J. Eur. Math. Soc. 23, 3193–3278 (2021) DOI 10.4171/JEMS/1081

Jinyeong Park · David Poyato · Juan Soler

Filippov trajectories and clustering in the Kuramoto model with singular couplings

Received September 28, 2018

Abstract. We study the synchronization of a generalized Kuramoto system in which the coupling weights are determined by the phase differences between oscillators. We employ the fast-learning regime in a Hebbian-like plasticity rule so that the interaction between oscillators is enhanced by the approach of phases. First, we study the well-posedness problem for the singular weighted Kuramoto systems in which the Lipschitz continuity fails to hold. We present the dynamics of the system equipped with singular weights in all the subcritical, critical and supercritical regimes of the singularity. A key fact is that solutions in the most singular cases must be considered in Filippov's sense. We characterize sticking of phases in the subcritical and critical case and we exhibit a continuation criterion for classical solutions after any collision state in the supercritical regime. Second, we prove that strong solutions to these systems of differential inclusions can be recovered as singular limits of regular weights. We also study the emergence of synchronous dynamics for the singular and regular weighted Kuramoto models.

Keywords. Kuramoto models, adaptive coupling, singular interactions, Hebbian learning, Filippov-type solutions, clustering, finite-time synchronization, sticking, Cucker–Smale

1. Introduction

Synchronization is the natural collective behavior arising from agent-based interactions given by periodic rules. These rhythmical motions can be easily observed in various biological complex systems such as flashing of fireflies, beating of cardiac cells, etc. One of the most significant examples of synchronization appears in neurons. Associative or Hebbian learning [24] proposes an explanation for the adaptation of neurons in the brain during the learning process. That mechanism is founded on the assumption that synchronous activ-

Jinyeong Park: Department of Mathematics and Research Institute of Natural Sciences, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea; jinyeongpark@hanyang.ac.kr

David Poyato: Institut Camille Jordan (ICJ), UMR 5208 CNRS & Université Claude Bernard Lyon 1, 69100 Villeurbanne, France, and Research Unit "Modeling Nature" (MNat), Universidad de Granada, 18071 Granada, Spain; poyato@math.univ-lyon1.fr; davidpoyato@ugr.es

Juan Soler: Departamento de Matemática Aplicada and Research Unit "Modeling Nature" (MNat), Universidad de Granada, 18071 Granada, Spain; jsoler@ugr.es

Mathematics Subject Classification (2020): 35Q70, 35Q83, 34A60, 34A36, 34C15, 34D06, 35L81, 82C32, 92B20, 92B25

ation of cells (firing of neurons) leads to selectively pronounced increases in synaptic strength between those cells. A consequence is that the pattern of activity will become self-organized. In Hebb's words: *Any two cells or systems of cells that are repeatedly active at the same time will tend to become associated, so that activity in one facilitates activity in the other*. In neuroscience, these processes provide the neuronal basis of unsupervised learning of cognitive functions in neural networks and can explain the phenomena that arise in the development of the nervous system.

Since Kuramoto proposed a mathematical model for coupled oscillators in [28,29], the synchronization has received a lot of attention and has been extensively studied in various disciplines from this point of view [1]. In the classical Kuramoto model, the system of oscillators has an all-to-all coupling with uniform weights given by a constant coupling strength K:

$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \quad i = 1, \ldots, N, \tag{1.1}$$

where Ω_i's are the natural frequencies of the oscillators. However, the uniform and constant couplings are a bit restrictive to explain the complicatedness of phenomena. Thus, it is more interesting to consider a generalization of the Kuramoto model endowed with plastic couplings, introduced in [3,18,21,34,38,41,42]:

$$\dot{\theta}_i = \Omega_i + \frac{1}{N}\sum_{j=1}^{N}K_{ij}\sin(\theta_j - \theta_i), \quad i = 1, \ldots, N, \tag{1.2}$$

where K_{ij} is the coupling between the i-th and j-th oscillators which has its own dynamics depending on the phase configuration. The coupling K_{ij} is assumed to be

$$K_{ij} = Ka_{ij},$$

where $a_{ij} \in [0, 1]$ measures the degree of connectedness between the i-th and j-th oscillators. They will be allowed to vary adaptively relying on the associated phases θ_i and θ_j, via the dynamic learning law

$$\dot{a}_{ij} = \eta(\Gamma(\theta_j - \theta_i) - a_{ij}) \tag{1.3}$$

for some plasticity function Γ. Here, η is regarded as the learning rate parameter such that a small η delays the adaptation of weight a_{ij}. According to the choice of the function Γ, the dynamics of the system (1.2) follows various scenarios. In neural network systems, the Hebbian-type dynamics is considered for the learning algorithm of couplings between oscillators. That learning law amounts to saying that the weight of coupling increases if the phases of oscillators are close to each other. For example, in [21,34,42], Γ is assumed to be $\Gamma(\theta) = \cos\theta$ so that attraction between near oscillators is reinforced whereas repulsive interaction arises between phases that are apart. On the other hand, anti-Hebbian type is also considered, such as $\Gamma(\theta) = |\sin\theta|$ in [21,41]. In this case, synchronization emerges slowly due to the reduction of weight for nearby oscillators. Other types of adaptive rules are considered in [18,38].

We will consider a Hebbian-like Γ for the dynamics of adaptive coupling so that the coupling is enhanced by approach of phases. Assume that the Hebbian-like plasticity function Γ is given by

$$\Gamma(\theta) := \frac{\sigma^{2\alpha}}{(\sigma^2 + c_{\alpha,\zeta}|\theta|_o^2)^\alpha}, \tag{1.4}$$

where $\sigma \in (0, \pi)$, $\zeta \in (0, 1]$ and $|\theta|_o$ is the orthodromic distance (to zero) over the unit circle, which can be defined by

$$|\theta|_o := |\bar{\theta}| \quad \text{for } \bar{\theta} \equiv \theta \bmod 2\pi,\ \bar{\theta} \in (-\pi, \pi].$$

Here, the parameter $c_{\alpha,\zeta} := 1 - \zeta^{-1/\alpha}$ has been chosen so that whenever two phases θ_i and θ_j stay at orthodromic distance σ or larger, then the adaptive function Γ predicts a maximum degree of connectedness no larger than ζ between such oscillators.

Since the plasticity function Γ in (1.4) is Lipschitz-continuous, we can apply the Tikhonov's theorem [27] to (1.2)–(1.3) in order to rigorously derive the fast learning regime $\eta \to \infty$. Then, we arrive at the following Kuramoto model with weighted coupling structure:

$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^N \Gamma(\theta_j - \theta_i)\sin(\theta_j - \theta_i), \tag{1.5}$$

which will play a central role in our work. If either $\alpha = 0$ or $\zeta = 1$, then our plasticity function (1.4) becomes 1 everywhere. In that case, our system (1.5) reduces to the classical Kuramoto model (1.1). Hence, we will assume that $\alpha > 0$ and $\zeta \in (0, 1)$ from now on. Our main interest is to analyze the system (1.4)–(1.5) and compare it with the associated singular counterpart with singular plasticity function

$$\Gamma(\theta) := \frac{1}{c_{\alpha,\zeta}^\alpha |\theta|_o^{2\alpha}}. \tag{1.6}$$

In the next section we will derive this new singular model from the regular one through a singular limit of the parameters. In the regular case (1.4), Γ is a Lipschitz-continuous function and the system (1.5) becomes the Kuramoto model with regular weights depending on the phase configuration. Then the well-posedness of global-in-time classical solutions is standard. However, in the singular case (1.6), the system (1.5) has a singular weight and we must deal with non-Lipschitz right hand side, where the Cauchy–Lipschitz theorem cannot guarantee the existence and uniqueness of global-in-time solutions. We will deal with three different regimes of the singularity: $\alpha \in (0, 1/2)$, $\alpha = 1/2$ and $\alpha \in (1/2, 1)$, which we respectively call the subcritical, critical and supercritical cases.

The main results of this paper are as follows. First we study the well-posedness of the singular weighted system. Depending on the value of α, the properties of the right hand side of (1.5) vary. Specifically, in the subcritical regime, we deal with systems of ODEs with Hölder-continuous right hand side, while we face discontinuous right hand sides of both bounded and unbounded type in the critical and supercritical cases. In addition, the type of uniqueness that we can expect in these systems is one-sided. Namely, a cluster of phases may eventually arise after a finite-time collision and oscillators belonging to the cluster might stay stuck together. This is a phenomenon that was recently found in other types of agent-based systems like the Cucker–Smale model with singular weights [36,37].

Our second result characterizes the explicit conditions for sticking in the subcritical and critical regimes. In the former case, we show that only clusters of oscillators with the same natural frequencies can stick together. Nevertheless, in the latter case, a cluster of oscillators with different natural frequencies may stick together as long as the frequencies fulfill an appropriate condition. Regarding the supercritical case, the analogous sticking condition becomes trivial and we can show a procedure of continuation of classical solutions after finite-time collisions. Namely, after a cluster is formed in finite time, the cluster keeps stuck together no matter what are the natural frequencies of the oscillators involved.

The third result consists in showing that these singular weights are physically relevant. Specifically, we will show that the system (1.5)–(1.6) with singular weights can be obtained as a rigorous singular limit of the regular model (1.4)–(1.5). Again, the strategy will differ in each of the regimes. For the subcritical case, similar tools to those in [36,37] for the singular Cucker–Smale model can be adapted. What is more, we can even obtain an analogous gain of extra $W^{1,1}$ piecewise regularity of the frequencies of oscillators. For the critical and supercritical cases we cannot resort to the same ideas. Hence, we use the underlying gradient-flow structure to gain compactness of frequencies. Identifying the limit will be the heart of the matter in this part.

Our last result concerns the emergence of synchronization in each regime of the parameter α. For identical oscillators, we show the emergence of complete phase synchronization in finite time under appropriate assumptions on the initial diameters of phases. At least in the subcritical regime, where frequencies become more regular, we study the asymptotic emergence of complete frequency synchronization of non-identical oscillators. Also, we study the stability properties of collisionless phase-locked states in all the three regimes.

The techniques are initially inspired by a combination of results for the classical Kuramoto model, but they require a new perspective allowing for singular interactions. For this purpose, we introduce a well-posedness result “à la Filippov” that is valid for systems of ODEs with discontinuous right hand sides. Specifically, we will rely on the study of absolutely continuous solutions of the differential inclusions associated with the Filippov set-valued map. The values of that map are convex polytopes that are bounded and unbounded in the critical and supercritical case respectively. Hence, the classical theory can be used in the former case whereas new ideas are developed for the latter case. Also, we prove some one-sided uniqueness results for non-Lipschitzian interactions that rely on the structure of the interaction kernel near the points of loss of Lipschitz-continuity. For the stability of equilibria, Lyapunov’s first method entails a similar scenario to that of the classical Kuramoto model in the critical and supercritical regimes. On the other hand, the subcritical regime requires a center manifold approach that yields the stability of the corresponding equilibria. What is more interesting is that we can still get some accurate control of the diameter of the system of singularly weighted coupled oscillators. That control amounts to the corresponding finite-time and asymptotic synchronization for the identical and non-identical cases. Unfortunately, the emergence of phased-locked states independently of the initial configurations cannot be derived as in previous results for the classical Kuramoto model (see [19]) because it is not clear whether the Łojasiewicz gradient inequality [32] holds for non-analytic systems with gradient structure like this. Regarding the singular limit of the regular coupling weights, the main goal is to prove that solutions of the regularized system converge towards absolutely continuous trajectories that fulfill the differential inclusion. For that, an appropriate H-representation (half-space representation) of such convex polytopes is obtained through convex analysis techniques. Then, the gain of compactness of frequencies along with the geometric representation of the Filippov map will provide the necessary tools for the singular limit to work in the critical and supercritical regimes.

The rest of the paper is organized as follows. In Section 2, we present definitions, basic properties of the weighted Kuramoto model, the underlying gradient-flow structure, the passage from regular to singular plasticity function and the expected macroscopic equations. In Section 3, we study the system with singular weights and we prove the well-posedness in each regime. In Section 4, we prove the rigorous singular limit in every regime and compare the model with previous results derived in other agent-based systems, in particular we make a comparison with Cucker–Smale models. In Section 5, we show the synchronization for the singular weighted system. In Appendix A, we will recall some classical tools of the Kuramoto models that we apply to show the emergence of synchronization in the regular weighted system for the sake of clarity. Appendix B shows the proofs of the H-representation of the Filippov set-valued map in the critical and supercritical cases. Finally, Appendix C introduces an explicit characterization of the sticking conditions.

2. Preliminaries

2.1. Basic properties and definitions

In this section, we study the basic properties of the weighted Kuramoto system and introduce some related results that will be useful in the following sections. For simplicity, let us denote the interaction kernel by $h(\theta) := \Gamma(\theta)\sin\theta$ (here Γ can be any even function, e.g., (1.4) or (1.6)). Then the system (1.5) can be expressed as

$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^{N} h(\theta_j - \theta_i). \tag{2.1}$$

For simplicity, we shall sometimes use vector notation in (2.1). We define the vector field $H = H(\Theta) = (H_1(\Theta), \dots, H_N(\Theta))$ whose components read

$$H_i(\Theta) = \Omega_i + \frac{K}{N}\sum_{j=1}^{N} h(\theta_j - \theta_i). \tag{2.2}$$

Then (2.1) can be restated as

$$\dot{\Theta} = H(\Theta). \tag{2.3}$$

Since h is an odd function, by taking sums on both sides of (2.1), we have

$$\sum_{i=1}^{N} \dot{\theta}_i = \sum_{i=1}^{N} \Omega_i,$$

i.e., the average of frequencies is conserved. Thus, without loss of generality, we may assume that the average of the natural frequencies is zero, $N^{-1}\sum_{i=1}^{N}\Omega_i = 0$, in order to focus on the fluctuation from the constant average motion.

For the discussion in Section 4, we briefly introduce the second order augmentation of the Kuramoto model [16]. By taking one more derivative of the system (2.1), we get the second order model

$$\begin{cases} \dot{\theta}_i = \omega_i, \\ \dot{\omega}_i = \dfrac{K}{N}\displaystyle\sum_{j=1}^{N} h'(\theta_j - \theta_i)(\omega_j - \omega_i). \end{cases} \tag{2.4}$$

For both systems (2.1) and (2.4) we have the following equivalence.

Theorem 2.1. *The Kuramoto model* (2.1) *is equivalent to an augmented Kuramoto model* (2.4) *in the following sense.*

(1) *If* $\Theta = (\theta_1, \ldots, \theta_N)$ *is a solution to* (2.1) *with initial data* Θ_0*, then* $(\Theta, \omega := \dot{\Theta})$ *is a solution to* (2.4) *with well-prepared initial data* (Θ_0, ω_0)*:*

$$\omega_{i,0} := \Omega_i + \frac{\kappa}{N}\sum_{j=1}^{N} h(\theta_{j,0} - \theta_{i,0}).$$

(2) *If* (Θ, ω) *is a solution to* (2.4) *with initial data* (Θ_0, ω_0)*, then* Θ *is a solution to* (2.1) *with natural frequencies*

$$\Omega_i := \omega_{i,0} - \frac{\kappa}{N}\sum_{j=1}^{N} h(\theta_{j,0} - \theta_{i,0}).$$

For the regular cases (1.4), the proof can be found in [16]. However, one has to take special care with the time regularity of solutions in the singular cases (1.6) before we take derivatives in (2.1). In the case of $\alpha \in (0, 1/2)$, the type of solutions to be considered for (2.1) are absolutely continuous solutions, while for (2.4) solutions have to be taken in a weak sense with C^1 and piecewise $W^{2,1}$ regularity (see [36] for this concept of solution for the discrete Cucker–Smale model with singular influence function). The well-posedness of both singular systems (2.1) and (2.4) will be established in Sections 3 and 4 (see Theorems 3.1, 3.3, 4.1, 4.2 and Remark 4.1), and comparisons with Cucker–Smale models with singular influence function will be made in Subsection 4.4.

For completeness, we recall the different definitions of synchronization [15].

Definition 2.1. Let $\Theta(t) = (\theta_1(t), \dots, \theta_N(t))$ be the phase configuration of oscillators with dynamics governed by the system (1.5).

(1) The system shows *complete phase synchronization* asymptotically if

$$\lim_{t\to\infty} |\theta_i(t) - \theta_j(t)| = 0 \quad \text{for all } i \neq j.$$

(2) The system shows *complete frequency synchronization* asymptotically if

$$\lim_{t\to\infty} |\dot\theta_i(t) - \dot\theta_j(t)| = 0 \quad \text{for all } i \neq j.$$

(3) A *phase-locked state emerges* asymptotically in the system if there exist constants θ_{ij}^∞ such that

$$\lim_{t\to\infty} |\theta_i(t) - \theta_j(t)| = \theta_{ij}^\infty \quad \text{for all } i \neq j.$$

Analogous definitions of synchronization will be considered if, instead of asymptotically, the relevant dynamics takes place in some finite time T (which then replaces ∞ in the above conditions).

We note that complete phase synchronization is a special case of a phase-locked state. It is obvious that emergence of a phase-locked state implies complete frequency synchronization. However, the converse is valid when frequency synchronization occurs fast, i.e., we have integrable decay of frequency differences.

2.2. *Singular weighted model*

In this part, we formally derive the Kuramoto model with singular weights as a singular limit of the regular weighted model. We note that the regular weighted model is (2.1) with interaction kernel given by

$$h(\theta) := \frac{\sigma^{2\alpha}\sin\theta}{(\sigma^2 + c_{\alpha,\zeta}|\theta|_o^2)^\alpha}.$$

Recall that the degree of connectedness is smaller than ζ for interparticle distances larger than σ and α imposes the fall-off of the interactions. Consequently, σ measures the effective range of interactions. Similarly, the parameter K measures the maximum strength of interactions. Hence, one can propose the following scaling:

$$\sigma = \mathcal{O}(\varepsilon), \qquad K\sigma^{2\alpha} = \mathcal{O}(1), \qquad \text{as } \varepsilon \to 0.$$

Or more specifically, using the change of variables

$$\sigma \to \varepsilon \quad \text{and} \quad K \to K\varepsilon^{-2\alpha},$$

where ε is a dimensionless parameter, we arrive at the scaled system

$$\dot\theta_i = \Omega_i + \frac{K}{N}\sum_{j=1}^N h_\varepsilon(\theta_j - \theta_i), \tag{2.5}$$

where the scaled interaction kernel now reads

$$h_\varepsilon(\theta) := \frac{\sin\theta}{(\varepsilon^2 + c_{\alpha,\zeta}|\theta|_o^2)^\alpha}. \tag{2.6}$$

If we formally take limits as $\varepsilon \to 0$, we arrive at the desired singular weighted Kuramoto model, whose singular interaction kernel is

$$h(\theta) := \frac{\sin\theta}{c_{\alpha,\zeta}^\alpha |\theta|_o^{2\alpha}}.$$

All these arguments are heuristic. However they might become rigorous depending on the value of α. For a rigorous derivation of the singular limit in all the subcritical, critical and supercritical regimes, see Section 4.

2.3. Emergence of clusters: collision and sticking of oscillators

In this part we introduce some notation that will be used all along the paper. We will denote the set of pairwise collisions of the i-th and j-th oscillators by

$$\mathcal{C}_{ij} := \{\Theta \in \mathbb{R}^N : \bar{\theta}_i = \bar{\theta}_j\},$$

where $\bar{\theta}$ denotes again the representative of θ in $(-\pi, \pi]$. Then the set of collisions is

$$\mathcal{C} := \bigcup_{i \neq j} \mathcal{C}_{ij} = \{\Theta \in \mathbb{R}^N : \exists i \neq j \text{ such that } \bar{\theta}_i = \bar{\theta}_j\}.$$

Consider any phase configuration of N oscillators, i.e.,

$$\Theta = (\theta_1, \ldots, \theta_N) \in \mathbb{R}^N.$$

We will say that the i-th oscillator *collides* with the j-th oscillator when $\Theta \in \mathcal{C}_{ij}$, and we will say that Θ is a *collision state* when $\Theta \in \mathcal{C}$. To deal with collisions, we define the binary relation

$$i \overset{\Theta}{\sim} j \iff \Theta \in \mathcal{C}_{ij}.$$

This is an equivalence relation, and we denote the equivalence classes by

$$\mathcal{C}_i(\Theta) := \{j \in \{1, \ldots, N\} : i \overset{\Theta}{\sim} j\} = \{j \in \{1, \ldots, N\} : \Theta \in \mathcal{C}_{ij}\}. \tag{2.7}$$

As is apparent from the definition, $\mathcal{C}_i(\Theta)$ is the set of *indices of collision* with the i-th oscillator. Thus, Θ is a collision state when some of its equivalence classes is non-trivial. Consequently, each equivalence class can be regarded as a *cluster of oscillators*. Let us denote by $\mathcal{E}(\Theta)$ the family of all the different equivalence classes, that is, clusters. It is apparent that $\mathcal{E}(\Theta)$ establishes a partition of $\{1, \ldots, N\}$, which we will call the *collisional type* of Θ. For simplicity of notation, we will enumerate the equivalence classes

$$\mathcal{E}(\Theta) = \{E_1(\Theta), \ldots, E_{\kappa(\Theta)}(\Theta)\}$$

in such a way that the minimal representatives in each of them, i.e., $\iota_k(\Theta) := \min E_k(\Theta)$, are increasingly ordered. $\kappa(\Theta) := \#\mathcal{E}(\Theta)$ will denote the total number of clusters in the phase configuration Θ and we will denote the size of the k-th cluster, that is, the number of particles which form the k-th cluster, by $n_k(\Theta) := \#E_k(\Theta)$, for each $k = 1, \ldots, \kappa(\Theta)$.

Assume now that not only do we know some phase configuration at a particular time, but we also know the whole absolutely continuous trajectory $t \mapsto \Theta(t) = (\theta_1(t), \ldots, \theta_N(t)) \in \mathbb{R}^N$ governing the dynamics of the N oscillators. Then, as long as it is clear from the context, we will simplify the notation and write

$$\mathcal{C}_i(t) := \mathcal{C}_i(\Theta(t)), \quad \mathcal{E}(t) := \mathcal{E}(\Theta(t)), \quad \kappa(t) := \kappa(\Theta(t)), \quad n_k(t) := n_k(\Theta(t)).$$

Similarly, time may be omitted in our notation for simplicity. Apart from collisions and the subsequent formation of clusters, it is important to characterize when those clusters remain stuck together. If the i-th and j-th oscillators have collided at time t, we will say that they *stick together* when

$$\bar{\theta}_i(s) = \bar{\theta}_j(s) \quad \text{for all } s \geq t.$$

Then we can define the set of *indices of sticking* with the i-th oscillator by

$$S_i(t) := \{j \in \mathcal{C}_i(t) : \bar{\theta}_i(s) = \bar{\theta}_j(s) \text{ for all } s \geq t\}. \tag{2.8}$$

In Section 3 we will introduce some results about the clustering and sticking behavior of solutions to our singular weighted Kuramoto model (2.5) with $\varepsilon = 0$.

2.4. *Gradient flow structure*

Note that (2.1) can be equivalently turned into a gradient flow system:

$$\dot{\Theta} = -\nabla V(\Theta), \tag{2.9}$$

governed by a potential $V = V(\Theta)$ defined by

$$V(\Theta) = -\sum_{i=1}^{N} \Omega_i \theta_i + V_{\text{int}}(\Theta) := -\sum_{i=1}^{N} \Omega_i \theta_i + \frac{K}{2N} \sum_{i \neq j} W(\theta_j - \theta_i). \tag{2.10}$$

Here, W is the primitive function of h such that $W(0) = 0$, i.e.,

$$W(\theta) := \int_0^\theta h(\theta')\, d\theta'. \tag{2.11}$$

The function W can be regarded as the interaction potential of binary interactions, while V_{int} stands for the total interaction potential due to binary interactions. This approach is obviously formal and relies on specifying the regularity of the plasticity function Γ. For instance, if we choose Γ to be analytic, then (2.1) can be regarded as a gradient flow system with analytic potential V. In that particular case, one can simplify the proof of synchronization as in the classical Kuramoto model [17]. Specifically, some boundedness property of the trajectory is all we need to ensure exponential convergence towards a phase-locked state by the Łojasiewicz inequality for analytic functions. For the choices of plasticity function of interest in this paper, i.e., (1.4) and (1.6), analyticity is lacking and this approach does not necessarily work. Nevertheless, we will focus on values $\alpha \in (0,1)$, and consequently V will be a globally continuous function that is smooth outside the set of collisions. Since in general we lack either analyticity or convexity of V, the gradient flow structure will not be used in this paper, except in Subsections 4.2 and 4.3.

2.5. *Kinetic formulation of the problem*

We now formally introduce the expected kinetic models associated with (2.5). The classical arguments rigorously proving the mean field limit as $N \to \infty$ are based on the analysis of propagation of chaos in the system as the number N of particles becomes large [26,31]. On the one hand, for every $\varepsilon > 0$ the mean field limit $f_\varepsilon = f_\varepsilon(t,\theta,\Omega)$ for the distribution function of the oscillators is governed by the following Vlasov equation with regular kernels:

$$\frac{\partial f_\varepsilon}{\partial t} + \frac{\partial}{\partial \theta}[(\Omega - K(h_\varepsilon * \rho_\varepsilon) f_\varepsilon] = 0, \qquad t \in \mathbb{R}_0^+,\ \theta \in [0, 2\pi],\ \Omega \in \mathbb{R}, \tag{2.12}$$

where periodic boundary conditions in the variable θ are assumed. Here the macroscopic phase density ρ_ε is just

$$\rho_\varepsilon(t,\theta) := \int_{\mathbb{R}} f_\varepsilon \, d\Omega.$$

Similarly, when $\varepsilon = 0$ the corresponding mean field limit $f = f(t,\theta,\Omega)$ for the distribution function of the oscillators is subject to a Vlasov equation with singular kernel

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \theta}[(\Omega - K(h * \rho) f] = 0, \qquad t \in \mathbb{R}_0^+,\ \theta \in [0, 2\pi],\ \Omega \in \mathbb{R}, \tag{2.13}$$

with analogous periodic conditions in θ. The derivation of the mean field limit is much more involved in the latter case and requires a sharper analysis; see [8,23,33] for a related singular model like the Cucker–Smale model with weakly singular influence function. Let us briefly recall the main formal idea supporting the above mean field limit through the empirical measures approach. Fix the following empirical measure as initial condition in (2.13):

$$\mu_0^N(\theta,\Omega) = \frac{1}{N}\sum_{i=1}^N \delta_{\theta_{i,0}^N}(\theta)\delta_{\Omega_i^N}(\Omega),$$

associated to some discrete initial configuration $\Theta_0^N = (\theta_{1,0}^N, \ldots, \theta_{N,0}^N)$. By the results in this paper, the Filippov solution $\Theta^N(t) = (\theta_1^N(t), \ldots, \theta_N^N(t))$ to the singular discrete model allows considering the following measure-valued solution to (2.13):

$$\mu_t^N(\theta,\Omega) = \frac{1}{N}\sum_{i=1}^N \delta_{\theta_i^N(t)}(\theta)\delta_{\Omega_i^N}(\Omega).$$

The ultimate effort to make is to show that the weak limit f of μ^N as $N \to \infty$ is another measure-valued solution in some generalized sense to the singular macroscopic system. For a comprehensive analysis of the singular macroscopic model (2.13) see [40]. Also, see [33] for a close approach in the Cucker–Smale model with weakly singular influence kernel corresponding to the smaller range of parameters $\alpha \in (0, 1/4)$ of the subcritical regime. Analogous results in aggregation models and the classical Kuramoto model have been studied in [8, 11, 12] and [7, 31] respectively.

3. Well-posedness of singular interaction

We now consider the Kuramoto model with singular coupling Γ, which we introduced in Section 2 as a singular limit of regular weighted coupling

$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^{N}\frac{\sin(\theta_j-\theta_i)}{|\theta_j-\theta_i|_o^{2\alpha}}, \quad i = 1, \dots, N. \tag{3.1}$$

Recall that in the limit as $\varepsilon \to 0$ of the regular kernel h_ε we recover the singular interaction kernel of the model, i.e.,

$$h(\theta) := \frac{\sin\theta}{|\theta|_o^{2\alpha}}.$$

For simplicity, we will forget about the constant $c = c_{\alpha,\zeta} = 1 - \zeta^{-1/\alpha}$. Then we can rewrite the system (3.1) as

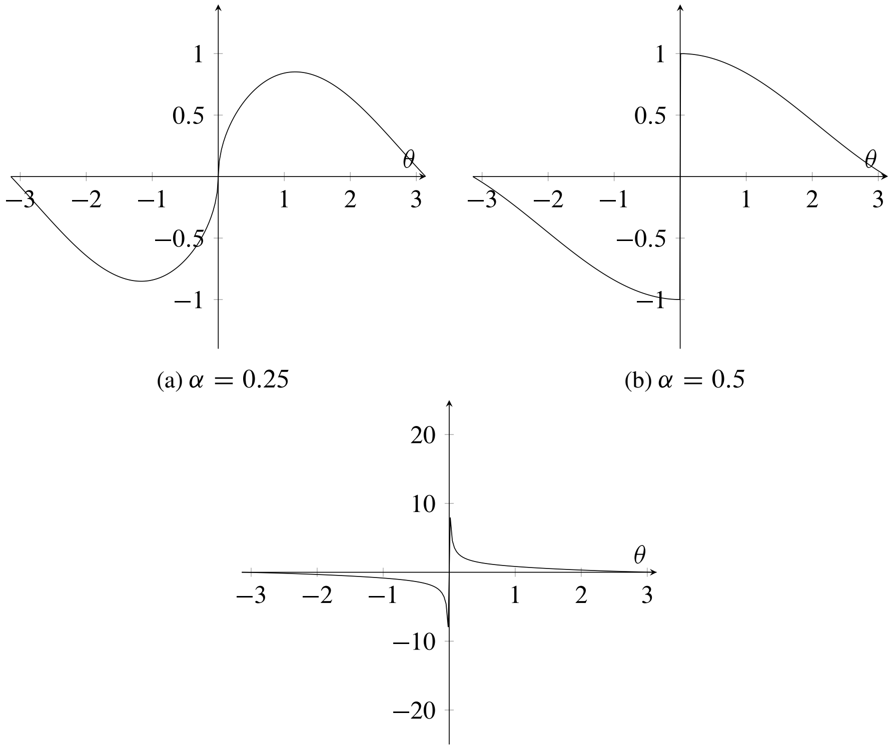
$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^{N} h(\theta_j - \theta_i), \quad i = 1, \dots, N. \tag{3.2}$$

Regarding the parameter α, it belongs to the interval $(0, 1)$ to allow for mild singularities. Note that the kernel is continuous for $\alpha \in (0, 1/2)$, it exhibits a jump discontinuity at $\theta \in 2\pi\mathbb{Z}$ for $\alpha = 1/2$, and it shows essential discontinuities for $\alpha \in (1/2, 1)$ (see Figure 1).

In this section, we will focus on developing the well-posedness theory of the system (3.1) of coupled ODEs. Note that uniqueness is not trivial even in the subcritical case. Indeed, due to the choice of singular plasticity function, the right hand side of the system (3.2) is not Lipschitz-continuous in any of the subcritical, critical and supercritical regimes. Thus, we need to explore existence and uniqueness of solution to the system (3.1) before we proceed to study synchronization. For the following discussion, we recall the definition of the vector field $H = H(\Theta)$ in (2.2) that allows dealing with the system (3.2) in the vector form (2.2).

3.1. *Well-posedness in the subcritical regime*

In the subcritical case, namely $\alpha \in (0, 1/2)$, the vector field $H = H(\Theta)$ in (2.2) is continuous. Therefore, it is a clear consequence of Peano's theorem that (3.1) has a local-in-time



(a) $\alpha = 0.25$ (b) $\alpha = 0.5$

(c) $\alpha = 0.75$

Fig. 1. Plot of the interaction kernel $h = h(\theta)$ in (3.1) for the values $\alpha = 0.25$, $\alpha = 0.5$ and $\alpha = 0.75$.

solution for every initial configuration $\Theta(0) = \Theta_0 \in \mathbb{R}^N$. Unfortunately, note that $h(\theta)$ exhibits an infinite slope at the phase values $\theta \in 2\pi\mathbb{Z}$ and so the classical Cauchy–Picard–Lindelöf theorem does not apply since $H = H(\Theta)$ is no longer a Lipschitz-continuous vector field. Nevertheless, one can still use an easy trick: it is enough to show that near the points of loss of Lipschitz-continuity our vector field can be locally split into the sum of a decreasing vector field and a Lipschitz-continuous vector field, thus ensuring the local one-sided Lipschitz condition that is enough to obtain a one-sided uniqueness result.

Lemma 3.1. *Let $F : \mathbb{R}^N \to \mathbb{R}^N$ be a bounded and continuous vector field and assume that for every $x^* \in \mathbb{R}^N$ there exists some open neighborhood $\mathcal{V} \subseteq \mathbb{R}^N$ and a positive constant M such that F satisfies the one-sided Lipschitz condition in $\mathcal{V}$,*

$$(F(x) - F(y)) \cdot (x - y) \leq M|x - y|^2$$

for every couple $x, y \in \mathcal{V}$. Then the following initial value problem associated with any initial configuration $x_0 \in \mathbb{R}^N$ enjoys a global-in-time solution, which is unique forward *in time:*

$$\begin{cases} \dot{x} = F(x), & t \geq 0, \\ x(0) = x_0. \end{cases}$$

Since the proof is classical, we omit it here. Let us now apply that result to our case of interest. To do so, it is enough to introduce a decomposition of the vector field $H = H(\Theta)$ in the Kuramoto model (3.2). We first set the following split of the interaction function $h = h(\theta)$. First, consider $\overline{h}$ and $\tilde{\theta} \in (0, \pi/2)$ such that

$$\overline{h} := \max_{0<r<\pi} h(r) \quad \text{and} \quad 2\alpha \sin\tilde{\theta} = \tilde{\theta}\cos\tilde{\theta}.$$

Note that $\tilde{\theta}$ is uniquely defined as the value in $(0, \pi)$ where h attains its maximum. Second, define a couple of functions $f = f(\theta)$ and $g = g(\theta)$ in $(-\pi, \pi)$ as follows:

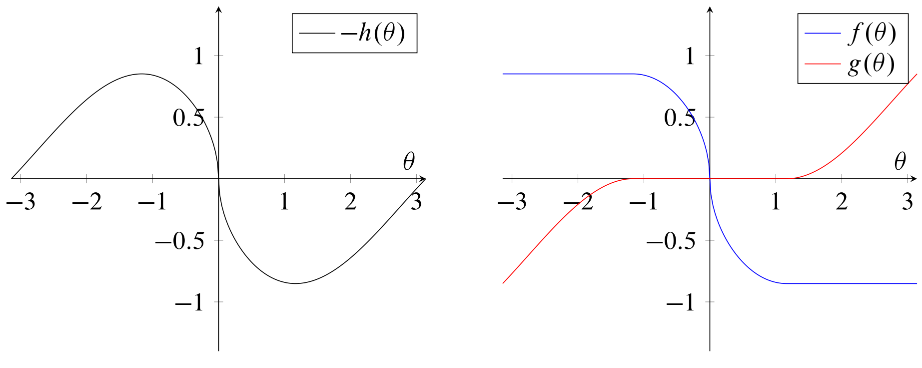
$$f(\theta) := \begin{cases} \overline{h} & \text{for } \theta \in (-\pi, -\tilde{\theta}), \\ -h(\theta) & \text{for } \theta \in [-\tilde{\theta}, \tilde{\theta}), \\ -\overline{h} & \text{for } \theta \in [\tilde{\theta}, \pi), \end{cases}$$

$$g(\theta) := \begin{cases} -\overline{h} - h(\theta) & \text{for } \theta \in (-\pi, -\tilde{\theta}), \\ 0 & \text{for } \theta \in [-\tilde{\theta}, \tilde{\theta}), \\ \overline{h} - h(\theta) & \text{for } \theta \in [\tilde{\theta}, \pi). \end{cases}$$

Notice that

$$-h(\theta) = f(\theta) + g(\theta) \quad \text{for all } \theta \in (-\pi, \pi), \tag{3.3}$$

as depicted in Figure 2.



(a) $-h(\theta)$ (b) $f(\theta)$ and $g(\theta)$

Fig. 2. Graph of the function $-h(\theta)$ and the functions $f(\theta)$ and $g(\theta)$ in the decomposition for the value $\alpha = 0.25$.

Remark 3.1. Note that although $-h(\theta)$ is not a Lipschitz-continuous function because of the infinite slope at $\theta \in 2\pi\mathbb{Z}$, one can locally decompose it around those values in terms of a decreasing function $f(\theta)$ and a Lipschitz-continuous function $g(\theta)$.

Finally, for any $\Theta^* = (\theta_1^*, \ldots, \theta_N^*) \in \mathbb{R}^N$ we will locally decompose H around Θ^*. For $\Theta = (\theta_1, \ldots, \theta_N)$ in a small enough neighborhood $\mathcal{V}$ of Θ^* in $\mathbb{R}^N$, we set

$$F_i(\Theta) := \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} f(\overline{\theta_i - \theta_j}), \tag{3.4}$$

$$G_i(\Theta) := \Omega_i + \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} g(\overline{\theta_i - \theta_j}) - \frac{K}{N} \sum_{j \notin \mathcal{C}_i(\Theta^*)} h(\theta_i - \theta_j), \tag{3.5}$$

where we recall that $\mathcal{C}_i(\Theta^*)$ stands for the set of indices of collision with the i-th oscillator in the phase configuration Θ^* (see Subsection 2.3).

Proposition 3.1. *Let $\Theta^* = (\theta_1^*, \ldots, \theta_N^*) \in \mathbb{R}^N$, and define the vector fields*

$$F, G : \mathcal{V} \to \mathbb{R}^N$$

via the formulas (3.4)–(3.5), for a small enough neighborhood $\mathcal{V}$ of Θ^* in $\mathbb{R}^N$. Then:

(1) *$H = F + G$ in $\mathcal{V}$.*

(2) *F is decreasing in $\mathcal{V}$.*

(3) *G is Lipschitz-continuous in $\mathcal{V}$.*

(4) *H is one-sided Lipschitz-continuous in $\mathcal{V}$.*

Proof. The decomposition of H is clear by the decomposition (3.3) and the definitions (3.4)–(3.5). Let us therefore focus on the last three properties.

First, consider $\Theta = (\theta_1, \ldots, \theta_N), \widetilde{\Theta} = (\widetilde{\theta}_1, \ldots, \widetilde{\theta}_N) \in \mathbb{R}^N$ in a small enough neighborhood of Θ^*. Without loss of generality, we will assume that $\theta_i - \theta_j$ and $\widetilde{\theta}_i - \widetilde{\theta}_j$ belong to $(-\pi, \pi]$. In the other case, we just need to work with representatives. On the one hand,

$$(F(\Theta) - F(\widetilde{\Theta})) \cdot (\Theta - \widetilde{\Theta}) = \frac{K}{N} \sum_{i=1}^N \sum_{j \in \mathcal{C}_i(\Theta^*)} \big(f(\theta_i - \theta_j) - f(\widetilde{\theta}_i - \widetilde{\theta}_j)\big)(\theta_i - \widetilde{\theta}_i).$$

Interchanging the indices i and j we obtain

$$\begin{aligned}(F(\Theta) - F(\widetilde{\Theta})) \cdot (\Theta - \widetilde{\Theta}) &= \frac{K}{N} \sum_{j=1}^N \sum_{i \in \mathcal{C}_j(\Theta^*)} \big(f(\theta_j - \theta_i) - f(\widetilde{\theta}_j - \widetilde{\theta}_i)\big)(\theta_j - \widetilde{\theta}_j) \\ &= -\frac{K}{N} \sum_{i=1}^N \sum_{j \in \mathcal{C}_i(\Theta^*)} \big(f(\theta_i - \theta_j) - f(\widetilde{\theta}_i - \widetilde{\theta}_j)\big)(\theta_j - \widetilde{\theta}_j),\end{aligned}$$

where the properties of the sets $\mathcal{C}_i(\Theta^*)$ along with the antisymmetry of f have been used in the last line. Taking the mean value of both expressions and using the fact that f is decreasing, we arrive at

$$
\begin{aligned}
&(F(\Theta) - F(\widetilde{\Theta})) \cdot (\Theta - \widetilde{\Theta}) \\
&\qquad = \frac{K}{2N} \sum_{i=1}^{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} \big(f(\theta_i - \theta_j) - f(\widetilde{\theta}_i - \widetilde{\theta}_j)\big)\big((\theta_i - \theta_j) - (\widetilde{\theta}_i - \widetilde{\theta}_j)\big) \leq 0,
\end{aligned}
$$

and, as a consequence, we get the monotonicity of F. On the other hand,

$$
\begin{aligned}
&|G_i(\Theta) - G_i(\widetilde{\Theta})| \\
&\qquad \leq \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} |g(\theta_i - \theta_j) - g(\widetilde{\theta}_i - \widetilde{\theta}_j)| + \frac{K}{N} \sum_{j \notin \mathcal{C}_i(\Theta^*)} |h(\theta_i - \theta_j) - h(\widetilde{\theta}_i - \widetilde{\theta}_j)|.
\end{aligned}
$$

Since g is Lipschitz-continuous in $(-\pi, \pi)$ and h is locally Lipschitz-continuous in $(-\pi, \pi) \setminus \{0\}$, there exists some constant $M = M(\mathcal{V})$ such that

$$
|G_i(\Theta) - G_i(\widetilde{\Theta})| \leq \frac{KM}{N} \sum_{j=1}^{N} |(\theta_i - \theta_j) - (\widetilde{\theta}_i - \widetilde{\theta}_j)| \leq \frac{N+1}{N} KM |\Theta - \widetilde{\Theta}|
$$

for every $i \in \{1, \ldots, N\}$, thus yielding the Lipschitz-continuity of G in $\mathcal{V}$. The last item is a simple consequence: consider $x, y \in \mathcal{V}$ and note that

$$
\begin{aligned}
(H(x) - H(y)) \cdot (x - y) &= (F(x) - F(y)) \cdot (x - y) + (G(x) - G(y)) \cdot (x - y) \\
&\leq \frac{N+1}{N} KM |x - y|^2,
\end{aligned}
$$

where the preceding two properties have been used along with the Cauchy–Schwarz inequality. ■

Finally, putting together Lemma 3.1 and Proposition 3.1, one concludes the following well-posedness property.

Theorem 3.1. *For any initial configuration, there is a global-in-time strong solution to the system (3.2) with $\alpha \in (0, 1/2)$; the solution is unique forwards in time.*

The next result is a simple consequence of the above well-posedness theorem and characterizes the eventual emergence of sticking in a cluster after a potential collision.

Theorem 3.2. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the global-in-time solution in Theorem 3.1. Assume that two oscillators collide at t^*, i.e., $\bar{\theta}_i(t^*) = \bar{\theta}_j(t^*) = \theta^*$ for some $i \neq j$. Then the following two statements are equivalent:*

(1) *θ_i and θ_j stick together at t^*.*

(2) *Their natural frequencies agree, i.e.,*

$$
\Omega_i = \Omega_j. \tag{3.6}
$$

Proof. Without loss of generality, assume that $i = 1$, $j = 2$ and $\theta_1(t^*) = \theta_2(t^*) \in (-\pi, \pi]$. Assume that the two particles keep stuck together after time t^*. Then, looking at the first two equations in (3.2), it is clear that $\Omega_1 = \Omega_2$. Conversely, assume that $\Omega_1 = \Omega_2 =: \Omega$ and consider the following system of $N - 1$ ODEs:

$$\dot{\vartheta} = \Omega + \frac{K}{N}\sum_{j=3}^{N} h(\vartheta_j - \vartheta),$$

$$\dot{\vartheta}_i = \Omega_i + \frac{2K}{N} h(\vartheta - \vartheta_i) + \frac{K}{N}\sum_{j=3}^{N} h(\vartheta_j - \vartheta_i), \quad i = 3, \ldots, N,$$

with initial data given by

$$(\vartheta(t^*), \vartheta_3(t^*), \ldots, \vartheta_N(t^*)) = (\theta^*, \theta_3(t^*), \ldots, \theta_N(t^*)).$$

A similar technique to that in Theorem 3.1 clearly yields a global-in-time solution to this initial value problem. Hence, the two trajectories in $\mathbb{R}^N$,

$$t \mapsto (\theta_1(t), \theta_2(t), \theta_3(t), \ldots, \theta_N(t)), \quad t \mapsto (\vartheta(t), \vartheta(t), \vartheta_3(t), \ldots, \vartheta_N(t)),$$

are both solutions to (3.2) that at $t = t^*$ take the value

$$(\theta^*, \theta^*, \theta_3(t^*), \ldots, \theta_N(t^*)).$$

By uniqueness they agree, and in particular $\theta_1(t) = \vartheta(t) = \theta_2(t)$ for all $t \geq t^*$. ∎

3.2. *Well-posedness in the critical regime*

In the critical case, i.e. $\alpha = 1/2$, the vector field $H = H(\Theta)$ is no longer continuous and the Peano existence theorem does not work. Nevertheless, in that case H is still a measurable and essentially bounded vector field. Consequently, one can apply Filippov's existence criterion [4, 14].

We introduce the necessary notation that will be used from now on: $2^{\mathbb{R}^N}$ stands for the power set of $\mathbb{R}^N$, $|\mathcal{N}|$ is the Lebesgue measure of any measurable set $\mathcal{N} \subseteq \mathbb{R}^N$, $\text{co}(A)$ is the convex hull of A and $\overline{\text{co}}(A) = \overline{\text{co}(A)}$ is its closure. For every convex set C we denote by $m(C)$ the element of minimal norm of C, i.e. $m(C) = \pi_C(0)$, where π_C is the orthogonal projection operator over the convex set C. The main ingredient will be the Filippov set-valued map of a given single-valued measurable map.

Definition 3.1. Let $F : \mathbb{R}^N \to \mathbb{R}^N$ be any measurable map. The *Filippov set-valued map* $\mathcal{F} : \mathbb{R}^N \to 2^{\mathbb{R}^N}$ of F is defined for any $x \in \mathbb{R}^N$ as follows:

$$\mathcal{F}(x) := \bigcap_{\delta>0} \bigcap_{|\mathcal{N}|=0} \overline{\text{co}}(F(B_\delta(x) \setminus \mathcal{N})).$$

The main interest in considering this map can be summarized in the next couple of results (see [4, Theorem 2.1.3, Theorem 2.1.4, Proposition 2.1.1]).

Lemma 3.2. *Let $F:\mathbb{R}^N\to\mathbb{R}^N$ be any measurable map and consider its Filippov set-valued map $\mathcal{F}$. Then:*

(1) *$\mathcal{F}(x)$ is a closed and convex set for every $x\in\mathbb{R}^N$.*

(2) *$F(x)\in\mathcal{F}(x)$ for almost every $x\in\mathbb{R}^N$.*

(3) *If F is continuous at $x\in\mathbb{R}^n$, then $\mathcal{F}(x)=\{F(x)\}$.*

(4) *If $\mathcal{F}$ takes non-empty values, then $\mathcal{F}$ has closed graph.*

(5) *If $\mathcal{F}$ has closed graph and $m(\mathcal{F})(U_x)$ lies in a compact set for some neighborhood U_x of each $x\in\mathbb{R}^N$, then $\mathcal{F}$ is upper semicontinuous.*

(6) *If F is locally essentially bounded, then $\mathcal{F}$ is upper semicontinuous, it takes non-empty values and $m(\mathcal{F})(U_x)$ lies in a compact set for some neighborhood U_x of each $x\in\mathbb{R}^N$.*

(7) *If F is essentially bounded, then $\mathcal{F}$ is upper semicontinuous, it takes non-empty values and $m(\mathcal{F})(\mathbb{R}^N)$ lies in a compact set.*

Here $m(\mathcal{F})$ stands for the map $m(\mathcal{F})(x):=m(\mathcal{F}(x))$ for every $x\in\mathbb{R}^N$.

Lemma 3.3. *Let $\mathcal{F}:\mathbb{R}^N\to 2^{\mathbb{R}^N}$ be any set valued-map with non-empty closed and convex values. Assume that $\mathcal{F}$ is upper semicontinuous and consider the following initial value problem associated with any given initial datum $x_0\in\mathbb{R}^N$:*

$$\begin{cases}\dot{x}\in\mathcal{F}(x),\\ x(0)=x_0.\end{cases}\tag{IVP}$$

(1) *If $m(\mathcal{F})(U_x)$ lies in a compact set for some neighborhood U_x of any $x\in\mathbb{R}^N$, then* (IVP) *has an absolutely continuous local-in-time solution.*

(2) *If $m(\mathcal{F})(\mathbb{R}^N)$ lies in a compact set, then* (IVP) *has an absolutely continuous global-in-time solution.*

Putting together Lemmas 3.2 and 3.4 we arrive at the next result.

Lemma 3.4. *Let $F:\mathbb{R}^N\to\mathbb{R}^N$ be any measurable map and $\mathcal{F}$ its Filippov set-valued map.*

(1) *If F is locally essentially bounded, then* (IVP) *has an absolutely continuous local-in-time solution.*

(2) *If, in addition, F is globally essentially bounded, then such a solution is indeed global.*

The solutions to the above differential inclusion are called *solutions in Filippov's sense* to the original discontinuous dynamical system. To deal with uniqueness we first introduce the next technical result.

Lemma 3.5. *Let $F:\mathbb{R}^N\to\mathbb{R}^N$ be a measurable and locally essentially bounded map and $\mathcal{F}:\mathbb{R}^N\to 2^{\mathbb{R}^N}$ its Filippov set-valued map. If F satisfies the one-sided Lipschitz* *condition a.e., then $\mathcal{F}$ also satisfies it in the set-valued sense: there exists a positive constant M such that*

$$(X - Y) \cdot (x - y) \leq M|x - y|^2$$

for all $x, y \in \mathbb{R}^N$ and all $X \in \mathcal{F}(x)$, $Y \in \mathcal{F}(y)$.

Proof. Fix $x, y \in \mathbb{R}^N$ and $X \in \mathcal{F}(x), Y \in \mathcal{F}(y)$. Also fix any $\delta > 0$ (assume $\delta < 1$ without loss of generality) and any negligible set $\mathcal{N}$. Using the definition of $\mathcal{F}$, we see that

$$X \in \overline{\mathrm{co}}(F(B_\delta(x) \setminus \mathcal{N})) \quad \text{and} \quad Y \in \overline{\mathrm{co}}(F(B_\delta(y) \setminus \mathcal{N})).$$

Thus, one can find sequences $\{X_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^N$ and $\{Y_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^N$ such that $X_n \to X$, $Y_n \to Y$ and

$$X_n \in \mathrm{co}(F(B_\delta(x) \setminus \mathcal{N})) \quad \text{and} \quad Y_n \in \mathrm{co}(F(B_\delta(y) \setminus \mathcal{N})),$$

for every $n \in \mathbb{N}$. Therefore, the Carathéodory theorem from convex analysis allows representing X_n and Y_n as convex combinations,

$$X_n = \sum_{i=1}^{N+1} \alpha_i^n F(x_i^n) \quad \text{and} \quad Y_n = \sum_{j=1}^{N+1} \beta_j^n F(y_j^n),$$

where $x_i^n \in B_\delta(x) \setminus \mathcal{N}$, $y_j^n \in B_\delta(y) \setminus \mathcal{N}$ and the coefficients $\alpha_i^n, \beta_j^n \in [0,1]$ satisfy

$$\sum_{i=1}^{N+1} \alpha_i^n = 1 = \sum_{j=1}^{N+1} \beta_j^n.$$

Note that

$$X_n = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \alpha_i^n \beta_j^n F(x_i^n) \quad \text{and} \quad Y_n = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \alpha_i^n \beta_j^n F(y_j^n).$$

By defining the constants

$$M_x := \operatorname*{ess\,sup}_{z\in B_1(x)} |F(z)| \quad \text{and} \quad M_y := \operatorname*{ess\,sup}_{z\in B_1(y)} |F(z)|,$$

we have

$$\begin{aligned}
(X_n - Y_n) \cdot (x - y) &= \Big(\sum_{i,j=1}^{N+1} \alpha_i^n \beta_j^n (F(x_i^n) - F(y_j^n))\Big) \cdot (x - y) \\
&= \sum_{i,j=1}^{N+1} \alpha_i^n \beta_j^n \big((F(x_i^n) - F(y_j^n)) \cdot (x - y)\big) \\
&= \sum_{i,j=1}^{N+1} \alpha_i^n \beta_j^n \big((F(x_i^n) - F(y_j^n)) \cdot (x_i^n - y_j^n) \\
&\qquad + (F(x_i^n) - F(y_j^n)) \cdot ((x - x_i^n) - (y - y_j^n))\big)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{i,j=1}^{N+1} \alpha_i^n \beta_j^n \big(M|x_i^n - y_j^n|^2 + 2(M_x + M_y)\delta\big) \\
&\leq \sum_{i,j=1}^{N+1} \alpha_i^n \beta_j^n \big(M(|x-y| + 2\delta)^2 + 2(M_x + M_y)\delta\big) \\
&= M(|x-y| + 2\delta)^2 + 2(M_x + M_y)\delta.
\end{aligned}
$$

Since the above property holds for arbitrary $n \in \mathbb{N}$ and $0 < \delta < 1$, we obtain

$$
(X - Y) \cdot (x - y) \leq M|x - y|^2. \qquad \blacksquare
$$

Lemma 3.6. *Let $F : \mathbb{R}^N \to \mathbb{R}^N$ be a measurable and essentially bounded vector field, and $\mathcal{F} : \mathbb{R}^N \to 2^{\mathbb{R}^N}$ its Filippov set-valued map. In addition, assume that F satisfies the local one-sided Lipschitz condition. Then the following initial value problem associated with any initial configuration $x_0 \in \mathbb{R}^N$ enjoys a global-in-time absolutely continuous solution, which is unique forwards in time:*

$$
\begin{cases}
\dot{x} \in \mathcal{F}(x), & t \geq 0, \\
x(0) = x_0.
\end{cases}
$$

Proof. The existence of global-in-time Filippov solutions follows from Lemma 3.4. Let us just discuss the uniqueness of solution. We consider two Filippov solutions $x_1 = x_1(t)$ and $x_2 = x_2(t)$ with the same initial datum x_0 and define

$$
T := \inf\{t > 0 : x_1(t) \neq x_2(t)\}.
$$

Our main goal is to prove that $T = \infty$. Assume $T < \infty$. Let $x^* := x_1(T) = x_2(T)$ and take a small enough neighborhood $\mathcal{V}$ of x^* such that F satisfies the one-sided Lipschitz condition in it. By continuity there is $\varepsilon > 0$ such that $x_1(t), x_2(t) \in \mathcal{V}$ for every $t \in [T, T + \varepsilon]$. Consequently,

$$
\frac{d}{dt}\frac{1}{2}|x_1 - x_2|^2 \in (\mathcal{F}(x_1(t)) - \mathcal{F}(x_2(t))) \cdot (x_1(t) - x_2(t)).
$$

By the one-sided Lipschitz condition, there is a constant M depending on x^* such that

$$
\frac{d}{dt}|x_1 - x_2|^2 \leq M|x_1 - x_2|^2
$$

for every $t \in [T, T + \varepsilon]$. By Grönwall's inequality, one then obtains $x_1(t) = x_2(t)$ for every $t \in [T, T + \varepsilon]$, and this contradicts the assumption $T < \infty$. $\blacksquare$

Let us now explicitly compute the Filippov set-valued map $\mathcal{H} = \mathcal{H}(\Theta)$ of our particular vector field $H = H(\Theta)$ for the critical case $\alpha = 1/2$. See Subsection 2.3 for the collision equivalence relation and the necessary notation to deal with clusters of oscillators.

Proposition 3.2. *In the critical regime $\alpha = 1/2$, the value $\mathcal{H}(\Theta)$ of the Filippov set-valued map associated with H is the convex and compact polytope consisting of the points $(\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ such that*

$$\omega_i = \Omega_i + \frac{K}{N} \sum_{j \notin \mathcal{C}_i(\Theta)} h(\theta_j - \theta_i) + \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta) \setminus \{i\}} y_{ij} \quad \text{for all } i = 1, \ldots, N,$$

for some $Y = (y_{ij})_{1 \le i,j \le N} \in \mathrm{Skew}_N([-1, 1])$.

Since the proof is clear by definition of the Filippov map, we omit it.

Remark 3.2. Notice that for every $(\omega_1, \ldots, \omega_N) \in \mathcal{H}(\Theta)$,

$$\sum_{i=1}^N \omega_i = \sum_{i=1}^N \Omega_i.$$

In particular, every Filippov solution $(\theta_1, \ldots, \theta_N)$ to (3.2), in the case $\alpha = 1/2$, satisfies

$$\sum_{i=1}^N \dot{\theta}_i(t) = \sum_{i=1}^N \Omega_i \quad \text{for a.e. } t \ge 0.$$

Hence, the Filippov solutions in the critical case still preserve the average frequency like classical solutions do for the subcritical case or in the original Kuramoto model.

Example 3.1. To gain some intuition about those sets, let us exhibit some examples:

(1) For every $N \in \mathbb{N}$, if $\Theta \notin \mathcal{C}$, then $\mathcal{H}(\Theta) = \{H(\Theta)\}$.

(2) For $N = 2$, if $\Theta = (\theta_1, \theta_2) \in \mathcal{C}_{12}$, then $\mathcal{H}(\Theta)$ is the polytope consisting of the points $(\omega_1, \omega_2) \in \mathbb{R}^2$ such that

$$\omega_1 = \Omega_1 + \frac{K}{2} y_{12},$$
$$\omega_2 = \Omega_2 - \frac{K}{2} y_{12},$$

for some $y_{12} \in [-1, 1]$.

(3) For $N = 3$, if $\Theta = (\theta_1, \theta_2, \theta_3) \in \mathcal{C}_{12} \setminus \mathcal{C}_{13}$, then $\mathcal{H}(\Theta)$ is the polytope consisting of the points $(\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$ such that

$$\omega_1 = \Omega_1 + \frac{K}{3} h(\theta_3 - \theta_1) + \frac{K}{3} y_{12},$$
$$\omega_2 = \Omega_2 + \frac{K}{3} h(\theta_3 - \theta_2) - \frac{K}{3} y_{12},$$
$$\omega_3 = \Omega_3 + \frac{K}{3} h(\theta_1 - \theta_3) + \frac{K}{3} h(\theta_2 - \theta_3),$$

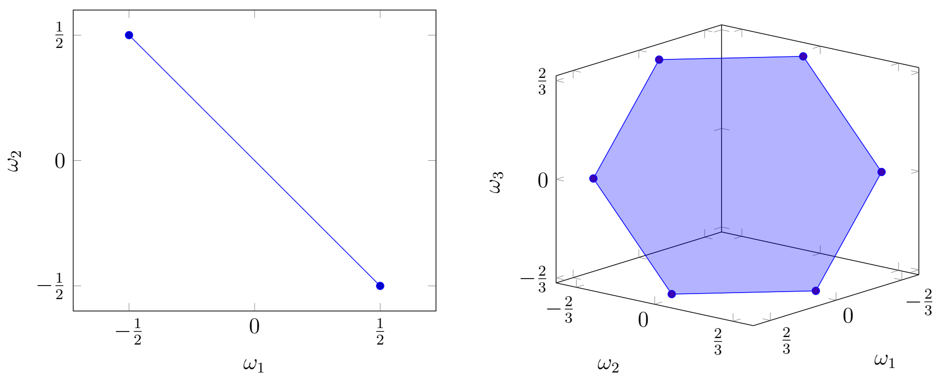
for some $y_{12} \in [-1, 1]$.

(4) For $N = 3$, if $\Theta = (\theta_1, \theta_2, \theta_3) \in \mathscr{C}_{12} \cap \mathscr{C}_{13}$, then $\mathscr{H}(\Theta)$ is the polytope consisting of the points $(\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$ such that

$$\omega_1 = \Omega_1 + \frac{K}{3} y_{12} + \frac{K}{3} y_{13},$$
$$\omega_2 = \Omega_2 - \frac{K}{3} y_{12} + \frac{K}{3} y_{23},$$
$$\omega_3 = \Omega_3 - \frac{K}{3} y_{13} - \frac{K}{3} y_{23},$$

for some $y_{12}, y_{13}, y_{23} \in [-1, 1]$.

The second and third cases yield line segments and the last one is a regular hexagon as depicted in Figure 3.



(a) $N = 2$ (b) $N = 3$

Fig. 3. Pictures of the Filippov set-valued map in the critical case at a total collision phase configuration. In (a), $N = 2$ and the polytope is a line segment joining $(\Omega_1 \pm K/2, \Omega_2 \mp K/2)$. In (b), $N = 3$ and the polytope is a regular hexagon with vertices $(\Omega_1 \pm 2K/3, \Omega_2 \mp 2K/3, \Omega_3)$, $(\Omega_1 \pm 2K/3, \Omega_2, \Omega_3 \mp 2K/3)$ and $(\Omega_1, \Omega_2 \pm 2K/3, \Omega_3 \mp 2K/3)$. For simplicity, the natural frequencies are set to zero and $K = 1$ in the figures.

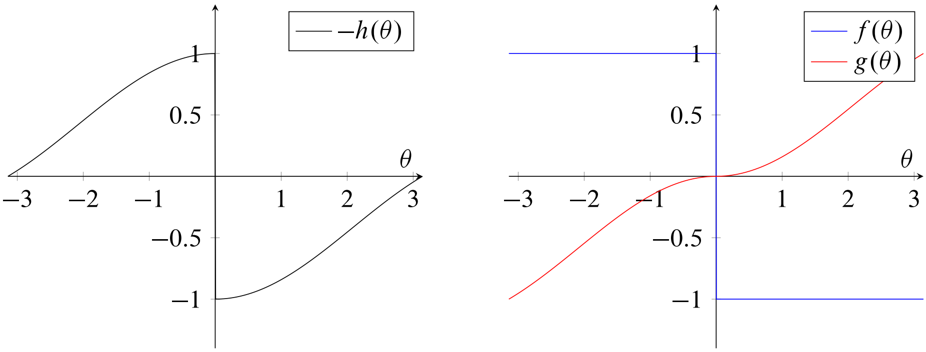
Finally, let us apply Lemma 3.6 to construct the unique Filippov solutions of our particular system (3.2) in the critical case $\alpha = 1/2$. The way to go is similar to that in Subsection 3.1 and relies on a good decomposition of $-h$. Define a couple of functions $f = f(\theta)$ and $g = g(\theta)$ in $(-\pi, \pi)$ as follows:

$$f(\theta) := \begin{cases} 1 & \text{for } \theta \in (-\pi, 0), \\ -1 & \text{for } \theta \in [0, \pi), \end{cases} \qquad g(\theta) := \begin{cases} -1 - h(\theta) & \text{for } \theta \in (-\pi, 0), \\ 1 - h(\theta) & \text{for } \theta \in [0, \pi). \end{cases}$$

Notice that

$$-h(\theta) = f(\theta) + g(\theta) \quad \text{for all } \theta \in (-\pi, \pi), \tag{3.7}$$

as depicted in Figure 4.



(a) $-h(\theta)$ (b) $f(\theta)$ and $g(\theta)$

Fig. 4. Graph of the function $-h(\theta)$ and the functions $f(\theta)$ and $g(\theta)$ in the decomposition for the value $\alpha = 0.5$.

Remark 3.3. Note that although $-h(\theta)$ is a discontinuous function because of the jump discontinuities at $\theta \in 2\pi\mathbb{Z}$, one can locally decompose it around such values in terms of a decreasing function $f(\theta)$ and a Lipschitz-continuous function $g(\theta)$.

Finally, for every $\Theta^* = (\theta_1^*, \ldots, \theta_N^*) \in \mathbb{R}^N$ we locally decompose H around Θ^* into

$$F_i(\Theta) := \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} f(\overline{\theta_i - \theta_j}), \tag{3.8}$$

$$G_i(\Theta) := \Omega_i + \frac{K}{N} \sum_{j \in \mathcal{C}_i(\Theta^*)} g(\overline{\theta_i - \theta_j}) - \frac{K}{N} \sum_{j \notin \mathcal{C}_i(\Theta^*)} h(\theta_i - \theta_j), \tag{3.9}$$

where the above functions are defined almost everywhere (f does not make sense at 0, so F_i just makes sense a.e.). Recall that for any $\theta \in \mathbb{R}$, $\bar{\theta}$ is its representative modulo 2π in the interval $(-\pi, \pi]$.

Proposition 3.3. *Let $\Theta^* = (\theta_1^*, \ldots, \theta_N^*) \in \mathbb{R}^N$ and define the vector fields*

$$F, G : \mathcal{V} \to \mathbb{R}^N$$

via the formulas (3.8)–(3.9), for a small enough neighborhood $\mathcal{V}$ of Θ^* in $\mathbb{R}^N$. Then:

(1) *$H = F + G$ in $\mathcal{V}$.*

(2) *F is decreasing in $\mathcal{V}$.*

(3) *G is Lipschitz-continuous in $\mathcal{V}$.*

(4) *H is one-sided Lipschitz continuous in $\mathcal{V}$.*

Proof. The proof is analogous to that of Proposition 3.1. ∎

Finally, putting Lemmas 3.4 and 3.6 and Proposition 3.3 together, one concludes the following well-posedness result.

Theorem 3.3. *For any initial configuration, there is a global-in-time Filippov solution to the system* (3.2) *with* $\alpha = 1/2$*; the solution is unique forwards in time.*

Again, we can characterize the eventual emergence of sticking of a cluster after a potential collision in a similar way as we did in Theorem 3.2. We require the following notation. For any $N \in \mathbb{N}$, each $1 \leq m \leq N$ and every permutation σ of $\{1, \dots, N\}$ we define the following couple of $m \times m$ matrices:

$$\begin{aligned} M_m^\sigma(\Omega) &:= (\Omega_{\sigma_i} - \Omega_{\sigma_j})_{1 \leq i,j \leq m}, \\ \mathbf{J}_m &:= (1)_{1 \leq i,j \leq m}, \end{aligned} \tag{3.10}$$

i.e., $M_m^\sigma(\Omega)$ stands for the matrix of relative natural frequencies of the m oscillators with indices $i = \sigma_1, \dots, \sigma_m$ and $\mathbf{J}_m$ is the $m \times m$ matrix whose all components are 1.

Theorem 3.4. *Let* $\Theta = (\theta_1, \dots, \theta_N)$ *be the global-in-time Filippov solution of Theorem* 3.3. *Assume that* t^* *is some collision time and fix any cluster* $E_k(t^*) \equiv E_k$ *with* $k = 1, \dots, \kappa(t^*)$. *Then the following two statements are equivalent:*

(1) *The* $n_k(t^*) = \# E_k(t^*)$ *oscillators in the cluster all stick together at time* t^*.

(2) *There exists a bijection* $\sigma : \{1, \dots, n_k\} \to E_k$ *and* $Y \in \mathrm{Skew}_{n_k}([-1, 1])$ *such that*

$$M_{n_k}^\sigma(\Omega) = \frac{K}{N}(Y \cdot \mathbf{J}_{n_k} + \mathbf{J}_{n_k} \cdot Y). \tag{3.11}$$

Proof. Write $n := n_k$ for simplicity and assume that the oscillators in the cluster are the first n oscillators, i.e., $E_k = \{1, \dots, n\}$. By continuity, take some small $\varepsilon > 0$ such that $\bar{\theta}_j(t) \neq \bar{\theta}_i(t)$ for every $t \in [t^*, t^* + \varepsilon]$, any $i \in E_k$ and each $j \notin E_k$. First, assume that (1) holds true. Without loss of generality we may assume that $\theta_1(t) = \dots = \theta_n(t)$ for all $t \geq t^*$ and we define $\theta(t) := \theta_1(t) = \dots = \theta_n(t)$ for all $t \geq t^*$. Then, looking at the explicit expression in Proposition 3.2 of the Filippov set-valued map $\mathscr{H}$ we get

$$\dot{\theta}_i = \Omega_i + \frac{K}{N} \sum_{j=n+1}^{N} h(\theta_j(t) - \theta(t)) + \frac{K}{N} \sum_{j=1}^{n} y_{ij}(t)$$

for a.e. $t \in [t^*, t^* + \varepsilon]$ and every $i = 1, \dots, n$, where $y_{ij} \in L^\infty(t^*, t^* + \varepsilon)$ and $Y(t) = (y_{ij}(t))_{1 \leq i,j \leq n} \in \mathrm{Skew}_n([-1, 1])$ for almost all $t \in [t^*, t^* + \varepsilon]$. Since $\dot{\theta}_i = \dot{\theta}_j$ a.e., for all $i, j = 1, \dots, n$, we obtain the system of equations

$$\Omega_i - \Omega_j = -\frac{K}{N} \sum_{\substack{l=1 \\ l \neq i}}^{n} y_{il}(t) + \frac{K}{N} \sum_{\substack{l=1 \\ l \neq j}}^{n} y_{jl}(t)$$

for a.e. $t \in [t^*, t^* + \varepsilon]$. In particular, (3.11) holds.

Conversely, assume that (3.11) is satisfied for some $Y \in \mathrm{Skew}_n([-1,1])$. Then

$$\Omega_i + \frac{K}{N}\sum_{\substack{l=1\\ l\neq i}}^{n} y_{il} = \Omega_j + \frac{K}{N}\sum_{\substack{l=1\\ l\neq j}}^{n} y_{jl} =: \widehat{\Omega}.$$

Now consider the vector field

$$\widehat{H}^n = (\widehat{H}^n_0, \widehat{H}^n_{n+1}, \ldots, \widehat{H}^n_N) : \mathbb{R}^{N-n+1} \to \mathbb{R}^{N-n+1}$$

given by

$$\widehat{H}^n_0(\vartheta, \vartheta_{n+1}, \ldots, \vartheta_N) = \widehat{\Omega} + \frac{K}{N}\sum_{j=n+1}^{N} h(\vartheta_j - \vartheta),$$

$$\widehat{H}^n_i(\vartheta, \vartheta_{n+1}, \ldots, \vartheta_N) = \Omega_i + \frac{nK}{N} h(\vartheta - \vartheta_i) + \frac{K}{N}\sum_{j=n+1}^{N} h(\vartheta_j - \vartheta_i)$$

for $i = n+1, \ldots, N$. Also, consider the associated Filippov set-valued map $\widehat{\mathscr{H}}^n$ and the associated differential inclusion

$$(\dot{\vartheta}, \dot{\vartheta}_{n+1}, \ldots, \dot{\vartheta}_N) \in \widehat{\mathscr{H}}^n(\vartheta, \vartheta_{n+1}, \ldots, \vartheta_N),$$

with initial datum

$$(\vartheta(t^*), \vartheta_{n+1}(t^*), \ldots, \vartheta_N(t^*)) = (\theta^*, \theta_{n+1}(t^*), \ldots, \theta_N(t^*)).$$

A well-posedness result similar to that in Theorem 3.3 shows that the resulting IVP enjoys a global-in-time solution. In addition, by definition it is apparent that whenever we pick $(\omega, \omega_{n+1}, \ldots, \omega_N) \in \widehat{\mathscr{H}}^n(\vartheta, \vartheta_{n+1}, \ldots, \vartheta_N)$, we obtain

$$(\underbrace{\omega, \ldots, \omega}_{n \text{ pairs}}, \omega_{n+1}, \ldots, \omega_N) \in \mathscr{H}(\underbrace{\vartheta, \ldots, \vartheta}_{n \text{ pairs}}, \vartheta_{n+1}, \vartheta_N).$$

Consequently, the two trajectories in $\mathbb{R}^N$,

$$t \mapsto (\theta_1(t), \ldots, \theta_n(t), \theta_{n+1}(t), \ldots, \theta_N(t)),$$
$$t \mapsto (\underbrace{\vartheta(t), \ldots, \vartheta(t)}_{n \text{ pairs}}, \vartheta_{n+1}(t), \ldots, \vartheta_N(t)),$$

are Filippov solutions to (3.2) with the same value at $t = t^*$, namely,

$$(\underbrace{\theta^*, \ldots, \theta^*}_{n \text{ pairs}}, \theta_{n+1}(t^*), \ldots, \theta_N(t^*)).$$

By uniqueness they agree, and in particular

$$\theta_i(t) = \vartheta(t) \quad \text{for all } t \geq t^* \text{ and every } i = 1, \ldots, n. \qquad \blacksquare$$

The sticking condition (3.11) can be characterized in a much more explicit manner by convex analysis techniques supported by *Farkas' alternative*: see Appendix C, and in particular, the characterization of condition (3.11) in Lemma C.2. These ideas lead to the next result.

Corollary 3.1. *Under the assumptions of Theorem 3.4, the following two assertions are equivalent:*

(1) *The n_k oscillators in the cluster E_k stick together at time t^*.*

(2) *We have*

$$\left|\frac{1}{m}\sum_{i\in I}\Omega_i - \frac{1}{n_k}\sum_{i\in E_k}\Omega_i\right| \le \frac{K}{N}(n_k - m) \tag{3.12}$$

for every $1 \le m \le n_k$ and every $I \subseteq E_k$ such that $\#I = m$.

Remark 3.4. Notice that in Theorem 3.4 and Corollary 3.1 we have characterized when the whole cluster E_k remains stuck together, but not when a subcluster of a given size instantaneously splits from the remaining oscillators of the cluster. The main problem in extending the above proof is that it is hard to quantify the way in which an oscillator splits from the subcluster. Specifically, it is possible that an oscillator departs from the cluster exhibiting a left accumulation of switches of state where it instantaneously splits and collides with the subcluster. Although this accumulation phenomenon will cause some problems throughout the paper, we will show how we can overcome them.

Let us mention that this phenomenon is called *left Zeno behavior* in the literature. It appears in the Filippov solutions of some systems like the reversed bouncing ball. For instance, in [14, p. 116] Filippov proposed a discontinuous first order system with solutions exhibiting Zeno behavior. In [14, Theorem 2.10.4], he considered absence of Zeno behavior as part of sufficient (but not necessary) conditions guaranteeing forward uniqueness. We skip the analysis of Zeno behavior here and will address it in a future work.

3.3. Well-posedness in the supercritical regime

Recall that in the supercritical regime, i.e., $\alpha > 1/2$, the vector field $H = H(\Theta)$ is not only discontinuous at the collision states but also unbounded near those points (see Figure 1). Thus, the classical theory of well-posedness cannot be applied either and one might look for a notion of generalized solutions in the same sense as in the critical case $\alpha = 1/2$ (see Subsection 3.2). Hence, one strategy could be to turn again the differential equation of interest into an augmented differential inclusion given by the associated Filippov set-valued map. A similar analysis to that in Proposition 3.2 yields the following characterization of the Filippov set-valued map for the supercritical regime.

Proposition 3.4. *In the supercritical regime $\alpha > 1/2$, the value $\mathcal{H} = \mathcal{H}(\Theta)$ of the Filippov set-valued map associated with H is the convex and unbounded polytope consisting of the points $(\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ such that for some $Y = (y_{ij})_{1\le i,j\le N} \in \mathrm{Skew}_N(\mathbb{R})$,*

$$\omega_i = \Omega_i + \frac{K}{N}\sum_{j\notin\mathcal{C}_i(\Theta)} h(\theta_j - \theta_i) + \frac{K}{N}\sum_{j\in\mathcal{C}_i(\Theta)\setminus\{i\}} y_{ij} \quad \textit{for all } i = 1, \ldots, N.$$

The Filippov set-valued map enjoys similar expressions in the critical and supercritical regimes except for a "slight" change. In the former case, the coefficients y_{ij} range in the interval $[-1,1]$ whereas in the latter case they take values in the whole $\mathbb{R}$. Indeed, the examples for $\alpha = 1/2$ in Example 3.1 can also be considered for $\alpha > 1/2$. For instance, similar polytopes to those in Figure 3 are obtained at the total collision phase configurations when the corresponding polygon is replaced by its affine envelope. Those similarities ensure that any Filippov solution to (3.2) with $\alpha > 1/2$ also conserves the average frequency as in Remark 3.2. What is more, since $\mathcal{H}(\Theta)$ is apparently non-empty, Lemma 3.2 shows that $\mathcal{H}$ takes values in the non-empty, closed and convex sets and it has closed graph in the set-valued sense.

However, the unboundedness in y_{ij} entails a severe change of behavior. Specifically, it breaks the local compactness of $m(\mathcal{H})$ and, as a consequence, the existence result in Lemma 3.3 fails to work. The loss of compactness is fatal and implies that in the supercritical regime $\alpha > 1/2$ all the "classical" assumptions ensuring global existence and one-sided uniqueness do not hold. The literature about the abstract analysis of unbounded differential inclusions is rare [25, 45]. In addition, all those results require some sort of relaxed set-valued Lipschitz condition and linear growth that do not hold in our problem. Nevertheless, we will show that in some cases we can still construct a Filippov solution which is unique under some conditions.

Remark 3.5. Notice that, despite the lack of uniqueness in the supercritical case, the approach in Theorem 3.4 may still be used to obtain a partial answer. Namely, it might give a sufficient condition on the natural frequencies to ensure that after a collision of a classical solution, we can continue a Filippov solution with sticking of the resulting cluster. Since we elaborate on this idea later, we skip it here and will just focus on a necessary condition of sticking as in (3.11). Indeed, consider some Filippov solution $\Theta = (\theta_1, \ldots, \theta_N)$ to (3.2) with $\alpha > 1/2$ and assume that it is defined in an interval $[0, T)$ and that $t^* \in (0, T)$ is some collision time. Then we might fix a cluster $E_k(t^*) \equiv E_k$ and assume that the $n_k(t^*) \equiv n_k$ oscillators in the cluster stick together at time t^*. Then a proof similar to that of Theorem 3.4 would entail the existence of some bijection $\sigma : \{1, \ldots, n_k\} \to E_k$ and some $Y \in \text{Skew}_{n_k}(\mathbb{R})$ such that

$$M_m^\sigma(\Omega) = \frac{K}{N}(Y \cdot \mathbf{J}_{n_k} + \mathbf{J}_{n_k} \cdot Y). \tag{3.13}$$

One might want to obtain again a more explicit characterization of that condition. We can resort to similar ideas coming from Farkas' alternative. Lemma C.1 ensures that (3.13) is perfectly equivalent to the condition (C.2),

$$m_{ij} + m_{jk} + m_{ki} = 0$$

for all $i, j, k = 1, \ldots, n_k$, where m_{ij} denotes the (i, j) component of the matrix $M_m^\sigma(\Omega)$. Let us look into the structure of $M_{n_k}^\sigma(\Omega)$ to restate the above condition (see (3.10)):

$$m_{ij} + m_{jk} + m_{ki} = (\Omega_{\sigma_i} - \Omega_{\sigma_j}) + (\Omega_{\sigma_j} - \Omega_{\sigma_k}) + (\Omega_{\sigma_k} - \Omega_{\sigma_i}).$$

Thus, the necessary sticking condition is automatically satisfied for every given configuration of natural frequencies. This suggests that, independently of the chosen natural frequencies, any classical solution in the supercritical case that stops at a collision state might always be continued as a Filippov solution with sticking of the cluster. For this, we will need some precise control of the behavior of the classical solutions at the maximal time of existence.

Lemma 3.7. *Let $\Theta = (\theta_1, \dots, \theta_N)$ be any classical solution to* (3.2) *with $\alpha \in (1/2, 1)$ that is defined in a finite maximal existence interval $[0, t^*)$. Then:*

(1) *The solution does not blow up at t^*, i.e.,*

$$\lim_{t \to t^*} |\Theta(t)| \neq \infty.$$

(2) *The solution converges to a collision state, i.e., there exists $\Theta^* \in \mathscr{C}$ such that*

$$\lim_{t \to t^*} \Theta(t) = \Theta^*.$$

In addition, the trajectory $t \mapsto \Theta(t)$ remains absolutely continuous up to the collision time $t = t^*$; specifically, $\dot{\Theta} \in L^2((0, t^*), \mathbb{R}^N)$.

Proof. We split the proof into three parts. The first part is devoted to showing that the classical trajectories satisfy the following fundamental inequalities:

$$\frac{1}{2} \int_0^t |\dot{\Theta}(s)|^2 \, ds \leq V_{\text{int}}(\Theta_0) + \frac{C_\Omega^2}{2} t, \tag{3.14}$$

$$|\Theta(t)| \leq |\Theta_0| + \int_0^t |\dot{\Theta}(s)| \, ds, \tag{3.15}$$

for every $t \in [0, t^*)$. Here, $V_{\text{int}}(\Theta)$ is the second term of the potential $V(\Theta)$ in (2.10) and we set

$$C_\Omega := \Big(\sum_{i=1}^N \Omega_i^2 \Big)^{1/2}.$$

In the second step we will show that the inequalities (3.14) and (3.15) imply

$$\frac{1}{2} \int_0^{t^*} |\dot{\Theta}(s)|^2 \, ds \leq V_{\text{int}}(\Theta_0) + \frac{C_\Omega^2}{2} t^* < \infty, \tag{3.16}$$

$$\int_0^{t^*} |\dot{\Theta}(s)| \, ds \leq V_{\text{int}}(\Theta_0) + \frac{1 + C_\Omega^2}{2} t^* < \infty, \tag{3.17}$$

$$|\Theta(t)| \leq |\Theta_0| + V_{\text{int}}(\Theta_0) + \frac{1 + C_\Omega^2}{2} t^*, \tag{3.18}$$

for every $t \in [0, t^*)$. Finally, in the third part we will deduce the assertions of the lemma from the inequalities (3.14)–(3.18).

• *Step 1.* Recall that in Section 2, the classical solution $t \mapsto \Theta(t)$ of (3.2) equivalently solves the gradient flow system (2.9), i.e.,

$$\dot{\Theta}(t) = -\nabla V(\Theta(t))$$

for all $t \in [0, t^*)$, where V is given in (2.10). Hence,

$$\frac{d}{dt} V(\Theta(t)) = \nabla V(\Theta(t)) \cdot \dot{\Theta}(t) = -|\dot{\Theta}(t)|^2$$

for every $t \in [0, t^*)$. Taking integrals in time, we obtain

$$\int_0^t |\dot{\Theta}(s)|^2 \, ds = V(\Theta_0) - V(\Theta(t)) = \sum_{i=1}^N \Omega_i (\theta_{i,0} - \theta_i(t)) + V_{\text{int}}(\Theta_0) - V_{\text{int}}(\Theta(t)) \tag{3.19}$$

for every $t \in [0, t^*)$. Recall that the function W in (2.11) involved in the potential (2.10) is a primitive function of h. Thus, $W \geq 0$, as a consequence of the antisymmetry of h and our choice $W(0) = 0$, and in particular $V_{\text{int}} \geq 0$. This, together with the Cauchy–Schwarz inequality, yields

$$\int_0^t |\dot{\Theta}(s)|^2 \, ds \leq C_\Omega \int_0^t |\dot{\Theta}(s)| \, ds + V_{\text{int}}(\Theta_0) \tag{3.20}$$

for every $t \in [0, t^*)$. Using Young's inequality in the first term of (3.20), we arrive at the first fundamental inequality (3.14). The inequality (3.15) is standard, but let us sketch it for the sake of clarity:

$$\frac{d}{dt} \frac{|\Theta|^2}{2} = \Theta \cdot \dot{\Theta} \leq |\Theta| \, |\dot{\Theta}|$$

for all $t \in [0, t^*)$. Hence, we arrive at

$$\frac{d}{dt} |\Theta(t)| \leq |\dot{\Theta}(t)|$$

for every $t \in [0, t^*)$, and integrating with respect to time yields (3.15).

• *Step 2.* First, taking limits as $t \to t^*$ in (3.14), we clearly obtain (3.16). Also, the finite length of the trajectory (3.17) holds true by the Cauchy–Schwarz inequality and Young's inequality both applied to (3.16). Finally, inequalities (3.15) and (3.17) entail (3.18).

• *Step 3.* The classical trajectory $t \mapsto \Theta(t)$ is defined up to a finite maximal time t^*. Hence, classical results show that either it blows up at $t = t^*$ or there exists some sequence $\{t_n\}_{n \in \mathbb{N}} \nearrow t^*$ and some $\Theta^* \in \mathcal{C}$ such that $\{\Theta(t_n)\}_{n \in \mathbb{N}} \to \Theta^*$. Since the former option is prevented by (3.18), the latter must hold true. Let us prove that the whole trajectory converges to the collision state Θ^*. Indeed, in the other case, there exists another sequence $\{s_n\}_{n \in \mathbb{N}} \nearrow t^*$ and some $\varepsilon_0 > 0$ such that

$$|\Theta(s_n) - \Theta^*| \geq \varepsilon_0 \tag{3.21}$$

for all $n \in \mathbb{N}$. Without loss of generality we can assume that the sequences $\{t_n\}_{n\in\mathbb{N}}$ and $\{s_n\}_{n\in\mathbb{N}}$ are ordered as $t_1 < s_1 < t_2 < s_2 < \cdots$ and that

$$|\Theta(t_n) - \Theta^*| \leq \varepsilon_0/2^n \tag{3.22}$$

for every $n \in \mathbb{N}$. Therefore,

$$|\Theta(t_n) - \Theta(s_n)| \geq |\Theta(s_n) - \Theta^*| - |\Theta(t_n) - \Theta^*| \geq \varepsilon_0 - \varepsilon_0/2^n,$$
$$|\Theta(s_n) - \Theta(t_{n+1})| \geq |\Theta(s_n) - \Theta^*| - |\Theta(t_{n+1}) - \Theta^*| \geq \varepsilon_0 - \varepsilon_0/2^{n+1},$$

for all $n \in \mathbb{N}$. Then it is clear that

$$\begin{aligned}\int_0^{t^*} |\dot{\Theta}(t)|\,dt \geq \int_{t_1}^{t^*} |\dot{\Theta}(t)|\,dt &= \sum_{n=1}^{\infty} \int_{t_n}^{s_n} |\dot{\Theta}(t)|\,dt + \sum_{n=1}^{\infty} \int_{s_n}^{t_{n+1}} |\dot{\Theta}(t)|\,dt \\ &\geq \sum_{n=1}^{\infty} |\Theta(t_n) - \Theta(s_n)| + \sum_{n=1}^{\infty} |\Theta(s_n) - \Theta(t_{n+1})| \\ &\geq \sum_{n=1}^{\infty} \varepsilon_0\left(1 - \frac{1}{2^n}\right) + \sum_{n=1}^{\infty} \varepsilon_0\left(1 - \frac{1}{2^{n+1}}\right) = \infty,\end{aligned}$$

which contradicts (3.17). Hence, $\lim_{t\to t^*} \Theta(t) = \Theta^*$. ∎

This conclusion shows that, as expected, it is possible to continue classical solutions by Filippov solutions (hence absolutely continuous) after a possible collision. The explicit method of continuation is exhibited in the following result.

Theorem 3.5. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be any classical solution to (3.2) with $\alpha \in (1/2, 1)$ that is defined in a finite maximal existence interval $[0, t^*)$ and, according to Lemma 3.7, let the collision state $\Theta^* \in \mathscr{C}$ be such that*

$$\lim_{t\to t^*} \Theta(t) = \Theta^*.$$

Then there exists $\varepsilon > 0$ such that the classical trajectory $t \mapsto \Theta(t)$ can be continued to a Filippov solution to (3.2) in $[t^*, t^* + \varepsilon)$ in such a way that oscillators belonging to the same cluster of the collision state Θ^* remain stuck together after t^*.

Proof. Let E_k be the k-th cluster of oscillators with $n_k = \#E_k$ for $k = 1, \ldots, \kappa$. We consider a bijection $\sigma^k : \{1, \ldots, n_k\} \to E_k$ for every $k = 1, \ldots, \kappa$. Since the necessary condition (3.13) is automatically satisfied as discussed in Remark 3.5, there exists a matrix $Y^k \in \mathrm{Skew}_{n_k}(\mathbb{R})$ such that

$$\Omega_{\sigma_i^k} + \frac{K}{N}\sum_{\substack{l=1\\ l\neq i}}^{n_k} y_{il}^k = \Omega_{\sigma_j^k} + \frac{K}{N}\sum_{\substack{l=1\\ l\neq j}}^{n_k} y_{jl}^k =: \widehat{\Omega}_k \tag{3.23}$$

for any $i, j \in \{1, \ldots, n_k\}$. Consider the system of κ differential equations

$$\dot{\vartheta}_k = \widehat{H}_k(\vartheta_1, \ldots, \vartheta_k) := \widehat{\Omega}_k + \frac{K}{N} \sum_{\substack{m=1 \\ m \neq k}}^{\kappa} n_m h(\vartheta_m - \vartheta_k) \tag{3.24}$$

for $k = 1, \ldots, \kappa$, with initial data

$$(\vartheta_1(t^*), \ldots, \vartheta_\kappa(t^*)) = (\theta^*_{\iota_1}, \ldots, \theta^*_{\iota_\kappa}). \tag{3.25}$$

Since the initial datum is a non-collision state in a lower-dimensional space $\mathbb{R}^\kappa$ of phase configurations, there exists a unique classical solution to the problem that is defined in a maximal existence interval $[t^*, t^{**})$ and such that if $t^{**} < \infty$, then $(\vartheta_1, \ldots, \vartheta_\kappa)$ converges to a new collision state by Lemma 3.7 (merging of clusters). The same result ensures that

$$[0, t^*) \ni t \mapsto (\theta_1(t), \ldots, \theta_N(t)),$$
$$[t^*, t^{**}) \ni t \mapsto (\vartheta_1(t), \ldots, \vartheta_\kappa(t)),$$

belong to $W^{1,2}((0, t^*), \mathbb{R}^N)$ and $W^{1,2}((t^*, t^{**}), \mathbb{R}^\kappa)$, respectively. Let us set the prolongation of $t \mapsto \Theta(t)$ in $[t^*, t^{**})$ in such a way that

$$\theta_{\sigma_i^k}(t) := \vartheta_k(t), \qquad \forall t \in [t^*, t^{**}),$$

for every $i \in E_k$ and $k = 1, \ldots, \kappa$. The two trajectories glue together in a $W^{1,2}$ way and it is clear, by the definition of $\widehat{H}^k$ in (3.24) and $\widehat{\Omega}_k$ in (3.23) along with the explicit expression of the Filippov map in Proposition 3.4, that $[0, t^{**}) \ni t \mapsto \Theta(t)$ becomes a Filippov solution to (3.2) in $[0, t^{**})$. ∎

Remark 3.6. The above procedure can be repeated as many times as needed after each collision time of the corresponding classical solutions to the reduced systems (3.24)–(3.25). Indeed, by Remark 3.5 the necessary condition (3.13) is automatically satisfied. Since there can only be $N - 1$ collisions of oscillators with sticking, we may apply Theorem 3.5 finitely many times to obtain global-in-time Filippov solutions to (3.2) in the supercritical case. However, one may wonder whether this global-in-time continuation procedure is unique or whether oscillators may also be allowed to split instantaneously after a collision. Although answering the general question for any number N of oscillators and any collision state is really difficult, let us give a particular answer for $N = 2$:

$$\dot{\theta}_1 = \Omega_1 + \frac{K}{2} h(\theta_2 - \theta_1), \tag{3.26}$$

$$\dot{\theta}_2 = \Omega_2 + \frac{K}{2} h(\theta_1 - \theta_2). \tag{3.27}$$

Consider the relative phase $\theta := \theta_2 - \theta_1$ and relative natural frequency $\Omega := \Omega_2 - \Omega_1$. Then the associated dynamics of a classical solution is governed by the equation

$$\dot{\theta} = \Omega - K h(\theta)$$

in the maximal interval of existence $[0, t^*)$. According to Lemma 3.7, we infer that $t^* = \infty$ if $\theta(0) = \bar{\theta}$, whereas $t^* < \infty$ if $\theta(0) \notin \{0, \bar{\theta}\}$. Here, $\bar{\theta}$ stands for the unique (unstable) equilibrium of the system (see Proposition 5.2). Without loss of generality, we will fix the initial relative phase so that $\theta(0) \in (0, \bar{\theta})$ (the other cases are similar). Thus, we arrive at a collision of oscillators at $t = t^*$, i.e., $\lim_{t\to t^*} \theta(t) = 0$.

(1) Assume for contradiction that there is another Filippov solution in $[t^*, t^{**})$ consisting of two particles that instantaneously split again after $t = t^*$. The split can arise in only two different manners:

   (a) (Sharp split) There exists a small $\varepsilon > 0$ such that $\theta(t) \neq 0$ for every $t \in (t^*, t^* + \varepsilon)$. In that case, either $\theta(t) > 0$ for all $t \in (t^*, t^* + \varepsilon)$, or $\theta(t) < 0$ for all $t \in (t^*, t^* + \varepsilon)$.

   (b) (Zeno split) There exist sequences $\{t_n\}_{n\in\mathbb{N}} \searrow t^*$ and $\{s_n\}_{n\in\mathbb{N}} \searrow t^*$ such that $\theta(s_n) = 0$ but $\theta(t_n) \neq 0$, for every $n \in \mathbb{N}$. Recall Remark 3.4 for the left accumulations of switches or Zeno behavior and see Figure 6.

   Replacing t^* by a suitable time, it is apparent that the second type of split at t^* guarantees the first one at a (possibly) later time. Let us focus on just the first case. Looking at the profile of $\Omega - kh(\theta)$ in Figure 5, we would then arrive at the following conclusion: either $\dot{\theta}(t) < 0$ and $\theta(t) > 0$ for all $t \in (t^*, t^* + \varepsilon)$, or $\dot{\theta}(t) > 0$ and $\theta(t) < 0$ for all $t \in (t^*, t^* + \varepsilon)$. In any case, we obtain a contradiction.

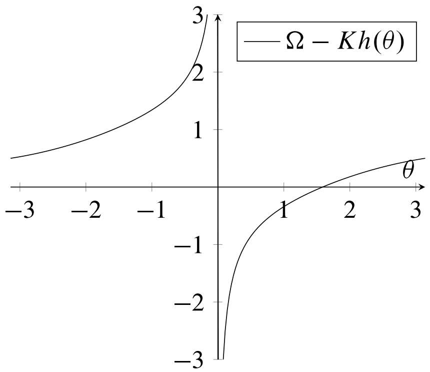


Fig. 5. Profile of $\Omega - Kh(\theta)$ for $\Omega = 0.25$, $K = 1$ and $\alpha = 0.75$.

(2) Hence, the only choice for the oscillators after the collision state is to stick together. Let us define the phase of the reduced system (see (3.23))

$$\widehat{\Omega} := \Omega_1 + y_{12} = \Omega_2 + y_{21},$$

   where $Y \in \mathrm{Skew}_2(\mathbb{R})$ is any matrix satisfying the necessary condition (3.13). Indeed, there exists just one such matrix Y, with $y_{12} = -y_{21} = (\Omega_2 - \Omega_1)/2$. Then $\widehat{\Omega} = (\Omega_1 + \Omega_2)/2$ and the reduced system (3.24) looks like

$$\dot{\vartheta} = \widehat{\Omega}, \qquad t \in [t^*, \infty).$$

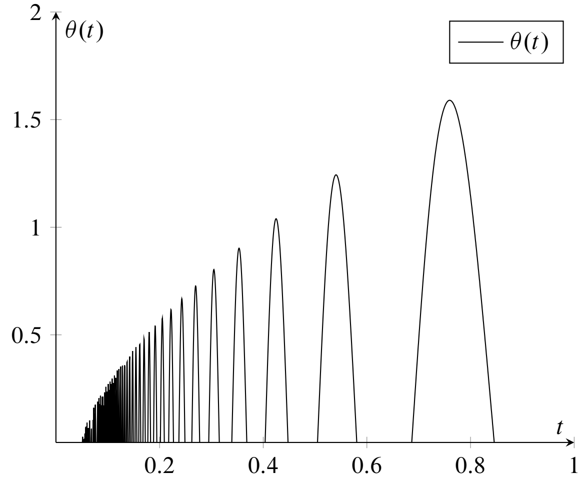


Fig. 6. Left-Zeno behavior in the relative phase $\theta(t) = \theta_2(t) - \theta_1(t)$ of two oscillators.

Consequently, the only Filippov solution to (3.2) evolves through (3.26)–(3.27) up to the collision time t^*. Afterwards, both oscillators stick together and they move with constant frequency equal to the average natural frequency.

For general N, it is not clear whether (b) in the above first item can be reduced to (a) in a similar way. Namely, we cannot guarantee that along the whole time interval $(t^*, t^* + \varepsilon)$ all the subclusters splitting from the given cluster remain a positive distance apart or they actually merge and split instantaneously with eventual switches of collision type in a similar way to Figure 6 in Zeno behavior. Also, studying the higher-dimensional phase portrait in the same spirit as we have done for $N = 2$ is not easy and we shall address it in future work.

4. Rigorous limit towards singular weights

In the previous section, we studied the existence and one-sided uniqueness of absolutely continuous solutions to the singular weighted first order Kuramoto model in all the subcritical, critical and supercritical cases. Because of the continuity of the kernel for $\alpha \in (0, 1/2)$, we can show that in that case the solutions are indeed C^1, although we can say the same neither for the critical case $\alpha = 1/2$ nor for the supercritical case $\alpha \in (1/2, 1)$. Also, these results do not necessarily provide any extra regularity of the frequencies $\omega_i = \dot{\theta}_i$ for an augmented second order model to make sense.

Let us recall that in Subsection 2.2, the singular Kuramoto model was formally obtained as the singular $\varepsilon \to 0$ limit of the scaled regular model (2.5)–(2.6). Notice that if we rigorously proved the $\varepsilon \to 0$ limit, we would achieve an alternative existence result for the singular models. In this section, we shall inspect to what extent this idea works and how many exponents we can obtain with that technique. In particular, we will recover the existence results of Section 3. Indeed, this technique will yield a gain of piecewise $W^{1,1}$ regularity of the frequencies ω_i in the subcritical case and will provide an equation for them in a weak sense that will be discussed and related to similar models in Subsection 4.4. However, this idea fails for the more singular cases, where the compactness of frequencies is very weak. While the singular limit for the subcritical case is straightforward, we need new ideas to deal with the limiting set-valued Filippov map in the critical and supercritical cases along with the loss of strong compactness of the frequencies in those cases.

4.1. Limit in the subcritical case and augmented flocking model

The following result provides a list of a priori estimates for the global-in-time classical solutions of the regularized system (2.5)–(2.6), for any $\varepsilon > 0$:

Lemma 4.1. *Let $\Theta_0 = (\theta_{1,0}, \dots, \theta_{N,0}) \in \mathbb{R}^N$ be any initial data and consider the unique global-in-time classical solution $\Theta^\varepsilon = (\theta_1^\varepsilon, \dots, \theta_N^\varepsilon)$ to (2.5)–(2.6) in the subcritical case $\alpha \in (0, 1/2)$, for every $\varepsilon > 0$. Then there exists a non-negative constant C such that*

$$\|\dot{\Theta}^\varepsilon\|_{C^{0,1-2\alpha}([0,\infty),\mathbb{R}^N)} \le C, \qquad \|\Theta^\varepsilon\|_{C^{1,1-2\alpha}([0,T],\mathbb{R}^N)} \le |\Theta_0| + CT,$$

for all $T > 0$ and $\varepsilon > 0$. As a consequence, there exists a subsequence of $\{\Theta^\varepsilon\}_{\varepsilon>0}$, denoted in the same way for simplicity, and $\Theta \in C^1([0,\infty), \mathbb{R}^N)$ such that $\dot{\Theta} \in C^{0,1-2\alpha}([0,\infty), \mathbb{R}^N)$, Θ satisfies the same estimates as above and

$$\{\Theta^\varepsilon\}_{\varepsilon>0} \to \Theta \quad \text{in } C^1([0,T], \mathbb{R}^N)$$

for every $T > 0$.

Proof. All the properties follow directly from the first one along with the Ascoli–Arzelà theorem. Recall that there is a constant $M > 0$ such that

$$|h_\varepsilon(\theta)| \le M \quad \text{and} \quad |h_\varepsilon(\theta_1) - h_\varepsilon(\theta_2)| \le M|\theta_1 - \theta_2|_o^{1-2\alpha}$$

for all $\theta, \theta_1, \theta_2 \in \mathbb{R}$ and all $\varepsilon > 0$. Thus, the first property is also a straightforward consequence of the uniform-in-ε boundedness and Hölder-continuity of the kernel. ■

The following result holds true as a clear consequence of the uniform equicontinuity of the sequence h_ε along with the compactness of the sequence $\{\Theta^\varepsilon\}_{\varepsilon>0}$.

Theorem 4.1. *The limit function Θ of $\{\Theta^\varepsilon\}_{\varepsilon>0}$ in Lemma 4.1 is a classical global-in-time solution of the singular model (2.5)–(2.6) with $\varepsilon = 0$ in the subcritical case $\alpha \in (0, 1/2)$.*

Notice that we have arrived at a construction of classical global-in-time solutions of the singular problem with $0 < \alpha < 1/2$ through two different techniques: Theorems 3.1 and 4.1. However, both techniques are actually closely related to each other, since originally, Filippov theory relies on a similar regularizing procedure. In what follows, we will see that this procedure provides us with extra a priori estimates for the "accelerations" (derivatives of frequencies). Also, the procedure will allow us to derive a "piecewise weak equation" for them. This is the rest of the content of this subsection.

Note that a necessary and sufficient condition for two oscillators θ_i and θ_j that collide at some time to stick together is that $\Omega_i = \Omega_j$, by Theorem 3.2. In some sense, those two oscillators are identified in a unique cluster with a bigger "mass". Further, we can quantify the times of "pure collisions" as follows. Starting with $T_0 = 0$, we define

$$T_k := \inf\{t > T_{k-1} : \exists i, j \in S_i(T_{k-1})^c \text{ such that } \bar{\theta}_i(t) = \bar{\theta}_j(t)\} \tag{4.1}$$

for every $k \in \mathbb{N}$. Recall the notation in Subsection 2.3, and see [37] for related notation in the discrete Cucker–Smale model with singular influence function. Then taking derivatives in (2.5)–(2.6) we can obtain the split

$$\begin{aligned}\ddot{\theta}_i^\varepsilon = &\frac{K}{N} \sum_{j \notin \mathcal{C}_i(T_{k-1})} h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)(\dot{\theta}_j^\varepsilon - \dot{\theta}_i^\varepsilon)\\ &+ \frac{K}{N} \sum_{j \in (\mathcal{C}_i \setminus S_i)(T_{k-1})} h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)(\dot{\theta}_j^\varepsilon - \dot{\theta}_i^\varepsilon)\\ &+ \frac{K}{N} \sum_{j \in S_i(T_{k-1})} h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)(\dot{\theta}_j^\varepsilon - \dot{\theta}_i^\varepsilon),\end{aligned} \tag{4.2}$$

where $t \in [T_{k-1}, T_k)$. The idea is to show that we can pass to the limit in the above expressions in $L^1([T_{k-1}, \tau])$-weak, for every $k \in \mathbb{N}$ and every $\tau \in (T_{k-1}, T_k)$. This is the content of the next theorem. Before going on, let us discuss the possible scenarios for the sequence $\{T_k\}_{k \in \mathbb{N}}$ and how we can cover the whole interval $[0, \infty)$ with them in any case so that our dynamics can be reduced to each of them:

(1) It might happen that there exists $k_0 \in \mathbb{N}$ such that $T_{k_0+1} = \infty$ (then $T_k = \infty$ for every $k > k_0$). This is the case when either all particles have stuck together in finite time or after some finite time there is no more collision. In this case
$$[0, \infty) = \bigcup_{0 \le k \le k_0 - 1} [T_k, T_{k+1}) \cup [T_{k_0}, \infty),$$
and there is no collision in any of these intervals.

(2) Also it might happen that the sequence $\{T_k\}_{k \in \mathbb{N}}$ is infinite and unbounded, $T_k \nearrow \infty$. Then
$$[0, \infty) = \bigcup_{k \ge 0} [T_k, T_{k+1}),$$
and there is no collision in any of these intervals.

(3) Finally, it might also be the case that the sequence $\{T_k\}_{k \in \mathbb{N}}$ is infinite but bounded. In that case, there exists $T^\infty \in \mathbb{R}^+$ with right Zeno behavior, i.e. $T_k \nearrow T^\infty$. Then

a straightforward argument involving the mean value theorem shows that T^∞ is a sticking point. Then we can split the dynamics up to time T^∞ through

$$[0, T^\infty) = \bigcup_{k\geq 0} [T_k, T_{k-1}).$$

Taking T^∞ as our initial time, we can repeat each of Steps 1, 2 and 3 above so that we can globally recover the whole dynamics. Notice that since there can be just $N-1$ times of sticking, there can be no more than $N-1$ times like T^∞.

For simplicity of argument, we will assume that we are in case (2), although the same results apply to any of the other cases. Before going to the heart of the result, let us summarize some good properties of the kernel h'_ε.

Lemma 4.2. *Let $\alpha \in (0, 1/2)$. Then the following properties hold true:*

(1) *Formula for the derivative:*

$$h'_\varepsilon(\theta) = \frac{1}{(\varepsilon^2 + c|\theta|_o^2)^\alpha}\left[\cos\theta - 2\alpha c \frac{\sin|\theta|_0}{|\theta|_0}\frac{|\theta|_0^2}{\varepsilon^2 + c|\theta|_o^2}\right].$$

(2) *Upper bound by an $L^1(\mathbb{T})$-function:*

$$|h'_\varepsilon(\theta)|, |h'(\theta)| \leq M\frac{1}{|\theta|_o^{2\alpha}}.$$

(3) *Strong $L^1(\mathbb{T})$ convergence:*

$$h'_\varepsilon \to h' \quad \text{in } L^1(\mathbb{T}).$$

(4) *Weighted Hölder-continuity:*

$$|h'_\varepsilon(\theta_1) - h'_\varepsilon(\theta_2)| \leq M\frac{|\theta_1 - \theta_2|_o^\beta}{\min\{|\theta_1|_o, |\theta_2|_o\}^\gamma}$$

for any exponents $\beta, \gamma \in (0, 1)$ such that $\gamma = 2\alpha + \beta$.

(5) *Weighted $L^\infty(\mathbb{T})$ convergence:*

$$|h'_\varepsilon(\theta) - h'(\theta)| \leq M\frac{\varepsilon^{1-2\alpha}}{|\theta|_o}.$$

Proof. The first two results are straightforward and the third one is a clear consequence of the dominated convergence theorem. The fourth property follows from an obvious application of the mean value theorem and the fifth one is a standard property of mildly singular kernels (one can show that $M = \alpha/\beta$) . ■

Theorem 4.2. *For any initial datum $\Theta_0 \in \mathbb{R}^N$, let Θ^ε be the classical global-in-time solution of (2.5)–(2.6) in the subcritical case $\alpha \in (0, 1/2)$. Also, consider the limiting Θ in Theorem 4.1 and the collision times $\{T_k\}_{k\in\mathbb{N}}$ in (4.1). Then:*

(1) *For every* $i \in \{1, \ldots, N\}$ *and* $j \notin \mathscr{C}_i(T_{k-1})$,

$$h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) \to h'(\theta_j - \theta_i) \quad \textit{as } \varepsilon \to 0, \quad \textit{in } C([T_{k-1}, \tau]).$$

(2) *For every* $i \in \{1, \ldots, N\}$ *and* $j \in \mathscr{C}_i(T_{k-1}) \setminus S_i(T_{k-1})$,

$$h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) \rightharpoonup h'(\theta_j - \theta_i) \quad \textit{as} \quad \varepsilon \to 0, \quad \textit{in } L^1([T_{k-1}, \tau]).$$

(3) *For every* $i \in \{1, \ldots, N\}$ *and* $j \in S_i(T_{k-1})$,

$$\frac{d}{dt} h_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) \to 0 \quad \textit{as } \varepsilon \to 0, \quad \textit{in } W^{-1,\infty}([T_{k-1}, \tau]).$$

Proof. (1) Fix any $i \in \{1, \ldots, N\}$ and $j \notin \mathscr{C}_i(T_{k-1})$. There exists (by definition) some positive constant $\delta_0 = \delta_0(k, \tau) < \pi$ such that

$$|\theta_i(t) - \theta_j(t)|_o \geq \delta_0 \quad \text{for all } t \in [T_{k-1}, \tau].$$

Then, by the uniform convergence in Lemma 4.1 there exists some $\varepsilon_0 > 0$ such that

$$|\theta_i^\varepsilon(t) - \theta_j^\varepsilon(t)|_o \geq \delta_0/2 \quad \text{for all } t \in [T_{k-1}, \tau], \tag{4.3}$$

for every $\varepsilon \in (0, \varepsilon_0)$. Consequently, by crossing terms we have

$$\begin{aligned}&|h'_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)) - h'(\theta_j(t) - \theta_i(t))| \\ &\quad \leq |h'_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)) - h'(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t))| + |h'(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)) - h'(\theta_j(t) - \theta_i(t))|\end{aligned}$$

for every $t \in [T_{k-1}, \tau]$. Hence, both terms converge to zero uniformly in $[T_{k-1}, \tau]$ as $\varepsilon \to 0$. This is due to (4.3), the third property in Lemma 4.2, the uniform continuity of h' in compact sets away from $2\pi\mathbb{Z}$ and the uniform convergence of the phases in Lemma 4.1. This ends the proof of the first part.

(2) Now fix $i \in \{1, \ldots, N\}$ and $j \in \mathscr{C}_i(T_{k-1}) \setminus S_i(T_{k-1})$. Then

$$\bar{\theta}_j(T_{k-1}) = \bar{\theta}_i(T_{k-1}) \quad \text{but} \quad \dot{\theta}_j(T_{k-1}) \neq \dot{\theta}_i(T_{k-1}).$$

Thus, it is clear that we again have $|\theta_j(t) - \theta_i(t)|_o > 0$ for $t \in [\tau^*, \tau]$ and for every $\tau^* \in (T_{k-1}, \tau)$. This amounts to saying that the preceding argument again holds in $[\tau^*, \tau]$, and consequently

$$h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) \to h'(\theta_j - \theta_i) \quad \text{in } C([\tau^*, \tau])$$

for every $\tau^* \in (T_{k-1}, \tau)$. Thus, we just need to prove the weak convergence in some interval $[T_{k-1}, \tau^*]$. Fix τ^*. Since $\dot{\theta}_j(T_{k-1}) \neq \dot{\theta}_i(T_{k-1})$, we can assume without loss of generality that

$$\delta_0 := \dot{\theta}_j(T_{k-1}) - \dot{\theta}_i(T_{k-1}) > 0.$$

By continuity of $\dot{\theta}_j$ and $\dot{\theta}_i$, there exists some small $\tau^* \in (T_{k-1}, \tau)$ such that

$$\dot{\theta}_i(t) - \dot{\theta}_j(t) \geq \delta_0/2 \quad \text{for all } t \in [T_{k-1}, \tau^*]. \tag{4.4}$$

Then, by the uniform convergence of the frequencies (see Lemma 4.1), we can take a small enough $\varepsilon_0 > 0$ such that if $\varepsilon \in (0, \varepsilon_0)$ then

$$\dot{\theta}_i^\varepsilon(t) - \dot{\theta}_j^\varepsilon(t) \geq \delta_0/4 \quad \text{for all } t \in [T_{k-1}, \tau^*]. \tag{4.5}$$

In particular, we have well defined inverses of $\theta_j - \theta_i$ and $\theta_j^\varepsilon - \theta_i^\varepsilon$ in $[T_{k-1}, \tau^*]$ for every $\varepsilon \in (0, \varepsilon_0)$. Indeed, the inverse function theorem states that

$$((\theta_j - \theta_i)^{-1})' = \frac{1}{(\dot{\theta}_j - \dot{\theta}_i) \circ (\theta_j - \theta_i)^{-1}}, \tag{4.6}$$

and a similar statement holds for $\theta_j^\varepsilon - \theta_i^\varepsilon$. In order to show the weak convergence in $L^1([T_{k-1}, \tau^*])$, we equivalently claim that the following assertions are true:

(a) Uniform-in-ε L^1 bound of $h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)$ and $h(\theta_j - \theta_i)$ in $[T_{k-1}, \tau^*]$: there exists a constant $M > 0$ such that

$$\|h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)\|_{L^1([T_{k-1},\tau^*])}, \|h'(\theta_j - \theta_i)\|_{L^1([T_{k-1},\tau^*])} \leq M$$

for every $\varepsilon \in (0, \varepsilon_0)$.

(b) Convergence of the mean values over finite intervals:

$$\lim_{\varepsilon\to 0} \int_{T_{k-1}}^{\tau^{**}} (h'_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)) - h'(\theta_j(t) - \theta_i(t)))\, dt = 0$$

for every $\tau^{**} \in (T_{k-1}, \tau^*)$.

Let us prove this claim. Regarding (a), we just focus on $h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon)$ (the other case is similar). By a simple change of variables $\theta = (\theta_j^\varepsilon - \theta_i^\varepsilon)(t)$ and (4.5)–(4.6),

$$\begin{aligned}\int_{T_{k-1}}^{\tau^{**}} |h'_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t))|\, dt &= \int_{\theta_j^\varepsilon(T_{k-1})-\theta_i^\varepsilon(T_{k-1})}^{\theta_j^\varepsilon(\tau^{**})-\theta_i^\varepsilon(\tau^{**})} \frac{|h'_\varepsilon(\theta)|\, d\theta}{(\dot{\theta}_j^\varepsilon - \dot{\theta}_i^\varepsilon)((\theta_j^\varepsilon - \theta_i^\varepsilon)^{-1}(\theta))} \\ &\leq \|h'_\varepsilon\|_{L^1(\mathbb{T})} \frac{4}{\delta_0}.\end{aligned}$$

Then (a) follows from Lemma 4.2(2).

Regarding (b), we split

$$\int_{T_{k-1}}^{\tau^{**}} (h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) - h'(\theta_j - \theta_i))\, dt = I_\varepsilon + II_\varepsilon,$$

where

$$\begin{aligned} I_\varepsilon &:= \int_{T_{k-1}}^{\tau^{**}} (h'_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) - h'(\theta_j^\varepsilon - \theta_i^\varepsilon))\, dt, \\ II_\varepsilon &:= \int_{T_{k-1}}^{\tau^{**}} (h'(\theta_j^\varepsilon - \theta_i^\varepsilon) - h'(\theta_j - \theta_i))\, dt. \end{aligned}$$

The same change of variables as above allows us to rewrite I_ε in the following way:

$$I_\varepsilon = \int_{\theta_j^\varepsilon(T_{k-1})-\theta_i^\varepsilon(T_{k-1})}^{\theta_j^\varepsilon(\tau^{**})-\theta_i^\varepsilon(\tau^{**})} (h'_\varepsilon(\theta) - h'(\theta)) \frac{d\theta}{(\dot\theta_j^\varepsilon - \dot\theta_i^\varepsilon)((\theta_j^\varepsilon - \theta_i^\varepsilon)^{-1}(\theta))}.$$

Then estimate (4.5) along with the strong $L^1(\mathbb{T})$ convergence of the kernels in Lemma 4.2(3) shows that I_ε vanishes when $\varepsilon \to 0$:

$$\begin{aligned} |I_\varepsilon| &\le \frac{4}{\delta_0} \int_{\theta_j^\varepsilon(T_{k-1})-\theta_i^\varepsilon(T_{k-1})}^{\theta_j^\varepsilon(\tau^{**})-\theta_i^\varepsilon(\tau^{**})} |h'_\varepsilon(\theta) - h'(\theta)|\, d\theta \\ &= \frac{4}{\delta_0} \|h'_\varepsilon(\theta) - h'(\theta)\|_{L^1(\mathbb{T})} \to 0 \quad \text{as } \varepsilon \to 0. \end{aligned}$$

For the term II_ε, we use Lemma 4.2(4) to show

$$\begin{aligned} |II_\varepsilon| &\le M \int_{T_{k-1}}^{\tau^{**}} \frac{|(\theta_j^\varepsilon - \theta_j) - (\theta_i^\varepsilon - \theta_i)|_o^\beta}{\min\{|\theta_j^\varepsilon - \theta_i^\varepsilon|_o, |\theta_j - \theta_i|_o\}^\gamma}\, dt \\ &\le 2^\beta M \|\Theta^\varepsilon - \Theta\|_{C([T_{k-1},\tau^{**}],\mathbb{R}^N)}^\beta \int_{T_{k-1}}^{\tau^{**}} \frac{1}{\min\{|\theta_j^\varepsilon - \theta_i^\varepsilon|_o, |\theta_j - \theta_i|_o\}^\gamma}\, dt \\ &\le 2^\beta M \|\Theta^\varepsilon - \Theta\|_{C([T_{k-1},\tau^{**}],\mathbb{R}^N)}^\beta \int_{T_{k-1}}^{\tau^{**}} \max\left\{\frac{1}{|\theta_j^\varepsilon - \theta_i^\varepsilon|_o^\gamma}, \frac{1}{|\theta_j - \theta_i|_o^\gamma}\right\} dt \\ &\le 2^\beta M \|\Theta^\varepsilon - \Theta\|_{C([T_{k-1},\tau^{**}],\mathbb{R}^N)}^\beta \int_{T_{k-1}}^{\tau^{**}} \left(\frac{1}{|\theta_j^\varepsilon - \theta_i^\varepsilon|_o^\gamma} + \frac{1}{|\theta_j - \theta_i|_o^\gamma}\right) dt. \end{aligned}$$

Then, a new change of variables along with the equations (4.5)–(4.6) and the local integrability in one dimension of an inverse power of order γ entail the existence of a non-negative constant C that does not depend on ε such that

$$|II_\varepsilon| \le C \|\Theta^\varepsilon - \Theta\|_{C([T_{k-1},\tau^{**}],\mathbb{R}^N)}.$$

Now (2) follows from the uniform convergence of the phases in Lemma 4.1.

(3) Consider $i \in \{1, \ldots, N\}$ and $j \in S_i(T_{k-1})$. By the uniqueness in Theorem 3.1, we can ensure that $\theta_j(t) = \theta_i(t)$ for all $t \ge T_{k-1}$. Then the uniform convergence of the kernels h_ε along with the uniform convergence of the phases in Lemma 4.1 shows that

$$h_\varepsilon(\theta_j^\varepsilon - \theta_i^\varepsilon) \to 0 \quad \text{in } C([T_{k-1}, \tau]),$$

and hence the result holds true by definition of the norm in $W^{-1,\infty}([T_{k-1}, \tau])$. ■

Remark 4.1. The preceding results show that the unique global-in-time solution Θ to the problem (3.2), with $\alpha \in (0, 1/2)$, which we constructed in Theorem 3.1, satisfies $\theta_i \in C^{1,1-2\alpha}([0, \infty), \mathbb{R}^N)$ and the frequencies $\dot\theta_i$ exhibit higher regularity. Indeed, they are piecewise $W^{1,1}$ in the sense that $\dot{\theta}_i \in W^{1,1}([T_{k-1}, \tau])$ for every $k \in \mathbb{N}$ and every $\tau \in (T_{k-1}, T_k)$. In addition, they satisfy the following equation in the weak sense:

$$\ddot{\theta}_i = \frac{K}{N} \sum_{j \notin S(i)(T_{k-1})} h'(\theta_j - \theta_i)(\dot{\theta}_j - \dot{\theta}_i) \tag{4.7}$$

in $[T_{k-1}, \tau]$. Throughout the proof of the above result we have just used the local integrability in one dimension of any inverse power of order smaller than 1. However, one might have tried to use the fact that such inverse powers actually belong to L^p_{loc} in order to show that in Step 2 the convergence takes place in $L^p([T_{k-1}, \tau])$-weak for any $1 \leq p < 1/(2\alpha)$. In this way, the gain of regularity is in reality higher, namely $\dot{\theta}_i \in W^{1,p}([T_{k-1}, \tau])$ for every $1 \leq p < 1/(2\alpha)$.

In the following, we shall discuss the corresponding singular limit in the critical and supercritical cases. Since the Filippov set-valued map is relatively simpler in the latter case, we will start from it. Later, we will adapt the ideas therein to show a parallel result in the critical regime.

4.2. *Limit in the supercritical case*

Using vector notation similar to that in (2.3) for the singular weighted model, our regularized system (2.5)–(2.6) can be restated as

$$\begin{cases} \dot{\Theta}^\varepsilon = H^\varepsilon(\Theta^\varepsilon), \\ \Theta^\varepsilon(0) = \Theta_0, \end{cases}$$

where the components of the vector field H^ε read

$$H_i^\varepsilon(\Theta) = \Omega_i + \frac{K}{N} \sum_{j \neq i} h_\varepsilon(\theta_j - \theta_i)$$

for every $\Theta \in \mathbb{R}^N$ and every $i \in \{1, \ldots, N\}$. Then one can mimic the ideas in Section 2 to show that the regularized system can also be written as a gradient flow

$$\begin{cases} \dot{\Theta}^\varepsilon = -\nabla V^\varepsilon(\Theta^\varepsilon), \\ \Theta^\varepsilon(0) = \Theta_0, \end{cases} \tag{4.8}$$

where the regularized potential now reads

$$V^\varepsilon(\Theta) := -\sum_{i=1}^N \Omega_i \theta_i + V^\varepsilon_{\mathrm{int}}(\Theta) := -\sum_{i=1}^N \Omega_i \theta_i + \frac{K}{2N} \sum_{i \neq j} W_\varepsilon(\theta_i - \theta_j) \tag{4.9}$$

for every $\Theta \in \mathbb{R}^N$. Again, W_ε is the antiderivative of h_ε such that $W_\varepsilon(0) = 0$, i.e.,

$$W_\varepsilon(\theta) := \int_0^\theta h_\varepsilon(\theta')\, d\theta'.$$

Also, it is clear that in the supercritical case, $W_\varepsilon \geq 0$ for every $\varepsilon > 0$. Thus, the following result holds true.

Lemma 4.3. *In the supercritical case $\alpha \in (1/2, 1)$, let Θ^ε be the unique global-in-time classical solution to the regularized system* (4.8). *Then*

$$\frac{1}{2}\int_0^t |\dot{\Theta}^\varepsilon(s)|^2\,ds \leq \frac{C_\Omega^2}{2}t + V_{\rm int}(\Theta_0)$$

for every $t > 0$ and every $\varepsilon > 0$, where $C_\Omega := (\sum_{i=1}^N \Omega_i^2)^{1/2}$.

The above result shows that $\{\Theta^\varepsilon\}_{\varepsilon>0}$ is bounded in $H^1((0,T),\mathbb{R}^N)$ for every $T > 0$. Hence, there exists a subsequence (not relabeled) such that $\{\Theta^\varepsilon\}_{\varepsilon>0}$ weakly converges to some $\Theta \in H^1_{\rm loc}((0,\infty),\mathbb{R}^N)$ in $H^1((0,T),\mathbb{R}^N)$ for every $T > 0$. The Sobolev embedding and the definition of weak convergence ensure that

$$\Theta^\varepsilon \to \Theta \quad \text{in } C([0,T],\mathbb{R}^N), \qquad \dot{\Theta}^\varepsilon \rightharpoonup \dot{\Theta} \quad \text{in } L^2((0,T),\mathbb{R}^N),$$

for every $T > 0$. Before we obtain the desired convergence of (4.8) towards a Filippov solution, let us introduce the following split of the frequencies:

$$\dot{\Theta}^\varepsilon(t) = x^\varepsilon(t) + y^\varepsilon(t), \tag{4.10}$$

where, componentwise, each term reads as follows:

$$x_i^\varepsilon(t) = \frac{K}{N}\sum_{j\notin\mathcal{C}_i(t)} (h_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)) - h(\theta_j(t) - \theta_i(t))),$$

$$y_i^\varepsilon(t) = \frac{K}{N}\sum_{j\notin\mathcal{C}_i(t)} h(\theta_j(t) - \theta_i(t)) + \frac{K}{N}\sum_{j\in\mathcal{C}_i(t)} h_\varepsilon(\theta_j^\varepsilon(t) - \theta_i^\varepsilon(t)).$$

Then it is clear by definition that

$$x^\varepsilon \to 0 \quad \text{in } C([0,T],\mathbb{R}^N), \qquad y^\varepsilon \rightharpoonup \dot{\Theta} \quad \text{in } L^2((0,T),\mathbb{R}^N),$$

for every $T > 0$, and $y^\varepsilon(t) \in \mathcal{H}(\Theta(t))$ for every $t \geq 0$. As a consequence, Θ^ε becomes a Filippov approximate solution in the following sense:

$$\dot{\Theta}^\varepsilon(t) \in \mathcal{H}(\Theta(t)) + x^\varepsilon(t). \tag{4.11}$$

Remark 4.2. Recall that $\mathcal{H}(\Theta(t))$ is a closed set for every $t \geq 0$ (see Proposition 3.2). Consequently, in order to prove that the limiting $\Theta(t)$ yields a Filippov solution, it would be enough to show the almost everywhere convergence of the sequence $\{\dot{\Theta}^\varepsilon\}_{\varepsilon>0}$ to $\dot{\Theta}$. Unfortunately, it is well known that weak convergence in L^2 is not enough for that purpose. Hence, we must deal only with that weak convergence.

Before going to the heart of the matter, we need to exhibit another characterization of the Filippov set-valued map in terms of implicit equations. The next technical lemma will be used for that. For the proof, see Lemma B.1 of Appendix B.

Lemma 4.4. *Let $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$. Then the following assertions are equivalent:*

(1) *There exists* $Y \in \mathrm{Skew}_n(\mathbb{R})$ *such that*

$$x = Y \cdot \mathbf{j}.$$

(2) *The following implicit equation holds true:*

$$x \cdot \mathbf{j} = 0,$$

where $\mathbf{j}$ *stands for the vector of ones.*

Hence, we are ready to obtain the above-mentioned characterization.

Proposition 4.1. *In the supercritical regime* $\alpha > 1/2$*, the value* $\mathcal{H}(\Theta)$ *of the Filippov set-valued map associated with* H *is the affine subspace of dimension* $N - \kappa$ *of points* $(\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ *obeying the following implicit equations (recall Subsection 2.3):*

$$\frac{1}{n_k} \sum_{i \in E_k} \omega_i = \frac{1}{n_k} \sum_{i \in E_k} \left(\Omega_i + \frac{K}{N} \sum_{j \notin \mathcal{C}_i} h(\theta_j - \theta_i) \right) \tag{4.12}$$

for every $k = 1, \ldots, \kappa$.

Proof. By Proposition 3.4, $\mathcal{H}(\Theta)$ is the set of $(\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ such that for every $k = 1, \ldots, \kappa$ there exist $Y^k \in \mathrm{Skew}_{n_k}(\mathbb{R})$ and a bijection $\sigma^k : \{1, \ldots, n_k\} \to E_k$ such that

$$\omega_{\sigma_i^k} = \Omega_{\sigma_i^k} + \frac{K}{N} \sum_{\substack{m=1 \\ m \neq k}}^{\kappa} n_m h(\theta_{\iota_m} - \theta_{\iota_k}) + \frac{K}{N} \sum_{j=1}^{n_k} y_{ij}^k$$

for every $i = 1, \ldots, n_k$. Then the result follows by applying Lemma 4.4 to each of the above sets of n_k equations to the vectors $x^k \in \mathbb{R}^{n_k(\Theta)}$ with components

$$x_i^k := \omega_{\sigma_i^k} - \Omega_{\sigma_i^k} - \frac{K}{N} \sum_{\substack{m=1 \\ m \neq k}}^{\kappa} n_m h(\theta_{\iota_m} - \theta_{\iota_k}), \quad i = 1, \ldots, n_k,$$

when we equivalently restate it using the notation of Subsection 2.3. ■

Remark 4.3. Here we discuss why the same approach as in Subsection 4.1 to decompose the dynamics for $\alpha \in [1/2, 1)$ into subintervals (T_k, T_{k+1}) of the same collisional type cannot be taken:

(1) Recall that in the subcritical case $\alpha \in (0, 1/2)$ in Subsection 4.1, any strong limit Θ already yielded a solution to the limiting system (3.2). Indeed, there can be just one such strong limit by the one-sided uniqueness of the limiting system (3.2). Also, in that case one can find a nice split of the dynamics into a sequence of intervals where no collision happens. Thus, on every such interval, the kind of collisional state of our trajectory remains unchanged. Let us remember that the reason why that sequence fills the whole half-line in the subcritical case relies on the following facts: first, by

uniqueness we can characterize the sticking of oscillators and once they stick during some time they remain stuck for all time. In particular, only $N-1$ sticking times can exist. Second, when an accumulation of collisions takes place, it has to be at a sticking time. Hence, there can be just $N-1$ such accumulations of collisions, thus covering the whole half-line.

(2) Unfortunately, for $\alpha \in (1/2, 1)$ or $\alpha \in 1/2$ we still do not know at this point whether any limit Θ becomes a Filippov solution to the limiting system (3.2). Thus, despite the fact that we have clear characterizations of sticking of such solutions, we cannot apply them to any such limit Θ. In addition, the behavior of any H^1 weak limit can be very wild. Specifically, a possible scenario of an H^1_{loc} trajectory is that sticking might happen just for a short period of time, after which the cluster detaches. Also, "pure collisions" might accumulate at a non-sticking time exhibiting Zeno behavior (recall Remark 3.4 and Figure 6). Therefore, a split of the dynamics into countably many intervals (T_k, T_{k+1}) as in Subsection 4.1, where the collisional state remains unmodified, is not viable.

As discussed in the above remark, it is not clear how to achieve a split of the dynamics into countably many time intervals covering the whole half-line, each of them exhibiting unvaried collisional state. Hence we need a new approach, where the above explicit H-representation of the Filippov set-valued map at any collision state will play a role. One of our main tools will be the *Kuratowski–Ryll-Nardzewski measurable selection theorem* [30], which applies to set-valued Effros-measurable maps. For completeness we include the statement of that result, adapted to a finite-dimensional setting.

Lemma 4.5 (Kuratowski–Ryll-Nardzewski). *For any $n, m \in \mathbb{N}$, consider any set-valued map $\mathcal{F} : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ with non-empty and closed values. Assume that $\mathcal{F}$ is Effros-measurable, that is, for every open set $U \subseteq \mathbb{R}^m$, the set*

$$\{x \in \mathbb{R}^n : \mathcal{F}(x) \cap U \neq \emptyset\}.$$

is measurable. Then $\mathcal{F}$ has a measurable selection, i.e., there exists a measurable function $F : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$F(x) \in \mathcal{F}(x), \quad \text{a.e. } x \in \mathbb{R}^n.$$

Sometimes, it is helpful to control how many of these single-valued measurable selections of the Effros-measurable set-valued map are needed in order to essentially have the whole set-valued map "represented" in some sense. This is the content of an intimately related result: the *Castaing representation theorem* [13, Theorem III.30].

Lemma 4.6 (Castaing). *For any $n, m \in \mathbb{N}$, consider any set-valued map $\mathcal{F} : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ with non-empty and closed values. Assume that $\mathcal{F}$ is Effros-measurable. Then $\mathcal{F}$ has a Castaing representation, i.e., there exists a sequence $\{F^n\}_{n\in\mathbb{N}}$ of measurable maps $F^n : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\mathcal{F}(x) = \overline{\{F^n(x) : n \in \mathbb{N}\}}, \quad \text{a.e. } x \in \mathbb{R}^n.$$

These results will be directly applied to the critical case in Subsection 4.3. However, for the supercritical case, we will need a refinement of the above theorem to allow for integrable representations of the set-valued map. The Effros-measurability has to be improved to some integrability condition for set-valued maps.

Lemma 4.7. *For any $n, m \in \mathbb{N}$, consider any set-valued map $\mathcal{F} : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ with non-empty and closed values. Assume that $\mathcal{F}$ is Effros-measurable and strongly integrable, that is, the single-valued map $|\mathcal{F}|$ is integrable, where*

$$|\mathcal{F}|(x) := \sup\{|y| : y \in \mathcal{F}(x)\}, \quad a.e.\ x \in \mathbb{R}^n.$$

Then every measurable selection of $\mathcal{F}$ is integrable. In particular, $\mathcal{F}$ enjoys a Castaing representation consisting of integrable selections.

Proof. Take any measurable selection F of $\mathcal{F}$, which exists by Lemma 4.5. Then, by definition of $|\mathcal{F}|$,

$$|F(x)| \leq |\mathcal{F}|(x), \quad \text{a.e. } x \in \mathbb{R}^n.$$

Since $|\mathcal{F}|$ is integrable, the first part of the result holds true. The second one is a simple consequence of the first one along with Lemma 4.6. ■

Remark 4.4. The same ideas as in Lemma 4.7 also yield similar statements for the spaces $L^1_{\text{loc}}(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$:

(1) If $\mathcal{F}$ is locally strongly integrable, i.e., $|\mathcal{F}| \in L^1_{\text{loc}}(\mathbb{R}^n)$, then every measurable selection belongs to $L^1_{\text{loc}}(\mathbb{R}^n)$.
(2) If $\mathcal{F}$ is strongly essentially bounded, i.e., $|\mathcal{F}| \in L^\infty(\mathbb{R}^n)$, then each measurable selection belongs to $L^\infty(\mathbb{R}^n)$.

Theorem 4.3. *Consider the classical solutions $\{\Theta^\varepsilon\}_{\varepsilon>0}$ to the regularized system (4.8) with $\alpha \in (1/2, 1)$ and any weak H^1_{loc} limit Θ. Then*

$$\dot{\Theta}(t) \in \mathcal{H}(\Theta(t)), \quad a.e.\ t \geq 0.$$

Proof. • *Step 1: H-representation of the Filippov map.* By Proposition 4.1,

$$\mathcal{H}(\Theta(t)) = \bigcap_{l=1}^{\kappa(t)} \mathcal{P}_l(t), \tag{4.13}$$

where

$$\mathcal{P}_l(t) := \{x \in \mathbb{R}^N : a_l(t) \cdot x = b_l(t)\}$$

with

$$a_l(t) := \frac{1}{n_l(t)} \sum_{i \in E_l(t)} e_i, \quad b_l(t) := \frac{1}{n_l(t)} \sum_{i \in E_l(t)} \left(\Omega_i + \frac{K}{N} \sum_{j \notin \mathcal{C}_i(t)} h(\theta_j - \theta_i)\right).$$

• *Step 2: Castaing representation of coefficients.* Define $\mathcal{A}:\mathbb{R}_0^+\to 2^{\mathbb{R}^N}$ and $\mathcal{B}:\mathbb{R}_0^+\to 2^{\mathbb{R}}$ by

$$\mathcal{A}(t):=\{a_l(t)\,:\,l=1,\ldots,\kappa(t)\}\quad\text{and}\quad\mathcal{B}(t):=\{b_l(t)\,:\,l=1,\ldots,\kappa(t)\}.$$

It is clear that both maps take closed non-empty values and they are Effros-measurable. Thus, Lemma 4.6 yields their Castaing representations. On the one hand, $\mathcal{A}$ is strongly essentially bounded (see Remark 4.4), so there exists a sequence $\{A^n\}_{n\in\mathbb{N}}\subseteq L^\infty(0,\infty)$ such that

$$\mathcal{A}(t)=\overline{\{A^n(t)\,:\,n\in\mathbb{N}\}}$$

for almost every $t\geq 0$. By the finiteness of $\mathcal{A}(t)$ we equivalently have

$$\{a_l(t)\,:\,l=1,\ldots,\kappa(t)\}=\{A^n(t)\,:\,n\in\mathbb{N}\},\tag{4.14}$$

for almost every $t\geq 0$. However, it is not clear whether $\mathcal{B}$ is strongly locally integrable since we expect possible switches of the collisional type of the limiting $\Theta(t)$, thus on the coefficients $b_l(t)$.

• *Step 3: Strong local integrability of $\mathcal{B}$.* Let us show that the above wild behavior still does not prevent us from achieving our goal. Consider the regularized coefficients

$$b_l^\varepsilon(t):=\frac{1}{n_l(t)}\sum_{i\in E_l(t)}\left(\Omega_i+\frac{K}{N}\sum_{j\notin\mathcal{C}_i(t)}h_\varepsilon(\theta_j^\varepsilon-\theta_i^\varepsilon)\right),\qquad l=1,\ldots,\kappa(t).$$

We can define a similar set-valued map $\mathcal{B}^\varepsilon:\mathbb{R}_0^+\to 2^{\mathbb{R}}$ by

$$\mathcal{B}^\varepsilon(t)=\{b_l^\varepsilon(t)\,:\,l=1,\ldots,\kappa(t)\}.$$

By definition it is clear that

$$\lim_{\varepsilon\to 0}b_l^\varepsilon(t)=b_l(t)$$

for every $l=1,\ldots,\kappa(t)$ since $j\notin\mathcal{C}_i(t)$ in the definitions and, at those $\theta_j(t)-\theta_i(t)$, the limiting kernel h is continuous. Since both $\mathcal{B}(t)$ and $\mathcal{B}^\varepsilon(t)$ consist of finitely many terms, we deduce that

$$|\mathcal{B}^\varepsilon(t)|\to|\mathcal{B}(t)|,\qquad\text{a.e. } t\in\mathbb{R}_0^+.\tag{4.15}$$

Then Fatou's lemma on any finite time interval $[0,T]\subseteq\mathbb{R}_0^+$ with $T>0$ entails

$$\int_0^T|\mathcal{B}|(t)\,dt\leq\liminf_{\varepsilon\to 0}\int_0^T|\mathcal{B}^\varepsilon|(t)\,dt.\tag{4.16}$$

By definition, it is clear that

$$\dot{\Theta}^\varepsilon(t)\cdot a_l(t)=\frac{1}{n_l(t)}\sum_{i\in E_l(t)}\left(\Omega_i+\frac{K}{N}\sum_{j=1}^N h_\varepsilon(\theta_j^\varepsilon-\theta_i^\varepsilon)\right)=b_l^\varepsilon(t),$$

where we have canceled the terms with $j \in E_l(t)$ in the last step by the antisymmetry of h_ε. Hence, our set-valued maps are strongly dominated as follows:

$$|\mathcal{B}^\varepsilon|(t) \leq |\dot{\Theta}^\varepsilon(t)|, \qquad \text{a.e. } t \geq 0. \tag{4.17}$$

Putting (4.17) into (4.16) we obtain

$$\begin{aligned}
\int_0^T |\mathcal{B}|(t)\,dt &= \int_0^T \liminf_{\varepsilon\to 0} |\mathcal{B}^\varepsilon|(t)\,dt \\
&\leq \liminf_{\varepsilon\to 0} \int_0^T |\dot{\Theta}^\varepsilon(t)|\,dt \\
&\leq T^{1/2} \liminf_{\varepsilon\to 0} \left(\int_0^T |\dot{\Theta}^\varepsilon(t)|^2\,dt\right)^{1/2} \\
&\leq T^{1/2}(C_\Omega^2 T + 2V_{\text{int}}(\Theta_0))^{1/2} < \infty.
\end{aligned}$$

Here, we have used the Cauchy–Schwarz inequality in the second step and the a priori bound of Lemma 4.3 in the last one. Then Remark 4.4 yields the existence of a Castaing representation $\{B^n\}_{n\in\mathbb{N}} \subseteq L^1_{\text{loc}}(0,\infty)$ of the map $\mathcal{B}$. Again, we conclude that

$$\{b_l(t) : l = 1, \ldots, \kappa(t)\} = \{B^n(t) : n \in \mathbb{N}\} \tag{4.18}$$

for almost every $t \geq 0$.

• *Step 4: Conclusion.* Since $y^\varepsilon(t) \in \mathcal{H}(\Theta(t))$ for every $\varepsilon > 0$ and every $t \geq 0$, the H-representation (4.13) along with the essentially bounded and locally integrable representations (4.14) and (4.18) yield

$$A^n(t) \cdot y^\varepsilon(t) = B^n(t), \qquad n \in \mathbb{N},$$

for almost every $t \geq 0$. In particular,

$$\int_0^\infty A^n(t) \cdot y^\varepsilon(t)\,\varphi(t)\,dt = \int_0^\infty B^n(t)\varphi(t)\,dt$$

for every $\varepsilon > 0$, each $\varphi \in C_c(\mathbb{R}^+)$ and any $n \in \mathbb{N}$. Notice that the boundedness and local integrability of our selectors ensures that such expressions make sense. We can now use the weak L^2 convergence of y^ε towards $\dot{\Theta}$ to obtain

$$\int_0^\infty A^n(t) \cdot \dot{\Theta}(t)\varphi(t)\,dt = \int_0^\infty B^n(t)\varphi(t)\,dt$$

for every $\varphi \in C_c(\mathbb{R}^+)$ and each $n \in \mathbb{N}$. The fundamental lemma of calculus of variations along with the Castaing representations in (4.14) and (4.18) and the H-representation in (4.13) allow us to deduce the desired result. ■

4.3. Limit in the critical case

In this subsection, we will address the singular limit of the regularized system (2.5)–(2.6) towards a Filippov solution to (3.2) in the critical regime $\alpha = 1/2$. We will mostly apply a similar approach to that in the supercritical regime. Nevertheless, there are several novelties. First, we will show that we actually enjoy a better $W^{1,\infty}$ a priori estimate, apart from the above H^1 bound of Lemma 4.3. Second, the explicit expression of the Filippov map in Proposition 4.1 in terms of the intersection of hyperplanes will be adapted to this case.

Lemma 4.8. *In the critical regime $\alpha = 1/2$, let Θ^ε be the unique global-in-time solution to the regularized system (4.8). Then*

$$\|\dot{\Theta}^\varepsilon\|_{L^\infty((0,\infty),\mathbb{R}^N)} \le C_\Omega + K$$

for every $\varepsilon > 0$, where $C_\Omega := (\sum_{i=1}^N \Omega_i^2)^{1/2}$.

We omit the proof since it is a clear consequence of the boundedness of h in the critical case. As a consequence of Lemma 4.8, we infer the existence of a subsequence of $\{\Theta^\varepsilon\}_{\varepsilon>0}$ (not relabeled) that weak-* converges to some $\Theta \in W^{1,\infty}_{\text{loc}}((0,\infty),\mathbb{R}^N)$ in $W^{1,\infty}((0,T),\mathbb{R}^N)$ for every $T > 0$. In particular,

$$\Theta^\varepsilon \to \Theta \quad \text{in } C([0,T],\mathbb{R}^N), \qquad \dot{\Theta}^\varepsilon \overset{*}{\rightharpoonup} \dot{\Theta} \quad \text{in } L^\infty((0,T),\mathbb{R}^N),$$

for every $T > 0$. In addition, the same split as in (4.10) can be considered and we obtain

$$x^\varepsilon \to 0 \quad \text{in } C([0,T],\mathbb{R}^N), \qquad y^\varepsilon \overset{*}{\rightharpoonup} \dot{\Theta} \quad \text{in } L^\infty((0,T),\mathbb{R}^N),$$

and $y^\varepsilon(t) \in \mathscr{H}(\Theta(t))$ for every $t \ge 0$ and $\varepsilon > 0$. Hence, Θ^ε becomes an approximate solution in the same sense as in (4.11). What is more, the same Remark 4.2 is in force. Thus, again we cannot ensure pointwise convergence of $\dot{\Theta}^\varepsilon$. In order to obtain an analogous characterization of the Filippov map, we will need the next technical lemma.

Lemma 4.9. *For any $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$, the following two assertions are equivalent:*

(1) *There exists $Y \in \mathrm{Skew}_n([-1,1])$ such that*

$$x = Y \cdot \mathbf{j}.$$

(2) *We have*

$$\frac{1}{k}\sum_{i=1}^k x_{\sigma_i} \in [-(n-k),(n-k)]$$

for every permutation σ of $\{1,\ldots,n\}$ and any $k \in \mathbb{N}$.

For easier readability, we postpone the proof to Appendix B. The following result a straightforward consequence of Lemma 4.9 along with the explicit formula in Proposition 3.2.

Proposition 4.2. *In the critical regime $\alpha = 1/2$, the value $\mathcal{H}(\Theta)$ of the Filippov set-valued map associated with H is the compact and convex polytope of points $(\omega_1, \dots, \omega_N) \in \mathbb{R}^N$ whose H-representations consist of the affine inequalities (recall Subsection 2.3)*

$$\frac{1}{m}\sum_{i\in I}\omega_i \in \frac{1}{m}\sum_{i\in I}\Big(\Omega_i + \frac{K}{N}\sum_{j\notin\mathcal{C}_i} h(\theta_j - \theta_i)\Big) + \Big[-\frac{K}{N}(n_k - m), \frac{K}{N}(n_k - m)\Big] \tag{4.19}$$

for every $k = 1, \dots, \kappa$ and $I \subseteq E_k$ with $\#I = m$.

Now we turn to the main result, i.e., the convergence of the singular limit to a Filippov solution of the critical system.

Theorem 4.4. *Consider the classical solutions $\{\Theta^\varepsilon\}_{\varepsilon>0}$ to the regularized system (4.8) with $\alpha = 1/2$ and any weak-* limit Θ in $W^{1,\infty}_{\mathrm{loc}}$. Then*

$$\dot\Theta(t) \in \mathcal{H}(\Theta(t)), \qquad \text{a.e. } t \geq 0.$$

Proof. We mimic the proof of Theorem 4.3. Recall that by Proposition 4.2, an analogous H-representation to that in (4.13) holds. Specifically,

$$\mathcal{H}(\Theta(t)) = \bigcap_{l=1}^{\kappa(t)} \bigcap_{I\subseteq E_l} (\mathcal{S}^+_{l,I}(t) \cap \mathcal{S}^-_{l,I}(t)), \tag{4.20}$$

where the half-spaces are

$$\mathcal{S}^+_{l,I}(t) := \{x \in \mathbb{R}^N : a_{l,I}(t)\cdot x \leq b^+_{l,I}(t)\},$$
$$\mathcal{S}^-_{l,I}(t) := \{x \in \mathbb{R}^N : a_{l,I}(t)\cdot x \geq b^-_{l,I}(t)\},$$

for every $I \subseteq E_l(t)$. We set

$$a_{l,I}(t) := \frac{1}{m}\sum_{i\in I} e_i,$$
$$b^\pm_{l,I}(t) := \frac{1}{m}\sum_{i\in I}\Big(\Omega_i + \frac{K}{N}\sum_{j\notin\mathcal{C}_i(t)} h(\theta_j(t) - \theta_i(t))\Big) \pm \frac{K}{N}(n_l(t) - m),$$

where $m = \#I$. Now, the coefficients are clearly uniformly bounded. A straightforward application of Remark 4.4 leads to the existence of essentially bounded selectors for the coefficients. Namely, we can give an ordering such that

$$\{a_{l,I}(t) : l = 1, \dots, \kappa(t),\ I \subseteq E_l(t)\} = \{A^n(t) : n \in \mathbb{N}\} \tag{4.21}$$
$$\{b^\pm_{l,I}(t) : l = 1, \dots, \kappa(t),\ I \subseteq E_l(t)\} = \{B^{\pm,n}(t) : n \in \mathbb{N}\}, \tag{4.22}$$

for almost every $t \geq 0$. Recall that $y^\varepsilon(t) \in \mathcal{H}(\Theta(t))$ for every $\varepsilon > 0$ and every $t \geq 0$. Then, by (4.20)–(4.22), we equivalently have

$$A^n(t)\cdot y^\varepsilon(t) \leq B^{+,n}(t) \quad \text{and} \quad A^n(t)\cdot y^\varepsilon(t) \geq B^{-,n}(t),$$

for all $n \in \mathbb{N}$, each $\varepsilon > 0$ and almost every $t \geq 0$. In particular,

$$\int_0^\infty A^n(t) \cdot y^\varepsilon(t)\,\varphi(t)\,dt \leq \int_0^\infty B^{+,n}\,dt,$$
$$\int_0^\infty A^n(t) \cdot y^\varepsilon(t)\,\varphi(t)\,dt \geq \int_0^\infty B^{-,n}\,dt,$$

for all $n \in \mathbb{N}$, each $\varepsilon > 0$ and any non-negative $\varphi \in C_c(\mathbb{R}^+)$. Then, using the weak-* L^∞ convergence we obtain

$$\int_0^\infty A^n(t) \cdot \dot{\Theta}(t)\,\varphi(t)\,dt \leq \int_0^\infty B^{+,n}\,dt,$$
$$\int_0^\infty A^n(t) \cdot \dot{\Theta}(t)\,\varphi(t)\,dt \geq \int_0^\infty B^{-,n}\,dt,$$

for all $n \in \mathbb{N}$ and any non-negative $\varphi \in C_c(\mathbb{R}^+)$. Hence, the result follows from the fundamental lemma of calculus of variations along with the Castaing representations (4.21)–(4.22) and the H-representation (4.20). ∎

4.4. Comparison with previous results about singular weighted systems

We have studied the existence and one-sided uniqueness for the singular weighted first order Kuramoto model in all the subcritical, critical and supercritical regimes. We now compare our result with previous research on the singular weighted Cucker–Smale model, which is a second order system describing the flocking behavior of interacting particles. In Section 2, the first order Kuramoto model (2.1) was shown to be equivalent to its second order augmentation (2.4). On the one hand, this is clear for regular weights as studied in Theorem 2.1 (see [16, 22]). What is more, it remains true in our case, which is characterized by singular weights. However, we must be specially careful with time regularity in order for such heuristic arguments to become true. Let us focus on the subcritical regime, where the rigorous equivalence between (2.1) and (2.4) follows from Remark 3.6 by virtue of the one-sided uniqueness in both models. Indeed, in the subcritical case, the "influence function" of the augmented flocking-type model reads

$$h'(\theta) = \frac{1}{|\theta|_o^{2\alpha}}\left[\cos\theta - 2\alpha\frac{\sin|\theta|_o}{|\theta|_o}\right] \sim \frac{1-2\alpha}{|\theta|_o^{2\alpha}} \quad \text{near } \theta \in 2\pi\mathbb{Z}, \tag{4.23}$$

which enjoys mild singularities of order $2\alpha < 1$ in the subcritical case. The singular second order model (2.4), (4.23) shares some similarities with the Cucker–Smale model with singular weights,

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \dfrac{K}{N}\displaystyle\sum_{j=1}^N \psi(|x_j - x_i|)(v_j - v_i), \end{cases} \tag{4.24}$$

where the communication weight ψ is given by

$$\psi(r) := \frac{1}{r^{\beta}} \tag{4.25}$$

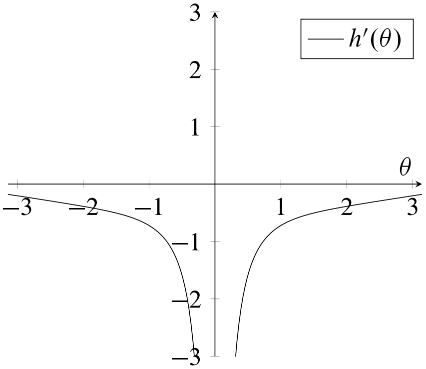
for $r > 0$ and $\beta > 0$. Although some results regarding the asymptotic behavior of that system have been established [20], its well-posedness has not been addressed until very recently in [36,37] for the microscopic model and [9,33,39,43] for some first and second order kinetic and macroscopic versions of the model. Regarding the microscopic system (4.24)–(4.25), the existence of global C^1 piecewise weak $W^{2,1}$ solutions $(x_1, \ldots, x_N)$ has been established in [36] for $\beta \in (0,1)$, which corresponds to $\alpha \in (0, 1/2)$ in our setting (see Theorem 3.1, Theorem 4.1 and Remark 4.1). Also, in the weakly singular regime $\beta \in (0, 1/2)$ (i.e., $\alpha \in (0, 1/4)$), the same author proved in [37] that the velocities $(v_1, \ldots, v_N)$ are indeed absolutely continuous. Consequently, the C^1 weak solutions $(x_1, \ldots, x_N)$ are actually $W^{2,1}_{\mathrm{loc}}$ in that case. This property was proved through a differential inequality.

The method of proof is similar to ours in Section 4 and relies on a regularization process of the second order model near the collision times. In our case, we have obtained a similar regularization process of the first order model, entailing the corresponding regularization of the augmented second order model. Indeed, the method has not only proved successful in the subcritical case, but also in the critical and supercritical cases. Also, we have obtained the well-posedness results in an alternative way based on the gain of continuity of the kernel in the first order model along with its particular structure near the points of loss of Lipschitz-continuity. Indeed, we have succeeded in introducing an analogous well-posedness theory in Filippov sense for the endpoint case $\alpha = 1/2$ and the supercritical case $\alpha > 1/2$.

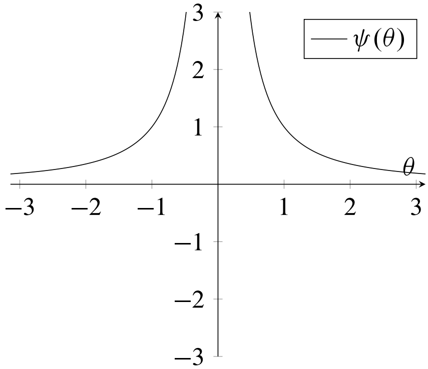
Regarding the more singular cases $\beta \geq 1$ (i.e., $\alpha \geq 1/2$), one can show that there exists some class of initial data for (4.24)–(4.25) such that one can avoid collisions and the solutions remain smooth for all time. Indeed, such solutions exhibit asymptotic flocking dynamics [2]. Very recently, it was shown in [10] that the loss of integrability of the kernel when $\beta \geq 1$ actually ensures the avoidance of collisions for general initial data. In that regime, the asymptotic flocking behavior is not guaranteed for any initial data. However, the ideas for (4.24)–(4.25) fail in our model (2.4)–(4.23) because the kernel h' with $\alpha \geq 1/2$ no longer behaves like the communication weight ψ with $\beta \geq 1$. Specifically, ψ is always a positive and decreasing function whereas h' is negative and increasing (see Figure 7). Thus, we expect our solutions to exhibit finite time collisions as shown in the results of Section 5. This is the reason for the generalized theory in Filippov's sense to come into play in the critical and supercritical cases.

5. Synchronization of the singular weighted system

We now analyze the collective behavior in the system (3.2). We first consider the system of two interacting oscillators. Then we extend the argument to the N-oscillator system.



(a) $h'(\theta)$



(b) $\psi(\theta)$

Fig. 7. Comparison of the functions $h'(\theta)$ and $\psi(\theta)$ with $\alpha = 0.75$.

5.1. Two-oscillator case

In this part, we consider the dynamics of two oscillators. The system (3.1) for two oscillators becomes

$$\dot{\theta}_1 = \Omega_1 + \frac{K}{2}\frac{\sin(\theta_2-\theta_1)}{|\theta_2-\theta_1|_o^{2\alpha}}, \qquad \dot{\theta}_2 = \Omega_2 + \frac{K}{2}\frac{\sin(\theta_1-\theta_2)}{|\theta_1-\theta_2|_o^{2\alpha}}. \tag{5.1}$$

Recall that in the critical and supercritical cases we do expect collisions (see Subsections 3.2 an 3.3). The above representation of the system is only valid before the first collision. After that time, the right hand side has to be replaced with the corresponding Filippov set-valued map. At this step, we shall focus on the dynamics before the first collision. Let us define the relative phase and the natural frequency by $\theta := \theta_2 - \theta_1$ and $\Omega := \Omega_2 - \Omega_1$. Then the system (5.1) can be rewritten as

$$\dot{\theta} = \Omega - K\frac{\sin\theta}{|\theta|_o^{2\alpha}}. \tag{5.2}$$

Proposition 5.1. *Let $\theta : [0,T) \to \mathbb{R}$ be a maximal classical solution to the differential equation (5.2) with $\alpha \in (0,1)$ such that the oscillators are identical, i.e., $\Omega = 0$, and with initial datum $0 < |\theta_0| < \pi$. Then the maximal time of existence T lies in the interval $[t_{\min}, t_{\max}]$, where*

$$t_{\min} = \frac{|\theta_0|^{2\alpha}}{2K\alpha} \quad \textit{and} \quad t_{\max} = \frac{|\theta_0|^{2\alpha+1}}{2K\alpha\sin|\theta_0|}.$$

In addition, the lower and upper estimates

$$|\theta_0|^{2\alpha} - 2K\alpha t \le |\theta|^{2\alpha} \le |\theta_0|^{2\alpha} - 2K\alpha t\frac{\sin|\theta_0|}{|\theta_0|}t$$

hold for all $t \in [0,T)$, and $\lim_{t\to T}\theta(t) = 0$. Hence, two identical oscillators confined to the half-circle exhibit finite-time phase synchronization.

Proof. First of all, in the identical case, $\pi + 2\pi\mathbb{Z}$ are equilibria of (5.2) where the interaction kernel is Lipschitz-continuous. Hence, the maximal solution θ cannot touch such values if started at θ_0. Therefore, $\theta(t) \in (-\pi, \pi)$ for every $t \in [0, T)$, and consequently $|\theta(t)|_o = |\theta(t)|$ for $t \in [0, T)$. Let us now multiply by $(2\alpha + 1)|\theta|^{2\alpha}\mathrm{sgn}(\theta)$ on both sides to obtain

$$\begin{aligned}\frac{d}{dt}|\theta|^{2\alpha+1} &= (2\alpha+1)|\theta|^{2\alpha}\mathrm{sgn}(\theta)\frac{d}{dt}\theta = -K(2\alpha+1)\sin\theta\,\mathrm{sgn}(\theta)\\ &= -K(2\alpha+1)\sin|\theta|.\end{aligned}$$

Denote $y = |\theta|^{2\alpha+1}$; then the equation becomes

$$\frac{d}{dt}y = -K(2\alpha+1)\sin y^{\frac{1}{2\alpha+1}}. \tag{5.3}$$

We now consider upper and lower estimates for (5.3) separately.

• *Lower estimate:* Since $|y| \ge \sin|y|$, we have

$$\frac{d}{dt}y \ge -K(2\alpha+1)y^{\frac{1}{2\alpha+1}}.$$

By multiplying by $\frac{2\alpha}{2\alpha+1}y^{-\frac{1}{2\alpha+1}}$ on both sides, we obtain

$$\frac{d}{dt}y^{\frac{2\alpha}{2\alpha+1}} \ge -2K\alpha.$$

This yields

$$y^{\frac{2\alpha}{2\alpha+1}} \ge y_0^{\frac{2\alpha}{2\alpha+1}} - 2K\alpha t.$$

Thus, we have a lower estimate

$$|\theta|^{2\alpha} \ge |\theta_0|^{2\alpha} - 2K\alpha t \qquad \text{for } 0 \le t < T.$$

In particular, the above lower estimate shows that

$$T \ge \frac{|\theta_0|^{2\alpha}}{2K\alpha} \equiv t_{\min}.$$

• *Upper estimate:* As long as $0 \le y < \pi^{2\alpha+1}$, the solution y is non-increasing, i.e., $\frac{d}{dt}y \le 0$. Since the initial data θ_0 satisfies $|\theta_0| < \pi$, we have $y_0 < \pi^{2\alpha+1}$, thus $y(t) \le y_0$ for $t > 0$. Hence,

$$\sin y^{\frac{1}{2\alpha+1}} \ge \frac{\sin y_0^{\frac{1}{2\alpha+1}}}{y_0^{\frac{1}{2\alpha+1}}}y^{\frac{1}{2\alpha+1}}. \tag{5.4}$$

Applying (5.4) to (5.3), we find

$$\frac{d}{dt}y \le -K(2\alpha+1)\frac{\sin y_0^{\frac{1}{2\alpha+1}}}{y_0^{\frac{1}{2\alpha+1}}}y^{\frac{1}{2\alpha+1}}.$$

Multiplying by $\frac{2\alpha}{2\alpha+1} y^{-\frac{1}{2\alpha+1}}$ on both sides, we obtain

$$\frac{d}{dt} y^{\frac{2\alpha}{2\alpha+1}} \leq -2K\alpha \frac{\sin y_0^{\frac{1}{2\alpha+1}}}{y_0^{\frac{1}{2\alpha+1}}},$$

which yields

$$y^{\frac{2\alpha}{2\alpha+1}} \leq y_0^{\frac{2\alpha}{2\alpha+1}} - 2K\alpha \frac{\sin y_0^{\frac{1}{2\alpha+1}}}{y_0^{\frac{1}{2\alpha+1}}} t.$$

This is equivalent to

$$|\theta|^{2\alpha} \leq |\theta_0|^{2\alpha} - 2K\alpha \frac{\sin|\theta_0|}{|\theta_0|} t \qquad \text{for } 0 \leq t < T.$$

Again, the upper estimate shows that

$$T \leq \frac{|\theta_0|^{2\alpha+1}}{2K\alpha \sin|\theta_0|} \equiv t_{\max}. \qquad \blacksquare$$

Assume that the oscillators are non-identical: $\Omega = \Omega_2 - \Omega_1 > 0$, and the system (5.1) has a phase-locked state $(\bar{\theta}_1, \bar{\theta}_2)$ satisfying $0 < \bar{\theta}_2 - \bar{\theta}_1 < \pi$. Then the equation (5.2) has an equilibrium $\bar{\theta} = \bar{\theta}_2 - \bar{\theta}_1 \in (0, \pi)$ such that

$$\Omega - K \frac{\sin\bar{\theta}}{|\bar{\theta}|_o^{2\alpha}} = 0. \tag{5.5}$$

To guarantee the existence of that equilibrium, we need the following conditions for the coupling strength K:

$$\begin{array}{llll} \text{if} & \alpha < 1/2, & \text{choose} & K \geq \Omega/\overline{h}, \\ \text{if} & \alpha = 1/2, & \text{choose} & K > \Omega, \end{array}$$

where $\overline{h} := \max_{0<r<\pi} h(r)$. Note that the equilibrium exists for of $\alpha > 1/2$ without any condition on $K > 0$. We now investigate the stabilities of the equilibria in each case.

Proposition 5.2. *Let θ be a solution of* (5.2).

(1) *For $\alpha \geq 1/2$, the equilibrium $\bar{\theta}$ is unstable. Furthermore, if*

$$\theta_0 \neq 0 \quad \textit{and} \quad \theta_0 \neq \bar{\theta},$$

then the solution θ reaches 0 or 2π in finite time.

(2) *For $\alpha < 1/2$, there are a stable equilibrium $\bar{\theta} \in (0, \tilde{\theta})$ and an unstable equilibrium $\bar{\theta}^* \in (\tilde{\theta}, \pi)$, where $\tilde{\theta} \in (0, \pi/2)$ is the solution to $\tilde{\theta} = 2\alpha \tan\tilde{\theta}$. Moreover, if $\theta_0 \in (-2\pi + \bar{\theta}^*, \bar{\theta}^*)$, then the solution θ converges to $\bar{\theta}$ asymptotically.*

Proof. We linearize the equation (5.2) near $\bar{\theta}$ as

$$\dot{\theta} = -kh'(\bar{\theta})(\theta - \bar{\theta}) + R(\bar{\theta}).$$

When $\alpha \geq 1/2$, we have $h'(\bar{\theta}) < 0$ for $\theta \in (0, \pi)$. Thus, the equilibrium $\bar{\theta}$ is unstable. For $\alpha < 1/2$, if the equilibrium $\bar{\theta}$ is located in $(0, \tilde{\theta})$, we have $h'(\bar{\theta}) > 0$, i.e., it is stable. By a similar argument, since $h'(\bar{\theta}^*) < 0$, the equilibrium $\bar{\theta}^*$ located in $(\tilde{\theta}, \pi)$ is unstable. We now investigate the convergence of the solution.

• *Step 1*: Critical and supercritical cases, $\alpha \geq 1/2$.

∘ *Case 1 ($\theta_0 > \bar{\theta}$)*: Since the function h is decreasing in $(0, 2\pi)$, we have $h(\theta) < h(\bar{\theta})$ for $\theta \in (\bar{\theta}, 2\pi)$. Thus,

$$\dot{\theta} = \Omega - Kh(\theta) > \Omega - Kh(\bar{\theta}) = 0 \quad \text{for } \theta \in (\bar{\theta}, 2\pi).$$

Moreover, due to the monotonic increase of θ, we obtain the lower estimate for the frequency:

$$\dot{\theta} = \Omega - Kh(\theta) > \Omega - Kh(\theta_0) > 0 \quad \text{for } \theta \in (\bar{\theta}, 2\pi).$$

Hence, there exists a finite time $t_1 < \frac{2\pi - \theta_0}{\Omega - Kh(\theta_0)}$ at which the solution converges to 2π.

∘ *Case 2 ($\theta_0 < \bar{\theta}$):* We apply an analogous argument. Since h is decreasing, we deduce $h(\theta) > h(\bar{\theta})$ for $\theta \in (0, \bar{\theta})$. Thus,

$$\dot{\theta} = \Omega - Kh(\theta) < \Omega - Kh(\bar{\theta}) = 0 \quad \text{for } \theta \in (0, \bar{\theta}).$$

This monotonic decrease of the phase yields the upper estimate for the frequency:

$$\dot{\theta} = \Omega - Kh(\theta) < \Omega - Kh(\theta_0) < 0 \quad \text{for } \theta \in (0, \bar{\theta}).$$

So, there exists a finite time $t_2 < \frac{\theta_0}{|\Omega - Kh(\theta_0)|}$ at which the solution converges to zero.

• *Step 2:* Subcritical case $\alpha < 1/2$. We consider two steps for the asymptotic convergence to the equilibrium:

∘ *Step 2a:* We first show the solution moves into the interval $(0, \tilde{\theta})$ in finite time when the initial datum θ_0 is in $(-2\pi + \bar{\theta}^*, 0] \cup [\tilde{\theta}, \bar{\theta}^*)$. As long as the solution θ is in $[\tilde{\theta}, \bar{\theta}^*)$, we have $h(\theta) > h(\bar{\theta})$. Thus, the solution is non-increasing:

$$\dot{\theta} = \Omega - Kh(\theta) < \Omega - Kh(\bar{\theta}) = 0 \quad \text{for } \theta \in [\tilde{\theta}, \bar{\theta}^*).$$

Moreover, the non-increase of solution $\theta(t) \leq \theta_0$ gives an upper bound on the frequency:

$$\dot{\theta} = \Omega - Kh(\theta) < \Omega - Kh(\theta_0) < 0,$$

while θ is in $[\tilde{\theta}, \bar{\theta}^*)$. So, there exists a finite time $t_3 := \frac{\theta_0 - \tilde{\theta}}{|\Omega - Kh(\theta_0)|}$ such that the solution satisfies $\theta(t) < \tilde{\theta}$ for $t > t_3$. Analogously, if the initial datum θ_0 is in $(-2\pi + \bar{\theta}^*, 0]$, then we have $h(\theta) < h(\bar{\theta})$, the solution is non-decreasing:

$$\dot{\theta} = \Omega - Kh(\theta) > \Omega - Kh(\bar{\theta}) = 0,$$

and the frequency has a lower bound

$$\dot{\theta} = \Omega - Kh(\theta) > \Omega - Kh(\theta_0) > 0,$$

as long as $\theta \in (-2\pi + \bar{\theta}^*, 0]$. Thus, there exists a finite time $t_4 := \frac{|\theta_0|}{|\Omega - Kh(\theta_0)|}$ such that the solution satisfies $\theta(t) > 0$ for $t > t_4$.

∘ *Step 2b:* We will show that the solution converges to the stable equilibrium $\bar{\theta}$ asymptotically when the initial datum is in $(0, \tilde{\theta})$. Suppose the initial datum is in $(0, \bar{\theta})$. Then

$$\frac{h(\bar{\theta})}{\bar{\theta}}\theta < h(\theta) < h'(\bar{\theta})(\theta - \bar{\theta}) + h(\bar{\theta}).$$

Thus, the solution satisfies the differential inequality

$$\Omega - K\big(h'(\bar{\theta})(\theta - \bar{\theta}) + h(\bar{\theta})\big) < \dot{\theta} < \Omega - \frac{Kh(\bar{\theta})}{\bar{\theta}}\theta.$$

By Grönwall's lemma, we obtain

$$\bar{\theta} - (\bar{\theta} - \theta_0)e^{-kh'(\bar{\theta})t} < \theta(t) < \bar{\theta} - (\bar{\theta} - \theta_0)e^{-\frac{Kh(\bar{\theta})}{\bar{\theta}}t}.$$

Similarly, if the initial datum θ_0 is in $(\bar{\theta}, \tilde{\theta})$, the function h satisfies

$$\frac{h(\tilde{\theta}) - h(\bar{\theta})}{\tilde{\theta} - \bar{\theta}}(\theta - \bar{\theta}) + h(\bar{\theta}) < h(\theta) < h'(\bar{\theta})(\theta - \bar{\theta}) + h(\bar{\theta}).$$

Then we have the following differential inequality:

$$\Omega - K\big(h'(\bar{\theta})(\theta - \bar{\theta}) + h(\bar{\theta})\big) < \dot{\theta} < \Omega - K\left(\frac{h(\tilde{\theta}) - h(\bar{\theta})}{\tilde{\theta} - \bar{\theta}}(\theta - \bar{\theta}) + h(\bar{\theta})\right).$$

Hence, by Grönwall's lemma, we find

$$\bar{\theta} - (\theta_0 - \bar{\theta})e^{-Kh'(\bar{\theta})t} < \theta(t) < \bar{\theta} - (\theta_0 - \bar{\theta})e^{-K\frac{h(\tilde{\theta}) - h(\bar{\theta})}{\tilde{\theta} - \bar{\theta}}t}.$$ ■

Remark 5.1. In the subcritical case $\alpha \in (0, 1/2)$, for two non-identical oscillators a phase-locked state emerges asymptotically (see Proposition 5.2), whereas for two identical oscillators phase synchronization holds in finite time (see Proposition 5.1). However, in the critical and supercritical cases $\alpha \in [1/2, 1)$, phase synchronization always appears in finite time as shown in Propositions 5.2 and 5.1 as long as the initial phase configuration does not agree with the unstable phase-locked state $\bar{\theta}$. Namely, in the supercritical case both oscillators stick together into a unique cluster moving at constant frequency $\widehat{\Omega} = (\Omega_1 + \Omega_2)/2$, independently of the chosen natural frequencies. However, in the critical case, the same only happens under the assumption $|\Omega_1 - \Omega_2| \leq K$. In the other case, the resulting cluster will instantaneously split.

5.2. N-oscillator case

In this subsection, we consider the system of N interacting oscillators. We will first focus on the dynamics in the simpler subcritical case $\alpha \in (0, 1/2)$, where solutions have been proved to be classical (see Theorem 3.1). The reason to start with this case is that the right hand side of (3.2) can then be considered in the single-valued sense. The dynamics in the critical case $\alpha = 1/2$ and some intuition about the dynamics in the supercritical regime $\alpha \in (1/2, 1)$ will be provided at the end of this subsection.

Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to the system (3.2). We first study phase synchronization for identical oscillators. Let the indices M and m satisfy

$$\theta_M(t) := \max\{\theta_1(t), \ldots, \theta_N(t)\} \quad \text{and} \quad \theta_m(t) := \min\{\theta_1(t), \ldots, \theta_N(t)\}, \tag{5.6}$$

for each time $t \geq 0$. Then we can define the phase diameter to be

$$D(\Theta) := \theta_M - \theta_m. \tag{5.7}$$

Theorem 5.1. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to (3.2) with $\alpha \in (0, 1/2)$ for identical oscillators ($\Omega_i = 0$ for $i = 1, \ldots, N$). Assume that the initial configuration Θ_0 is confined to a half-circle, i.e., $0 < D(\Theta_0) < \pi$. Then there is complete phase synchronization at a finite time not larger than*

$$T_c = \frac{D(\Theta_0)^{1-2\alpha}}{2\alpha K h(D(\Theta_0))}.$$

Proof. We consider the dynamics of the phase diameter:

$$\frac{d}{dt} D(\Theta) = \frac{K}{N} \sum_{j=1}^{N} \big(h(\theta_j - \theta_M) - h(\theta_j - \theta_m) \big).$$

Since $h(\theta_j - \theta_M) < 0$ and $h(\theta_j - \theta_m) > 0$ as long as $D(\Theta) < \pi$, we have

$$\frac{d}{dt} D(\Theta) \leq 0 \quad \text{and} \quad D(\Theta(t)) \leq D(\Theta_0) < \pi \quad \text{for } t > 0.$$

Due to the contraction of phase, and the fact that $(0, \pi) \ni \theta \mapsto h(\theta)/\theta$ is decreasing, we have

$$h(\theta_j - \theta_M) \leq \frac{h(D(\Theta_0))}{D(\Theta_0)} (\theta_j - \theta_M) \quad \text{and} \quad h(\theta_j - \theta_m) \geq \frac{h(D(\Theta_0))}{D(\Theta_0)} (\theta_j - \theta_m).$$

Thus, we obtain the following differential inequality:

$$\begin{aligned}
\frac{d}{dt} D(\Theta) &\leq \frac{K}{N} \sum_{j=1}^{N} \left(\frac{h(D(\Theta_0))}{D(\Theta_0)} (\theta_j - \theta_M) - \frac{h(D(\Theta_0))}{D(\Theta_0)} (\theta_j - \theta_m) \right) \\
&= \frac{K}{N} \frac{h(D(\Theta_0))}{D(\Theta_0)} \sum_{j=1}^{N} \big((\theta_j - \theta_M) - (\theta_j - \theta_m) \big) = -K \frac{h(D(\Theta_0))}{D(\Theta_0)} D(\Theta).
\end{aligned}$$

By Grönwall's lemma, we obtain

$$D(\Theta) \leq D(\Theta_0) e^{-K\frac{h(D(\Theta_0))}{D(\Theta_0)}t} \quad \text{for } t \geq 0.$$

Notice that $h(\theta)$ behaves like $\theta^{1-2\alpha}$ near the origin. Indeed, it is easy to prove that for every $\theta_* \in (0, \pi)$,

$$h(\theta) \geq \frac{h(\theta_*)}{\theta_*^{1-2\alpha}}\,\theta^{1-2\alpha}, \qquad \forall \theta \in [0, \theta_*].$$

The main idea is to show that the mapping $\theta \mapsto h(\theta)/\theta^{1-2\alpha}$ is non-increasing in $[0, \pi]$. Since the phase diameter $D(\Theta)$ is bounded above by $D(\Theta_0)$ we can take $\theta_* = D(\Theta_0)$ and apply the above lower estimate for h to obtain

$$\begin{aligned}
\frac{d}{dt}D(\Theta) &= \frac{K}{N}\sum_{j=1}^{N}\big(h(\theta_j - \theta_M) - h(\theta_j - \theta_m)\big)\\
&\leq \frac{K}{N}\,\frac{h(D(\Theta_0))}{D(\Theta_0)}\sum_{j=1}^{N}\big(-(\theta_M - \theta_j)^{1-2\alpha} - (\theta_j - \theta_m)^{1-2\alpha}\big)\\
&\leq -\frac{K}{N}\,\frac{h(D(\Theta_0))}{D(\Theta_0)}\sum_{j=1}^{N}\big((\theta_M - \theta_j) + (\theta_j - \theta_m)\big)^{1-2\alpha}\\
&= -\frac{Kh(D(\Theta_0))}{D(\Theta_0)}D(\Theta)^{1-2\alpha}
\end{aligned}$$

for every $t \geq 0$. In the last inequality we have used that $1 - 2\alpha \in (0, 1)$, and consequently

$$(a + b)^{1-2\alpha} \leq a^{1-2\alpha} + b^{1-2\alpha}$$

for any $a, b \geq 0$. Integrating the above differential inequality implies

$$D(\Theta(t)) \leq \left(D(\Theta_0)^{2\alpha} - 2\alpha K\frac{h(D(\Theta_0))}{D(\Theta_0)}t\right)^{\frac{1}{2\alpha}}$$

for all $t \geq 0$. This implies the convergence to zero at a finite time not larger than T_c. ■

We now consider the system for non-identical oscillators. The next proposition yields the structure of a phase-locked state of (3.2) for non-identical oscillators with mutually distinct natural frequencies in the subcritical regime.

Proposition 5.3. *Let* $\alpha \in (0, 1/2)$ *and* $\bar{\Theta} = (\bar{\theta}_1, \ldots, \bar{\theta}_N)$ *be an equilibrium of the system* (3.2) *such that* $\max_{i,j} |\bar{\theta}_i - \bar{\theta}_j| < \tilde{\theta}$ *where* $\tilde{\theta} \in (0, \pi/2)$ *is the solution to*

$$\tilde{\theta} = 2\alpha \tan \tilde{\theta}.$$

Assume the natural frequencies satisfy $\Omega_1 < \cdots < \Omega_N$*. Then the phase-locked state* $\bar{\Theta}$ *satisfies* $\bar{\theta}_1 < \cdots < \bar{\theta}_N$*.*

Proof. First, we show that the equilibria $\bar{\theta}_i$ are mutually distinct, i.e., $\bar{\theta}_i \neq \bar{\theta}_j$ for $i \neq j$. Since $\bar{\Theta}$ is an equilibrium, it satisfies

$$\Omega_i + \frac{K}{N}\sum_{k\neq i} h(\bar{\theta}_k - \bar{\theta}_i) = 0 \tag{5.8}$$

for every $i = 1, \ldots, N$. If there existed two oscillators having the same equilibria $\bar{\theta}_i = \bar{\theta}_j$, then we would have

$$\frac{K}{N}\sum_{k\neq i} h(\bar{\theta}_k - \bar{\theta}_i) = \frac{K}{N}\sum_{k\neq j} h(\bar{\theta}_k - \bar{\theta}_j),$$

which contradicts $\Omega_i \neq \Omega_j$.

We now show the ordering property. From (5.8), we have

$$\begin{aligned}
\Omega_{i+1} - \Omega_i &= -\frac{K}{N}\sum_{j\neq i+1} h(\bar{\theta}_j - \bar{\theta}_{i+1}) + \frac{K}{N}\sum_{j\neq i} h(\bar{\theta}_j - \bar{\theta}_i) \\
&= \frac{K}{N}\sum_{j\neq i,i+1} \big(h(\bar{\theta}_{i+1} - \bar{\theta}_j) - h(\bar{\theta}_i - \bar{\theta}_j)\big) - \frac{K}{N}\big(h(\bar{\theta}_i - \bar{\theta}_{i+1}) - h(\bar{\theta}_{i+1} - \bar{\theta}_i)\big) \\
&= \frac{K}{N}\sum_{j\neq i,i+1} \big(h(\bar{\theta}_{i+1} - \bar{\theta}_j) - h(\bar{\theta}_i - \bar{\theta}_j)\big) + \frac{2K}{N} h(\bar{\theta}_{i+1} - \bar{\theta}_i) \\
&= \frac{K}{N}\sum_{j\neq i,i+1} c_j (\bar{\theta}_{i+1} - \bar{\theta}_i) + \frac{2K}{N} h(\bar{\theta}_{i+1} - \bar{\theta}_i),
\end{aligned}$$

where

$$c_j := \frac{h(\bar{\theta}_{i+1} - \bar{\theta}_j) - h(\bar{\theta}_i - \bar{\theta}_j)}{\bar{\theta}_{i+1} - \bar{\theta}_i}.$$

They are properly defined because all the equilibria are mutually distinct, and they are positive because h is strictly increasing in $(-\tilde{\theta}, \tilde{\theta})$. Thus, the order $\Omega_{i+1} > \Omega_i$ yields the order of equilibria $\bar{\theta}_{i+1} > \bar{\theta}_i$. ∎

In the subcritical case, we can attain the uniform boundedness of phase differences under sufficiently large coupling strength.

Lemma 5.1. *Let Θ be the solution to (3.2) for $\alpha \in (0, 1/2)$ and non-identical oscillators with initial data Θ_0 satisfying $D(\Theta_0) < D^\infty < \tilde{\theta}$. If*

$$K > \frac{D(\dot{\Theta}_0)}{h'(D^\infty)(D^\infty - D(\Theta_0))},$$

then the phase diameter $D(\Theta)$ is uniformly bounded by D^∞:

$$D(\Theta(t)) < D^\infty \quad \text{for } t \geq 0.$$

Proof. Suppose there exists a finite time $t^* > 0$ such that

$$t^* := \sup\{t : D(\Theta(s)) < D^\infty \text{ for } 0 \le s \le t\} \quad \text{and} \quad D(\Theta(t^*)) = D^\infty.$$

Let indices F and S be such that

$$\dot\theta_F := \max\{\dot\theta_1, \dots, \dot\theta_N\} \quad \text{and} \quad \dot\theta_S := \min\{\dot\theta_1, \dots, \dot\theta_N\},$$

for each time t. We define the frequency difference so that

$$D(\dot\Theta(t)) := \dot\theta_F - \dot\theta_S.$$

We note that

$$D(\dot\Theta(t)) - D(\dot\Theta_0) = \int_0^t \frac{d}{dt} D(\dot\Theta(s))\, ds. \tag{5.9}$$

By taking the time derivative of $D(\dot\Theta)$, we obtain

$$\frac{d}{dt} D(\dot\Theta) = \frac{K}{N} \sum_{j=1}^N \big(h'(\theta_j - \theta_F)(\dot\theta_j - \dot\theta_F) - h'(\theta_j - \theta_S)(\dot\theta_j - \dot\theta_m)\big).$$

As long as $D(\Theta) < D^\infty$, we have $h'(\theta_j - \theta_i) \ge h'(D^\infty) > 0$. Thus, we get

$$\frac{d}{dt} D(\dot\Theta) \le \frac{K}{N} \sum_{j=1}^N h'(D^\infty)\big((\dot\theta_j - \dot\theta_F) - (\dot\theta_j - \dot\theta_S)\big) = -Kh'(D^\infty) D(\dot\Theta). \tag{5.10}$$

We combine (5.9) and (5.10) to obtain

$$D(\dot\Theta(t)) \le D(\dot\Theta_0) - Kh'(D^\infty) \int_0^t D(\dot\Theta(s))\, ds. \tag{5.11}$$

Setting $y(s) := \int_0^t D(\dot\Theta(s))\, ds$, we can rewrite (5.11) as

$$y'(t) \le y'(0) - Kh'(D^\infty) y(t).$$

Hence,

$$y(t) \le \frac{y'(0)}{Kh'(D^\infty)}(1 - e^{-Kh'(D^\infty)t}) \le \frac{y'(0)}{Kh'(D^\infty)}.$$

Since $D(\Theta(t^*)) = D^\infty$ and $K > \frac{D(\dot\Theta_0)}{h'(D^\infty)(D^\infty - D(\Theta_0))}$, we get

$$\begin{aligned} D^\infty &= D(\Theta_0) + \int_0^{t^*} \frac{d}{ds} D(\Theta(s))\, ds \le D(\Theta_0) + \int_0^{t^*} D(\dot\Theta(s))\, ds \\ &\le D(\Theta_0) + \frac{D(\dot\Theta_0)}{Kh'(D^\infty)} < D^\infty, \end{aligned}$$

which is a contradiction. Thus, we have the desired uniform bound $D(\Theta(t)) < D^\infty$ for $t \ge 0$. ∎

Remark 5.2. In the preceding proof, the solution $\Theta = \Theta(t)$ is C^1 but not necessarily C^2 because of the essential discontinuity of h'. Hence, one cannot directly argue with two time derivatives in the computation of $\frac{d}{dt}D(\dot{\Theta})$. However, the preceding arguments can be made rigorous because the C^1 solution of (3.2) is a piecewise $W^{2,1}$ solution of the augmented model (2.4)–(4.23) as discussed in Remark 4.1. Another possible approach is to directly show the Grönwall inequality (5.11) in integral form.

In the following result, we show the collision avoidance when the oscillators are initially well-ordered.

Lemma 5.2. *Let Θ be the solution to (3.2) with $\alpha \in (0, 1/2)$ and with initial datum Θ_0 satisfying $D(\Theta_0) < D^\infty < \tilde{\theta}$. Assume the natural frequencies and the initial configuration satisfy $\Omega_1 < \cdots < \Omega_N$ and $\theta_{1,0} < \cdots < \theta_{N,0}$. Assume that*

$$K > \frac{D(\dot{\Theta}_0)}{h'(D^\infty)(D^\infty - D(\Theta_0))}.$$

Then there is no collision between oscillators, i.e.,

$$\theta_i(t) \neq \theta_j(t) \quad \text{for } i \neq j,\ t > 0.$$

Proof. From Lemma 5.1, we have a uniform bound $D(\Theta(t)) < D^\infty$ for $t \geq 0$. Let ℓ be an index such that

$$\theta_{\ell+1}(t) - \theta_\ell(t) = \min_{j=1,\dots,N-1} (\theta_{j+1}(t) - \theta_j(t))$$

for each $t \geq 0$. For notational simplicity, we set

$$\Delta := \theta_{\ell+1} - \theta_\ell.$$

Then

$$\begin{aligned} \dot{\Delta} &= \Omega_{\ell+1} - \Omega_\ell + \frac{K}{N}\sum_{j=1}^N \big(h(\theta_j - \theta_{\ell+1}) - h(\theta_j - \theta_\ell)\big) \\ &\geq \Omega_\delta + \frac{K}{N}\sum_{j=1}^N \big(h(\theta_j - \theta_{\ell+1}) - h(\theta_j - \theta_\ell)\big), \end{aligned} \tag{5.12}$$

where $\Omega_\delta := \min_{j=1,\dots,N-1}(\Omega_{j+1} - \Omega_j) > 0$. We define

$$\mathcal{S}_1(\ell) := \{j : j < \ell\} \quad \text{and} \quad \mathcal{S}_2(\ell) := \{j : j > \ell + 1\}.$$

Note that $h(\theta)$ is convex increasing for $\theta \in (-\tilde{\theta}, 0)$ and is concave increasing for $\theta \in (0, \tilde{\theta})$. Thus, we have

$$\begin{aligned} &0 < h'(b) \leq \frac{h(b) - h(a)}{b - a} \leq h'(a) \quad \text{for } 0 \leq a < b \leq \tilde{\theta}, \\ &0 < h'(c) \leq \frac{h(d) - h(c)}{d - c} \leq h'(d) \quad \text{for } -\tilde{\theta} \leq c < d \leq 0. \end{aligned} \tag{5.13}$$

From (5.12) and (5.13), we obtain

$$
\begin{aligned}
\dot{\Delta} &\geq \Omega_\delta + \frac{K}{N}\sum_{j\in\mathcal{S}_1(\ell)}\big(h(\theta_j-\theta_{\ell+1})-h(\theta_j-\theta_\ell)\big)+\frac{K}{N}h(\theta_\ell-\theta_{\ell+1})\\
&\quad -\frac{K}{N}h(\theta_{\ell+1}-\theta_\ell)+\frac{K}{N}\sum_{j\in\mathcal{S}_2(\ell)}\big(h(\theta_j-\theta_{\ell+1})-h(\theta_j-\theta_\ell)\big)\\
&\geq \Omega_\delta-\frac{K}{N}\sum_{j\in\mathcal{S}_1(\ell)}h'(\theta_j-\theta_\ell)\Delta-\frac{K}{N}\sum_{j\in\mathcal{S}_2(\ell)}h'(\theta_j-\theta_{\ell+1})\Delta-\frac{2K}{N}h(\Delta)\\
&\geq \Omega_\delta-\frac{K|\mathcal{S}_1(\ell)|}{N}h'(\Delta)\Delta-\frac{K|\mathcal{S}_2(\ell)|}{N}h'(\Delta)\Delta-\frac{2K}{N}h(\Delta)\\
&\geq \Omega_\delta-Kh'(\Delta)\Delta-\frac{2K}{N}h(\Delta)\geq\Omega_\delta-C\Delta^\gamma=:q(\Delta),
\end{aligned}
$$

where we have used $h(\theta)\leq C_1\theta^\gamma$ and $h'(\theta)\theta\leq C_2\theta^\gamma$ for $\theta\geq 0$ and $0<\gamma<1-2\alpha$. Since $\lim_{\theta\to 0+}q(\theta)=\Omega_\delta>0$ and $q(\theta)$ is continuous for $\theta>0$, there exists a positive $\varepsilon>0$ such that $q(\theta)>0$ for $\theta\in(0,\varepsilon)$. Hence, the distance Δ has a positive lower bound. ■

In what follows, we study the stability of the phase-locked state for the system of non-identical oscillators. We use the center manifold theorem to investigate the stability of the linearized system.

Lemma 5.3 (Center Manifold Theorem [6]). *Consider the system*

$$
\dot{x}=Ax+f_A(x,y),\qquad \dot{y}=By+f_B(x,y) \tag{5.14}
$$

where $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$ and A and B are constant matrices such that all the eigenvalues of A have zero real parts while all the eigenvalues of B have negative real parts. Assume that the functions f_A and f_B are C^2 with $f_A(0,0)=0$, $\nabla f_A(0,0)=0$, $f_B(0,0)=0$, $\nabla f_B(0,0)=0$. Then:

(1) *There exists a center manifold for (5.14), $y=\phi(x)$, $|x|<\delta$, where $\phi=\phi(x)$ is C^2. The flow on the center manifold is governed by the n-dimensional system*

$$
\dot{u}=Au+f_A(u,\phi(u)). \tag{5.15}
$$

(2) *Assume the zero solution of (5.15) is stable (respectively asymptotically stable/unstable). Then so is the zero solution of (5.14).*

Theorem 5.2. *Let $\bar{\Theta}:=(\bar{\theta}_1,\ldots,\bar{\theta}_N)\notin\mathcal{C}$ be a collisionless equilibrium of (3.2).*

(1) *If $\alpha\geq 1/2$, then the phase-locked state $\bar{\Theta}$ is unstable.*

(2) *If $\alpha<1/2$, then the phase-locked state $\bar{\Theta}$ is stable.*

Proof. (1) In the critical and supercritical regimes $\alpha\in[1/2,1)$, we first linearize the system (3.2):

$$
\dot{\Theta}=A(\Theta-\bar{\Theta})+R(\bar{\Theta}), \tag{5.16}
$$

where the elements of the matrix $A = (a_{ij})_{1\le i,j\le N}$ are determined by

$$\begin{aligned} a_{ij} &= \frac{\cos(\bar\theta_j - \bar\theta_i)}{|\bar\theta_j - \bar\theta_i|_o^{2\alpha}} - 2\alpha \frac{\sin|\bar\theta_j - \bar\theta_i|_o}{|\bar\theta_j - \bar\theta_i|_o^{2\alpha+1}} \quad \text{for } i \neq j, \\ a_{ii} &= -\sum_{j\neq i} a_{ij}. \end{aligned} \tag{5.17}$$

If $\alpha \ge 1/2$, we find $a_{ij} < 0$ for $i \neq j$, and hence $a_{ii} > 0$ for $i = 1, \ldots, N$. Hence A is a Laplacian type matrix with all eigenvalues non-negative. Since A represents all-to-all connected network, there exists a zero eigenvalue of multiplicity 1 and all the other eigenvalues are positive, which implies the unstability of the equilibrium.

(2) In the subcritical regime $\alpha \in (0, 1/2)$, since the equilibrium satisfies

$$\max_{i,j} |\bar\theta_i - \bar\theta_j| < \tilde\theta$$

and $\bar\theta_i \neq \bar\theta_j$ for $i \neq j$, the elements of the matrix have $a_{ij} > 0$ for $i \neq j$ and $a_{ii} < 0$ for $i = 1, \ldots, N$. By a similar argument to the above, the eigenvalues of A are non-positive and there is a zero eigenvalue with multiplicity 1. Let $\lambda_1 = 0$ and $\lambda_2, \ldots, \lambda_N < 0$ be the eigenvalues of A and let $v_1, \ldots, v_N$ be the corresponding left eigenvectors such that

$$v_i A = \lambda_i v_i \quad \text{for } i = 1, \ldots, N.$$

We note that $v_1 = (1, \ldots, 1)$. Consider the matrices

$$P^{-1} := \begin{pmatrix} 1 & \cdots & 1 \\ & v_2 & \\ & \vdots & \\ & v_N & \end{pmatrix} \quad \text{and} \quad D := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}.$$

Then P diagonalizes the matrix A:

$$P^{-1} A P = D. \tag{5.18}$$

We change variables from $\Theta = (\theta_1, \ldots, \theta_N)$ to $X = (x_1, \ldots, x_N)$ such that

$$X := P^{-1}\Theta. \tag{5.19}$$

Then the system (5.16) can be transformed into

$$\dot X = D(X - \bar X) + \tilde R(X). \tag{5.20}$$

Let $\hat x_1 := (x_2, \ldots, x_N)$ and

$$\hat D := \begin{pmatrix} \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N \end{pmatrix}.$$

Then we can rewrite the system (5.20) as

$$\begin{pmatrix} x_1 \\ \hat{x}_1 \end{pmatrix}' = \begin{pmatrix} 0 & 0 \\ 0 & \hat{D} \end{pmatrix} \begin{pmatrix} x_1 - \bar{x}_1 \\ \hat{x}_1 - \hat{\bar{x}}_1 \end{pmatrix} + \begin{pmatrix} \tilde{R}_1(x_1, \hat{x}_1) \\ \hat{\tilde{R}}_1(x_1, \hat{x}_1) \end{pmatrix}. \tag{5.21}$$

Consider the center manifold in Lemma 5.3, which can be written as

$$W_c := \{(x, y) \in \mathbb{R} \times \mathbb{R}^{N-1} : y = c(x) \text{ for } |x| < \varepsilon,\ \phi(\bar{x}_1) = 0,\ D\phi(\bar{x}_1) = 0\},$$

and consider the equation

$$\dot{x}_1 = \tilde{R}_1(x_1, \phi(x_1)). \tag{5.22}$$

By the Center Manifold Theorem, the stability of (5.22) implies the stability of (5.21). Since (5.19) yields $x_1 = \theta_1 + \cdots + \theta_N$, we have

$$\dot{x}_1 = \sum_{i=1}^{N} \dot{\theta}_i = \sum_{i=1}^{N} \Omega_i = 0.$$

Thus, $\tilde{R}_1 \equiv 0$ and the dynamics of (5.22) is stable. Therefore, the phase-locked state $\bar{\Theta}$ is stable for $\alpha < 1/2$. ■

Finally, we are ready to show the emergence of a phase-locked state for non-identical oscillators.

Theorem 5.3. *Let Θ be a solution to (3.2) for $\alpha \in (0, 1/2)$ with initial datum Θ_0 satisfying*

$$D(\Theta_0) < D^\infty < \tilde{\theta}.$$

If

$$K > \frac{D(\dot{\Theta}_0)}{h'(D^\infty)(D^\infty - D(\Theta_0))},$$

then we can show the emergence of a phase-locked state. Moreover, if the oscillators have distinct natural frequencies, i.e., $\Omega_i \neq \Omega_j$ for $i \neq j$, then synchronization occurs asymptotically.

Proof. By applying Grönwall's lemma to (5.10), we have exponential decay of the upper estimate on the frequency diameter:

$$D(\dot{\Theta}(t)) \leq D(\dot{\Theta}_0)e^{-Kh'(D^\infty)t}.$$

This exponential decay implies the emergence of a phase-locked state.

Assume the oscillators have mutually distinct natural frequencies. Since Proposition 5.3 gives the structure of a phase-locked state, the oscillators become sorted in increasing order of natural frequencies in finite time. After this time, by Lemma 5.2, we have a positive lower bound $\varepsilon_\Delta > 0$ for the distances between the oscillators. Consequently,

$$\frac{d}{dt}D(\dot{\Theta}) = \frac{K}{N}\sum_{j=1}^{N}\big(h'(\theta_j-\theta_F)(\dot\theta_j-\dot\theta_F) - h'(\theta_j-\theta_S)(\dot\theta_j-\dot\theta_S)\big)$$
$$\geq \frac{K}{N}\sum_{j=1}^{N}\big(h'(\varepsilon_\Delta)(\dot\theta_j-\dot\theta_F) - h'(\varepsilon_\Delta)(\dot\theta_j-\dot\theta_S)\big) = -Kh'(\varepsilon_\Delta)D(\dot\Theta).$$

By Grönwall's lemma, we have a lower estimate on the frequency diameter:

$$D(\dot\Theta(t)) \geq D(\dot\Theta_0)e^{-Kh'(\varepsilon_\Delta)t}.$$ ∎

Let us now get some insight into the behavior of the Filippov solutions to (3.2) (see Theorems 3.3 and 3.5) in the most singular cases $\alpha = 1/2$ and $\alpha \in (1/2, 1)$. Looking at Remark 5.1 for the dynamics of two oscillators, we expect global synchronization in finite time for N oscillators. Specifically, in the supercritical case, the emerging global cluster is expected to stay stuck independently of the chosen natural frequencies. In the critical case, the sticking conditions (3.12) are required for the cluster to remain stuck. To begin, let us prove finite-time global phase synchronization of identical oscillators in the critical and supercritical cases. To that end, we need the following technical results.

Lemma 5.4. *Let* $\alpha \in [1/2, 1)$, $\beta \in (0, 2\alpha]$ *and* $\theta_* \in (0, \pi)$ *and define*

$$c(\alpha,\beta) = \left(\frac{2\alpha-\beta}{\beta}\right)^{1/2}.$$

Then

$$h_\varepsilon(\theta) \geq \frac{h_\varepsilon(\theta_*)}{\theta_*^\beta}\theta^\beta, \qquad \forall \theta \in [c(\alpha,\beta)\varepsilon, \theta_*],$$

for every $0 < \varepsilon < c(\alpha,\beta)^{-1}\theta_*$.

Proof. Define

$$g_\varepsilon(\theta) := \frac{h_\varepsilon(\theta)}{\theta^\beta} = \frac{\frac{\sin\theta}{(\varepsilon^2+\theta^2)^\alpha}}{\theta^\beta}, \qquad \theta \in (0,\pi).$$

We claim that g_ε is non-increasing in $(c(\alpha,\beta)\varepsilon, \pi)$ for every $\varepsilon \in (0, c(\alpha,\beta)^{-1}\theta_*)$; then the result is apparent. Indeed, taking derivatives we have

$$g_\varepsilon'(\theta) = \frac{1}{\theta^{\beta+1}(\varepsilon^2+\theta^2)^\alpha}\left[\theta\cos\theta - \left(2\alpha\frac{\theta^2}{\varepsilon^2+\theta^2}+\beta\right)\sin\theta\right]$$
$$= \frac{1}{\theta^{\beta+1}(\varepsilon^2+\theta^2)^\alpha}\left[\theta\cos\theta - \left(2\alpha + \frac{\beta\theta^2-(2\alpha-\beta)\varepsilon^2}{\theta^2+\varepsilon^2}\right)\sin\theta\right]$$

for every $\theta \in (0, \pi/2)$. Notice that $2\alpha \geq 1$ and $\beta \leq 2\alpha$. Then, by virtue of the definition of $c(\alpha, \beta)$ one checks that

$$\theta \cos\theta - \left(2\alpha + \frac{\beta\theta^2 - (2\alpha-\beta)\varepsilon^2}{\theta^2+\varepsilon^2}\right)\sin\theta \leq \theta\cos\theta - \sin\theta \leq 0$$

for every $\theta \in (c(\alpha,\beta)\varepsilon, \pi)$ and the monotonicity of g_ε becomes clear. ■

Lemma 5.5. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to (3.2) with $\alpha \in [1/2, 1)$ for identical oscillators, $\Omega_i = 0$ for $i = 1, \ldots, N$, obtained in Theorems 4.3 and 4.4 as singular limits. Suppose the initial configuration Θ_0 is confined to a half-circle, i.e., $0 < D(\Theta_0) < \pi$. Then*

$$D(\Theta(t)) \leq D(\Theta_0) e^{-K\frac{h(D(\Theta_0))}{D(\Theta_0)}t} \qquad \text{if } \alpha = 1/2,$$

$$D(\Theta(t)) \leq \left(D(\Theta_0)^{1-2\alpha} + (2\alpha-1)2^{2\alpha-1}K\frac{h(D(\Theta_0))}{D(\Theta_0)^{2\alpha}}t\right)^{-\frac{1}{2\alpha-1}} \qquad \text{if } \alpha \in (1/2, 1),$$

for every $t \geq 0$.

Proof. The main idea is to handle the approximate sequence $\{\Theta^\varepsilon\}_{\varepsilon>0}$ obtained as solutions to the regularized system (4.8) and to take limits as $\varepsilon \to 0$ in the phase diameter estimates. First, notice that by the assumed initial condition on the diameter one has

$$\frac{d}{dt}D(\Theta^\varepsilon) \leq 0 \quad \text{and} \quad D(\Theta^\varepsilon(t)) \leq D(\Theta_0) < \pi \quad \text{for } t > 0.$$

Indeed, note that we can obtain an explicit decay rate for the diameter by mimicking the ideas in Theorem 5.1. Namely, choosing $\theta_* = D(\Theta_0)$ and $\beta = 2\alpha$ in Lemma 5.4, we notice that $c(\alpha, \beta) = 0$. Consequently, the lower bound of the kernel h_ε is valid in the whole interval $[0, D(\Theta_0)]$. Hence,

$$\begin{aligned}
\frac{d}{dt}D(\Theta^\varepsilon) &= \frac{K}{N}\sum_{j=1}^N \left(h_\varepsilon(\theta_j^\varepsilon - \theta_M^\varepsilon) - h_\varepsilon(\theta_j^\varepsilon - \theta_m^\varepsilon)\right) \\
&\quad - \frac{K}{N}\sum_{j=1}^N \left(h_\varepsilon(\theta_M^\varepsilon - \theta_j^\varepsilon) + h_\varepsilon(\theta_j^\varepsilon - \theta_m^\varepsilon)\right) \\
&\leq -\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^{2\alpha}}\sum_{j=1}^N \left((\theta_M^\varepsilon - \theta_j^\varepsilon)^{2\alpha} + (\theta_j^\varepsilon - \theta_m^\varepsilon)^{2\alpha}\right) \\
&\leq -\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^{2\alpha}}2^{2\alpha-1}\sum_{j=1}^N \left((\theta_M^\varepsilon - \theta_j^\varepsilon) + (\theta_j^\varepsilon - \theta_m^\varepsilon)\right)^{2\alpha} \\
&= -K\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^{2\alpha}}2^{2\alpha-1}D(\Theta)^{2\alpha}.
\end{aligned}$$

Let us integrate the above differential inequality. We need to distinguish the cases $\alpha = 1/2$ and $\alpha \in (1/2, 1)$:

$$D(\Theta^\varepsilon(t)) \leq D(\Theta_0) e^{-K\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)}t} \qquad \text{if } \alpha = 1/2,$$

$$D(\Theta^\varepsilon(t)) \leq \left(D(\Theta_0)^{1-2\alpha} + (2\alpha - 1)2^{2\alpha-1} K \frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^{2\alpha}} t \right)^{-\frac{1}{2\alpha-1}} \qquad \text{if } \alpha \in (1/2, 1),$$

for every $t \geq 0$. Recall that by Lemmas 4.3 and 4.8,

$$\Theta^\varepsilon \overset{*}{\rightharpoonup} \Theta \quad \text{in } H^1((0,T), \mathbb{R}^N).$$

In particular,

$$\Theta^\varepsilon \to \Theta \quad \text{in } C([0,T], \mathbb{R}^N).$$

Finally, we can take the limit as $\varepsilon \to 0$ in the above estimates to obtain the desired result. ∎

Under the assumptions of Lemma 5.5 one obtains exponential decay of the diameter in the critical case and algebraic decay in the supercritical regime. However, finite-time global synchronization is expected. This is the content of the following result.

Theorem 5.4. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to (3.2) with $\alpha \in [1/2, 1)$ for identical oscillators, $\Omega_i = 0$ for $i = 1, \ldots, N$, obtained in Theorems 4.3 and 4.4 as singular limit of the regularized solutions Θ^ε to (4.8). Assume that the initial configuration Θ_0 is confined to a half-circle, i.e., $0 < D(\Theta_0) < \pi$. Then for every $\beta \in (0,1)$ there exist two oscillators that collide at some time not larger than*

$$T_c^1 = \frac{D(\Theta_0)}{(1-\beta)Kh(D(\Theta_0))}.$$

Proof. Assume the contrary. Then by continuity there exists $T > T_c^1$ such that there is no collision between oscillators along the time interval $[0, T]$. Again, by continuity there exists $\delta_T \in (0, D(\Theta_0))$ such that

$$|\theta_i(t) - \theta_j(t)| \geq \delta_T/2$$

for all $t \in [0, T]$ and all $i \neq j$. Since $\Theta^\varepsilon \to \Theta$ in $C([0,T], \mathbb{R}^N)$, there exists $\varepsilon_0 > 0$ such that

$$|\theta_i^\varepsilon(t) - \theta_j^\varepsilon(t)| \geq \delta_T$$

for all $t \in [0, T]$, all $i \neq j$ and all $\varepsilon \in (0, \varepsilon_0)$. Take $\theta_* = D(\Theta_0)$ and let

$$0 \leq \varepsilon_1 < \min\{\varepsilon_0, c(\alpha,\beta)\theta_*^{-1}, c(\alpha,\beta)^{-1}\delta_T\}.$$

Then it is clear that

$$|\theta_i^\varepsilon(t) - \theta_j^\varepsilon(t)| \in [c(\alpha,\beta)\varepsilon, \theta_*]$$

for every $t \in [0, T]$, any $\varepsilon \in (0, \varepsilon_1)$ and $i \neq j$. Applying Lemma 5.4 we obtain

$$\begin{aligned}
\frac{d}{dt} D(\Theta^\varepsilon) &= \frac{K}{N} \sum_{j=1}^{N} \big(h_\varepsilon(\theta_j^\varepsilon - \theta_M^\varepsilon) - h_\varepsilon(\theta_j^\varepsilon - \theta_m^\varepsilon)\big) \\
&\quad - \frac{K}{N} \sum_{j=1}^{N} \big(h_\varepsilon(\theta_M^\varepsilon - \theta_j^\varepsilon) + h_\varepsilon(\theta_j^\varepsilon - \theta_m^\varepsilon)\big) \\
&\leq -\frac{K}{N} \frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta} \sum_{j=1}^{N} \big((\theta_M^\varepsilon - \theta_j^\varepsilon)^\beta + (\theta_j^\varepsilon - \theta_m^\varepsilon)^\beta\big) \\
&\leq -\frac{K}{N} \frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta} \sum_{j=1}^{N} \big((\theta_M^\varepsilon - \theta_j^\varepsilon) + (\theta_j^\varepsilon - \theta_m^\varepsilon)\big)^\beta \\
&= -K \frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta} D(\Theta)^\beta
\end{aligned}$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_1)$. Integrating the differential inequality yields

$$D(\Theta(t)^\varepsilon) \leq \left(D(\Theta_0)^{1-\beta} - (1-\beta) K \frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta} t \right)^{\frac{1}{1-\beta}}$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_1)$. Letting $\varepsilon \to 0$ leads to

$$D(\Theta(t)) \leq \left(D(\Theta_0)^{1-\beta} - (1-\beta) K \frac{h(D(\Theta_0))}{D(\Theta_0)^\beta} t \right)^{\frac{1}{1-\beta}}$$

for each $t \in [0, T]$. However, this clearly contradicts $T > T_c^1$ due to the definition of T_c^1. ■

The above result leads to a time estimate for the first collision between a couple of oscillators in the critical and supercritical cases. However, this idea can be repeated and improved in the critical case to give a total collision in finite time. The key ideas will be the uniqueness in Theorem 3.3 or, more specifically, the characterization of sticking of oscillators in Corollary 3.1.

Theorem 5.5. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to (3.2) with $\alpha = 1/2$ for identical oscillators, $\Omega_i = 0$ for $i = 1, \ldots, N$. Assume that the initial configuration Θ_0 is confined to a half-circle, i.e., $0 < D(\Theta_0) < \pi$. Then there is complete phase synchronization in a finite time not larger than*

$$T_c = \frac{D(\Theta_0)}{K h(D(\Theta_0))}.$$

Proof. Assume the contrary, i.e., complete synchronization does not arise along $[0, T_c]$. By continuity there exists $T > T_c$ such that it does not happen along $[0, T]$ either. Recall that by Corollary 3.1, sticking of oscillators takes place in the critical case after any collision. Then the collision classes $\mathcal{C}_i(t)$ and sticking classes $S_i(t)$ in Subsection 2.3 agree with each other. Let us list the family of collision (or sticking) classes, i.e., the different clusters at time t:

$$\mathcal{E}(t) = \{\mathcal{C}_1(t), \ldots, \mathcal{C}_N(t)\} = \{E_1(t), \ldots, E_{\kappa(t)}(t)\}.$$

As a consequence of the assumed hypothesis $\kappa(t)$ is non-increasing with respect to t and bounded below by 2. Coming back to the initial configuration, we define i_M and i_m to be such that

$$\max_{1\leq j\leq N} \theta_{j,0} = \theta_{i_M,0} \quad \text{and} \quad \min_{1\leq j\leq N} \theta_{j,0} = \theta_{i_m,0}.$$

Since the regularized system (4.8) enjoys uniqueness in the full sense, the oscillators θ_i^ε and θ_j^ε cannot cross. Similarly, by Corollary 3.1, the oscillators θ_i and θ_j cannot cross either unless they keep stuck together after that time. In any case, it is clear that

$$\begin{aligned}
\max_{1\leq j\leq N} \theta_j(t) &= \theta_{i_M}(t), \qquad \min_{1\leq j\leq N} \theta_j(t) = \theta_{i_m}(t),\\
\max_{1\leq j\leq N} \theta_j^\varepsilon(t) &= \theta_{i_M}^\varepsilon(t), \qquad \min_{1\leq j\leq N} \theta_j^\varepsilon(t) = \theta_{i_m}^\varepsilon(t),
\end{aligned}$$

for every $t \geq 0$ and any $\varepsilon > 0$. Then

$$D(\Theta^\varepsilon(t)) = \theta_{i_M}^\varepsilon(t) - \theta_{i_m}^\varepsilon(t) \quad \text{and} \quad D(\Theta(t)) = \theta_{i_M}(t) - \theta_{i_m}(t),$$

for all $t \geq 0$ and $\varepsilon > 0$. All the above remarks ensure that for every $t \in [0, T]$,

$$\begin{aligned}
\theta_j(t) - \theta_{i_m}(t) &> 0 \quad \text{for all } j \in \mathcal{C}_{i_M}(t),\\
\theta_{i_M}(t) - \theta_j(t) &> 0 \quad \text{for all } j \in \mathcal{C}_{i_m}(t),\\
\theta_{i_M}(t) - \theta_j(t) &> 0 \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t),\\
\theta_j(t) - \theta_{i_m}(t) &> 0 \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t).
\end{aligned}$$

Since $\Theta^\varepsilon \to \Theta$ in $C([0, T], \mathbb{R}^N)$, by continuity we can find $\varepsilon_0, \delta_T > 0$ such that

$$\begin{aligned}
\theta_j^\varepsilon(t) - \theta_{i_m}^\varepsilon(t) &> \delta_T \quad \text{for all } j \in \mathcal{C}_{i_M}(t),\\
\theta_{i_M}^\varepsilon(t) - \theta_j^\varepsilon(t) &> \delta_T \quad \text{for all } j \in \mathcal{C}_{i_m}(t),\\
\theta_{i_M}^\varepsilon(t) - \theta_j^\varepsilon(t) &> \delta_T \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t),\\
\theta_j^\varepsilon(t) - \theta_{i_m}^\varepsilon(t) &> \delta_T \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t),
\end{aligned} \tag{5.23}$$

for all $t \in [0, T]$ and all $\varepsilon \in (0, \varepsilon_0)$. Take $\theta_* = D(\Theta_0)$, fix $\beta \in (0, 1)$ and consider

$$0 \leq \varepsilon_1 < \min\{\varepsilon_0, c(\alpha, \beta)\theta_*^{-1}, c(\alpha, \beta)^{-1}\delta_T\}.$$

Then it is clear that

$$\begin{aligned}
\theta_j^\varepsilon(t) - \theta_{i_m}^\varepsilon(t) &\in [c(\alpha, \beta)\varepsilon, \theta_*] \quad \text{for all } j \in \mathcal{C}_{i_M}(t),\\
\theta_{i_M}^\varepsilon(t) - \theta_j^\varepsilon(t) &\in [c(\alpha, \beta)\varepsilon, \theta_*] \quad \text{for all } j \in \mathcal{C}_{i_m}(t),\\
\theta_{i_M}^\varepsilon(t) - \theta_j^\varepsilon(t) &\in [c(\alpha, \beta)\varepsilon, \theta_*] \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t),\\
\theta_j^\varepsilon(t) - \theta_{i_m}^\varepsilon(t) &\in [c(\alpha, \beta)\varepsilon, \theta_*] \quad \text{for all } j \notin \mathcal{C}_{i_M} \cup \mathcal{C}_{i_m}(t),
\end{aligned} \tag{5.24}$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_1)$. Now, let us split as follows:

$$\begin{aligned}\frac{d}{dt}D(\Theta^\varepsilon) = &-\frac{K}{N}\sum_{j\in\mathcal{C}_{i_M}(t)}\left(h_\varepsilon(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)+h_\varepsilon(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})\right)\\ &-\frac{K}{N}\sum_{j\in\mathcal{C}_{i_m}(t)}\left(h_\varepsilon(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)+h_\varepsilon(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})\right)\\ &-\frac{K}{N}\sum_{j\notin\mathcal{C}_{i_M}(t)\cup\mathcal{C}_{i_m}(t)}\left(h_\varepsilon(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)+h_\varepsilon(\theta^\varepsilon_j-\theta^\varepsilon_m)\right)\\ \leq &-\frac{K}{N}\sum_{j\in\mathcal{C}_{i_M}(t)}h_\varepsilon(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})-\frac{K}{N}\sum_{j\in\mathcal{C}_{i_m}(t)}h_\varepsilon(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)\\ &-\frac{K}{N}\sum_{j\notin\mathcal{C}_{i_M}(t)\cup\mathcal{C}_{i_m}(t)}\left(h_\varepsilon(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)+h_\varepsilon(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})\right)\end{aligned}$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_1)$. By Lemma 5.4 and the estimates in (5.24), the above chain of inequalities implies

$$\begin{aligned}\frac{d}{dt}D(\Theta^\varepsilon) \leq &-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\sum_{j\in\mathcal{C}_{i_M}(t)}(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})^\beta\\ &-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\sum_{j\in\mathcal{C}_{i_m}(t)}(\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)^\beta\\ &-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\sum_{j\notin\mathcal{C}_{i_M}(t)\cup\mathcal{C}_{i_m}(t)}\left((\theta^\varepsilon_{i_M}-\theta^\varepsilon_j)^\beta+(\theta^\varepsilon_j-\theta^\varepsilon_{i_m})^\beta\right).\end{aligned}$$

Let us integrate this differential inequality to obtain

$$\begin{aligned}&D(\Theta^\varepsilon(t))\\ &\quad\leq D(\Theta^0)-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t\sum_{j\in\mathcal{C}_{i_M}(s)}(\theta^\varepsilon_j(s)-\theta^\varepsilon_{i_m}(s))^\beta\,ds\\ &\quad-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t\sum_{j\in\mathcal{C}_{i_m}(s)}(\theta^\varepsilon_{i_M}(s)-\theta^\varepsilon_j(s))^\beta\,ds\\ &\quad-\frac{K}{N}\frac{h_\varepsilon(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t\sum_{j\notin\mathcal{C}_{i_M}(s)\cup\mathcal{C}_{i_m}(s)}\left((\theta^\varepsilon_{i_M}(s)-\theta^\varepsilon_j(s))^\beta+(\theta^\varepsilon_j(s)-\theta^\varepsilon_{i_m}(s))^\beta\right)ds\end{aligned}$$

for all $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_1)$. Letting $\varepsilon \to 0$ we obtain

$$\begin{aligned} D(\Theta(t)) & \\ \leq D(\Theta^0) & - \frac{K}{N}\frac{h(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t \sum_{j\in\mathcal{C}_{i_M}(s)} (\theta_{i_M}(s)-\theta_{i_m}(s))^\beta \, ds \\ & - \frac{K}{N}\frac{h(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t \sum_{j\in\mathcal{C}_{i_m}(s)} (\theta_{i_M}(s)-\theta_{i_m}(s))^\beta \, ds \\ & - \frac{K}{N}\frac{h(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t \sum_{j\notin\mathcal{C}_{i_M}(s)\cup\mathcal{C}_{i_m}(s)} \left((\theta_{i_M}(s)-\theta_j(s))^\beta + (\theta_j(s)-\theta_{i_m}(s))^\beta\right) ds \end{aligned}$$

for every $t \in [0, T]$. To sum up, we obtain

$$D(\Theta(t)) \leq D(\Theta_0) - K\frac{h(D(\Theta_0))}{D(\Theta_0)^\beta}\int_0^t D(\theta(s))^\beta \, ds.$$

Hence,

$$D(\Theta(t)) \leq \left(D(\Theta_0)^{1-\beta} - (1-\beta)K\frac{h(D(\Theta_0))}{D(\Theta_0)^\beta} t\right)^{\frac{1}{1-\beta}}$$

for all $t \in [0, T]$. Then, it is clear that

$$T < \frac{D(\Theta_0)}{(1-\beta)Kh(D(\Theta_0))}$$

for all $\beta \in (0, 1)$. Letting $\beta \to 0$ shows that $T \leq T_c$, a contradiction. ■

Remark 5.3. Notice that Theorem 5.4 also works in the supercritical case. However, the same proof as in Theorem 5.5 is not valid to show finite-time complete phase synchronization of identical oscillators for $\alpha \in (1/2, 1)$. The reason is that at this point we cannot guarantee that the Filippov solution in Θ obtained as the singular limit of the regularized solutions Θ^ε to system (4.8) in Theorem 4.3 agrees with the solution obtained in Remark 3.6 via the "sticking after collision" continuation procedure of classical solutions. However, if the limiting Θ obtained in Theorem 4.3 has the "sticking after collision" property, we can mimic Theorem 5.5 to show that it exhibits complete phase synchronization no later than at time

$$T_c = \frac{D(\Theta_0)}{Kh(D(\Theta_0))}.$$

Appendix A. Regular interactions

In this appendix, we study the Kuramoto model with regular coupling weights:

$$\dot{\theta}_i = \Omega_i + \frac{K}{N}\sum_{j=1}^N \frac{\sigma^{2\alpha}}{(\sigma^2 + c|\theta_j - \theta_i|_o^2)^\alpha}\sin(\theta_j - \theta_i) \quad \text{for } i = 1, \ldots, N, \qquad \text{(A.1)}$$

where we denote $c \equiv c_{\alpha,\zeta} = 1 - \zeta^{-1/\alpha}$ for simplicity. Recall that this model comes from the choice (1.4) of Γ as the Hebbian plasticity function in (1.5). Since the right hand side of (A.1) is Lipschitz-continuous, the system (A.1) has a unique solution by Cauchy–Lipschitz theory in this case.

For positive σ, we get the following bounds for Γ:

$$\varepsilon_\sigma := \frac{\sigma^{2\alpha}}{(\sigma^2 + c\pi^2)^\alpha} \le \Gamma(\theta) \le 1, \qquad \Gamma(0) = \Gamma(2\pi) = 1.$$

Note that ε_σ converges to zero as $\sigma \to 0$. We will study the emergence of synchronization for identical and non-identical oscillators and we will use the idea of [15] to prove synchronization.

A.1. Identical oscillators

Consider the Kuramoto model (A.1) for identical oscillators which have the same natural frequency. Without loss of generality, we may assume $\Omega_i = 0$ for all $i = 1, \ldots, N$. The system (A.1) becomes

$$\dot\theta_i = \frac{K}{N}\sum_{j=1}^N \frac{\sigma^{2\alpha}}{(\sigma^2 + c|\theta_j - \theta_i|_o^2)^\alpha}\sin(\theta_j - \theta_i), \qquad i = 1, \ldots, N. \tag{A.2}$$

We can show complete phase synchronization asymptotically for (A.2) with a constraint on the initial configuration. Let us recall the notation $\theta_M(t)$ and $\theta_m(t)$ in (5.6) for the indices of largest and shortest phases and $D(\Theta)$ for the phase diameter defined in (5.7).

Theorem A.1. *Let $\Theta = (\theta_1, \ldots, \theta_N)$ be the solution to (A.2). Assume that the initial configuration is confined to a half-circle, i.e., $D(\Theta_0) < \pi$, and the coupling strength K is positive. Then the solution Θ shows complete phase synchronization asymptotically:*

$$D(\Theta_0)e^{-Kt} \le D(\Theta) \le D(\Theta_0)e^{-\frac{K\Gamma(D(\Theta_0))\sin D(\Theta_0)}{D(\Theta_0)}t}.$$

Proof. We consider the dynamics of the phase diameter,

$$\frac{d}{dt}D(\Theta) = \frac{K}{N}\sum_{j=1}^N \big(\Gamma(\theta_j - \theta_M)\sin(\theta_j - \theta_M) - \Gamma(\theta_j - \theta_m)\sin(\theta_j - \theta_m)\big). \tag{A.3}$$

Since $\sin(\theta_j - \theta_M) \le 0$ and $\sin(\theta_j - \theta_m) \ge 0$, as long as $D(\Theta) \le \pi$ we have

$$\frac{d}{dt}D(\Theta) \le 0 \quad \text{and} \quad D(\Theta(t)) \le D(\Theta_0) < \pi \quad \text{for } t > 0.$$

By this contraction of phase difference, we have

$$\sin(\theta_j - \theta_M) \le \frac{\sin D(\Theta_0)}{D(\Theta_0)}(\theta_j - \theta_M) \quad \text{and} \quad \sin(\theta_j - \theta_m) \ge \frac{\sin D(\Theta_0)}{D(\Theta_0)}(\theta_j - \theta_m). \tag{A.4}$$

On the other hand,

$$\varepsilon_\sigma < \Gamma(D(\Theta_0)) \le \Gamma(D(\Theta)) \le 1. \tag{A.5}$$

By applying (A.4) and (A.5) to (A.3), we obtain the differential inequality

$$\begin{aligned}
&\frac{d}{dt} D(\Theta) \\
&\quad \le \frac{K}{N} \sum_{j=1}^{N} \left(\Gamma(\theta_j - \theta_M) \frac{\sin D(\Theta_0)}{D(\Theta_0)} (\theta_j - \theta_M) - \Gamma(\theta_j - \theta_m) \frac{\sin D(\Theta_0)}{D(\Theta_0)} (\theta_j - \theta_m) \right) \\
&\quad = -\frac{K}{N} \frac{\sin D(\Theta_0)}{D(\Theta_0)} \sum_{j=1}^{N} \big(\Gamma(\theta_j - \theta_M)(\theta_M - \theta_j) + \Gamma(\theta_j - \theta_m)(\theta_j - \theta_m) \big) \\
&\quad \le -\frac{K}{N} \frac{\Gamma(D(\Theta_0)) \sin D(\Theta_0)}{D(\Theta_0)} \sum_{j=1}^{N} \big((\theta_M - \theta_j) + (\theta_j - \theta_m) \big) \\
&\quad = -\frac{K \Gamma(D(\Theta_0)) \sin D(\Theta_0)}{D(\Theta_0)} D(\Theta).
\end{aligned}$$

Grönwall's lemma yields the desired upper estimate. Similarly, from (A.5) and $\sin x \le x$ for $0 \le x \le \pi$, we have

$$\frac{d}{dt} D(\Theta) \ge \frac{K}{N} \sum_{j=1}^{N} \big((\theta_j - \theta_M) - (\theta_j - \theta_m) \big) = -K D(\Theta),$$

which gives the lower estimate. ∎

A.2. Non-identical oscillators

We assume that the diameter of the initial configuration is less than $D^\infty < \pi/2$. We first show that the phase diameter is less than D^∞ for all time $t \ge 0$ for sufficiently large coupling strength K. Recall that for $\theta \in (-\pi, \pi)$ the plasticity function is $\Gamma(\theta) = \frac{\sigma^{2\alpha}}{(\sigma^2 + c\theta^2)^\alpha}$. Thus,

$$\Gamma'(\theta) = -\frac{2\sigma^{2\alpha} \alpha c \theta}{(\sigma^2 + c\theta^2)^{\alpha+1}}, \qquad \Gamma''(\theta) = -\frac{2\sigma^{2\alpha} \alpha c [\sigma^2 - (2\alpha + 1) c\theta^2]}{(\sigma^2 + c\theta^2)^{\alpha+2}}.$$

If we set

$$\theta_\pm := \pm \frac{\sigma}{\sqrt{c(2\alpha + 1)}},$$

then Γ' attains its global extrema at those points:

$$\Gamma'(\theta_-) = \max_{\theta \in (-\pi, \pi)} \Gamma'(\theta) > 0 \quad \text{and} \quad \Gamma'(\theta_+) = \min_{\theta \in (-\pi, \pi)} \Gamma'(\theta) < 0.$$

Indeed,

$$\Gamma'(\theta_-) = -\Gamma'(\theta_+) = \frac{2\alpha\sqrt{c}}{\sigma\sqrt{2\alpha+1}\,(1+\frac{1}{2\alpha+1})^{\alpha+1}}.$$

We first show the boundedness of phase differences.

Lemma A.1. *Assume that $D(\Theta_0) < D^\infty$ for some small $D^\infty < \pi/2$, and that the coupling strength is sufficiently large so that*

$$-\Gamma'(\theta_+) < \frac{\Gamma(D^\infty)}{\tan D^\infty}$$

and

$$K > \frac{D(\dot\Theta_0)}{[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty](D^\infty - D(\Theta_0))}.$$

Then

$$D(\Theta(t)) < D^\infty \quad \textit{for } t \geq 0.$$

Proof. Assume that there exists a time t for which $D(\Theta(t)) \geq D^\infty$. Then by continuity

$$t^* := \sup\{t > 0 : D(\Theta(s)) < D^\infty \text{ for } 0 \leq s \leq t\}$$

is positive and finite and $D(\Theta(t^*)) = D^\infty$. Let indices F and S be such that

$$\dot\theta_F(t) := \max\{\dot\theta_1(t), \ldots, \dot\theta_N(t)\} \quad \text{and} \quad \dot\theta_S(t) := \min\{\dot\theta_1(t), \ldots, \dot\theta_N(t)\},$$

for each time t, and define the frequency diameter by

$$D(\dot\Theta(t)) := \dot\theta_F(t) - \dot\theta_S(t).$$

Then

$$D(\dot\Theta(t)) - D(\dot\Theta_0) = \int_0^t \frac{d}{ds} D(\dot\Theta(s))\, ds. \tag{A.6}$$

By taking the time derivative, we get

$$\begin{aligned}\frac{d}{dt}D(\dot\Theta) = \frac{K}{N}\sum_{j=1}^N \big[\Gamma'(\theta_j-\theta_F)\sin(\theta_j-\theta_F) + \Gamma(\theta_j-\theta_F)\cos(\theta_j-\theta_F)\big](\dot\theta_j - \dot\theta_F)\\ - \frac{K}{N}\sum_{j=1}^N \big[\Gamma'(\theta_j-\theta_S)\sin(\theta_j-\theta_S) + \Gamma(\theta_j-\theta_S)\cos(\theta_j-\theta_S)\big](\dot\theta_j - \dot\theta_S).\end{aligned} \tag{A.7}$$

Then we get the upper and lower bounds

$$\Gamma'(\theta_+)\sin D^\infty \leq \Gamma'(\theta_j-\theta_i)\sin(\theta_j-\theta_i) \leq 0, \tag{A.8}$$

$$\Gamma(D^\infty)\cos D^\infty \leq \Gamma(\theta_j-\theta_i)\cos(\theta_j-\theta_i) \leq 1. \tag{A.9}$$

By applying (A.8) and (A.9) in (A.7), we deduce

$$\begin{aligned}\frac{d}{dt}D(\dot\Theta) &\le \frac{K}{N}\sum_{j=1}^{N}\big[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty\big]\big((\dot\theta_j-\dot\theta_F)-(\dot\theta_j-\dot\theta_S)\big)\\ &= -K\big[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty\big](\dot\theta_F-\dot\theta_S)\\ &\le -K\underbrace{\big[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty\big]}_{>0}D(\dot\Theta)\end{aligned} \tag{A.10}$$

for every $t\in[0,t^*]$. Combining (A.6) and (A.10), we obtain

$$D(\dot\Theta(t)) \le D(\dot\Theta_0) - K\big[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty\big]\int_0^t D(\dot\Theta(s))\,ds \tag{A.11}$$

for every $t\in[0,t^*]$. Define $y(t) := \int_0^t D(\dot\Theta(s))\,ds$. Then (A.11) can be rewritten as

$$y'(t) \le y'(0) - Cy(t).$$

Here, $C := K[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty]$ and $t\in[0,t^*]$. Then

$$y(t) \le \frac{y'(0)}{C}(1-e^{-Ct}) \le \frac{y'(0)}{C}$$

for all $t\in[0,t^*]$. However, since $D(\Theta(t^*)) = D^\infty$, we get

$$\begin{aligned}D^\infty &= D(\Theta_0) + \int_0^{t^*}\frac{d}{ds}D(\Theta(s))\,ds \le D(\Theta_0) + \int_0^{t^*} D(\dot\Theta(s))\,ds\\ &\le D(\Theta_0) + y(t^*) \le D(\Theta_0) + \frac{y'(0)}{C} < D^\infty\end{aligned}$$

when

$$K > \frac{D(\dot\Theta_0)}{[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty](D^\infty - D(\Theta_0))},$$

a contradiction. Thus, $D(\Theta(t)) < D^\infty$ for all $t \ge 0$. ■

We are ready to prove frequency synchronization for non-identical oscillators.

Theorem A.2. *Assume that $D(\Theta_0) < D^\infty$ for some small $D^\infty < \pi/2$, and that the coupling strength is sufficiently large so that*

$$-\Gamma'(\theta_+) < \frac{\Gamma(D^\infty)}{\tan D^\infty} \quad \text{and} \quad K > \frac{D(\dot\Theta_0)}{[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty](D^\infty - D(\Theta_0))}.$$

Then we have complete frequency synchronization

$$D(\dot\Theta(0))e^{-Kt} \le D(\dot\Theta(t)) \le D(\dot\Theta(0))e^{-K[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty]t}.$$

Proof. From (A.7)–(A.10), we obtain

$$\frac{d}{dt}D(\dot{\Theta}) \leq -K\big[\Gamma'(\theta_+)\sin D^\infty + \Gamma(D^\infty)\cos D^\infty\big]D(\dot{\Theta}).$$

On the other hand, from (A.7)–(A.9), we have

$$\frac{d}{dt}D(\dot{\Theta}) \geq -KD(\dot{\Theta}).$$

By Grönwall's lemma, we achieve the exponential estimates for frequency synchronization. ∎

Since the decay rate of the asymptotic frequency synchronization is exponential, the solution Θ shows the emergence of a phase-locked state.

Appendix B. H-representation of Filippov set-valued maps

In this appendix, we prove the technical Lemmas 4.4 and 4.9. Recall that they were respectively applied in Propositions 4.1 and 4.2 in order to explicitly characterize some H-representation of the Filippov set-valued map in the supercritical and critical cases. We introduce some notation.

Definition B.1. For $n \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$ we define the linear operators

$$\begin{aligned} L_{ij} &: \mathrm{Skew}_n(\mathbb{R}) \to \mathbb{R}, \quad && Y \mapsto y_{ij}, \\ L_i &: \mathrm{Skew}_n(\mathbb{R}) \to \mathbb{R}, \quad && Y \mapsto \sum_{k=1}^n y_{ik}, \\ \mathcal{L} &: \mathrm{Skew}_n(\mathbb{R}) \to \mathbb{R}^n, \quad && Y \mapsto Y \cdot \mathbf{j}. \end{aligned}$$

By definition, we have

$$L_i = \sum_{k=1}^n L_{ik} \quad \text{and} \quad \mathcal{L} = (L_1, \ldots, L_n).$$

First, we give the simpler proof of Lemma 4.4:

Lemma B.1. *For any $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$, the following assertions are equivalent:*

(1) *There exists $Y \in \mathrm{Skew}_n(\mathbb{R})$ such that $x = Y \cdot \mathbf{j}$.*

(2) *$x \cdot \mathbf{j} = 0$, where $\mathbf{j}$ stands for the vector of ones.*

Proof. Define the linear operator

$$\mathcal{L} : \mathrm{Skew}_n(\mathbb{R}) \to \mathbb{R}^n, \quad Y \mapsto Y \cdot \mathbf{j}.$$

Then the assertion is equivalent to

$$\mathcal{L}(\mathrm{Skew}_n(\mathbb{R})) = \mathbf{j}^\perp. \tag{B.1}$$

On the one hand, it is clear that the inclusion $\subseteq$ in (B.1) holds by the properties of skew-symmetric matrices. On the other hand, define the matrices

$$E_{ij} := \frac{1}{2}(e_i \otimes e_j - e_j \otimes e_i) \tag{B.2}$$

for every $i \neq j$, where $\{e_i : i = 1, \ldots, N\}$ is the standard basis of $\mathbb{R}^n$ and $\otimes$ denotes the Kronecker product. Notice that

$$\mathcal{L}(E_{ij}) = \frac{1}{2}(e_i \otimes e_j - e_j \otimes e_i) \cdot \mathbf{j} = e_i - e_j.$$

Hence,

$$\{\mathcal{L}(E_{i,i+1}) : i = 1, \ldots, n-1\} = \{e_i - e_{i+1} : i = 1, \ldots, n-1\}$$

consists of $n - 1$ linearly independent vectors. Consequently, $\mathcal{L}$ has rank larger than or equal to $n - 1$. Since $\mathbf{j}^\perp$ has rank $n - 1$, the identity in (B.1) holds true. ∎

Now, we turn to the proof of Lemma 4.9. Our main tool will be the *Farkas alternative* from convex analysis, which we now recall.

Lemma B.2 (Farkas alternative). *Consider any finite-dimensional vector space V, a finite family of linear operators $T_1, \ldots, T_k : V \to \mathbb{R}$ and $b = (b_1, \ldots, b_k) \in \mathbb{R}^k$. Then exactly one of the following statements holds true:*

(1) *There exists $v \in V$ such that*

$$T_i(v) \leq b_i, \quad i = 1, \ldots, k.$$

(2) *There exists $q \in \mathbb{R}^n$ with $q_i \geq 0$ for all $i = 1, \ldots, k$ such that*

$$\sum_{i=1}^k q_i T_i = 0 \quad \text{and} \quad q \cdot b < 0.$$

This result has several equivalent representations in the literature and it is sometimes called the *Theorem of Alternatives*. One reference where we can find our version is [44, Lemma 2.54]. We are now ready to give the proof of Lemma 4.9.

Lemma B.3. *For any $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$, the following assertions are equivalent:*

(1) *There exists $Y \in \mathrm{Skew}_n([-1, 1])$ such that $x = Y \cdot \mathbf{j}$.*

(2) *There exists $Y \in \mathrm{Skew}_n(\mathbb{R})$ such that*

$$L_{ij}(Y) \leq 1, \quad L_i(Y) \leq x_i \quad \text{and} \quad -L_i(Y) \leq -x_i,$$

for all $i, j \in \{1, \ldots, n\}$.

(3) *We have*

$$\sum_{i,j=1}^{n} q_{ij} + \lambda_i x_i \geq 0 \tag{B.3}$$

for all $Q \in \mathcal{M}_n(\mathbb{R}_0^+)$ *and* $\lambda \in \mathbb{R}^n$ *such that* $q_{ij} + \lambda_i = q_{ji} + \lambda_j$.

(4) *We have*

$$\sum_{i=1}^{k} x_{\sigma_i} \in [-k(n-k), k(n-k)]$$

for every permutation σ *of* $\{1, \ldots, n\}$ *and any* $k \in \mathbb{N}$.

Proof. We split the proof into two parts. First, we establish the equivalence of the first three assertions. The main tool here is Lemma B.2. Secondly, we focus on the more convoluted equivalence between the first group of assertions and the last assertion.

• *Step 1: Equivalence of the first three assertions.* The first two assertions are equivalent by Definition B.1. In (2), we have a system of affine inequalities in the vector space $\mathrm{Skew}_n(\mathbb{R})$ of skew-symmetric matrices. Hence, by the Farkas alternative (see Lemma B.2), (2) amounts to saying that whenever q_{ij}, q_i^+, q_i^- are non-negative coefficients satisfying

$$\sum_{i,j=1}^{n} q_{ij} L_{ij} + \sum_{i=1}^{n} q_i^+ L_i - \sum_{i=1}^{n} q_i^- L_i \equiv 0 \quad \text{in } \mathrm{Skew}_n(\mathbb{R}),$$

then

$$\sum_{i,j=1}^{n} q_{ij} + \sum_{i=1}^{n} q_i^+ x_i - \sum_{i=1}^{n} q_i^- x_i \geq 0.$$

Defining $\lambda_i = q_i^+ - q_i^-$, we can simplify the equivalent assertion: for every $Q \in \mathcal{M}_n(\mathbb{R}_0^+)$ and $\lambda \in \mathbb{R}^n$ such that

$$\sum_{i,j=1}^{n} q_{ij} L_{ij} + \sum_{i=1}^{n} \lambda_i L_i \equiv 0 \quad \text{in } \mathrm{Skew}_n(\mathbb{R}), \tag{B.4}$$

we have

$$\sum_{i,j=1}^{n} q_{ij} + \sum_{i=1}^{n} \lambda_i x_i \geq 0.$$

Thus, the equivalence with the third assertion follows by evaluating the identity (B.4) on every matrix in the canonical basis of $\mathrm{Skew}_n(\mathbb{R})$, i.e.,

$$\{e_i \otimes e_j - e_j \otimes e_i : 1 \leq i < j \leq n\},$$

and noticing that we obtain the condition $q_{ij} + \lambda_i = q_{ji} + \lambda_j$ in the third assertion.

• *Step 2: Equivalence with the last assertion.* Assume that (1) is satisfied, i.e., $x = Y \cdot \mathbf{j}$ for some $Y \in \text{Skew}_n([-1, 1])$. Taking any permutation σ of $\{1, \ldots, n\}$ and any $1 \le k \le n$ we obtain

$$\sum_{i=1}^{k} x_{\sigma_i} = \sum_{i=1}^{k}\sum_{j=1}^{n} y_{\sigma_i\sigma_j} = \sum_{i=1}^{k}\sum_{j=1}^{k} y_{\sigma_i\sigma_j} + \sum_{i=1}^{k}\sum_{j=k+1}^{n} y_{\sigma_i\sigma_j}.$$

Since the first term is zero (by antisymmetry) and the second term consists of $n(n-k)$ terms with values in $[-1, 1]$, it follows that

$$\sum_{i=1}^{k} x_{\sigma_i} \in [-k(n-k), k(n-k)].$$

Conversely, assume that (4) is true and let us prove (B.3). Consider $Q \in \mathcal{M}_n(\mathbb{R}_0^+)$ and $\lambda \in \mathbb{R}^n$ such that

$$q_{ij} - q_{ji} = \lambda_j - \lambda_i. \tag{B.5}$$

Without loss of generality we will assume that $q_{ii} = 0$ for every $i = 1, \ldots, n$ (in the other case, (B.3) is even larger), and we split

$$I := \sum_{i\neq j} q_{ij} + \sum_{i=1}^{n} \lambda_i x_i =: I_1 + I_2.$$

Notice that

$$I_2 = \sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} (\lambda_i - \lambda_j) x_i + \lambda_j \sum_{i=1}^{n} x_i$$

for every $j = 1, \ldots, n$. Since the sum of all the x_i is zero by hypothesis, taking averages with respect to all the indices $j = 1, \ldots, n$ we obtain

$$I_2 = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} (\lambda_i - \lambda_j) x_i.$$

Finally, interchanging i and j and taking the average of the resulting two expressions we can equivalently write

$$I_2 = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n} (\lambda_i - \lambda_j)(x_i - x_j) = \frac{1}{n}\sum_{i<j} (\lambda_j - \lambda_i)(x_j - x_i).$$

Thus, substituting (B.5) into I_2 and putting it together with I_1 we can write

$$I = \sum_{i\neq j} q_{ij} + \frac{1}{2n}\sum_{i\neq j} (q_{ij} - q_{ji})(x_j - x_i) = \sum_{i\neq j} q_{ij}\left(1 + \frac{1}{n}(x_j - x_i)\right). \tag{B.6}$$

Let us consider a permutation σ of $\{1, \ldots, n\}$ that puts the coefficients λ_i in increasing order:

$$\lambda_{\sigma_1} \leq \cdots \leq \lambda_{\sigma_n}. \tag{B.7}$$

Then

$$\begin{aligned} I &= \sum_{i \neq j} q_{\sigma_i \sigma_j} \left(1 + \frac{1}{n}(x_{\sigma_j} - x_{\sigma_i})\right) \\ &= \sum_{i<j} (q_{\sigma_i \sigma_j} - q_{\sigma_j \sigma_i}) \left(1 + \frac{1}{n}(x_{\sigma_j} - x_{\sigma_i})\right) + 2 \sum_{i<j} q_{\sigma_j \sigma_i} \\ &=: I_3 + I_4. \end{aligned}$$

It is clear that I_4 is non-negative. We will show that so is I_3. By (B.5), it is easy to show that

$$q_{\sigma_i \sigma_j} - q_{\sigma_j} \sigma_i = \sum_{k=i}^{j-1} (q_{\sigma_k \sigma_{k+1}} - q_{\sigma_{k+1} \sigma_k})$$

for any $i < j$. Therefore,

$$\begin{aligned} I_3 &= \sum_{i<j} \sum_{k=i}^{j-1} (q_{\sigma_k \sigma_{k+1}} - q_{\sigma_{k+1} \sigma_k}) \left(1 + \frac{1}{n}(x_{\sigma_j} - x_{\sigma_i})\right) \\ &= \sum_{k=1}^{n-1} a_k (q_{\sigma_k \sigma_{k+1}} - q_{\sigma_{k+1} \sigma_k}) \\ &= \sum_{k=1}^{n-1} a_k (\lambda_{\sigma_{k+1}} - \lambda_{\sigma_k}), \end{aligned} \tag{B.8}$$

where in the last step we have used (B.5) again and the coefficients a_k are

$$\begin{aligned} a_k &:= \sum_{\substack{i \leq k \\ j \geq k+1}} \left(1 + \frac{1}{n}(x_{\sigma_j} - x_{\sigma_i})\right) \\ &= k(n-k) + \frac{k}{n} \sum_{j=k+1}^{n} x_{\sigma_j} - \frac{n-k}{n} \sum_{i=1}^{k} x_{\sigma_i}. \end{aligned}$$

Bearing in mind that the sum of all the x_i vanishes by hypothesis, we have

$$a_k = k(n-k) - \sum_{i=1}^{k} x_{\sigma_i}.$$

Thus, $a_k \geq 0$ by hypothesis. Since we have chosen σ so that (B.7) takes place, the result follows from the expression (B.8) for I_3. ∎

Appendix C. Characterizing the sticking conditions

Our purpose in this appendix is to give explicit conditions on the weights yielding necessary and sufficient conditions for sticking of particles (3.12) and (3.13) in Subsections 4.3 and 4.1 respectively. The first part is devoted to the supercritical case and the second part will focus on the critical case.

Apart from the linear operators in Definition B.1 we will need the following ones.

Definition C.1. For any $n \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$ we define the linear operator

$$T_{ij} : \text{Skew}_n(\mathbb{R}) \to \mathbb{R}, \qquad Y \mapsto \sum_{k=1}^{n} (y_{ik} - y_{jk}).$$

Notice that by definition we get the following relation to the operators in Definition B.1:

$$T_{ij} = \sum_{k=1}^{n} (L_{ik} - L_{jk}).$$

The next result yields a characterization of the sticking condition (3.13).

Lemma C.1. *For any $n \in \mathbb{N}$ and $M \in \text{Skew}_n(\mathbb{R})$, the following assertions are equivalent:*

(1) *There exists $Y \in \text{Skew}_n(\mathbb{R})$ such that*

$$M = Y \cdot \mathbf{J} + \mathbf{J} \cdot Y.$$

(2) *There exists $Y \in \text{Skew}_n(\mathbb{R})$ such that*

$$T_{ij}(Y) \leq m_{ij} \quad \text{and} \quad -T_{ij}(Y) \leq -m_{ij}.$$

(3) *We have*

$$m_{1i} + m_{ij} + m_{j1} = 0 \tag{C.1}$$

for all $2 \leq i < j \leq n$.

(4) *We have*

$$m_{ij} + m_{jk} + m_{ki} = 0 \tag{C.2}$$

for all $1 \leq i < j < k \leq n$.

Proof. First, it is clear that the first two assertions are equivalent. Second, let us briefly show that (C.1) and (C.2) are equivalent. On the one hand, it is clear that (C.1) is a particular case of (C.2). On the other hand, assume that (C.1) holds. Then in particular for $1 \leq i < j < k \leq n$ we have

$$\begin{aligned} m_{1i} + m_{ij} + m_{j1} &= 0, \\ m_{1j} + m_{jk} + m_{k1} &= 0, \\ m_{1k} + m_{ki} + m_{i1} &= 0. \end{aligned}$$

Taking the sum of these equations we obtain (C.2) by the skew-symmetry of M.

Hence, let us prove the equivalence between the second and third assertions. By Lemma B.2, (2) amounts to saying that whenever $\Lambda \in \mathcal{M}_n(\mathbb{R})$ satisfies

$$\sum_{i,j=1}^{n} \lambda_{ij} T_{ij} = 0,$$

then

$$\sum_{i,j=1}^{n} \lambda_{ij} m_{ij} \geq 0.$$

Evaluating on the basis $e_i \otimes e_j - e_j \otimes e_i$ we equivalently write the former condition as

$$\sum_{k=1}^{n} [(\lambda_{ik} - \lambda_{ki}) - (\lambda_{jk} - \lambda_{kj})] = 0.$$

Hence, if we define $p_{ij} = \lambda_{ij} - \lambda_{ji}$ we can conclude that (2) is equivalent to the fact that whenever $P \in \mathrm{Skew}_n(\mathbb{R})$ satisfies

$$\sum_{k=1}^{n} (p_{ik} - p_{jk}) = 0 \tag{C.3}$$

for all $i, j \in \{1, \ldots, n\}$, then

$$\sum_{i,j=1}^{n} p_{ij} m_{ij} \geq 0. \tag{C.4}$$

• *Step 1.* Here, we characterize the condition (C.3). Taking

$$x = \Big(\sum_{k=1}^{n} p_{jk}\Big)\mathbf{j}$$

in Lemma 4.4 shows that the matrices $P \in \mathrm{Skew}_n(\mathbb{R})$ fulfilling (C.3) coincide with those in the kernel of the operator $\mathcal{L} = (L_1, \ldots, L_n)$. Recall that $\mathcal{L}$ has rank $n-1$. Since $\mathrm{Skew}_n(\mathbb{R})$ is a vector space of dimension $d_1 := n(n-1)/2$, we find that

$$d_2 := \dim(\ker \mathcal{L}) = \frac{n(n-1)}{2} - (n-1) = \frac{(n-1)(n-2)}{2}.$$

Consider the matrices

$$P_{ij} := E_{1i} + E_{ij} + E_{j1} = E_{1i} - E_{ij} + E_{ij}, \tag{C.5}$$

where E_{ij} are the skew symmetric matrices in (B.2). Then

$$\mathcal{L}(P_{ij}) = \mathcal{L}(E_{1i}) + \mathcal{L}(E_{ij}) + \mathcal{L}(E_{j1}) = (e_1 - e_j) + (e_i - e_j) + (e_j - e_1) = 0.$$

Hence, the subset

$$\mathcal{P} := \{P_{ij} : 2 \leq i < j \leq n\} \subseteq \ker \mathcal{L}$$

consists of $(n-1)(n-2)/2$ different elements, which can be ordered according to the lexicographic order of the pairs (i, j). Let us show that they are linearly independent, thus generating the whole kernel. We first consider the basis of skew-symmetric matrices

$$\mathcal{B} := \{E_{ij} : 1 \le i < j \le n\},$$

and again we arrange them in lexicographic order. Let $\mathcal{M} \in \mathcal{M}_{d_2 \times d_1}(\mathbb{R})$ be the matrix of coordinates of the elements in $\mathcal{P}$ with respect to the basis $\mathcal{B}$. Then by (C.5) one infers that the $d_2 \times d_2$ identity matrix appears as the submatrix of $\mathcal{M}$ consisting of all the d_2 rows but just the last d_2 columns. Hence, rank $\mathcal{M} = d_2$ and consequently

$$\ker \mathcal{L} = \operatorname{span}(\mathcal{P}).$$

• *Step 2.* Here, we characterize the condition (C.4), which clearly amounts to

$$\sum_{i,j=1}^{n} p_{ij} m_{ij} = 0$$

for every $P \in \mathcal{P}$. Taking $P = P_{ij}$ for $2 \le i < j \le n$ we get

$$\sum_{i,j=1}^{n} p_{ij} m_{ij} = \tfrac{1}{2}(m_{1i} - m_{i1} + m_{ij} - m_{ji} + m_{j1} - m_{1j}) = m_{1i} + m_{ij} + m_{j1},$$

and this concludes the proof. ■

Finally, we focus on the sticking condition (3.11) in the critical case. The next result gives an explicit characterization that uses similar techniques to those in Lemma B.3.

Lemma C.2. *For any $n \in \mathbb{N}$ and $M \in \operatorname{Skew}_n(\mathbb{R})$, the following assertions are equivalent:*

(1) *There exists $Y \in \operatorname{Skew}_n([-1, 1])$ such that $M = Y \cdot \mathbf{J} + \mathbf{J} \cdot Y$.*

(2) *There exists $Y \in \operatorname{Skew}_n(\mathbb{R})$ such that*

$$T_{ij}(Y) \le m_{ij}, \quad -T_{ij}(Y) \le -m_{ij} \quad \text{and} \quad L_{ij}(Y) \le 1.$$

(3) *We have*

$$\sum_{i,j=1}^{n} q_{ij} + \frac{1}{2} \sum_{i,j=1}^{n} p_{ij} m_{ij} \ge 0$$

for any $i, j = 1, \ldots, n$, and for every $P \in \operatorname{Skew}_n(\mathbb{R})$ and $Q \in \mathcal{M}_n(\mathbb{R}_0^+)$ such that $\sum_{k=1}^{n}(p_{ik} - p_{jk}) + q_{ij} - q_{ji} = 0$.

(4) *The following two conditions hold:*

   (a) *Condition (C.2) holds true.*

   (b) *We have*

$$\sum_{i=1}^{m} \sum_{j=m+1}^{n} m_{\sigma_i \sigma_j} \in [-nm(n-m), nm(n-m)] \tag{C.6}$$

   for every permutation σ of $\{1, \ldots, n\}$ and any $1 \le m \le n$.

Proof. Assertions (1) and (2) are apparently equivalent due to the definition of the linear operators involved. Also, (2) and (3) are equivalent by an application of Lemma B.2 that is analogous to that in the proof of Lemma B.3; hence, we skip the proof for simplicity. Therefore, we will only focus on the equivalence with the last assertion. First, assume that for some $Y \in \mathrm{Skew}_n([-1,1])$ the first assertion holds true, i.e.,

$$m_{ij} = \sum_{k=1}^{n}(y_{ik} - y_{jk}).$$

By Lemma C.1 we arrive at (C.2). Moreover,

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=m+1}^{n} m_{\sigma_i\sigma_j} &= \sum_{i=1}^{m}\sum_{j=m+1}^{n}\sum_{k=1}^{n}(y_{\sigma_i\sigma_k} - y_{\sigma_j\sigma_k})\\ &= (n-m)\sum_{i=1}^{m}\sum_{k=m+1}^{n} y_{\sigma_i\sigma_k} - m\sum_{j=m+1}^{n}\sum_{k=1}^{m} y_{jk} = n\sum_{i=1}^{m}\sum_{k=m+1}^{n} y_{ik}.\end{aligned}$$

Since it is n times the sum of $m(n-m)$ numbers in $[-1,1]$, the condition (C.6) is also satisfied. Conversely, assume that both (C.2) and (C.6) hold and take any $P \in \mathrm{Skew}_n(\mathbb{R})$ and $Q \in \mathcal{M}_n(\mathbb{R}_0^+)$ such that

$$\sum_{k=1}^{n}(p_{ik} - p_{jk}) + q_{ij} - q_{ji} = 0 \tag{C.7}$$

for any $i, j = 1, \ldots, n$. Without loss of generality we can assume that $q_{ii} = 0$ for every $i = 1, \ldots, n$. Also, define $\lambda_i := \sum_{k=1}^{n} p_{ik}$ and consider a permutation σ of $\{1, \ldots, n\}$ such that the λ_{σ_i} are in non-decreasing order,

$$\lambda_{\sigma_1} \le \cdots \le \lambda_{\sigma_n}. \tag{C.8}$$

Let us split

$$I := \sum_{i,j=1}^{n} q_{\sigma_i\sigma_j} + \frac{1}{2}\sum_{i,j=1}^{n} p_{\sigma_i\sigma_j} m_{\sigma_i\sigma_j} =: I_1 + I_2.$$

Using (C.2) in the second term we can write

$$I_2 = \frac{1}{2}\sum_{i,j=1}^{n} p_{\sigma_i\sigma_j}(m_{\sigma_i\sigma_k} - m_{\sigma_j\sigma_k})$$

for any $k = 1, \ldots, n$. Let us take the average with respect to k in the above expression:

$$\begin{aligned}I_2 &= \frac{1}{2n}\sum_{i=1}^{n}\Big(\sum_{k=1}^{n} m_{\sigma_i\sigma_k}\Big)\lambda_{\sigma_i} + \frac{1}{2n}\sum_{j=1}^{n}\Big(\sum_{k=1}^{n} m_{\sigma_j\sigma_k}\Big)\lambda_{\sigma_j}\\ &= \frac{1}{n}\sum_{i=1}^{n}\Big(\sum_{k=1}^{n} m_{\sigma_i\sigma_k}\Big)\lambda_{\sigma_i} = \frac{1}{n}\sum_{i=1}^{n}\Big(\sum_{k=1}^{n} m_{\sigma_i\sigma_k}\Big)(\lambda_{\sigma_j} + q_{\sigma_j\sigma_i} - q_{\sigma_i\sigma_j})\end{aligned}$$

for any $j = 1, \ldots, n$, where (C.7) has been used in the last step. Taking the average with respect to j we get

$$I_2 = \frac{1}{n^2} \sum_{i,j=1}^{n} \Big(\sum_{k=1}^{n} m_{\sigma_i \sigma_k}\Big)(q_{\sigma_j \sigma_i} - q_{\sigma_i \sigma_j})$$

$$= \frac{1}{2n^2} \sum_{i,j=1}^{n} \Big(\sum_{k=1}^{n} (m_{\sigma_i \sigma_k} - m_{\sigma_j \sigma_k})\Big)(q_{\sigma_j \sigma_i} - q_{\sigma_i \sigma_j})$$

$$= \frac{1}{n^2} \sum_{i<j} \Big(\sum_{k=1}^{n} (m_{\sigma_i \sigma_k} - m_{\sigma_j \sigma_k})\Big)(q_{\sigma_j \sigma_i} - q_{\sigma_i \sigma_j}). \tag{C.9}$$

On the other hand,

$$I_2 = \sum_{j>i} q_{\sigma_i \sigma_j} + \sum_{i<j} (q_{\sigma_i \sigma_j} - q_{\sigma_j \sigma_i}). \tag{C.10}$$

Putting (C.9)–(C.10) together we obtain

$$I = 2 \sum_{j>i} q_{ij} + \sum_{i<j} \Big(1 - \frac{1}{n^2} \sum_{k=1}^{n} (m_{\sigma_i \sigma_k} - m_{\sigma_j \sigma_k})\Big)(q_{\sigma_i \sigma_j} - q_{\sigma_j \sigma_i}).$$

Finally, notice that for any $i < j$, the condition (C.7) entails

$$q_{\sigma_i \sigma_j} - q_{\sigma_j \sigma_i} = \sum_{m=i}^{j-1} (q_{\sigma_m \sigma_{m+1}} - q_{\sigma_{m+1} \sigma_m}),$$

and consequently

$$I = 2 \sum_{j>i} q_{ij} + \sum_{k=1}^{n} a_m (q_{\sigma_m \sigma_{m+1}} - q_{\sigma_{m+1} \sigma_m}),$$

where the coefficients are

$$a_m = \sum_{i=1}^{m} \sum_{j=m+1}^{n} \Big(1 - \frac{1}{n^2} \sum_{k=1}^{n} (m_{\sigma_i \sigma_k} - m_{\sigma_j \sigma_k})\Big)$$

$$= m(n-m) - \frac{1}{n} \sum_{i=1}^{m} \sum_{j=m+1}^{n} m_{\sigma_i \sigma_j}.$$

Here, (C.2) has been used again. Since the a_m are all non-negative by (C.6) and the λ_{σ_i} are ordered as in (C.8), we conclude that $I \geq 0$ and this ends the proof. ∎

Funding. The work of J. Park was supported by the Basic Research Program through the National Research Foundation of Korea (NRF) funded by the MSIT(NRF-2020R1A4A3079066). The work of D. Poyato (D.P.) and J. Soler (J.S.) has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 639638), the MECD (Spain) research grant FPU14/06304 (D.P), the MINECO-Feder (Spain) research grant number RTI2018-098850-B-I00, the Junta de Andalucía (Spain) Projects PY18-RT-2422 & A-FQM-311-UGR18 (D.P., J.S.).

References

[1] Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F., Spigler, R.: The Kuramoto model: A simple paradigm for synchronization phenomena. Rev. Modern Phys. **77**, 137–185 (2005)

[2] Ahn, S. M., Choi, H., Ha, S.-Y., Lee, H.: On collision-avoiding initial configurations to Cucker–Smale type flocking models. Comm. Math. Sci. **10**, 625–643 (2012) Zbl 1321.92084 MR 2901323

[3] Aoki, T., Aoyagi, T.: Co-evolution of phases and connection strengths in a network of phase oscillators. Phys. Rev. Lett. **102**, art. 034101, 4 pp. (2009)

[4] Aubin, J.-P., Cellina, A.: Differential Inclusions. Grundlehren Math. Wiss. 264, Springer, Berlin (1984) Zbl 0538.34007 MR 755330

[5] Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., Soler, J.: Exactly solvable phase oscillator models with synchronization dynamics, Phys. Rev. Lett. **81**, art. 3643, 4 pp. (1998)

[6] Carr, J.: Applications of Centre Manifold Theory. Appl. Math. Sci. 35, Springer, New York (1981) Zbl 0464.58001 MR 635782

[7] Carrillo, J. A., Choi, Y.-P., Ha, S.-Y., Kang, M.-J., Kim, Y.: Contractivity of transport distances for the kinetic Kuramoto equation. J. Statist. Phys. **156**, 395–415 (2014) Zbl 1303.82016 MR 3215629

[8] Carrillo, J. A., Choi, Y.-P., Hauray, M.: The derivation of swarming models: mean-field limit and Wasserstein distances. In: Collective Dynamics from Bacteria to Crowds, CISM Courses and Lect. 553, Springer, Vienna, 1–46 (2014) MR 3331178

[9] Carrillo, J. A., Choi, Y.-P., Hauray, M.: Local well-posedness of the generalized Cucker–Smale model with singular kernels. In: MMCS, Mathematical Modelling of Complex Systems, ESAIM Proc. Surveys 47, EDP Sci., Les Ulis, 17–35 (2014) Zbl 1356.35252 MR 3419383

[10] Carrillo, J. A., Choi, Y.-P., Mucha, P. B., Peszek, J.: Sharp conditions to avoid collisions in singular Cucker–Smale interactions. Nonlinear Anal. Real World Appl. **37**, 317–328 (2017) Zbl 1383.34069 MR 3648384

[11] Carrillo, J. A., DiFrancesco, M., Figalli, A., Laurent, T., Slepčev, D.: Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. Duke Math. J. **156**, 229–271 (2011) Zbl 1215.35045 MR 2769217

[12] Carrillo, J. A., James, F., Lagoutière, F., Vauchelet, N.: The Filippov characteristic flow for the aggregation equation with mildly singular potentials. J. Differential Equations **260**, 304–338 (2016) Zbl 1323.35005 MR 3411674

[13] Castaing, C., Valadier, M.: Convex Analysis and Measurable Multifunctions. Lecture Notes in Math. 580, Springer, Berlin (1977) Zbl 0346.46038 MR 0467310

[14] Filippov, A. F.: Differential Equations with Discontinuous Right-Hand Sides. Math. Appl. (Soviet Ser.) 18, Kluwer, Dordrecht (1988) Zbl 0664.34001 MR 1028776

[15] Ha, S.-Y., Ha, T., Kim, J.-H.: On the complete synchronization of the Kuramoto phase model. Phys. D **239**, 1692–1700 (2010) Zbl 1213.34068 MR 2684624

[16] Ha, S.-Y., Kim, J., Park, J., Zhang, X.: Uniform stability and mean-field limit for the augmented Kuramoto model. Netw. Heterog. Media **13**, 297–322 (2018) Zbl 1405.70007 MR 3811564

[17] Ha, S.-Y., Kim, H. K., Ryoo, S. W.: Emergence of phase-locked states for the Kuramoto model in a large coupling regime. Comm. Math. Sci. **14**, 1073–1091 (2016) Zbl 1383.92016 MR 3491817

[18] Ha, S.-Y., Lee, J., Li, Z., Park, J.: Emergent dynamics of Kuramoto oscillators with adaptive couplings: conservation law and fast learning. SIAM J. Appl. Dynam. Systems **17**, 1560–1588 (2018) Zbl 1401.70030 MR 3805840

[19] Ha, S.-Y., Li, Z., Xue, X.: Formation of phase-locked states in a population of locally interacting Kuramoto oscillators. J. Differential Equations **255**, 3053–3070 (2013) Zbl 1325.34066 MR 3093356

[20] Ha, S.-Y., Liu, J.-G.: A simple proof of the Cucker–Smale flocking dynamics and mean-field limit. Comm. Math. Sci. **7**, 297–325 (2009) Zbl 1177.92003 MR 2536440

[21] Ha, S.-Y., Noh, S. E., Park, J.: Synchronization of Kuramoto oscillators with adaptive couplings. SIAM J. Appl. Dynam. Systems **15**, 162–194 (2016) Zbl 1393.34052 MR 3530389

[22] Ha, S.-Y., Park, J., Zhang, X.: A first-order reduction of the Cucker–Smale model on the real line and its clustering dynamics. Comm. Math. Sci. **16**, 1907–1931 (2018) Zbl 07034981 MR 3928053

[23] Hauray, M., Jabin, P.-E.: Particle approximation of Vlasov equations with singular forces: propagation of chaos. Ann. Sci. École Norm. Sup. (4) **48**, 891–940 (2015) Zbl 1329.35309 MR 3377068

[24] Hebb, D. O.: The Organization of Behavior. Wiley, New York (1949)

[25] Ioffe, A.: Existence and relaxation theorems for unbounded differential inclusions. J. Convex Anal. **13**, 353–362 (2006) Zbl 1104.34006 MR 2252236

[26] Jabin, P.-E., Wang, Z.: Mean field limit and propagation of chaos for Vlasov systems with bounded forces. J. Funct. Anal. **271**, 3588–3627 (2016) Zbl 1388.60163 MR 3558251

[27] Kalitin, B. S., Sari, T.: B-stability and its applications to the theorems of Tikhonov and Malkin–Gorshin. Differ. Uravn. **37**, 12–17, 140 (2001) (in Russian) Zbl 1002.34041 MR 1846393

[28] Kuramoto, Y.: Self-entrainment of a population of coupled non-linear oscillators. In: International Symposium on Mathematical Problems in Theoretical Physics, Lecture Notes in Phys. 39, Springer, Berlin, 420–422 (1975) Zbl 0335.34021 MR 0676492

[29] Kuramoto, Y.: Chemical Oscillations, Waves, and Turbulence. Springer Ser. Synergetics 19, Springer, Berlin (1984) Zbl 0558.76051 MR 762432

[30] Kuratowski, K., Ryll-Nardzewski, C.: A general theorem on selectors. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **13**, 397–403 (1965) Zbl 0152.21403 MR 188994

[31] Lancellotti, C.: On the Vlasov limit for systems of nonlinearly coupled oscillators without noise. Transport Theory Statist. Phys. **34**, 523–535 (2005) Zbl 1206.82107 MR 2265477

[32] Łojasiewicz, S.: Une propriété topologique des sous-ensembles analytiques réels. In: Les Équations aux Dérivées Partielles (Paris, 1962), Éditions du Centre National de la Recherche Scientifique, Paris, 87–89 (1963) Zbl 0234.57007 MR 0160856

[33] Mucha, P. B., Peszek, J.: The Cucker–Smale equation: singular communication weight, measure-valued solutions and weak-atomic uniqueness. Arch. Ration. Mech. Anal. **227**, 273–308 (2018) Zbl 1384.35138 MR 3740375

[34] Niyogi, R. K., English, L. Q.: Learning-rate-dependent clustering and self-development in a network of coupled phase oscillators. Phys. Rev. E. **80**, art. 066213 (2009)

[35] Osgood, W. F.: Beweis der Existenz einer Lösung der Differentialgleichung $\frac{dy}{dx} = f(x, y)$ ohne Hinzunahme der Cauchy–Lipschitz'schen Bedingung. Monatsh. Math. Phys. **9**, 331–345 (1898) JFM 29.0260.03 MR 1546565

[36] Peszek, J.: Existence of piecewise weak solutions of a discrete Cucker–Smale's flocking model with a singular communication weight. J. Differential Equations **257**, 2900–2925 (2014) Zbl 1304.34092 MR 3249275

[37] Peszek, J.: Discrete Cucker–Smale flocking model with a weakly singular weight. SIAM J. Math. Anal. **47**, 3671–3686 (2015) Zbl 1350.92062 MR 3403135

[38] Picallo, C. B., Riecke, H.: Adaptive oscillator networks with conserved overall coupling: sequential firing and near-synchronized states. Phys. Rev. E (3) **83**, art. 036206, 12 pp. (2011) MR 2788259

[39] Poyato, D., Soler, J.: Euler-type equations and commutators in singular and hyperbolic limits of kinetic Cucker–Smale models. Math. Models Methods Appl. Sci. **27**, 1089–1152 (2017) Zbl 1369.35050 MR 3659047

[40] Poyato, D.: Filippov flows and mean-field limits in the kinetic singular Kuramoto model. arXiv:1903.01305 (2019)

[41] Ren, Q., Zhao, J.: Adaptive coupling and enhanced synchronization in coupled phase oscillators. Phys. Rev. E **76**, art. 016207, 6 pp. (2007)

[42] Seliger, P., Young, S. C., Tsimring, L. S.: Plasticity and learning in a network of coupled phase oscillators. Phys. Rev. E (3) **65**, art. 041906, 7 pp. (2002) Zbl 1244.34078 MR 1917989

[43] Shvydkoy, R., Tadmor, E.: Eulerian dynamics with a commutator forcing II: Flocking. Discrete Contin. Dynam. Systems **37**, 5503–5520 (2017) Zbl 1372.35227 MR 3681948

[44] Tao, T.: Structure and Randomness: Pages from Year One of a Mathematical Blog. Amer. Math. Soc., Providence, RI (2008) Zbl 1245.00024 MR 2459552

[45] Tolstonogov, A. A.: Existence and relaxation of solutions of differential inclusions with an unbounded right-hand side in a Banach space. Sibirsk. Mat. Zh. **58**, 937–953 (2017) (in Russian); English transl.: Siberian Math. J. **58**, 727–742 (2017) Zbl 1384.34073 MR 3762367