A convergent finite volume method for the Kuramoto equation and related nonlocal conservation laws

N. Chatterjee and U. S. Fjordholm*

Department of Mathematics, University of Oslo, Postboks 1053 Blindern, 0316 Oslo, Norway
*Corresponding author: ulriksf@math.uio.no

We derive and study a Lax–Friedrichs-type finite volume method for a large class of nonlocal continuity equations in multiple dimensions. We prove that the method converges weakly to the measure-valued solution and converges strongly if the initial data is of bounded variation. Several numerical examples for the kinetic Kuramoto equation are provided, demonstrating that the method works well for both regular and singular data.

Keywords: nonlocal conservation laws; finite volume methods; Kuramoto equation.

1. Introduction

The purpose of the present paper is to derive a nonoscillatory and convergent numerical method for a large class of nonlocal continuity equations of the form

$$\partial_t \mu + \nabla \cdot \big(V[\mu]\mu\big) = 0, \qquad x \in \mathcal{K}, \ t > 0,$$
$$\mu(x,0) = \mu_0(x), \qquad x \in \mathcal{K} \tag{1.1}$$

on some domain $\mathcal{K} \subset \mathbb{R}^d$. Here $\mu(x,t)$ is interpreted as a density of particles at time t, and $V[\mu] : \mathcal{K} \to \mathbb{R}^d$ depends nonlocally on μ, typically the convolution of μ with some kernel. The initial data is assumed to be some probability measure $\mu_0 \in \mathcal{P}(\mathcal{K})$. Conservation laws of this form arise in a large variety of applications, such as the simulation of crowd dynamics, traffic flow (Colombo *et al.*, 2011), microbiology (Buck & Buck, 1966), in the analytical study of chemotaxis (James & Vauchelet, 2013), the consensus problem (Bouchut & James, 1998; James & Vauchelet, 2015) and more (Buck & Buck, 1966; Winfree, 1967; Kuramoto, 1975). One particular instance of (1.1) which has been studied extensively is the so-called *Kuramoto–Sakaguchi equation*, also called the *kinetic Kuramoto equation*,

$$\partial_t f + \partial_\theta(\omega[f]f) = 0, \qquad (\theta, \Omega) \in \mathbb{T} \times \mathbb{R}, \ t > 0,$$
$$f(\theta, \Omega, 0) = f_0(\theta, \Omega), \qquad (\theta, \Omega) \in \mathbb{T} \times \mathbb{R}, \tag{1.2}$$

where $\mathbb{T} = \mathbb{R}/2\pi$ is the one-dimensional torus and

$$\omega[f](\theta, \Omega, t) := \Omega - K \int_{\mathbb{T}} \sin(\theta - \theta^*)\rho(\theta^*, t) \, \mathrm{d}\theta^*, \qquad \rho(\theta, t) := \int_{\mathbb{R}} f(\theta, \Omega, t) \, \mathrm{d}\Omega.$$

This equation arises as the mean-field limit of the *Kuramoto equation*, a system of ordinary differential equations for coupled oscillators that was first studied by Winfree and Kuramoto (Winfree, 1967; Kuramoto, 1975, 1984). The unknowns in the Kuramoto equation are the phase $\theta \in \mathbb{T}$ and the natural frequency $\Omega \in \mathbb{R}$ of each oscillator, and the interaction between the oscillators depends on the coupling strength $K > 0$ and the relative difference in phase $\theta - \theta^*$ between pairs of oscillators. The mean-field limit as the number of oscillators goes to infinity is a probability distribution $f = f(\theta, \Omega, t)$ obeying the above nonlocal continuity equation; see e.g., Dobrushin (1979), Neunzert (1984), Chiba (2013). See also the recent paper Carrillo *et al.* (2014) for some qualitative properties of (1.2). We will also let $g(\Omega, t) := \int_{\mathbb{T}} f(\theta, \Omega, t) \, \mathrm{d}\theta$ denote the distribution function for natural frequencies. From (1.2) it is easily seen that g is constant in time: $g(\Omega, t) \equiv g(\Omega)$. Here and elsewhere we will assume that f_0 (and hence also g) is compactly supported.

The Kuramoto equation (1.2) is valid for oscillators with nonidentical natural frequencies. The situation is somewhat simpler in the case of identical oscillators, i.e., oscillators whose natural frequencies coincide. This corresponds to g being a Dirac measure, and without loss of generality we can assume that $g = \delta$, the Dirac measure located at $\Omega = 0$. Consequently,

$$f(\theta, \Omega, t) = \rho(\theta, t)\delta(\Omega), \qquad \omega[f]f = L[\rho]\rho, \qquad L[\rho](\theta, t) := -K \int_{\mathbb{T}} \sin(\theta - \theta^*)\rho(\theta^*, t) \, \mathrm{d}\theta^*.$$

Therefore, (1.2) reduces to the following equation for ρ:

$$\begin{aligned} \partial_t \rho + \partial_\theta (L[\rho]\rho) &= 0, \\ \rho(\theta, 0) &= \rho_0(\theta). \end{aligned} \tag{1.3}$$

Both equations (1.2) and (1.3) are instances of general nonlocal conservation laws of the form (1.1).

The purpose of the present paper is to derive and analyze a finite volume numerical method for a large class of equations of the form (1.1) (including the kinetic Kuramoto equations (1.2) and (1.3)). In particular, we prove that the scheme converges strongly to the unique weak solution μ whenever $\mu_0 \in L^1 \cap \mathrm{BV}$, and in all other cases converges in the sense of measures to the so-called *measure-valued* solution μ. We emphasize that the stability and convergence properties of the scheme are valid even when μ_0 (and hence also the exact solution) has point-mass singularities.

In the recent works Amadori *et al.* (2017a; 2017b) a front-tracking numerical method for the Kuramoto–Sakaguchi equation for identical oscillators (1.3) was designed and analyzed. In contrast to their method our finite volume method works in any number of dimensions, does not require any regularity on the initial data and does not impose 'entropy conditions' on solutions. We also mention the work by Crippa & Lécureux-Mercier (2013) where the well-posedness of a *system* of nonlocal continuity equations of the form (1.1) is established. The extension of our finite volume scheme to such systems should be straightforward.

1.1 *Nonlocal conservation laws*

Nonlocal conservation laws of the form (1.1) were first studied by Neunzert (1984) and Dobrushin (1979). Using techniques that by now are standard, they showed existence and uniqueness of solutions of (1.1). We will consider solutions that are weakly continuous maps $\mu : \mathbb{R}_+ \to \mathcal{M}(\mathcal{K})$, mapping time t to probability measures $\mu_t \in \mathcal{P}(\mathcal{K})$. Here $\mathcal{M}(\mathcal{K}) = (C_0(\mathcal{K}))^*$ is the space of bounded Radon measures on $\mathcal{K}$, $\mathcal{P}(\mathcal{K})$ is the subset of probability measures and by weakly continuous we mean that

$t \mapsto \langle \mu_t, \varphi \rangle := \int_{\mathcal{K}} \varphi(x) \, \mathrm{d}\mu_t(x)$ is continuous for every $\varphi \in C_0(\mathcal{K})$. We define the 1-*Wasserstein* metric

$$d_W(\nu_1, \nu_2) := \sup \left\{ \int_{\mathcal{K}} \psi(x) \, \mathrm{d}(\nu_2 - \nu_1)(x) \, : \, \psi \in C_b(\mathcal{K}), \, \|\psi\|_{\mathrm{Lip}(\mathcal{K})} \leqslant 1 \right\}, \quad \nu_1, \nu_2 \in \mathcal{P}(\mathcal{K}).$$

It can be shown that d_W metrizes the topology of weak (or *narrow*) convergence in $\mathcal{P}_1(\mathcal{K})$, the set of probability measures ν with finite first moment, $\int_{\mathcal{K}} |x| \, \mathrm{d}\nu(x) < \infty$ (see e.g., Villani, 2003).

DEFINITION 1.1   A map $\mu \in C_w(\mathbb{R}_+; \mathcal{M}(\mathcal{K}))$ is a *measure-valued solution* of (1.1) if it satisfies (1.1) in the sense of distributions, i.e.,

$$\int_0^\infty \int_{\mathcal{K}} \partial_t \varphi(x, t) + V[\mu_t](x) \cdot \nabla \varphi(x, t) \, \mathrm{d}\mu_t(x) \, \mathrm{d}t + \int_{\mathcal{K}} \varphi(x, 0) \, \mathrm{d}\mu_0(x) = 0 \tag{1.4}$$

for every $\varphi \in C_c^\infty(\mathcal{K} \times \mathbb{R}_+)$.

Since μ is assumed to be weakly continuous in time one can show that μ is a measure-valued solution *if and only if*

$$\langle \mu_t, \psi \rangle - \langle \mu_s, \psi \rangle = \int_s^t \int_{\mathcal{K}} V[\mu_\tau](x) \cdot \nabla \psi(x) \, \mathrm{d}\mu_\tau(x) \, \mathrm{d}\tau \tag{1.5}$$

for every $0 \leqslant s < t$ and every $\psi \in C_c^\infty(\mathcal{K})$. We say that μ is a *weak solution* of (1.1) if it is a measure-valued solution that is absolutely continuous with respect to Lebesgue measure.

We will henceforth assume that $\mathcal{K}$ is of the form $\mathcal{K} = \mathbb{K}_1 \times \cdots \times \mathbb{K}_d$, where each $\mathbb{K}_i$ is either the torus $\mathbb{T}$ or the whole real line $\mathbb{R}$. For the well-posedness of (1.1) and the convergence of our numerical scheme, we also need some of the following assumptions on V:

(A1) $\exists \, C_1 > 0$ such that $\|V[\nu]\|_{L^\infty(\mathcal{K})} \leqslant C_1$ for all $\nu \in \mathcal{P}(\mathcal{K})$;

(A2) $\exists \, C_2 > 0$ such that $|V[\nu](x) - V[\nu](y)| \leqslant C_2 |x - y|$ for all $x, y \in \mathcal{K}$ and $\nu \in \mathcal{P}(\mathcal{K})$;

(A3) $\exists \, C_3 > 0$ such that $\|V[\nu_1] - V[\nu_2]\|_{L^\infty(\mathcal{K})} \leqslant C_3 d_W(\nu_1, \nu_2)$ for all $\nu_1, \nu_2 \in \mathcal{P}(\mathcal{K})$;

(A4) $\exists \, C_4 > 0$ such that $\|\nabla V[\nu]\|_{\mathrm{Lip}(\mathcal{K})} \leqslant C_4$ for all $\nu \in \mathcal{P}(\mathcal{K})$.

The main well-posedness result for the nonlocal conservation law is the following.

THEOREM 1.2 (Neunzert, 1984, Dobrushin, 1979). Let $\mu_0 \in \mathcal{P}_1(\mathcal{K})$ and let V satisfy assumptions (A1)–(A3). Then there exists a unique measure-valued solution of (1.1). This solution is Lipschitz in time, $d_W(\mu_t, \mu_s) \leqslant C_1 |t - s|$. If $\mu_0 \in L^1(\mathcal{K})$ then $\mu_t \in L^1(\mathcal{K})$ for all $t > 0$.

REMARK 1.3   It is easy to see that both kinetic Kuramoto equations (1.2) and (1.3) satisfy assumptions (A1)–(A4) and are therefore well posed by Theorem 1.2. Indeed, in the case of (1.2) we let $d = 2$, $\mathcal{K} = \mathbb{T} \times \mathbb{R}$ and $V[\nu] = (V^1[\nu], V^2[\nu])$, where $V^2[\nu] \equiv 0$ and

$$V^1[\nu](\theta, \Omega) = \begin{cases} \Omega - K \int_{\mathcal{K}} \sin(\theta - \theta^*) \, \mathrm{d}\nu(\theta^*, \Omega^*) & \text{for } \Omega \in \mathrm{supp} \, g, \\ 0 & \text{otherwise.} \end{cases}$$

We can then choose $C_1 = K + \max\{|\Omega| : \Omega \in \text{supp } g\}$, $C_2 = 1 + K$ and $C_3 = C_4 = K$. In the case of (1.3) we let $d = 1$, $\mathcal{K} = \mathbb{T}$, $V = L$ and $C_1 = C_2 = C_3 = C_4 = K$.

'Kinetic' partial differential equations (PDEs) (1.1) under our assumptions (A1)–(A3) are rather well behaved under approximations. For instance, in a *particle approximation* one approximates the solution as a convex combination of Dirac measures, $\mu_t^M = \sum_{i=1}^M \delta_{x_i(t)} \mu_i$, where $x_i(t)$ is the position of the *i*th particle and $\mu_i \in (0, \infty)$ is its mass. It is straightforward to see that μ^M is in fact a measure-valued solution of (1.1) provided $x_i(t)$ satisfies the system of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}x_i}{\mathrm{d}t}(t) = V[\mu^M](x_i(t)). \tag{1.6}$$

Assumptions (A1) and (A2) guarantee that this system has a unique solution. Taking an approximating sequence μ_0^M converging weakly to μ_0, assumption (A3) guarantees that the limit is a measure-valued solution. Thus, if one can solve the system of ODEs (1.6) then the question of convergence boils down to the approximation of the initial data by Dirac measures (see Neunzert, 1978, 1984; Dobrushin, 1979; Lancelotti, 2005).

Although the dynamical system (1.6) can be studied qualitatively in certain simple cases, it is computationally infeasible to solve (1.6) in more realistic settings. In the next section we proceed to construct a simple, computationally efficient numerical approximation to the nonlocal PDE (1.1).

2. The Lax–Friedrichs scheme

2.1 *Derivation of the method*

For the sake of simplicity we derive the method in one space dimension with either $\mathcal{K} = \mathbb{R}$ or $\mathcal{K} = \mathbb{T}$ and then simply state the method in multiple space dimensions (Section 2.5). We start by deriving a *staggered* version of the method but—again for the sake of simplicity—we will analyze only the unstaggered version of this method.

Consider a mesh $\cdots < x_{i-1/2} < x_{i+1/2} < \cdots$, where *i* runs over $\mathcal{I} := \mathbb{Z}$ in the unbounded case $\mathcal{K} = \mathbb{R}$ or over some finite set $\mathcal{I} := \{1, \ldots, N_x\}$ in the periodic case $\mathcal{K} = \mathbb{T}$ and such that $\bigcup_i [x_{i-1/2}, x_{i+1/2}) = \mathcal{K}$. For simplicity we assume a uniform mesh, $x_{i+1/2} - x_{i-1/2} \equiv \Delta x$. We will denote $x_i := \frac{x_{i-1/2} + x_{i+1/2}}{2}$. Let $0 = t^0 < t^1 < \cdots < t^N = T$ be a partition of the time interval $[0, T]$ with uniform step size $t^{n+1} - t^n \equiv \Delta t$. Assuming that we are given an approximate solution $\mu^n = \sum_i \delta_{x_i} \mu_i^n$ at time t^n we compute an approximation μ^{n+1} at $t = t^{n+1}$ as follows. Let μ_t be the exact solution of

$$\partial_t \mu + \partial_x \big(V[\mu]\mu\big) = 0, \qquad t^n < t < t^{n+1},$$
$$\mu_{t^n} = \mu^n. \tag{2.1}$$

Define $\mu_{i+1/2}^{n+1} := \big\langle \mu_{t^{n+1}}, \psi_{i+1/2} \big\rangle$, where ψ are the usual 'witch's hat' finite element basis functions,

$$\psi_{i+1/2}(x) := \begin{cases} \frac{x - x_{i-1/2}}{\Delta x} & x_{i-1/2} \leqslant x < x_{i+1/2}, \\ \frac{x_{i+3/2} - x}{\Delta x} & x_{i+1/2} \leqslant x < x_{i+3/2}, \\ 0 & \text{otherwise}, \end{cases}$$

and the index *i* is taken over all $i \in \mathcal{I}$ if *n* is even and over $i \in \mathcal{I} - 1/2$ if *n* is odd. The approximation μ^{n+1} at time t^{n+1} is then defined to be the projected solution $\mu^{n+1} := \sum_i \mu_{i+1/2}^{n+1} \delta_{x_{i+1/2}}$.

We can derive a simplified expression of $\mu_{i+1/2}^{n+1}$ as follows. Since the initial data μ^n in (2.1) is a convex combination of Dirac measures, the solution of (2.1) can be written as $\mu_t = \sum_i \delta_{x_i(t)} \mu_i^n$, where x_i solves the system of ODEs (1.6). If we assume the *Courant–Friedrichs–Lewy (CFL) condition*

$$\frac{\Delta t}{\Delta x} \| V[\mu] \|_{L^\infty(\mathcal{K})} \leqslant \frac{1}{2}, \tag{2.2}$$

then the particle with position $x_i(t)$ stays in the interval $(x_{i-1/2}, x_{i+1/2})$ for all $t \in [t^n, t^{n+1})$. In particular, the atoms $\delta_{x_i(t)}$ in μ_t stay away from the kinks in $\psi_{i+1/2}$, and hence we may use the (nondifferentiable) function $\psi_{i+1/2}$ as a test function in the weak formulation (1.5). Using the fact that $\langle \mu^n, \psi_{i+1/2} \rangle = \frac{\mu_i^n + \mu_{i+1}^n}{2}$, we obtain

$$
\begin{aligned}
\mu_{i+1/2}^{n+1} &= \frac{\mu_i^n + \mu_{i+1}^n}{2} + \int_{t^n}^{t^{n+1}} \int_{\mathbb{R}} V[\mu_t](x) \partial_x \psi(x) \, \mathrm{d}\mu_t(x) \mathrm{d}t \\
&= \frac{\mu_i^n + \mu_{i+1}^n}{2} - \frac{1}{\Delta x} \int_{t^n}^{t^{n+1}} V[\mu_t]\big(x_{i+1}(t)\big)\mu_{i+1}^n - V[\mu_t]\big(x_i(t)\big)\mu_i^n \, \mathrm{d}t.
\end{aligned}
$$

Approximating $V[\mu_t](x_i(t)) \approx V[\mu^n](x_i)$ yields our final scheme,

$$\mu_{i+1/2}^{n+1} = \frac{\mu_i^n + \mu_{i+1}^n}{2} - \Delta t \frac{V[\mu^n]_{i+1}\mu_{i+1}^n - V[\mu^n]_i \mu_i^n}{\Delta x} \tag{2.3}$$

(where we denote $V[\mu^n]_i := V[\mu^n](x_i)$). The initial data is set as $\mu_i^0 = \langle \mu_0, \psi_i \rangle$. We refer to (2.3) as the *staggered Lax–Friedrichs method*, after Nessyahu & Tadmor (1990).

For the sake of simplicity, we will consider only an unstaggered version of (2.3), obtained by inserting an extra mesh point between all pairs of neighboring mesh points. The *unstaggered Lax–Friedrichs method* is then

$$
\begin{aligned}
\mu_i^0 &= \langle \mu_0, \psi_i \rangle, \\
\mu_i^{n+1} &= \frac{\mu_{i-1}^n + \mu_{i+1}^n}{2} - \Delta t \frac{V[\mu^n]_{i+1}\mu_{i+1}^n - V[\mu^n]_{i-1}\mu_{i-1}^n}{2\Delta x},
\end{aligned}
\tag{2.4}
$$

where $\mu^n = \sum_i \mu_i^n \delta_{x_i}$.

2.2 *Properties of the method*

In this section we prove several stability properties of the unstaggered Lax–Friedrichs method. Since the method shares the same properties as the one-dimensional method, we will prove these properties for the one-dimensional method (2.4) and merely state the properties for the two-dimensional method (2.9) in Section 2.5.

PROPOSITION 2.1 Let $\mu_0 \in \mathcal{P}(\mathcal{K})$. Consider the unstaggered, one-dimensional Lax–Friedrichs method (2.4) and define the piecewise linear measure-valued map

$$\mu_t^{\Delta x} := \frac{t^{n+1} - t}{\Delta t} \mu^n + \frac{t - t^n}{\Delta t} \mu^{n+1} \qquad \text{for } t \in [t^n, t^{n+1}), \tag{2.5}$$

where $\mu^n := \sum_i \mu_i^n \delta_{x_i}$. Assume that V satisfies condition (A1) and that the CFL condition

$$\lambda_0 \leqslant \frac{\Delta t}{\Delta x} \leqslant C_1^{-1} \tag{2.6}$$

is satisfied for some $\lambda_0 > 0$ (where C_1 is the constant in assumption (A1)). Then for all $t \geqslant 0$,

  (i) $\mu_t^{\Delta x} \geqslant 0$;

  (ii) $\|\mu_t^{\Delta x}\|_{\mathcal{M}(\mathbb{K})} = 1$;

  (iii) if supp $\mu_0 \subset B_M(0)$ then supp $\mu_t^{\Delta x} \subset B_{M+\lambda_0^{-1}t}(0)$;

  (iv) $d_W\big(\mu_t^{\Delta x}, \mu_s^{\Delta x}\big) \leqslant \lambda_0^{-1}|t-s|$ for all $s \geqslant 0$.

Proof. Writing

$$\mu_i^{n+1} = \mu_{i-1}^n \frac{1 + \frac{\Delta t}{\Delta x}V[\mu^n]_{i-1}}{2} + \mu_{i+1}^n \frac{1 - \frac{\Delta t}{\Delta x}V[\mu^n]_{i+1}}{2},$$

it is clear that (2.6) and assumption (A1) ensure that μ_i^{n+1} is a linear combination of non-negative terms, whence (i) follows. Property (ii) follows from the fact that the method (2.4) is conservative. From (2.4) we see that if $\mu_{i-1}^n = \mu_i^n = \mu_{i+1}^n = 0$ then $\mu_i^{n+1} = 0$. Hence, after n time steps the support of $\mu_t^{\Delta x}$ can have grown at most a distance $n\Delta x \leqslant n\frac{\Delta t}{\lambda_0} = \lambda_0^{-1}t^n$, and (iii) follows. For (iv) it is clear that by the definition (2.5) of $\mu^{\Delta x}$, it is enough to prove the claim for $s = t^n, t = t^{n+1}$. Let $\psi \in C_b(\mathbb{K})$ be Lipschitz continuous and write $\psi_i = \psi(x_i)$. Then

$$\langle \mu^{n+1} - \mu^n, \psi \rangle = \sum_i \psi_i \left(\frac{\mu_{i-1}^n - 2\mu_i^n + \mu_{i+1}^n}{2} - \frac{\Delta t}{2\Delta x}\big(V[\mu^n]_{i+1}\mu_{i+1}^n - V[\mu^n]_{i-1}\mu_{i-1}^n\big) \right)$$

(summation by parts)

$$= -\frac{1}{2}\sum_i (\psi_{i+1} - \psi_i)\left(\mu_{i+1}^n - \mu_i^n - \frac{\Delta t}{\Delta x}\big(V[\mu^n]_i\mu_i^n + V[\mu^n]_{i+1}\mu_{i+1}^n\big) \right)$$

$$= \frac{1}{2}\sum_i (\psi_{i+1} - \psi_i)\left(\mu_i^n\left(1 + \frac{\Delta t}{\Delta x}V[\mu^n]_i\right) + \mu_{i+1}^n\left(\frac{\Delta t}{\Delta x}V[\mu^n]_{i+1} - 1\right) \right)$$

$$\leqslant \frac{1}{2}\Delta x\|\psi\|_{\mathrm{Lip}}\sum_i \mu_i^n\left(1 + \frac{\Delta t}{\Delta x}V[\mu^n]_i\right) + \mu_{i+1}^n\left(1 - \frac{\Delta t}{\Delta x}V[\mu^n]_{i+1}\right)$$

$$= \Delta x\|\psi\|_{\mathrm{Lip}},$$

the last two steps following from the CFL condition and (ii). Taking the supremum over ψ with $\|\psi\|_{\mathrm{Lip}} \leqslant 1$ yields (iv). □

2.3 *Weak convergence of the method*

We split the proof of convergence of the numerical method to the (unique) measure-valued solution into two parts. First, we show that our method is consistent (Lemma 2.2) in the sense that *if* the method

converges then the limit is the measure-valued solution. Next we show that the method indeed converges either weakly (Theorem 2.4) or strongly (Theorem 2.6), depending on the assumptions on the velocity field V and the initial data μ_0.

LEMMA 2.2 Assume that V satisfies conditions (A2), (A3) and define $\mu^{\Delta x}$ by (2.5). Assume that $d_W(\mu_t^{\Delta x}, \mu_t) \to 0$ as $\Delta x \to 0$ uniformly on bounded intervals $t \in [0, T]$, for some $\mu \in C_w([0, \infty); \mathcal{M}(\mathbb{K}))$. Then μ is the measure-valued solution of (1.1).

Proof. The convergence $d_W(\mu_t^{\Delta x}, \mu_t) \to 0$ implies that $\mu_t \in \mathcal{P}(\mathcal{K})$ for all t. Let $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R})$; we want to show that μ satisfies (1.4) with $\mathcal{K} = \mathbb{K}$. Define the measure $\rho^{\Delta x}$ by

$$\rho_t^{\Delta x} = \mu^n \quad \text{for} \quad t \in [t^n, t^{n+1}).$$

Due to the Lipschitz continuity in time of $\mu^{\Delta x}$ (Proposition 2.1 (iv)), we have that also $d_W(\rho_t^{\Delta x}, \mu_t) \to 0$ as $\Delta x \to 0$ uniformly on bounded intervals $t \in [0, T]$.

Let $\varphi \in C_c^\infty(\mathbb{K} \times \mathbb{R})$, and define $\varphi_i^n := \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \varphi(x_i, t) \mathrm{d}t$. Then

$$\int_0^\infty \int_{\mathbb{K}} \varphi(x, t) \, \mathrm{d}\rho_t^{\Delta x}(x) \mathrm{d}t = \sum_i \sum_{n=0}^\infty \int_{t^n}^{t^{n+1}} \varphi(x_i, t) \mu_i^n \, \mathrm{d}t = \sum_i \sum_{n=0}^\infty \varphi_i^n \mu_i^n \Delta t. \tag{2.7}$$

Multiplying (2.4) by φ_i^n and summing over i, n we get

$$0 = \sum_{n=0}^\infty \sum_i \left(\varphi_i^n \mu_i^{n+1} - \varphi_i^n \frac{\mu_{i-1}^n + \mu_{i+1}^n}{2} + \Delta t \varphi_i^n \frac{V[\mu^n]_{i+1} \mu_{i+1}^n - V[\mu^n]_{i-1} \mu_{i-1}^n}{2\Delta x} \right)$$

(summation by parts)

$$= -\Delta t \sum_{n=0}^\infty \sum_i \frac{\frac{\varphi_{i-1}^n + \varphi_{i+1}^n}{2} - \varphi_i^{n-1}}{\Delta t} \mu_i^n - \Delta t \sum_{n=0}^\infty \sum_i \frac{\varphi_{i+1}^n - \varphi_{i-1}^n}{2\Delta x} V[\mu^n]_i \mu_i^n - \sum_i \varphi_i^{-1} \mu_i^0$$

$$= -\int_0^\infty \int_{\mathbb{K}} \frac{\frac{\varphi(x-\Delta x, t) + \varphi(x+\Delta x, t)}{2} - \varphi(x, t - \Delta t)}{\Delta t} + \frac{\varphi(x + \Delta x, t) - \varphi(x - \Delta x, t)}{2\Delta x} V[\mu_t^{\Delta x}](x) \, \mathrm{d}\rho_t^{\Delta x}(x) \mathrm{d}t$$

$$- \int_{\mathbb{K}} \varphi(x, -\Delta t) \, \mathrm{d}\mu_0^{\Delta x}(x),$$

where in the last step we have used (2.7). By condition (A3) we know that $V[\mu_t^{\Delta x}] \to V[\mu_t]$ uniformly on $\mathbb{R} \times [0, T]$, and by condition (A2), the functions $V[\mu_t^{\Delta x}]$ are uniformly Lipschitz. Moreover, the divided differences in the integrand converge to $\partial_t \varphi$ and $\partial_x \varphi$, respectively, by the mean value theorem. Finally, a straightforward projection error estimate in d_W shows that $\mu_0^{\Delta x}$ converges weakly to μ_0 as $\Delta x \to 0$. It follows that the above integral converges to

$$-\int_0^\infty \int_{\mathbb{K}} \partial_t \varphi(x, t) + \partial_x \varphi(x, t) V[\mu_t](x) \, \mathrm{d}\mu_t(x) \, \mathrm{d}t - \int_{\mathbb{K}} \varphi(x, 0) \, \mathrm{d}\mu_0(x),$$

and the proof is complete. □

For the convergence proof we use the following compactness lemma, whose proof is postponed to the appendix.

LEMMA 2.3   Let $K \subset \mathcal{P}_1(\mathbb{R}^d)$ be a bounded set, i.e., $\sup_{\mu \in K} d_W(\mu, \bar{\mu}) < \infty$ for some $\bar{\mu} \in \mathcal{P}(\mathbb{R}^d)$. Then K is relatively compact in the metric space $(\mathcal{P}_1(\mathbb{R}^d), d_W)$.

THEOREM 2.4   Assume that V satisfies conditions (A1)–(A3) and let $\mu_0 \in \mathcal{P}(\mathbb{K})$ have compact support. Assume that the CFL condition (2.6) is satisfied. Then the Lax–Friedrichs method (2.4) converges weakly to the measure-valued solution of (1.1). More precisely, if $\mu^{\Delta x}$ is given by (2.5) then

$$\sup_{t \in [0,T]} d_W\big(\mu_t^{\Delta x}, \mu_t\big) \to 0 \qquad \text{as } \Delta x \to 0,$$

for any $T > 0$.

Proof.   By Proposition 2.1 (i)–(iii) the measures $\mu_t^{\Delta x}$ stay in $\mathcal{P}(\mathbb{R})$ and have uniformly bounded support for all $\Delta x > 0$, $t \in [0,T]$, and by (iv), the maps $\mu^{\Delta x} : [0,T] \to \mathcal{P}(\mathbb{R}^d)$ are Lipschitz. Noting that the projection map $\mu_0 \mapsto \mu_0^{\Delta x}$ moves mass a distance of at most Δx leads to the estimate $d_W(\mu_0, \mu_0^{\Delta x}) \leqslant \Delta x$, we have that $d_W(\mu_0, \mu_t^{\Delta x}) \leqslant d_W(\mu_0, \mu_0^{\Delta x}) + d_W(\mu_0^{\Delta x}, \mu_t^{\Delta x}) \leqslant \Delta x + \frac{1}{\lambda_0} t$, so by Lemma 2.3 the set

$$K := \big\{ \mu_t^{\Delta x} \, : \, 0 < \Delta x < 1, \ t \in [0,T] \big\}$$

is relatively compact with respect to d_W-convergence. Hence, by Ascoli's theorem there exists a subsequence $\Delta x_k \to 0$ as $k \to \infty$ and some Lipschitz map $\mu : [0,T] \to \mathcal{P}(\mathbb{R}^d)$ such that $d_W(\mu_t^{\Delta x_k}, \mu_t) \to 0$ as $k \to \infty$, uniformly in $t \in [0,T]$. By Lemma 2.2, the limit μ is the measure-valued solution of (1.1). But this solution is unique, which implies that the whole sequence $\mu^{\Delta x}$ converges to μ.   □

2.4   *Strong convergence of the method*

If the initial data and the velocity field are sufficiently smooth then we can show that the numerical method in fact converges *strongly*. We assume that the initial data is absolutely continuous with respect to Lebesgue measure, and hence has a density function $\mu_0(x)$. This data is sampled by its cell averages,

$$\hat{\mu}_i^0 := \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mu_0(x) \, \mathrm{d}x$$

and the numerical solution is realized as the linear-in-time, piecewise constant L^1 function

$$\hat{\mu}_t^{\Delta x}(x) = \frac{t^{n+1} - t}{\Delta t} \hat{\mu}^n + \frac{t - t^n}{\Delta t} \hat{\mu}^{n+1} \qquad \text{for } t \in [t^n, t^{n+1}), \tag{2.8}$$

where

$$\hat{\mu}^n(x) = \sum_i \mu_i^n \mathbb{1}_{[x_{i-1/2}, x_{i+1/2})}(x),$$

and μ_i^n is the solution of the Lax–Friedrichs scheme (2.4) with initial data $\hat{\mu}_i^0$.

PROPOSITION 2.5 Consider the unstaggered, one-dimensional Lax–Friedrichs method (2.4). Assume that V satisfies conditions (A1), (A2) and (A4) and that the CFL condition (2.6) is satisfied. Then for all $t \in [0, T]$,

(i) $\hat{\mu}_t^{\Delta x} \geqslant 0$;

(ii) $\|\hat{\mu}_t^{\Delta x}\|_{L^1(\mathbb{K})} = 1$;

(iii) $\mathrm{TV}(\hat{\mu}_t^{\Delta x}) \leqslant \mathrm{TV}(\mu_0) e^{C_2 t} + (e^{C_2 t} - 1) \frac{C_4}{2C_2}$;

(iv) $\|\hat{\mu}_t^{\Delta x} - \hat{\mu}_s^{\Delta x}\|_{L^1(\mathbb{K})} \leqslant C_T |t - s|$ for all $s \geqslant 0$.

(In (iv), $C_T = C_2 + \lambda_0^{-1} c_T$, where c_T is the upper bound from (iii): $c_T := \mathrm{TV}(\mu_0) e^{C_2 T} + (e^{C_2 T} - 1) \frac{C_4}{2C_2}$.)

Proof. Properties (i) and (ii) obviously follow from Proposition 2.1(i) and (ii). To show (iii) we split $\mathrm{TV}(\hat{\mu}^{n+1})$ as

$$
\mathrm{TV}(\hat{\mu}^{n+1}) \leqslant \frac{1}{2} \sum_i \left| (\mu_{i+2}^n - \mu_{i+1}^n) \left(1 - \frac{\Delta t}{\Delta x} V[\hat{\mu}^n]_{i+2} \right) + (\mu_i^n - \mu_{i-1}^n) \left(1 + \frac{\Delta t}{\Delta x} V[\hat{\mu}^n]_i \right) \right.
$$
$$
\left. - \frac{\Delta t}{\Delta x} \mu_{i+1}^n \left(V[\hat{\mu}^n]_{i+2} - V[\hat{\mu}^n]_{i+1} \right) + \frac{\Delta t}{\Delta x} \mu_{i-1}^n \left(V[\hat{\mu}^n]_i - V[\hat{\mu}^n]_{i-1} \right) \right|
$$
$$
\leqslant A^n + B^n,
$$

where

$$
A^n = \frac{1}{2} \sum_i \left| \mu_{i+2}^n - \mu_{i+1}^n \right| \left(1 - \frac{\Delta t}{\Delta x} V[\hat{\mu}^n]_{i+2} \right) + \left| \mu_i^n - \mu_{i-1}^n \right| \left(1 + \frac{\Delta t}{\Delta x} V[\hat{\mu}^n]_i \right) \leqslant \mathrm{TV}(\hat{\mu}^n)
$$

and

$$
B^n = \frac{\Delta t}{2 \Delta x} \sum_i \left| \mu_{i+1}^n \left(V[\hat{\mu}^n]_{i+2} - V[\hat{\mu}^n]_{i+1} \right) - \mu_{i-1}^n \left(V[\hat{\mu}^n]_i - V[\hat{\mu}^n]_{i-1} \right) \right|
$$
$$
\leqslant \frac{\Delta t}{2} \sum_i \left[\left| \mu_{i+1}^n - \mu_i^n \right| + \left| \mu_i^n - \mu_{i-1}^n \right| \right] \frac{\left| V[\hat{\mu}^n]_{i+2} - V[\hat{\mu}^n]_{i+1} \right|}{\Delta x}
$$
$$
+ \frac{\Delta t}{2} \sum_i \mu_{i-1}^n \frac{\left| V[\hat{\mu}^n]_{i+2} - V[\hat{\mu}^n]_{i+1} - V[\hat{\mu}^n]_i + V[\hat{\mu}^n]_{i-1} \right|}{\Delta x^2} \Delta x
$$
$$
\leqslant C_2 \Delta t \, \mathrm{TV}(\hat{\mu}^n) + \frac{C_4}{2} \Delta t.
$$

Iterating over all n gives

$$
\mathrm{TV}(\hat{\mu}^{n+1}) \leqslant (1 + C_2 \Delta t) \mathrm{TV}(\hat{\mu}^n) + \frac{C_4}{2} \Delta t \leqslant \cdots \leqslant (1 + C_2 \Delta t)^{n+1} \left(\mathrm{TV}(\mu_0) + \frac{C_4}{2C_2} \right) - \frac{C_4}{2C_2}
$$

and (iii) follows, after applying the inequality $1 + \varepsilon \leqslant e^\varepsilon$.

Finally, we show the time continuity (iv):

$$
\begin{aligned}
\|\hat{\mu}^{n+1} - \hat{\mu}^n\|_{L^1(\mathbb{K})} &= \sum_i \left|\mu_i^{n+1} - \mu_i^n\right| \Delta x \\
&= \frac{\Delta x}{2} \sum_i \left| \left(\mu_{i+1}^n - \mu_i^n\right)\left(1 - \frac{\Delta t}{\Delta x}V[\hat{\mu}^n]_{i+1}\right) - \left(\mu_i^n - \mu_{i-1}^n\right)\left(1 + \frac{\Delta t}{\Delta x}V[\hat{\mu}^n]_{i-1}\right) \right. \\
&\qquad \left. - \frac{\Delta t}{\Delta x}\mu_i^n\left(V[\hat{\mu}^n]_{i+1} - V[\hat{\mu}^n]_{i-1}\right) \right| \\
&\leqslant \frac{\Delta x}{2} \sum_i \left|\mu_{i+1}^n - \mu_i^n\right|\left(1 - \frac{\Delta t}{\Delta x}V[\hat{\mu}^n]_{i+1}\right) + \left|\mu_i^n - \mu_{i-1}^n\right|\left(1 + \frac{\Delta t}{\Delta x}V[\hat{\mu}^n]_{i-1}\right) \\
&\qquad + \frac{\Delta t}{\Delta x}\mu_i^n\left|V[\hat{\mu}^n]_{i+1} - V[\hat{\mu}^n]_{i-1}\right| \\
&\leqslant \Delta x \mathrm{TV}(\hat{\mu}^n) + C_2 \Delta t.
\end{aligned}
$$

Iterating over n yields (iv). $\qquad\qquad\square$

THEOREM 2.6 Assume that V satisfies conditions (A1)–(A4) and let $\mu_0 \in \mathcal{P}(\mathbb{K})$ be compactly supported and absolutely continuous with respect to Lebesgue measure with density $\mu_0 \in \mathrm{BV}(\mathbb{K})$. Assume that the CFL condition (2.6) holds. Then the measure-valued solution μ of (1.1) is absolutely continuous, and the one-dimensional Lax–Friedrichs method (2.4) converges strongly to μ. More precisely, if $\hat{\mu}^{\Delta x}$ is given by (2.8) then

$$
\sup_{t \in [0,T]} \left\|\hat{\mu}_t^{\Delta x} - \mu_t\right\|_{L^1(\mathbb{K})} \to 0 \qquad \text{as } \Delta x \to 0,
$$

for any $T > 0$.

Proof. By Proposition 2.5(ii) and (iii) the set

$$
K := \left\{\hat{\mu}_t^{\Delta x} \ : \ 0 < \Delta x < 1, \ t \in [0,T]\right\}
$$

is uniformly bounded in $\mathrm{BV}(\mathbb{K})$, so by Helly's theorem K is relatively compact in $L^1(\mathbb{K})$. Hence, by Ascoli's theorem there is a subsequence $\Delta x_k \to 0$ as $k \to \infty$ and some Lipschitz map $\tilde{\mu} : [0,T] \to L^1(\mathbb{K})$ such that $\hat{\mu}_t^{\Delta x_k} \to \tilde{\mu}_t$ in $L^1(\mathbb{K})$ as $k \to \infty$ uniformly in $t \in [0,T]$. Since convergence in L^1 implies convergence in d_W, also $d_W(\hat{\mu}_t^{\Delta x_k}, \tilde{\mu}_t) \to 0$. On the other hand, Theorem 2.4 implies that $d_W(\mu_t^{\Delta x_k}, \mu_t) \to 0$, where μ is the measure-valued solution of (1.1). Noting that $d_W(\hat{\mu}_t^{\Delta x_k}, \mu_t^{\Delta x_k}) \leqslant \frac{\Delta x}{2}$, we can conclude $\mu = \tilde{\mu}$. The conclusion follows. $\qquad\square$

2.5 *Multiple dimensions*

The extension to multiple space dimensions is done in a tensorial fashion; we limit ourselves to the two-dimensional variant for the sake of notational simplicity. As before, let $\mathbb{K}_1, \mathbb{K}_2$ be either the torus $\mathbb{T}$ or the real line $\mathbb{R}$. Let $(x_{i+1/2})_i$ and $(y_{j+1/2})_j$ be discretizations of $\mathbb{K}_1$ and $\mathbb{K}_2$ with mesh lengths $x_{i+1/2} - x_{i-1/2} \equiv \Delta x$ and $y_{j+1/2} - y_{j-1/2} \equiv \Delta y$, respectively. With obvious notation we get the unstaggered,

two-dimensional Lax–Friedrichs method

$$\mu_{i,j}^0 = \langle \mu_0, \varphi_{i,j} \rangle,$$

$$\mu_{i,j}^{n+1} = \frac{\mu_{i-1,j}^n + \mu_{i+1,j}^n + \mu_{i,j-1}^n + \mu_{i,j+1}^n}{4} - \Delta t \frac{V^1[\mu^n]_{i+1,j}\mu_{i+1,j}^n - V^1[\mu^n]_{i-1,j}\mu_{i-1,j}^n}{2\Delta x}$$

$$- \Delta t \frac{V^2[\mu^n]_{i,j+1}\mu_{i,j+1}^n - V^2[\mu^n]_{i,j-1}\mu_{i,j-1}^n}{2\Delta y}, \tag{2.9}$$

where we denote $\mu^n = \sum_i \sum_j \mu_{i,j}^n \delta_{(x_i,y_j)}$. As before, μ is interpolated linearly between time steps,

$$\mu_t^{\Delta x, \Delta y} := \frac{t^{n+1} - t}{\Delta t}\mu^n + \frac{t - t^n}{\Delta t}\mu^{n+1} \qquad \text{for } t \in [t^n, t^{n+1}). \tag{2.10}$$

We omit the proofs of the following stability and convergence results, as they are straightforward generalizations of their one-dimensional counterparts.

PROPOSITION 2.7 Consider the unstaggered, two-dimensional Lax–Friedrichs method (2.9). Assume that V satisfies condition (A1) and that the CFL condition

$$\lambda_0 \leqslant \frac{\Delta t}{\min(\Delta x, \Delta y)} C_1 \leqslant \frac{1}{2} \tag{2.11}$$

is satisfied for some $\lambda_0 > 0$. Then for all $t \geqslant 0$,

(i) $\mu_t^{\Delta x, \Delta y} \geqslant 0$;

(ii) $\|\mu_t^{\Delta x, \Delta y}\|_{\mathcal{M}(\mathcal{K})} = 1$;

(iii) $d_W\big(\mu_t^{\Delta x, \Delta y}, \mu_s^{\Delta x, \Delta y}\big) \leqslant \frac{\Delta x + \Delta y}{2\Delta t}|t - s|$ for all $s \geqslant 0$.

THEOREM 2.8 Consider the two-dimensional equation (1.1). Assume that V satisfies conditions (A1)–(A3) and let $\mu_0 \in \mathcal{P}(\mathbb{K}_1 \times \mathbb{K}_2)$ have compact support. Assume that the CFL condition (2.11) is satisfied. Then the Lax–Friedrichs method (2.9) converges weakly to the measure-valued solution of (1.1). More precisely, if $\mu^{\Delta x, \Delta y}$ is given by (2.10) then

$$\sup_{t \in [0,T]} d_W\big(\mu_t^{\Delta x, \Delta y}, \mu_t\big) \to 0 \qquad \text{as } \Delta x, \Delta y \to 0,$$

for any $T > 0$.

As in the one-dimensional case, sufficient smoothness for the initial data and the velocity field will yield *strong* convergence of the numerical method. Assume the initial datum is absolutely continuous with respect to the Lebesgue measure with density function $\mu_0(x, y)$. This initial data is sampled by its cell averages

$$\hat{\mu}_{i,j}^0 := \frac{1}{\Delta x \Delta y} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \mu_0(x, y) \, dx \, dy.$$

Similarly, we define the numerical solution as a piecewise linear L^1 function as

$$\hat{\mu}_t^{\Delta x, \Delta y} = \frac{t^{n+1} - t}{\Delta t} \hat{\mu}^n + \frac{t - t^n}{\Delta t} \hat{\mu}^{n+1} \qquad \text{for } t \in [t^n, t^{n+1}), \tag{2.12}$$

where

$$\hat{\mu}^n(x, y) = \sum_{i,j} \hat{\mu}_{i,j}^n \mathbb{1}_{[x_{i-1/2}, x_{i+1/2}) \times [y_{j-1/2}, y_{j+1/2})}(x, y).$$

PROPOSITION 2.9   Consider the unstaggered, one-dimensional LaxFriedrichs method (2.4). Assume that V satisfies the conditions (A1), (A2) and (A4) and that the CFL condition (2.11) is satisfied. Then for all $t \in [0, T]$,

  (i) $\hat{\mu}_t^{\Delta x, \Delta y} \geqslant 0$;

  (ii) $\|\hat{\mu}_t^{\Delta x, \Delta y}\|_{L^1(\mathbb{K}_1 \times \mathbb{K}_2)} = 1$;

  (iii) $\mathrm{TV}(\hat{\mu}_t^{\Delta x, \Delta y}) \leqslant c_T := \mathrm{TV}(\hat{\mu}_0) e^{C_2 T} + (e^{C_2 T} - 1) \frac{C_4}{2 C_2}$;

  (iv) $\|\hat{\mu}_t^{\Delta x, \Delta y} - \hat{\mu}_s^{\Delta x, \Delta y}\|_{L^1(\mathbb{K}_1 \times \mathbb{K}_2)} \leqslant C_T |t - s|$ for all $s \geqslant 0$.

(In (iv), $C_T = C_2 + \lambda_0^{-1} c_T$, where c_T is the upper bound from (iii).)

So now we can finally state the following.

THEOREM 2.10   Assume that V satisfies conditions (A1)–(A4) and let $\mu_0 \in \mathcal{P}(\mathbb{K}_1 \times \mathbb{K}_2)$ be compactly supported and absolutely continuous with respect to Lebesgue measure with density $\mu_0 \in \mathrm{BV}(\mathbb{K}_1 \times \mathbb{K}_2)$. Assume that the CFL condition (2.11) holds. Then the measure-valued solution μ of (1.1) takes values in $L^1(\mathbb{K}_1 \times \mathbb{K}_2)$, and the two-dimensional Lax–Friedrichs method (2.9) converges strongly to μ. More precisely, if $\hat{\mu}^{\Delta x, \Delta y}$ is given by (2.12) then

$$\sup_{t \in [0, T]} \|\hat{\mu}_t^{\Delta x, \Delta y} - \mu_t\|_{L^1(\mathbb{K}_1 \times \mathbb{K}_2)} \to 0 \qquad \text{as } \Delta x, \Delta y \to 0,$$

for any $T > 0$.

3. Numerical experiments

In this section we illustrate our analytical results by performing several numerical experiments for the Kuramoto equation with identical (1.3) and nonidentical (1.2) natural frequencies. Recall that in these equations, θ takes values in the periodic domain $\mathbb{T}$, while Ω lies in $\mathbb{R}$.

3.1 One-dimensional simulations

We consider first the one-dimensional Kuramoto equation with identical oscillators (1.3). In all experiments we have used the unstaggered Lax–Friedrichs method (2.4) with CFL number 0.4 and with $T = 0.5$ as the final time. All simulations were computed on a sequence of meshes with $N = 32, 64, 128, 256, 512$ grid points as well as a reference solution with $N = 4096$ points. The tables in each subsection show the number of grid points as well as the error and the experimental order of

convergence (EOC)

$$\mathrm{EOC}_{\Delta x, \Delta x'} := \frac{\log\left(E_{\Delta x}/E_{\Delta x'}\right)}{\log\left(\Delta x/\Delta x'\right)}, \qquad E_{\Delta x} = d\left(\mu_T^{\Delta x}, \mu_T\right)$$

computed both in the two metrics $d = d_W$ and $d(u, v) = \|u - v\|_{L^1(\mathcal{K})}$.

The experiments in this section solve the Kuramoto equation with progressively more singular initial data. As will be seen from the tables the EOC in the 1-Wasserstein distance is close to 1 in all cases, while in L^1 it is close to 1 only for smooth data.

3.1.1 *Polynomial initial data.* The initial data is taken to be the continuous and piecewise parabolic function

$$\mu_0(\theta) = \begin{cases} \frac{6}{\pi^3}(\frac{3\pi}{2} - \theta)(\theta - \frac{\pi}{2}) & \text{if } \theta \in [\frac{\pi}{2}, \frac{3\pi}{2}), \\ 0 & \text{otherwise.} \end{cases}$$

As shown in Fig. 1, the numerical solution is a nonoscillatory and reasonable approximation at all mesh resolutions. Table 1 shows that the numerical method seems to converge at a rate close to 1, both in the 1-Wasserstein and the L^1 distances.
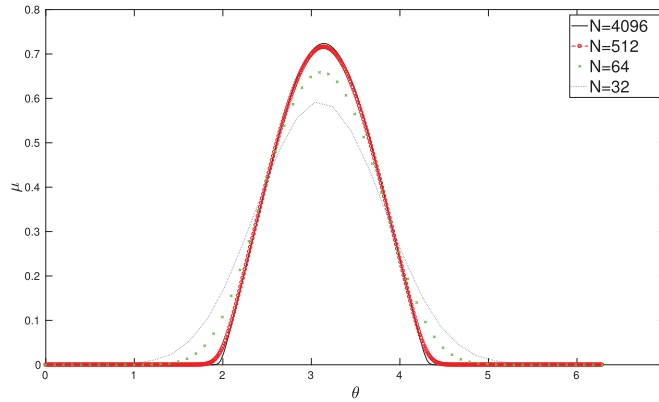

FIG. 1. Comparing solutions on different meshes at $t = 0.5$ for smooth initial data.

TABLE 1   *Errors and EOC in the* 1-*Wasserstein and* L^1
distances at $t = 0.5$ *for smooth initial data*

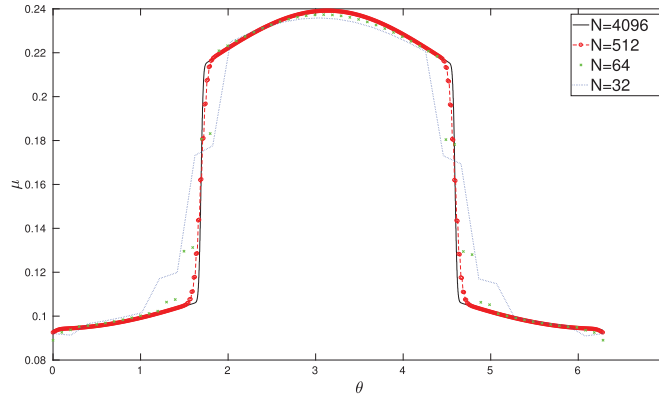| N	1-Wasserstein		L^1	
	Error	EOC	Error	EOC
32	0.1351		0.2811	
64	0.0634	1.09	0.1336	1.07
128	0.0303	1.07	0.0663	1.01
256	0.0153	0.99	0.0336	0.98
512	0.0073	1.07	0.0157	1.10

FIG. 2. Comparing solutions on different meshes for piecewise constant initial data.

TABLE 2   *Errors and EOC in the* 1-*Wasserstein and* L^1
distances at $t = 0.5$ *for piecewise constant initial data*

	1-Wasserstein		L^1	
N	Error	EOC	Error	EOC
32	0.0517		0.0610	
64	0.0258	1.00	0.0315	0.95
128	0.0131	0.98	0.0196	0.69
256	0.0067	0.97	0.0123	0.67
512	0.0034	0.98	0.0071	0.79

3.1.2 *Piecewise constant initial data.*   This experiment was taken from Amadori *et al.* (2017) and uses piecewise constant initial data
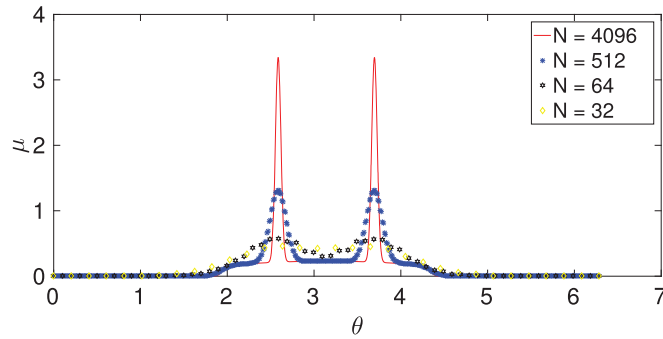
$$\mu_0(\theta) = \begin{cases} \frac{2}{3\pi} & \text{if } \theta \in [\frac{\pi}{2}, \frac{3\pi}{2}), \\ \frac{1}{3\pi} & \text{otherwise.} \end{cases}$$

Figure 2 shows again that the approximation is reasonably accurate even on coarse meshes. Table 2 shows that the rate of convergence in 1-Wasserstein is again 1, while it is around 0.7 in L^1.

3.1.3 *Singular initial data.*   The final experiment uses the singular initial data

$$\mu_0 = \frac{1}{4}\big(\delta_{\frac{3\pi}{4}} + \delta_{\frac{5\pi}{4}}\big) + \frac{1}{2}\chi_{[\frac{\pi}{2}, \frac{3\pi}{2}]},$$

where $\chi_{[a,b]}$ is the indicator function of the interval $[a, b]$. As shown in Fig. 3, the numerical approximation is nonoscillatory and seems to converge, even in the presence of Dirac singularities. This is confirmed in Table 3: the method seems to converge at a rate of around 3/4 in d_W, while it does not converge at all in L^1, as expected.

FIG. 3. Comparing solutions on different meshes at $t = 0.5$ for initial data with Dirac masses.

TABLE 3   *Errors and EOC in the* 1-*Wasserstein and* L^1
distances at $t = 0.5$ *for singular initial data*

N	1-Wasserstein		L^1	
	Error	EOC	Error	EOC
32	0.1737		0.8703	
64	0.1119	0.63	0.7945	0.13
128	0.0687	0.70	0.7131	0.16
256	0.0433	0.67	0.6153	0.21
512	0.0260	0.74	0.4861	0.34

3.2 *Polynomial initial data in two dimensions*

The final numerical experiment approximates the two-dimensional Kuramoto equation (1.2). We consider the piecewise linear initial data

$$\mu_0(\theta, \Omega) = \begin{cases} \frac{64}{3\pi^2}\theta\Omega & \text{if } \theta \in [\frac{\pi}{4}, \frac{\pi}{2}) \text{ and } \Omega \in [0, 1], \\ 0 & \text{otherwise.} \end{cases}$$
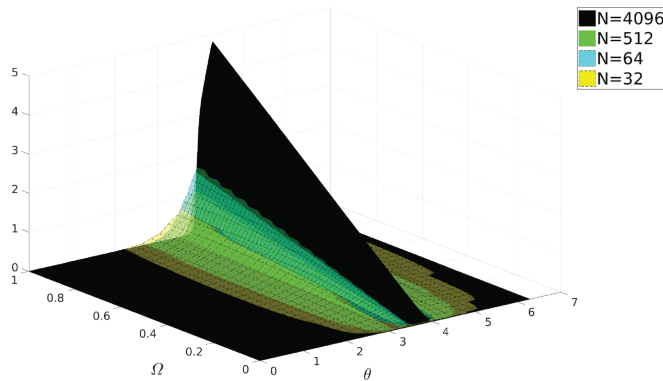

FIG. 4. Comparing solutions on different meshes for two-dimensional polynomial initial data.

TABLE 4   *Table of L^1 errors for two-dimensional initial data*

	L^1	
N	Error	EOC
32	1.2912	
64	1.0054	0.36
128	0.7346	0.45
256	0.4925	0.58
512	0.2999	0.72
1024	0.1638	0.87

Figure 4 shows the solution at $t = 0.5$, and Table 4 shows the L^1 errors. (Due to the complexities of computing the Wasserstein distance for multi-dimensional measures, we only compute the L^1 errors in this experiment.) The convergence rate in L^1 is seen to be about the same as in the piecewise constant one-dimensional example in Section 3.1.2.

Funding

REFERENCES

AMADORI, D., HA, S. Y. & PARK, J. (2017a) On the global well-posedness of BV weak solutions to the Kuramoto–Sakaguchi equation. *J. Differ. Equ.*, **262**, 978–1022.

AMADORI, D., HA, S. Y. & PARK, J. (2017b) Wave-front tracking analysis for the Kuramoto–Sakaguchi equation. *Innovative Algorithms and Analysis*. Springer INdAM Series, Springer International Publishing ISBN: 978-3-319-49261-2.

BOUCHUT, F. & JAMES, F. (1998) One-dimensional transport equations with discontinuous coeffiecients. *Nonlin. Anal.*, **32**, 891–933.

BUCK, J. & BUCK, E. (1966) Biology of synchronous flashing of fireflies. *Nature*, **211**, 562.

CARRILLO, J. A., CHOI, Y. P., HA, S. Y., KANG, M. J. & KIM, Y. (2014) Contractivity of transport distances for the kinetic Kuramoto equation. *J. Stat. Phys.*, **156**, 395–415.

CHIBA, H. (2013) Continuous limit of the moments system for the globally coupled phase oscillators. *Discrete Contin. Dyn. Syst. Ser. A*, **5**, 1891–1903.

COLOMBO, R. M., HERTY, M. & MERCIER, M. (2011) Control of the continuity equation with a non local flow. *ESAIM Control Optim. Calc. Var.*, **17**, 353–379.

CRIPPA, G. & LÉCUREUX-MERCIER, M. (2013) Existence and uniqueness of measure solutions for a system of continuity equations with non-local flow. *Nonlinear Differ. Equ. Appl.*, **20**, 523–537.

DOBRUSHIN, R. L. (1979) Vlasov equations. *Funct. Anal. Appl.*, **13**, 115–123.

JAMES, F. & VAUCHELET, F. (2013) Chemotaxis: from kinetic equations to aggregate dynamics. *Nonlin. Differ. Equ. Appl.*, **20**, 101–127.

JAMES, F. & VAUCHELET, N. (2015) Numerical methods for one-dimensional aggregation equations. *SIAM J. Numer. Anal.*, **53**, 895–916.

KURAMOTO, Y. (1975) Self-entrainment of a population of coupled non-linear oscillators. *International Symposium on Mathematical Problems in Mathematical Physics*. Lecture Notes in Theoretical Physics, vol. 30. Berlin, Heidelberg: Springer, pp. 420–422.

KURAMOTO, Y. (1984) *Chemical Oscillations Waves, and Turbulence*. Berlin: Springer.

LANCELOTTI, C. (2005) On the Vlasov limit for systems of non-linearly coupled oscillators without noise. *Transp. Theory Stat. Phys.*, **34**, 523–535.

NESSYAHU, H. & TADMOR, E. (1990) Non-oscillatory central differencing for hyperbolic conservation laws. *J. Computat. Phys.*, **87**, 408–463.

NEUNZERT, H. (1978). Mathematical investigations on particle-in-cell methods. *Fluid Dyn. Trans.* **9**, 229–254.

NEUNZERT, H. (1984) An introduction to the non-linear Boltzman–Vlasov equation. *Kinetic Theories and the Boltzman Equation*. Lecture Notes in Mathematics, vol. 1048. Berlin, Heidelberg: Springer, pp. 60–110.

VILLANI, C. (2003). *Topics in Optimal Transportation*. Providence, RI: American Mathematical Society.

WINFREE, A. T. (1967). Biological rhythms and the behaviour of populations of coupled oscillators. *J. Theor. Biol.*, **16**, 15–42.

A. Appendix

Proof of Lemma 2.3. We claim that K is tight. Define

$$\chi(x) := \begin{cases} 1, & |x| \leqslant 1, \\ 2 - |x|, & 1 < |x| < 2, \\ 0, & 2 \leqslant |x|, \end{cases} \qquad \chi_M(x) := \chi(x/M).$$

For an $\varepsilon > 0$, let $M > 0$ be such that $\bar{\mu}(\mathbb{R} \setminus B_M(0)) \leqslant \varepsilon$, and let $C := \sup_{\mu \in K} d_W(\mu, \bar{\mu})$. If $\mu \in K$ then

$$\mu\big(\mathbb{R} \setminus B_{2M}(0)\big) \leqslant \langle \mu, 1 - \chi_M \rangle = \langle \bar{\mu}, 1 - \chi_M \rangle + \langle \mu - \bar{\mu}, 1 - \chi_M \rangle$$

$$\leqslant \bar{\mu}(\mathbb{R} \setminus B_M(0)) + \|1 - \chi_M\|_{\text{Lip}} d_W(\mu, \bar{\mu})$$

$$\leqslant \varepsilon + \frac{C}{M},$$

which can be made arbitrarily small, independently of μ. Hence, K is tight, so by Prokhorov's theorem, K is relatively compact in the topology of narrow (or 'weak') convergence. But narrow convergence is equivalent to d_W-convergence. This completes the proof. □