &\leq \sum_{\bar{j} \in [n]^d} n^{2d} \int_{Q_{n, \bar{i}} \times Q_{n, \bar{j}}} (\alpha_n^{-k} \wedge W(x, y)^k) dx dy \\ &\leq n^{2d} \sum_{\bar{j} \in [n]^d} \int_{Q_{n, \bar{i}} \times Q_{n, \bar{j}}} W(x, y)^k dx dy \\ &= n^d \left(n^d \int_{Q_{n, \bar{i}}} W^k dx \right) \\ &\leq n^d \bar{W}_k, \end{aligned} \tag{8.5}$$

where we used Jensen's inequality in the second line and (W-1s) in the last line. Thus,

$$\sup_{n \in \mathbb{N}} \max_{\bar{i} \in [n]^d} n^{-d} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i}\bar{j}}^k \leq \bar{W}_k, \quad k \in [k].$$

The bound for $\sup_{n \in \mathbb{N}} \max_{\bar{j} \in [n]^d} n^{-d} \sum_{\bar{i} \in [n]^d} W_{n, \bar{i}\bar{j}}^k$ is proved similarly. $\square$

LEMMA 8.3. For $K \geq 2\bar{W}_1$, we have

$$(8.6) \quad \mathbb{P} \left(\max_{\bar{i} \in [n]^d} \sum_{\bar{j} \in [n]^d} \left| \frac{a_{n, \bar{i}\bar{j}}}{\alpha_n} - W_{n, \bar{i}\bar{j}} \right| \geq K n^d \right) \leq n^d \exp \left\{ \frac{-\frac{1}{2} (K - 2\bar{W}_1)^2 \alpha n^d}{\bar{W}_1 + O(\alpha_n) + K} \right\}.$$

In particular, with probability 1 there exists $n_0 \in \mathbb{N}$ such that

$$(8.7) \quad \max \left\{ \max_{\bar{i} \in [n]^d} \sum_{\bar{j} \in [n]^d} \left| \frac{a_{n, \bar{i}\bar{j}}}{\alpha_n} - W_{n, \bar{i}\bar{j}} \right|, \max_{\bar{i} \in [n]^d} \sum_{\bar{j} \in [n]^d} \left| \frac{a_{n, \bar{j}\bar{i}}}{\alpha_n} - W_{n, \bar{j}\bar{i}} \right| \right\} \leq K n^d$$

for all $n \geq n_0$.

For the next lemma, we will need the following notation:

$$(8.8) \quad Z_{n, \bar{i}}(t) = n^{-d} \sum_{\bar{j} \in [n]^d} b_{n, \bar{i}\bar{j}}(t) \eta_{n, \bar{i}\bar{j}},$$

$$(8.9) \quad b_{n, \bar{i}\bar{j}}(t) = D(v_{n, \bar{j}}(t) - v_{n, \bar{i}}(t)),$$

$$(8.10) \quad \eta_{n, \bar{i}\bar{j}} = a_{n, \bar{i}\bar{j}} - \alpha_n W_{n, \bar{i}\bar{j}},$$

and $Z_n = (Z_{n, \bar{i}}, \bar{i} \in [n]^d)$.

LEMMA 8.4. For arbitrary $\epsilon > 0$, we have

$$(8.11) \quad \alpha_n^{-2} \int_0^\infty e^{-Ls} \|Z_n(s)\|_{2, n^d}^2 ds \leq C(\alpha_n n^d)^{1-\epsilon},$$

where C is a positive constant independent of n and

$$(8.12) \quad \|Z_n(s)\|_{2, n^d} = \left(n^{-1} \sum_{\bar{j} \in [n]^d} Z_{n, \bar{j}}(s)^2 \right)^{1/2}.$$

Proof of Theorem 8.1. Recall that $f(u, x, t)$ and D are Lipschitz continuous function in u with Lipschitz constants L_f and L_D , respectively.

Further, $a_{n, \bar{i}\bar{j}}$, are Bernoulli random variables

$$(8.13) \quad \mathbb{P}(a_{n, \bar{i}\bar{j}} = 1) = \alpha_n W_{n, \bar{i}\bar{j}}.$$

Denote $\psi_{n, \bar{i}} := v_{n, \bar{i}} - u_{n, \bar{i}}$. By subtracting (2.3) from (3.4), multiplying the result by $n^{-d} \psi_{n, \bar{i}}$, and summing over $\bar{i} \in [n]^d$, we obtain

$$\begin{aligned}
(8.14) \quad \frac{1}{2} \frac{d}{dt} \|\psi_n\|_{2,n^d}^2 &= n^{-d} \underbrace{\sum_{\bar{i} \in [n]^d} (f(v_{n,\bar{i}}, t) - f(u_{n,\bar{i}}, t)) \psi_{n,\bar{i}}}_{I_1} \\
&\quad + n^{-2d} \alpha_n^{-1} \underbrace{\sum_{\bar{i}, \bar{j} \in [n]^d} (\alpha_n W_{n,\bar{i}\bar{j}} - a_{n,\bar{i}\bar{j}}) D(v_{n,\bar{j}} - v_{n,\bar{i}}) \psi_{n,\bar{i}}}_{I_2} \\
&\quad + n^{-2d} \alpha_n^{-1} \underbrace{\sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} [D(v_{n,\bar{j}} - v_{n,\bar{i}}) - D(u_{n,\bar{j}} - u_{n,\bar{i}})] \psi_{n,\bar{i}}}_{I_3} \\
&=: I_1 + I_2 + I_3,
\end{aligned}$$

where $\|\cdot\|_{2,n^d}^2$ is the discrete L^2 -norm (cf. (8.12)).

Using Lipschitz continuity of f in u , we have

$$(8.15) \quad |I_1| \leq L_f \|\psi_n\|_{2,n^d}^2.$$

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

$$\begin{aligned}
(8.16) \quad |I_3| &\leq L_D n^{-2d} \alpha_n^{-1} \sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} (|\psi_{n,\bar{i}}| + |\psi_{n,\bar{j}}|) |\psi_{n,\bar{i}}| \\
&\leq L_D n^{-2d} \alpha_n^{-1} \left(\frac{3}{2} \sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \psi_{n,\bar{i}}^2 + \frac{1}{2} \sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \psi_{n,\bar{j}}^2 \right).
\end{aligned}$$

Using Lemma 8.3 and (8.4), we obtain

$$\begin{aligned}
(8.17) \quad \alpha_n n^{-2d} \sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \psi_{n,\bar{i}}^2 &\leq n^{-d} \sum_{\bar{i} \in [n]^d} \left\{ n^{-d} \sum_{\bar{j} \in [n]^d} \left(\left| \frac{a_{n,\bar{i}\bar{j}}}{\alpha_n} - W_{n,\bar{i}\bar{j}} \right| + W_{n,\bar{i}\bar{j}} \right) \psi_{n,\bar{i}}^2 \right\} \\
&\leq n^{-d} \sum_{\bar{i} \in [n]^d} (K + \bar{W}_1) \psi_{n,\bar{i}}^2 = (K + W_1) \|\psi_n\|_{2,n^d}^2.
\end{aligned}$$

Similarly,

$$(8.18) \quad n^{-2d} \alpha_n^{-1} \sum_{\bar{i}, \bar{j} \in [n]^d} a_{n,\bar{i}\bar{j}} \psi_{n,\bar{j}}^2 \leq (K + \bar{W}_2) \|\psi_n\|_{2,n^d}^2.$$

By plugging (8.17) and (8.18) into (8.16), we have

$$(8.19) \quad |I_3| \leq L_D \left(2K + \frac{3}{2} \bar{W}_1 + \frac{1}{2} \bar{W}_2 \right) \|\psi\|_{n^d, 2}^2.$$

It remains to bound I_2 :

$$(8.20) \quad |I_2| = |n^{-d} \alpha_n^{-1} \sum_{\bar{i} \in [n]^d} Z_{n,\bar{i}} \psi_{n,\bar{i}}| \leq 2^{-1} \alpha_n^{-2} \|Z_n\|_{2,n^d}^2 + 2^{-1} \|\psi_n\|_{2,n^d}^2.$$

The combination of (8.14), (8.15), (8.19), and (8.20) yields

$$(8.21) \quad \frac{d}{dt} \|\psi_n(t)\|_{2,n}^2 \leq L \|\psi_n(t)\|_{2,n}^2 + \frac{1}{\alpha_n^2} \|Z_n(t)\|_{2,n}^2,$$

where $L = L_f + L_D (2K + \frac{3}{2}\bar{W}_1 + \frac{1}{2}\bar{W}_2) + \frac{1}{2}$.

Using the Gronwall's inequality and Lemma 8.4, we have

$$(8.22) \quad \sup_{t \in [0, T]} \|\psi_n(t)\|_{2,n^d}^2 \leq \alpha_n^{-2} e^{LT} \int_0^\infty e^{-Ls} \|Z_n(s)\|_{2,n^d}^2 ds \\ \leq \alpha_n^{-2} e^{LT} (n^d \alpha_n)^{-1+\epsilon}.$$

□

Proof of Lemma 8.3. Let

$$(8.23) \quad \xi_{n,\bar{i}\bar{j}} = \left| \frac{a_{n,\bar{i}\bar{j}}}{\alpha_n} - W_{n,\bar{i}\bar{j}} \right| - 2W_{n,\bar{i}\bar{j}} (1 - \alpha_n W_{n,\bar{i}\bar{j}}), \quad \bar{i}, \bar{j} \in [n]^d.$$

Note that for fixed $\bar{i} \in [n]^d$, $\{\xi_{n,\bar{i}\bar{j}}, \bar{j} \in [n]^d\}$ are mean zero independent random variables. Further, using the definition of $\xi_{n,\bar{i}\bar{j}}$, it is straightforward to bound

$$(8.24) \quad |\xi_{n,\bar{i}\bar{j}}| \leq \alpha_n^{-1} + 2W_{n,\bar{i}\bar{j}} \leq 3\alpha_n^{-1} =: M,$$

$$(8.25) \quad \mathbb{E} \xi_{n,\bar{j}}^2 \leq 2\alpha_n^{-1} W_{n,\bar{i}\bar{j}} + 2W_{n,\bar{i}\bar{j}}^2 + 4\alpha_n W_{n,\bar{i}\bar{j}}^2 + 4\alpha_n W_{n,\bar{i}\bar{j}}^3.$$

From (8.25), we have

$$(8.26) \quad \mathbb{E} \left(\sum_{\bar{j} \in [n]^d} \xi_{n,\bar{i}\bar{j}}^2 \right) \leq \alpha_n^{-1} \sum_{\bar{j} \in [n]^d} \left(2W_{n,\bar{i}\bar{j}} + \alpha_n 2W_{n,\bar{i}\bar{j}}^2 + 4\alpha_n^2 W_{n,\bar{i}\bar{j}}^2 + 4\alpha_n^2 W_{n,\bar{i}\bar{j}}^3 \right) \\ \leq \alpha_n^{-1} n^d W_1 + O(\alpha_n).$$

Using Bernstein's inequality and the union bound, we have

$$(8.27) \quad \mathbb{P} \left(\max_{\bar{i} \in [n]^d} \sum_{\bar{j} \in [n]^d} \xi_{n,\bar{i}\bar{j}} \geq (K - 2\bar{W}_1) n^d \right) \\ \leq n^d \exp \left\{ \frac{\frac{-1}{2} (K - 2\bar{W}_1)^2 n^{2d}}{\sum_{\bar{j} \in [n]^d} \mathbb{E} \xi_{n,\bar{i}\bar{j}}^2 + (1/3) M (K - 2\bar{W}_1) n^d} \right\} \\ \leq n^d \exp \left\{ \frac{\frac{-1}{2} (K - 2\bar{W}_1)^2 n^{2d}}{\alpha_n^{-1} n^d (\bar{W}_1 + O(\alpha_n)) + \alpha_n^{-1} (K - 2\bar{W}_1) n^d} \right\} \\ \leq N \exp \left\{ \frac{\frac{-1}{2} (K - 2\bar{W}_1)^2 \alpha_n n^d}{\bar{W}_1 + O(\alpha_n) + K} \right\}.$$

Finally, the combination of (8.23) and (8.27) yields

$$\begin{aligned}
 & \mathbb{P} \left(\max_{\bar{\imath} \in [n]^d} \sum_{\bar{j} \in [n]^d} \left| \frac{a_{n, \bar{i} \bar{j}}}{\alpha_n} - W_{n, \bar{i} \bar{j}} \right| \geq K n^d \right) \\
 & \leq \mathbb{P} \left(\max_{\bar{\imath} \in [n]^d} \sum_{\bar{j} \in [n]^d} \xi_{n, \bar{i} \bar{j}} \geq \left(K - \frac{2}{n^d} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i} \bar{j}} \right) n^d \right) \\
 & \leq \mathbb{P} \left(\max_{\bar{\imath} \in [n]^d} \sum_{\bar{j} \in [n]^d} \xi_{n, \bar{i} \bar{j}} \geq (K - 2\bar{W}_1) n^d \right) \\
 & \leq n^d \exp \left\{ \frac{-1}{2} \frac{(K - 2\bar{W}_1)^2 \alpha n^d}{\bar{W}_1 + O(\alpha_n) + K} \right\}.
 \end{aligned}$$

This proves (8.6). By the Borel–Cantelli lemma, (8.7) follows. $\square$

Proof of Lemma 8.4. Recall (8.8)–(8.10) and rewrite

$$(8.28) \quad \int_0^\infty e^{-Ls} \|Z_n(s)\|_{2, n^d}^2 ds = n^{-3d} \sum_{\bar{i}, \bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}},$$

where

$$(8.29) \quad c_{n, \bar{i} \bar{k} \bar{l}} = \int_0^\infty e^{-Ls} b_{n, \bar{i} \bar{k}}(s) b_{n, \bar{i} \bar{l}}(s) ds \quad \text{and} \quad |c_{n, \bar{i} \bar{k} \bar{l}}| \leq L^{-1} =: \bar{c}.$$

By (8.1), one can choose a sequence $\delta_n \searrow 0$ such that

$$(8.30) \quad n^d \delta_n \gg \alpha_n^{-1}.$$

Specifically, let

$$(8.31) \quad \delta_n := \frac{1}{\sqrt{\ln n}}$$

and define events

$$(8.32) \quad \Omega_n = \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{i}, \bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} > \delta_n n^d \right\},$$

$$(8.33) \quad A_{n, \bar{i}} = \left\{ \sum_{\bar{j} \in [n]^d} \left| \frac{a_{n, \bar{i} \bar{j}}}{\alpha_n} - W_{n, \bar{i} \bar{j}} \right| > K n^d \right\}, \text{ and } A_n = \bigcup_{\bar{i} \in [n]^d} A_{n, \bar{i}}.$$

Clearly,

$$(8.34) \quad \mathbb{P}(\Omega_n) \leq \mathbb{P}(\Omega_n \cap A_n^c) + \mathbb{P}(A_n).$$

We want to show that $\mathbb{P}(\Omega_n \text{ infinitely often}) = 0$. By the Borel–Cantelli lemma, it is sufficient to show that

$$\sum_{n \geq 1} \mathbb{P}(\Omega_n) < \infty.$$

From Lemma 8.3, we know that $\sum_{n \geq 1} \mathbb{P}(A_n) < \infty$ for $K > 2\bar{W}_1$. In the remainder of the proof, we show that $\sum_{n \geq 1} \mathbb{P}(\Omega_n \cap A_n^c)$ is convergent.

Applying the exponential Markov inequality to $\mathbb{P}(\Omega_n | A_n^c)$, from $\mathbb{P}(\Omega_n \cap A_n^c) \leq \mathbb{P}(\Omega_n | A_n^c)$ and (8.32), we have

$$(8.35) \quad \mathbb{P}(\Omega_n \cap A_n^c) \leq \exp \left\{ -n^d \delta_n + \ln \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{i}, \bar{k}, \bar{l} \in [n]^d}^n c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right] \right\}.$$

Using the independence of $\eta_{n, \bar{i} \bar{k}}$ with respect to $\bar{l}$, we have

$$(8.36) \quad \begin{aligned} & \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{i}, \bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right] \\ &= \prod_{\bar{i} \in [n]^d} \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right]. \end{aligned}$$

Using

$$e^x \leq 1 + |x|e^{|x|}, \quad x \in \mathbb{R},$$

and the Cauchy–Schwarz inequality, we bound the right-hand side of (8.36) as follows:

$$(8.37) \quad \begin{aligned} & \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right] \\ & \leq 1 + \mathbb{E} \left[\mathbf{1}_{A_n^c} \left| (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right| \right] \\ & \quad \times \exp \left\{ \left| (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right| \right\} \\ & \leq 1 + \left(\mathbb{E} \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\}^2 \right)^{1/2} \\ & \quad \times \left(\mathbb{E} \left\{ \mathbf{1}_{A_n^c} \exp \left\{ 2(n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right\} \right)^{1/2}. \end{aligned}$$

From (8.10), (8.29), and under A_n^c (cf. (8.33)), we have

$$(8.38) \quad \mathbf{1}_{A_n^c} 2(n^d \alpha_n)^{-2} \left| \sum_{\bar{k}, \bar{l} \in [n]^d}^n c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right| \leq 2K\bar{c}.$$

Further,

$$\begin{aligned}
 (8.39) \quad & \mathbb{E} \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\}^2 \\
 & \leq (n^d \alpha_n)^{-4} \sum_{\bar{j}, \bar{k}, \bar{l}, \bar{p} \in [n]^d} \mathbb{E} (\eta_{n, \bar{i} \bar{j}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \eta_{n, \bar{i} \bar{p}}) c_{n, \bar{i} \bar{j} \bar{p}} c_{n, \bar{i} \bar{k} \bar{l}} \\
 & \leq \frac{(\bar{c})^2}{(n^d \alpha_n)^4} \left\{ \sum_{\bar{j} \in [n]^d} \mathbb{E} \eta_{n, \bar{i} \bar{j}}^4 + 6 \left(\sum_{\bar{j} \in [n]^d} \mathbb{E} \eta_{n, \bar{i} \bar{j}}^2 \right)^2 \right\}.
 \end{aligned}$$

Using (8.10), we estimate the sum of the fourth moments of $\eta_{n, \bar{i} \bar{j}}$,

$$\begin{aligned}
 (8.40) \quad & \sum_{\bar{j} \in [n]^d} \mathbb{E} \eta_{n, \bar{i} \bar{j}}^4 = \sum_{\bar{j} \in [n]^d} \left\{ \alpha_n W_{n, \bar{i} \bar{j}} (1 - \alpha_n W_{n, \bar{i} \bar{j}})^4 + \alpha_n^4 W_{n, \bar{i} \bar{j}}^4 (1 - \alpha_n W_{n, \bar{i} \bar{j}}) \right\} \\
 & \leq n^d \alpha_n \left(N^{-1} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i} \bar{j}} + \alpha_n^3 n^{-d} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i} \bar{j}}^4 \right) \\
 & \leq n^d \alpha_n (\bar{W}_1 + \alpha_n^3 \bar{W}_4) = O(\alpha_n n^d),
 \end{aligned}$$

where we also use (8.4). Similarly,

$$\begin{aligned}
 (8.41) \quad & \sum_{\bar{j} \in [n]^d} \mathbb{E} \eta_{n, \bar{i} \bar{j}}^2 = \sum_{\bar{j} \in [n]^d} \left\{ \alpha_n W_{n, \bar{i} \bar{j}} (1 - \alpha_n W_{n, \bar{i} \bar{j}})^2 + \alpha_n^2 W_{n, \bar{i} \bar{j}}^2 (1 - \alpha_n W_{n, \bar{i} \bar{j}}) \right\} \\
 & \leq n^d \alpha_n \left(N^{-1} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i} \bar{j}} + \alpha_n n^{-d} \sum_{\bar{j} \in [n]^d} W_{n, \bar{i} \bar{j}}^2 \right) \\
 & \leq n^d \alpha_n (\bar{W}_1 + \alpha_n \bar{W}_2) = O(\alpha_n N).
 \end{aligned}$$

By combining (8.39)–(8.41), we obtain

$$(8.42) \quad \mathbb{E} \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\}^2 = O((n^d \alpha_n)^{-2}).$$

By plugging (8.38) and (8.42) into (8.37), we obtain

$$(8.43) \quad \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right] \leq 1 + \frac{C_1}{n^d \alpha_n} e^{C_2}.$$

Using this bound on the right-hand side of (8.36), we further obtain

$$(8.44) \quad \mathbb{E} \left[\mathbf{1}_{A_n^c} \exp \left\{ (n^d \alpha_n)^{-2} \sum_{\bar{i}, \bar{k}, \bar{l} \in [n]^d} c_{n, \bar{i} \bar{k} \bar{l}} \eta_{n, \bar{i} \bar{k}} \eta_{n, \bar{i} \bar{l}} \right\} \right] \leq e^{C_3 \alpha_n^{-1}}.$$

Using (8.44), from (8.35) we obtain

$$(8.45) \quad \mathbb{P}(\Omega_n \cap A_n^c) \leq \exp \{ -n^d \delta_{n^d} + C_3 \alpha_n^{-1} \} \rightarrow 0, \quad n \rightarrow \infty.$$

Furthermore, using (8.31) it is straightforward to check that

$$\sum_{n=1}^{\infty} \mathbb{P}(\Omega_n \cap A_n^c) < \infty.$$

The statement of the lemma then follows from (8.32)–(8.34) via the Borel–Cantelli lemma. $\square$

REFERENCES

- [1] N. I. AKHIESER, *Theory of Approximation*, Dover Publications, New York, 1992.
- [2] F. ANDREU-VAILLO, J. M. MAZÓN, J. D. ROSSI, AND J. J. TOLEDO-MELERO, *Nonlocal Diffusion Problems*, Math. Surveys Monographs 165, AMS, Providence, RI, 2010.
- [3] K. E. ATKINSON, *An Introduction to Numerical Analysis*, 2nd ed., John Wiley & Sons, New York, 1989.
- [4] L. BANJAI, J. M. MELENK, R. H. NOCHETTO, E. OTÁROLA, A. J. SALGADO, AND C. SCHWAB, *Tensor FEM for spectral fractional diffusion*, *Found. Comput. Math.*, 19 (2019), pp. 901–962.
- [5] A. BONITO, J. P. BORTHAGARAY, R. H. NOCHETTO, E. OTÁROLA, AND A. J. SALGADO, *Numerical methods for fractional diffusion*, *Comput. Vis. Sci.*, 19 (2018), pp. 19–46.
- [6] 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*, *Trans. Amer. Math. Soc.*, 372 (2019), pp. 3019–3062.
- [7] N. BOURNAVEAS AND V. CALVEZ, *The one-dimensional Keller-Segel model with fractional diffusion of cells*, *Nonlinearity*, 23 (2010), pp. 923–935.
- [8] C. CARRILLO AND P. FIFE, *Spatial effects in discrete generation population models*, *J. Math. Biol.*, 50 (2005), pp. 161–188.
- [9] S. COOMBES, P. BEIM GRABEN, AND R. POTTHAST, *Tutorial on neural field theory*, in *Neural Fields*, Springer, Heidelberg, 2014, pp. 1–43.
- [10] F. COPPINI, H. DIETERT, AND G. GIACOMINI, *A Law of Large Numbers and Large Deviations for Interacting Diffusions on Erdos-Renyi Graphs*, *Stoch. Dyn.* 20 (2020), no. 2.
- [11] A. DE PABLO, F. QUIRÓS, AND A. RODRÍGUEZ, *Nonlocal filtration equations with rough kernels*, *Nonlinear Anal.*, 137 (2016), pp. 402–425.
- [12] F. DEL TESO, J. ENDAL, AND E. R. JAKOBSEN, *Uniqueness and properties of distributional solutions of nonlocal equations of porous medium type*, *Adv. Math.*, 305 (2017), pp. 78–143.
- [13] R. A. DEVORE AND G. G. LORENTZ, *Constructive Approximation*, Grundlehren Math. Wiss. 303, Springer-Verlag, Berlin, 1993.
- [14] Q. DU, *An invitation to nonlocal modeling, analysis, and computation*, in *Proceedings of the International Conference of Mathematicians*, Vol. 3, 2018, pp. 3523–3552.
- [15] Q. DU, L. JU, AND J. LU, *A discontinuous Galerkin method for one-dimensional time-dependent nonlocal diffusion problems*, *Math. Comp.*, 88 (2019), pp. 123–147.
- [16] L. C. EVANS AND R. F. GARIEPY, *Measure Theory and Fine Properties of Functions*, Textb. Math., revised ed., CRC Press, Boca Raton, FL, 2015.
- [17] K. FALCONER, *Fractal Geometry*, Math. Found. Appl., 3rd ed., John Wiley & Sons, Chichester, UK, 2014.
- [18] F. GOLSE, *On the dynamics of large particle systems in the mean field limit*, in *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, *Lect. Notes Appl. Math. Mech.* 3, Springer, Cham, 2016, pp. 1–144.
- [19] D. KALIUZHNYI-VERBOVETSKYI AND G. S. MEDVEDEV, *The semilinear heat equation on sparse random graphs*, *SIAM J. Math. Anal.*, 49 (2017), pp. 1333–1355.
- [20] C. R. LAING, *Chimeras in two-dimensional domains: Heterogeneity and the continuum limit*, *SIAM J. Appl. Dyn. Syst.*, 16 (2017), pp. 974–1014.
- [21] L. LOVÁSZ, *Large Networks and Graph Limits*, AMS, Providence, RI, 2012.
- [22] L. LOVÁSZ AND B. SZEGEDY, *Limits of dense graph sequences*, *J. Combin. Theory Ser. B*, 96 (2006), pp. 933–957.
- [23] G. S. MEDVEDEV, *The nonlinear heat equation on dense graphs and graph limits*, *SIAM J. Math. Anal.*, 46 (2014), pp. 2743–2766.
- [24] G. S. MEDVEDEV, *The nonlinear heat equation on W -random graphs*, *Arch. Ration. Mech. Anal.*, 212 (2014), pp. 781–803.

- [25] G. S. MEDVEDEV, *Small-world networks of Kuramoto oscillators*, Phys. D, 266 (2014), pp. 13–22.
- [26] G. S. MEDVEDEV, *The continuum limit of the Kuramoto model on sparse random graphs*, Commun. Math. Sci., 17 (2019), pp. 883–898.
- [27] G. S. MEDVEDEV AND X. TANG, *Stability of twisted states in the Kuramoto model on Cayley and random graphs*, J. Nonlinear Sci., 25 (2015), pp. 1169–1208.
- [28] G. S. MEDVEDEV AND X. TANG, *The Kuramoto model on power law graphs: Synchronization and contrast states*, J. Nonlinear Sci., 30 (2020), pp. 2405–2427.
- [29] G. S. MEDVEDEV AND J. D. WRIGHT, *Stability of twisted states in the continuum Kuramoto model*, SIAM J. Appl. Dyn. Syst., 16 (2017), pp. 188–203.
- [30] R. H. NOCHETTO, E. OTÁROLA, AND A. J. SALGADO, *A PDE approach to space-time fractional parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 848–873.
- [31] R. H. NOCHETTO, E. OTÁROLA, AND A. J. SALGADO, *A PDE approach to space-time fractional parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 848–873.
- [32] R. W. SCHWAB AND L. SILVESTRE, *Regularity for parabolic integro-differential equations with very irregular kernels*, Anal. PDE, 9 (2016), pp. 727–772.
- [33] W. SHEN AND X. XIE, *Approximations of random dispersal operators/equations by nonlocal dispersal operators/equations*, J. Differential Equations, 259 (2015), pp. 7375–7405.
- [34] J. L. VÁZQUEZ, *The mathematical theories of diffusion: Nonlinear and fractional diffusion*, in Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions, Lecture Notes in Math. 2186, Springer, Cham, 2017, pp. 205–278.
- [35] D. A. WILEY, S. H. STROGATZ, AND M. GIRVAN, *The size of the sync basin*, Chaos, 16 (2006), 015103.
- [36] D. WILLIAMS, *Probability with Martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, UK, 1991.