TWO PROBLEMS ABOUT COUPLED OSCILLATORS: TRANSIENT DYNAMICS AND SWARMING.

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kevin Philip O'Keeffe

August 2017

TWO PROBLEMS ABOUT COUPLED OSCILLATORS: TRANSIENT DYNAMICS AND SWARMING.

Kevin Philip O'Keeffe, Ph.D.

Cornell University 2017

The conduct of coupled oscillators has long beguiled scientists. Here we study two models of such oscillators. The first is Peskin's integrate-and-fire model. We focus on the transitory behavior, showing that in its infancy, synchrony looks much like aggregation. In the second model, we consider oscillators which adjust their positions in space as well as their phases. We show the coaction of these two effects produces novel spatiotemporal patterns, which we study both analytically and numerically.

BIOGRAPHICAL SKETCH

This bio was provided by the author's dear friend Conall O'Deasmhunaigh

* * *

Kevin Philip O'Keeffe was born in Cork, Ireland's second city, and grew up in the nearby small town of Cobh. The sleepy port town is perhaps best known for being one of the final stops made by the Titanic on its way to the United States. Certainly, Cobh is not known for giving rise to capable academics. Nonetheless, Kevin's interest and capability was evident from an early age; he obtained a perfect score in both the Junior and Leaving Certificate, Ireland's two state exams.

On the basis of these achievements, he was awarded a full entrance scholarship to study Physics and Applied Mathematics at University College Cork. In University, Kevin's high academic performance continued. He earned the honor of college scholar in every year of attendance, before graduating top of his class with a first class honors degree.

Soon, Kevin would leave Irish shores, like many before him, bound for the United States. The reason for Kevin's emigration marked him out from the majority of his countrymen however - he had been accepted to pursue a PhD in Applied Mathematics in Cornell University. Kevin flourished academically at Cornell under the stewardship of Professor Steven Strogatz, publishing various papers in prestigious academic journals and securing a post-doctoral position at MIT.

Cornell's scholarly influence can also be seen in Kevin's recreational pursuits.

In his downtime, he enjoys Ithaca's autumnal beauty while out walking, and has even flirted with vegetarianism. A competitive tennis player at a national level in his youth, Kevin continues to take to the court occasionally, although he has somehow always found an excuse to decline this author's challenge.

The mediocre teacher tells. The good teacher explains.
The superior teacher demonstrates. The great teacher inspires.

William Arthur Ward

To Steven Strogatz, a good, superior, and great teacher.

ACKNOWLEDGEMENTS

I never like thanking people. It's not because I lack gratitude. Rather, it's because I feel it forces me to *rank* the people in my life, a messy business. Will people assume, wrongly, that I'm most thankful to the first person I mention? Will sister *A* envy sister *B* if she gets the nod first? Will invitations to barbecues go mysteriously missing if I tell a joke about friend x, but not about friend y? The pitfalls are legion. The truth of course is that the order in which one acknowledges others confers no meaning. I'm grateful for, bettered by, and indebted to, all friends, family, colleagues, and mentors in an ineffable and unorderable way. I therefore simply list my thanks in the order they occur to me.

My advisor Steven Strogatz has been inestimably helpful. I regard him the consummate mentor and a quintessential scholar. His technical advice was always sound, and his creative perspective always fruitful. He exudes and provokes a positive, playful attitude, which endears him both to his students and the greater scientific community. He is also very *wise*, having a keen awareness of the saving role that guidance plays in the trying periods of academic research. He has been an overwhelming positive force in my life, and I cite our many joyful meetings as the highlights of my time at Cornell.

Chris Myers, another member of my committee, has also been a kind and generous mentor. At the start of my graduate career he introduced me to disease dynamics, and guided me through the rocky roads of my first research project. He remained a reliable ally throughout my PhD, being quick with insight, tips, and numerical tricks. For example his leadership during my A-exam led to the fruitful development of a project on swarmalators (which constitute the fourth chapter of this thesis). Richard Rand, the final member of my committee, was

also a warm and reliable mentor. In an invitingly thick New York City accent, he taught me all there is to know about nonlinear vibrations and equipped me with a formidable arsenal of perturbative techniques. I owe a lot to his clear and effective teaching.

A similar debt is owed to my collaborators. Pavel Krapivsky expertise was integral to the work on pulse-coupled oscillators (second and third chapters); he had the key idea of viewing the behavior of the system as an aggregation process, and pioneered the analysis. Hyunsuk Hong made valuable contributions to both the swarmalator work, and spearheaded two other projects on correlated disorder in the Kuramoto model (not contained in this thesis). I thank them both for their intellectual openness and benevolent mentorship.

I received an eclectic education from my dear friend Colin Clement. From practical domestic skills to efficient pythonic scripts, he was a fountain of knowledge, and always eager to share. Ismail El Baggari was a somewhat princely figure in my life, taking it upon himself to educate me in the pleasures of the cocktail, and to make sure I finally got a driver's license. I thank him unreservedly. I also thank Neal Reynolds, for trying, and failing, to deepen my appreciation of experimental physics, but always succeeding in making me laugh.

Irena Papst has been unendingly obliging. I think she twice supplied me with paper, after midnight, when I had run out, and was quick with advice or assistance whenever I needed it. She is also my organizational hero. From departmental titivations to github commits, she appalls me with her persistent precision. Given such talents, I expect our collaboration shall soon bear fruit.

I deeply admire the jocular Lindsay Mercer. We've known each other no more

than two semesters, but under the more appropriate metric of belly laughs, it feels like years. Disarmingly honest and conversationally brilliant, she always lightened my mood and touched me with her insightful kindness. I hold Emily DiAngelo in similar esteem. Part older sister, part co-moron, she was the perfect mirror for my odd brand of silly energy, and provided a bounty of much needed color to my increasingly monkish lifestyle.

It's said that failure humbles you. If this is true, then I'm indebted to the CAM lads for introducing me to the strategy game 'Coup'. Never in my life have I felt so thoroughly incompetent. Many a Friday night I spent being expertly outwitted by the likes of Andrew Loeb, Marc Gilles, and Matt Hin, who, with patience comparable to their strategic acumen, would repeatedly explain the rules to me. The fact that I sometimes won is a testament to dumb luck.

While Coup was a disaster, I did enjoy some strategic success in my choice of desk. I wisely positioned myself between the two numerical heavyweights, Heather Wilber and David Eriksson, from whom I profited immensely. Via little more than osmosis, I've become reasonably fluent in a miscellany of numerico-analytic topics. I thank them for their patient explanations, and for putting up with my cranky and distracted office demeanor.

I owe a perennial debt to my family, for their warm blanket of love and security. They tirelessly prop up my errable confidence, and coax me into being my best self with the most enduring and sedulous care. It clicheic to remark that I couldn't have gotten this far in life without them, but then again, things are often clicheic for a reason.

Lastly I thank Maggie Osinksi, for changing me forever.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Truth is ever to be found in the simplicity, and not in the multiplicity and confusion of things.

Isaac Newton

In my first year of grad school, Prof Chris Henley told me Ising was an average student. I was shocked – the Ising model is famous in physics, one of the cornerstones of statistical mechanics. It was originally introduced as a toy model of ferromagnetic materials, where the magnetic spin of the constituent particles was idealized to be either 'up' or 'down' (in reality the spin can point in any direction). But its full importance came later, when it was used to understand general critical phenomena, universality, and ultimately the renormalization group; well known triumphs of modern physics. Since then, it has found application in fields ranging from social dynamics to neurobiology.

I was curious how such a successful model could have such a humble origin. I found out that the supposedly modest ability of its progenitor was just one of the quirks of its curious history [2]. For instance, it was Ising's advisor, Lenz, who conceived the model, whereas Ising just carried out the calculations (to his credit, Ising tried to re-christen it the "Lenz-Ising" model, but it didn't stick [3]). Furthermore, Ising's original paper had a mistake! He correctly showed that in one dimension there was no ferromagnetic transition, but wrongly concluded that this result would hold in three dimensions.

But perhaps the most significant quirk is that the model wasn't taken seriously. The physicist Stephen Brush writes "it was scorned or ignored by most scientists" for being "greatly over simplified" [2]. They believed that idealizing the spin vector to a binary value, which made the model solvable, was too unrealistic. The Nobel prize winner Chen Yang said it was "considered an arcane exercise[s], narrowly interesting, mathematically seducing, but of little consequence" [4]. Amusingly, this view was shared by Ising, who was said to have been so disenchanted with the model's relevance that he gave up research [2]! Yet over time this view changed, as the broad utility of the model, stemming from its simplicity, was gradually realized.

I was fascinated that something as important as the Ising model had been so misjudged, and wondered if other scientific work had been too. In my third year, when I became interested in synchronization, I found out it had.

Collective synchronization occurs when a large population of oscillators spontaneously lock their cycles, in spite of dissimilarities in their natural frequencies. This can be thought of as a temporal analogue of the ferromagnetic phase transition in the Ising model. This effect spans many disciplines. In biology, it is seen in discharging pacemaker cells, in chemistry in metabolic cycles of yeast cell suspensions, and in physics in arrays of superconducting Josephson junctions. There are many other examples, too numerous to list here.

Art Winfree first discovered this sync phenomenon when studying biological rhythms [5]. In a seminal, eloquent, and inspiring paper he posited that various physiological rhythms are the product of 'the mutual synchronization of myri-

ads of individual oscillatory processes'. However, the individual oscillators he was imagining weren't of the usual, linear variety. Instead they were nonlinear, which made them oscillate with a preferred amplitude and waveform called a limit cycle. How a population of nonlinear oscillators behaved was unknown. Mathematically, this group behavior would be determined by a large system of nonlinear differential equations – a formidable challenge immune to the usual tools.

But Winfree wasn't deterred. Like Ising – or rather Lenz – his approach was to write down an extremely idealized model in the hope of capturing the essential features of the sync phenomenon. It worked. Simulations showed a transition to "striking community synchronization", which due to the simplicity of the model he was able to partially analyze. Kuramoto later made Winfree's model even simpler, resulting in his famous, eponymous model. In a beautiful analysis he solved the model exactly, ultimately leading to the plenitude of attention the model enjoys today. However, similar to the Ising story, both Kuramoto and the rest of the community were initially rather under-enthused with his results.

In my fourth year, I attended a conference organized in Kuramoto's honor. He was unable to travel in person, so instead recounted his model's tale via video [6]. He told us that when he developed the model he was new to nonlinear science, and so was unsure of its worth. He thus sought expert opinion before publishing his results. With endearing honesty and trademark understatement, he said the feedback was "not so encouraging". The experts "showed little interest in his [my] work, and were even critical". They said his results "didn't

seem new at all" and that "there were many similar works in mechanical and electrical engineering".

Dispirited by this negativity, Kuramoto didn't work on the model for five years. He didn't even publish the work beyond a two page article in a conference proceedings. Serendipitously however, this article was seen by Winfree who recognized its importance and described it in his famous book on biological rhythms [5]. This gave the Kuramoto model much exposure, ultimately leading to the wide popularity it enjoys today.

I love both this story and the Ising story. I like that Ising blundered when analyzing his model in three dimensions – it's encouraging that famous scientists make mistakes too – and I'm amused that Lenz didn't get titular credit. I also like that the experts' disdain made Kuramoto so reluctant to publish; it humanizes him, while also showing that importance can be gravely mismeasured. But what I love most is that the stories illustrate the power of minimal models. Stripping a model down to its barest features really can work. Simplicity makes things tractable, and tractability makes things understandable. Then, like an inverse series of Russian dolls, greater realism can be added to make things quantitative.

This thesis is an homage to this style of work. It explores minimal models of coupled oscillators. It consists of three projects, each of which is an autonomous publication with its own introduction and conclusion. The first two are about the transient behavior of pulse-coupled oscillators. These are idealizations of relaxation oscillators, which communicate with each other by firing sudden

pulses. They are often collectively referred to as the Peskin model, since Charlie Peskin first introduced them when studying cardiac pacemaker cells. Since then however they have found application in low-powered radios, sensor networks, and earthquakes dynamics.

As is typical is nonlinear dynamics, traditional research on the Peskin model focuses on the steady state – usually some form of synchrony. But the approach to the steady state is less studied. Our work focuses on this part of the dynamics: the prelude to synchrony. We show that this looks like an aggregation process, a link which gives us access to new mathematical tools. In spite of this, the nonlinearities in the system make progress difficult. In a barefaced attempt to mimic the minimalism credo, we eliminate these difficulties by making two idealizations. This makes the model exactly solvable, and its analysis constitutes Chapter one.

In the second Chapter, we show that our idealized model captures the essence of the original problem. This insight, along with the exact analysis of our idealized model, lets us derive good approximations to the transient behavior of the Peskin model, our original goal.

Chapter three is more ambitious. It seeks to wed two similar, but largely separate, fields: swarming and synchronization. In synchronization, the units are characterized by their phases, whose influence on each other promotes temporal order. In swarming systems, the degrees of freedom are the units' positions, whose influence on each other triggers spatial order. But what happens if these two effects interact? In what ways can units with both temporal and spatial

degrees of freedom organize themselves?

To answer this question, we study a generalized Kuramoto model where now the elements are free to move around in space as well as in phase. Further, there is a bidirection coupling between their spatial and phase dynamics. We call these entities 'swarmalators' since they swarm and oscillate.

The original motivation for this work was Janus particles [7]. These are micrometer sized particles which can be induced to oscillate about their centers of mass using magnetic fields. When in suspension, they are also free to move around in space. The dynamics of the phase (of oscillation about the center of mass) and position of each particle are coupled to each other, and so Janus particles satisfy our proposed definition of swarmalators.

As detailed in [7], populations of Janus particles display novel collective behavior. We first considered trying to model Janus particles directly. The equations of motions were however complicated, making analysis seem unlikely. Numerically solving them was an obvious alternative. But as pointed our by my committee member Chris Myers, the complexity of the governing equations would likely obscure the core phenomenology that we wanted to study, namely the interaction between synchronization and swarming. He made the key suggestion that we instead follow the minimalism principle. Thus we developed a minimal model, whose simplicity, as we show, lets us study some of the states of our system analytically.

* * *

A final word on minimal modeling. This approach to science, of course, is far from perfect. Not all phenomena are simple, or expressible with elementary mathematics, and so it will often fail. Of course another danger is that the modeler, in his zealotry for simplicity, will over do it. No one would remember even *seeing* a baby, it having been discharged so feverishly with the dishwater. Yet as as evidenced by my heroes Ising, Winfree, and Kuramoto, it pays to be daring. The skillful modeler therefore strikes a balance between intrepitude and overconfidence. As Emanuel Derman puts it, "A little north of common sense, but south of idolatry, lies the wise use of models".

CHAPTER 2

SYNCHRONIZATION AS AGGREGATION: CLUSTER KINETICS OF PULSE-COUPLED OSCILLATORS

One does not discover new continents without consenting to lose sight of the shore for a very long time .

Andre Gide

2.1 Introduction [1]

In one of the first experiments on firefly synchronization, the biologists John and Elisabeth Buck captured hundreds of male fireflies along a tidal river near Bangkok and then released them at night, fifty at a time, in their darkened hotel room [8]. They observed that "centers of synchrony began to build up slowly among the fireflies on the wall. In one area we would notice that a pair had begun to pulse in unison; in another part of the room a group of three would be flashing together, and so on." Synchronized groups continued to emerge and grow, until as many as a dozen fireflies were blinking on and off in concert. The Bucks realized that the fireflies were phase shifting each other with their flashes, driving themselves into sync.

Here we study stylized models of oscillators akin to the fireflies, in which syn-

[1] This chapter is reproduced from: O'Keeffe, Kevin P., Pavel L. Krapivsky, and Steven H. Strogatz. "Synchronization as aggregation: Cluster kinetics of pulse-coupled oscillators." Physical review letters 115.6 (2015): 064101.

chrony builds up stepwise, in expanding clusters. By borrowing techniques used to analyze aggregation phenomena in polymer physics, materials science, and related subjects [9, 10], we give the first analytical description of how these synchronized clusters emerge, coalesce, and grow. We hasten to add, however, that the models we discuss are not even remotely realistic descriptions of fireflies; they are merely intended as tractable first steps toward understanding how clusters evolve en route to synchrony.

Our work is part of a broader interdisciplinary effort [11, 12]. Oscillators coupled by sudden pulses have been used to model sensor networks [13–17], earthquakes [18, 19], economic booms and busts [20], firing neurons [21, 22], and cardiac pacemaker cells [23]. Diverse forms of collective behavior can occur in these pulse-coupled systems, depending on how the oscillators are connected in space. Systems with local coupling often display waves [24, 25] or self-organized criticality [18, 26, 27], with possible relevance to neural computation [22] and epilepsy [28]. In contrast, systems with global coupling, where every oscillator interacts equally with every other, tend to fall into perfect synchrony. Rigorous convergence results have been proven for this case [27, 29–32]. But the techniques used previously have not revealed much about the transient dynamics leading up to synchrony—the opening and middle game, as opposed to the end game. Aggregation theory offers a new set of tools to explore this prelude to synchrony.

Exact results for the transient dynamics can be obtained in at least two cases. In the Supplemental Material [33], we apply aggregation theory to the deterministic Peskin model [23], assuming the oscillators rise linearly to threshold and

fire pulses of size $1/N$, where $N \gg 1$ is the number of oscillators. A simplified stochastic version of this model yields similar results, but because it illustrates the main ideas more clearly we present it here in the main text. This toy model, which we call *scrambler oscillators*, consists of N identical integrate-and-fire oscillators coupled all to all. Each oscillator has a voltage-like state variable x that increases linearly according to $\dot{x} = 1$, rising from a baseline value of 0 to a threshold value of 1. Whenever any oscillator reaches threshold, it fires and does three things. (i) It kicks every oscillator (and every synchronous cluster of oscillators) to a new random voltage, independently and uniformly—in this sense, it scrambles the other oscillators. However, no scrambling occurs within a cluster; all oscillators at the same voltage get kicked to the same new voltage. Thus, clusters never desynchronize; once formed they are preserved by the scrambling procedure. (ii) The firing oscillator then "absorbs" any scrambled oscillators that lie within a distance $1/N$ of threshold, by bringing them to threshold and thereby synchronizing with them. To avoid the complications that would be caused by chain reactions of firings, we assume that the oscillators being brought to threshold do not get to fire until the *next* time they reach threshold. (iii) The oscillator that fired resets to $x = 0$ along with the oscillators it absorbed.

If a cluster of j oscillators does the firing, the same rules apply, except that now any oscillators within a distance j/N of threshold get absorbed. The assumed proportionality to j is natural, if each member of the cluster contributes to the pulse strength. We study other plausible coupling rules in [33].

The motivation for this scrambler model is that it leads to the simplest possi-

ble mean-field approximation. In the infinite-N limit, we would like clusters of every size to be uniformly distributed in voltage at all times. This convenient property would greatly ease the derivation of the rate equations for the cluster kinetics. As we will see below, the predictions that follow from this approximation agree reasonably well with simulations. (For finite-N, these assumptions break down at large times and for large clusters, limitations that we analyze in [33].)

Assume the initial voltages x_i, for $i = 1, \ldots, N$, are independent and uniformly distributed. At first, nothing interesting happens. The oscillators increase their voltages without interacting. But then one oscillator reaches threshold and fires. The remaining oscillators get scrambled, and perhaps some get absorbed. Then another oscillator fires, and so on. After a while, the system has formed clusters of various sizes.

Let $N_j(t)$ denote the number of clusters of size j at time t. Thus there are $N_1(t)$ singleton oscillators, $N_2(t)$ pairs of synchronized oscillators, $N_3(t)$ triplets, and so on. The N_j are correlated random quantities. They are correlated because oscillators belonging to clusters of one size are unavailable to clusters of another size, and they are random because of the randomness in the initial conditions and the scrambling procedure. It does not seem feasible to understand the time-evolution of the N_j unless they are so large that their fluctuations from one random realization to another are negligible.

So assume from now on that $N_j \gg 1$ for all j and replace these random quantities by their ensemble averages. Let $c_j = N^{-1}\langle N_j \rangle$ denote the average cluster densi-

ties. One hopes that relative fluctuations are small; more precisely, $N^{-1}N_j = c_j + O(N^{-1/2})$. An even stronger assumption is that the densities of different subpopulations are asymptotically uncorrelated: $N^{-2}N_iN_j = c_ic_j + O(N^{-1/2})$.

These c_j allow us to define a natural disorder parameter, given by the total density $c(t) = \sum_j c_j(t)$. It measures the extent of the system's fragmentation. To see this, note that at $t = 0$ each oscillator is alone; only clusters of size 1 exist. Accordingly $c_1(0) = 1$ and all other $c_j(0) = 0$ for $j > 1$. Hence $c(0) = 1$, correctly indicating that the system starts out maximally fragmented. At the opposite extreme, as $t \to \infty$ only one giant cluster of synchronized oscillators exists. The system is then minimally fragmented: $c(t) = 1/N \to 0$ as $N \to \infty$.

To derive a rate equation for the decline of $c(t)$, let R_i be the rate at which clusters of size i fire, for $i = 1, \ldots, N$, and let L_i be the number of clusters lost to absorption in each such firing. Then $\dot{c} = -\sum_i R_iL_i$.

To find L_i, recall that when a cluster of size i fires, all the other clusters get assigned a new voltage uniformly at random. Moreover, any clusters assigned to the interval $[1 - i/N, 1)$ get brought to threshold and absorbed. Since the voltages of these other clusters are uniformly distributed on $[0, 1]$, a fraction i/N of them will be absorbed. There are $\sum_j N_j$ clusters in total. Hence the number absorbed is $L_i = (i/N)\sum_j N_j = i\sum c_j = ic$.

The rate R_i takes more work to calculate. Since some clusters get absorbed, not every cluster gets the chance to fire. We must account for this depletion when calculating R_i. First consider the background rate of firing of clusters of

size i in the absence of absorptions. In other words, pretend for a moment that when an i-cluster fires, it simply scrambles every other cluster and restarts its own cycle without absorbing anyone. Call this background rate R_i^0. Since all oscillators move with velocity $v_i = \dot{x}_i = 1$, and since the cluster density is c_i, the corresponding background rate of firing is $R_i^0 = c_i v_i = c_i$. Next, to find the actual R_i, we must subtract from R_i^0 the rate at which clusters of size i are being absorbed and hence deprived of their chance at firing. Call this absorption rate R_i^a. Clusters of size i are absorbed when clusters of size j fire, for $j = 1, \ldots, N$, taking a fraction j/N of the uniformly distributed i-clusters along with them. Since there are N_i clusters of size i and the j-clusters fire at rate R_j, the total rate at which i-clusters are being absorbed is given by $R_i^a = \sum_j (j/N) N_i R_j = \sum_j j c_i R_j = c_i \sum j R_j$.

Putting all this together gives $R_i = R_i^0 - R_i^a = c_i - c_i \sum_j j R_j = c_i(1 - \sum_j j R_j)$. Let $\beta = 1 - \sum_j j R_j$. Note that β is the same for all i, which enables it to be determined self-consistently, as follows. From $R_i = \beta c_i$ we obtain $\beta = 1 - \sum_j j R_j = 1 - \sum j(\beta c_j)$. Now invoke the identity $\sum_j j c_j = j(N_j/N) = 1$, which expresses conservation of oscillators. Solving for β then gives $\beta = 1/2$ and therefore $R_i = c_i/2$.

Next, plug the expressions derived for R_i and L_i into the rate equation $\dot{c} = -\sum_i R_i L_i$. The result is $\dot{c} = -\sum_i (c_i/2)(ic) = -(c/2) \sum_i i c_i = -c/2$. Recalling that $c(0) = 1$, we conclude that

$$c(t) = \exp(-t/2). \tag{2.1}$$

Figure 2.1 shows this result matches simulations.

How do the individual cluster densities c_i behave? To derive their rate equa-

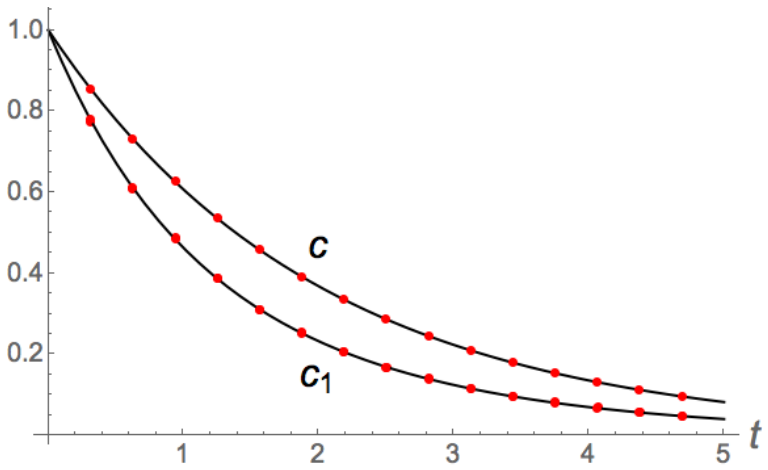

Figure 2.1: Theoretical and simulated $c(t)$ and $c_1(t)$. Solid lines show theoretical curves obtained analytically (see text). Data points show simulation results for $N = 10^4$ oscillators.

tions, note that since the voltage space is the interval $[0, 1]$, a segment of length N^{-1} contains on average $Nc \times N^{-1} = c$ clusters. In fact, the probability that it contains n clusters (of any sizes) is given by the Poisson distribution: $\Pi_n = c^n e^{-c}/n!$. This is the mathematical expression of the assumption that clusters are distributed randomly without correlations.

With this in mind, let us solve for $c_1(t)$, the density of singletons. It is the easiest $c_j(t)$ to analyze, since it can only decrease. Two mechanisms decrease $c_1(t)$: (i) The loss of a *firing* singleton when it absorbs a cluster of any size, and (ii) the loss of $p > 1$ *absorbed* singletons, when a cluster of any size fires.

Consider mechanism (i). Since $R_i = c_i/2$ as shown above, singletons fire at a rate $R_1 = c_1/2$. When they fire, they absorb any cluster lying in the voltage segment $[1 - 1/N, 1)$. The probability that this segment contains one or more clusters is, according to the Poisson distribution, $1 - e^{-c}$. In this case, the firing singleton is lost by its absorption of a cluster, thus decreasing N_1 by 1. Otherwise N_1

is unchanged. Hence singletons are lost by mechanism (i) at an expected rate $(c_1/2)\left[1 \times (1 - e^{-c(t)}) + 0 \times e^{-c(t)}\right] = (c_1/2)\left[1 - e^{-c(t)}\right]$. Note: we only account for the loss of the *firing* singleton here; any singletons it *absorbs* are accounted for in the following mechanism (ii).

Suppose p singletons lie in the interval $[1 - j/N, 1)$ when a cluster of size j fires, for $j = 1, \ldots, N$. This event happens with probability $e^{-jc_1}(jc_1)^p/p!$, and when it does, it consumes p singletons. (If a singleton did the firing, the loss would be p+1. However the loss of the firing singleton was already counted in mechanism (i). So the consumption factor of p for each firing j-cluster is valid even for $j = 1$.) As before, j-clusters fire at a rate $R_j = c_j/2$. Hence singletons are lost by mechanism (ii) at a rate

$$\sum_{j \geq 1} \frac{c_j}{2} \times \sum_{p \geq 1} p \frac{(jc_1)^p e^{-jc_1}}{p!} = c_1/2. \tag{2.2}$$

Summing the loss rates from (i) and (ii) gives

$$\frac{dc_1}{dt} = -\frac{c_1}{2}(2 - e^{-c(t)}). \tag{2.3}$$

This equation has a closed-form solution in terms of exponential integrals:

$$c_1(t) = \exp(-t + \mathrm{Ei}(-1) - \mathrm{Ei}(-e^{-t/2})), \tag{2.4}$$

where we have used the initial condition $c_1(0) = 1$. Figure 1 shows good agreement between the theoretical and numerical $c_1(t)$.

For $i > 1$, the rate equation for c_i includes gain terms as well as loss terms. Clusters of size $i > 1$ can be created when two or more smaller clusters coalesce, or destroyed when they themselves coalesce with at least one other clus-

ter. The loss term is a straightforward generalization of that for c_1, and is given by $(c_i/2)\left[2 - e^{-ic(t)}\right]$.

To find the gain term, imagine that a cluster of size k fires. The segment $[1 - k/N, 1)$ may contain a_1 clusters of size 1, a_2 clusters of size 2, etc. This event happens with probability $\frac{(kc_1)^{a_1}}{a_1!}e^{-kc_1} \times \frac{(kc_2)^{a_2}}{a_2!}e^{-kc_2} \times \ldots$ (where we are using the assumption that clusters of different sizes are independent as well as Poisson distributed). If the segment contains a combination of clusters such that $k + a_1 + 2a_2 + 3a_3 + \cdots = i$, then a cluster of size i will form. We sum over all such combinations for a cluster of size k firing, and then sum over all k, to get the rate at which clusters of size i are created:

$$\sum_{k=1}^{i-1} \frac{c_k}{2} e^{-kc} \sum_{a_1+2a_2+\cdots=i-k} \left(\prod_{p\geq 1} \frac{(kc_p)^{a_p}}{a_p!} \right). \tag{2.5}$$

Combining the loss and gain terms, and transferring $c_i e^{-ic}$ into the gain term, we finally obtain

$$\dot{c}_i = -c_i + \sum_{k=1}^{i} \frac{c_k}{2} e^{-kc} \sum_{\sum p a_p = i-k} \left(\prod_{p\geq 1} \frac{(kc_p)^{a_p}}{a_p!} \right). \tag{2.6}$$

We see from the sum that the equations (3.29) are recursive. They can be solved one by one, though not analytically, so we resort to numerical integration. Figure 2.2 shows that the theoretical and simulated c_i agree.

Although we cannot find all the $c_i(t)$ explicitly, we can get their moments $M_n(t) = \sum_j j^n c_j(t)$ through the use of a generating function. We already know two moments: $M_0(t) = c(t)$, given by Eq. (1), and $M_1 = 1$. A few of the higher

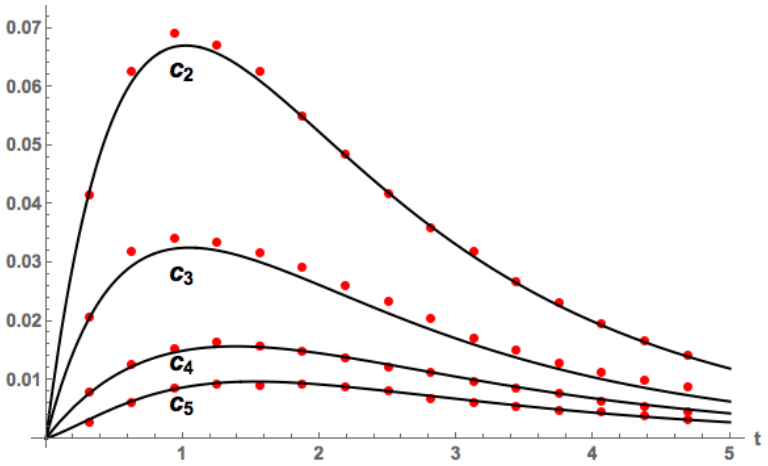

Figure 2.2: Theoretical and simulated cluster densities $c_2(t)$ through $c_5(t)$. Solid lines show theoretical predictions computed from numerical integration of Eq. (3.29). Data points show simulation results for $N = 5 \times 10^4$ oscillators.

moments are

$$\begin{aligned} M_2(t) &= e^{3t/2} \\ M_3(t) &= 7e^{7t/2} - 6e^{3t} \\ M_4(t) &= \frac{448}{5}e^{5t} - 128e^{9t/2} + \frac{217}{5}e^{15t/4} - 4e^{27t/8}. \end{aligned} \tag{2.7}$$

These and further results are discussed in [33].

We also studied two modifications of the scrambler model. For example, suppose that when a cluster of size j fires, it absorbs all oscillators within a distance kj/N of threshold, where $k > 0$ is a tunable coupling strength. Or suppose that the pulse strength is k/N, independent of the size j of the firing cluster. We discuss both cases in [33]. In the latter case the disorder parameter $c(t)$ decays algebraically rather than exponentially. This makes sense physically: by assuming that larger clusters no longer fire larger pulses proportional to their size, we cut the positive feedback loop underlying the exponential growth of synchrony in the original scrambler model.

The stochastic scrambler model approximates the deterministic models studied by Peskin [23] and Mirollo and Strogatz [29]. In those models, when a cluster of size j fires it adds a voltage pulse $j\epsilon$ to every other oscillator, or pulls it up to threshold, whichever is less. For the case where $\epsilon = 1/N$ and the charging curve is linear, we show in [33] that these deterministic systems can also be analyzed by the methods above. The main new feature is that $c(t)$ and the cluster densities $c_i(t)$ become piecewise linear. But their overall shapes still resemble those seen in the scrambler model.

Intuitively, the piecewise linearity in the deterministic case arises because the speed of each oscillator, and the effect of a pulse on each oscillator, is the same. Hence the oscillators, or clusters of oscillators, maintain their initial ordering; they all march forward through $[0, 1]$ in a line with no passing. This property then implies, in a mean-field sense discussed in [33], that the oscillators condense into clusters whose size doubles periodically. At the end of the first period, all oscillators synchronize into pairs spaced equally apart. At the end of the second period, those pairs merge into clusters of size 4, and so on. Moreover, the clusters begin each period evenly spaced from each other (again, in a mean-field sense where fluctuations are neglected), which yields the piecewise constant firing rate mentioned above.

One limitation of our analysis, for both the scrambler and deterministic models, is that each oscillator obeys $\dot{x}_i = 1$ between firings. Such linear sawtooth waveforms are reasonable for the oscillators used in sensor networks [15], but not for neurons or cardiac pacemaker cells. In [33] we show that our results for the deterministic model are robust to the addition of small concavity in the

charging curve. But large concavity introduces new effects, not yet understood theoretically. The analysis becomes more difficult because clusters are no longer uniformly distributed as we have assumed throughout.

There are many avenues to explore in future work. It would be interesting to study cluster kinetics in oscillator systems with local coupling, network structure, heterogeneity, delays, and other realistic features. Several of these features would break clusters apart, and so would require including fragmentation processes in the analysis. By incorporating suitable new loss and gain terms in the rate equations, one could perhaps derive useful estimates for synchronization speeds in more complex but random networks where synchronization is guaranteed but speed estimates are lacking [34].

Another possible application concerns the detection of network topology. Arenas et al. showed that in the Kuramoto model, the time course of cluster formation en route to synchronization can be used to shed light on a network's topology [35] and community structure [36]. While the mean-field approach used above is suitable for homogenous topologies, extensions of our approach using degree distributions might prove useful in probing a network's hidden structure.

2.2 Supplemental Material

Section I presents additional analytical results for the Scrambler model. These include asymptotic results about the long-time behavior of the individual cluster densities c_i described in the main text, as well as an exact solution for the time-dependence of the *moments* of the c_i.

Section II modifies the Scrambler oscillator model to allow alternative couplings between oscillators. Specifically, we change how close to threshold an oscillator has to be in order to be absorbed by a firing cluster. This only modestly changes the analysis.

Section III examines a more substantial alteration of the model, in which we change its character from stochastic to deterministic. Now, when an oscillator reaches threshold, it no longer scrambles every other oscillator to a new, random, voltage. Instead, it kicks every other oscillator up by a constant amount, or up to threshold, whichever is less. This deterministic resetting rule is in line with the simplest traditional models of pulse-coupled oscillators. As will be shown, this change makes for a more involved analysis, but the results are qualitatively similar to those obtained in the main text. This qualitative agreement is what originally motivated our use of the Scrambler model in the first place.

Section IV explores how some of our results break down if we relax the assumption $N_j \gg 1$ made in the main text.

2.3 I. Analytical Results

2.3.1 Asymptotic Behavior

In this subsection we investigate the long-time behavior of the cluster densities $c_i(t)$. For large t, we find that about 45% of all clusters in the system are singletons, while about 14% are doublets. Indeed, it seems likely that for any cluster size k, the fraction $c_k(t)/c(t)$ tends to a nonzero constant as $t \to \infty$. We were quite surprised by these results, since usually in aggregation the fraction of singletons decays to zero as the time grows (and the same holds true for all other species of clusters).

To derive these results, we begin by analyzing $c_1(t)$. Figure 1 in the main text indicates that $c_1(t)$ decays at a rate comparable to that of the disorder parameter $c(t)$. Thus it is natural to study their ratio $c_1(t)/c(t)$ for large t. Equations (1) and (4) of the main text showed that

$$c(t) = \exp(-t/2)$$

and

$$c_1(t) = \exp[-t + \mathrm{Ei}(-1) - \mathrm{Ei}(-e^{-t/2})],$$

where Ei denotes the exponential integral. If one expands $\mathrm{Ei}(-e^{-t/2})$ in the large-t limit, one finds that

$$\lim_{t\to\infty} \frac{c_1(t)}{c(t)} = \exp[\mathrm{Ei}(-1) - \gamma] = 0.45\ldots,$$

where γ is the Euler constant. Thus, asymptotically about 45% of all clusters are singletons.

To calculate the asymptotic fraction of doublets, we rewrite Eq.(6) of the main text as

$$\dot{c_2} = -c_2\left(1 - \frac{1}{2}\,e^{-2c}\right) + \frac{1}{2}\,c_1^2\,e^{-c}. \tag{2.8}$$

Next we solve this equation subject to $c_2(0) = 0$. The resulting expression for $c_2(t)$ is cumbersome, but it looks slightly simpler if we use c rather than t as the independent variable:

$$\frac{c_2}{c} = \exp\left(-\int_c^1 dx\,\frac{1 - e^{-2x}}{x}\right)\mathcal{E}(c) \tag{2.9}$$

with

$$\mathcal{E}(c) \equiv \int_c^1 dy\,\exp\left[-y - \int_y^1 dx\,\frac{(1 - e^{-x})^2}{x}\right].$$

Finally, since $c(t) \to 0$ as $t \to \infty$, we replace the c's in the lower limits of the integrals with $c = 0$ and thereby obtain (after numerical quadratures)

$$\lim_{t\to\infty} \frac{c_2(t)}{c(t)} = 0.14\ldots,$$

which shows that asymptotically, about 14% of all clusters are doublets.

Figure 2.3 shows that these predictions agree reasonably well with simulations.

2.3.2 Moments

In the main text, we derived the following rate equation for the individual cluster densities c_i:

$$\dot{c_i} = -c_i + \sum_{k=1}^{i} \frac{c_k}{2} e^{-kc} \sum_{\sum p a_p = i-k} \left(\prod_{p\geq 1} \frac{(kc_p)^{a_p}}{a_p!}\right). \tag{2.10}$$

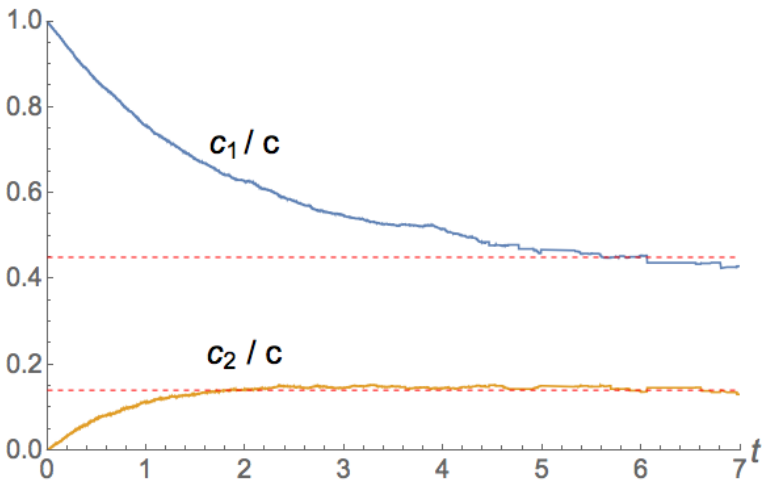

Figure 2.3: Relative fraction of singletons and doublets, $c_1(t)/c(t)$ and $c_2(t)/c(t)$, for $N = 10^4$ oscillators. Red dotted lines show predicted asymptotic values, derived above.

Only c_1 and c_2 had closed-form solutions, so we resorted to numerical integration for the higher c_i. While we cannot analytically solve for all these higher c_i, we can solve for their moments, defined as

$$M_n(t) = \sum_j j^n c_j(t).$$

The first two moments are trivial: $M_0 = \sum_j c_j(t) = c(t)$ and $M_1(t) = \sum_j j c_j(t) = 1$, from conservation of oscillators. The higher moments can be obtained from a generating function. Let $G(z,t) = \sum_{k\geq 1} c_k(t) e^{kz}$. Then the infinite set of differential equations (3.29) transforms into

$$\frac{\partial G(z,t)}{\partial t} + G(z,t) = \frac{1}{2} G[z - c(t) + G(z,t), t]. \tag{2.11}$$

This equation looks neat, but it is far from trivial, as the right-hand side involves G in a very nonlinear manner. Using the identity $M_n(t) = \frac{\partial^n G}{\partial z^n}|_{z=0}$, we can however

derive the following equations for the moments:

$$\begin{aligned}\dot{M}_2 &= \frac{3}{2}M_2 \\ \dot{M}_3 &= \frac{7}{2}M_3 + 3(M_2)^2 \\ \dot{M}_4 &= \frac{15}{4}M_4 + 16M_2M_3 + \frac{3}{2}(M_2)^3.\end{aligned} \tag{2.12}$$

Like the c_i equations (Eq. (6) in the main text), these moment equations are recursive and can be solved in succession, except that here it is possible to do so explicitly. We find

$$\begin{aligned}M_2(t) &= e^{3t/2} \\ M_3(t) &= 7e^{7t/2} - 6e^{3t} \\ M_4(t) &= \frac{448}{5}e^{5t} - 128e^{9t/2} + \frac{217}{5}e^{15t/4} - 4e^{27t/8}.\end{aligned} \tag{2.13}$$

Figure 2.4 plots theoretical and simulated values of the M_i. The agreement is good for M_2 but worse for M_3 and M_4. This is to be expected. Each $c_i(t)$ is a stochastic process, subject to fluctuations dominated by the chance formation of big clusters. Since $M_n(t) = \sum_j j^n c_j(t)$, the higher moments amplify these fluctuations more and are therefore noisier themselves.

2.4 II. Alternative Couplings

We first restate the dynamics of the original Scrambler model, and then describe the variations.

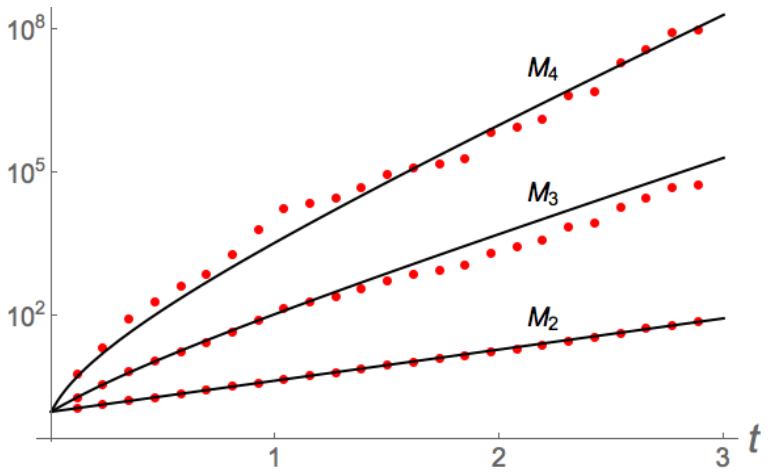

Figure 2.4: Log plot of the first three nontrivial moments M_2, M_3, M_4. Black curves, theoretical predictions obtained from (2.12); red dots, average simulation results for 100 realizations of $N = 10^4$ oscillators.

Recall that in the main text we considered a population of $N \gg 1$ integrate-and-fire oscillators coupled all-to-all. Each oscillator was characterized by a voltage-like state variable x, which increased linearly according to $\dot{x}_i = 1$. When a cluster of j oscillators reached a threshold value set to 1, they fired and then instantly did three things: (i) they reassigned every other oscillator (or cluster of oscillators) a new voltage uniformly at random (they scrambled the oscillators) (ii) they absorbed any oscillators within a distance j/N of threshold and (iii) they reset their voltage to 0 along with any oscillators they absorbed.

We now modify event (ii) in either of two ways: when a cluster of size j fires, it either absorbs all oscillators within a new distance of threshold given by (a) kj/N or (b) k/N. Modification (a) generalizes the original model by including an adjustable coupling k. Modification (b) assumes that the absorption region is independent of j, the size of the firing cluster.

These generalizations change the analysis of the Scrambler model only slightly. For instance, to find the disorder parameter $c(t)$, we again use the rate equation $\dot{c} = -\sum_i R_i L_i$, where R_i denotes the rate at which a cluster of size i fires, and L_i denotes the number of oscillators absorbed when a cluster of size i fires. In the original model, to find L_i and R_i, we made liberal use of the fact that all oscillators on the interval $[1 - j/N, 1)$ were captured when a cluster of size j fired. With the generalized couplings (absorption distances), this interval simply changes to $[1 - kj/N, 1)$ and $[1 - k/N, 1)$. This change propagates through the analysis straightforwardly. Hence, we state the results in the following table without derivation. (In the table, $W(x)$ refers to Lambert's W function.)

	j/N	kj/N	k/N
L_i	ic	kic	kc
R_i	$c_i/2$	$c_i/(1+k)$	$c_i/(1+kc)$
$c(t)$	$e^{-t/2}$	$e^{-\frac{k}{k+1}t}$	$1/(kW(k^{-1}e^{(k^{-1}+t)}))$

The rows of the table give the results for L_i, R_i and $c(t)$; the columns show how the results vary for the three coupling schemes: original, (a), and (b). For coupling scheme (a), where the absorption distance is kj/N, the exponential decay constant in $c(t)$ is predicted to be $-k/(k+1)$. To test this, we simulated $c(t)$ for various k, and found the exponents of best fit. Figure 2.5 shows the results along with the theoretical curve.

We find similarly good agreement between theory and simulations for coupling scheme (b), as shown in Figure 2.6. The chief difference in this case is that the disorder parameter $c(t)$ decays algebraically as opposed to exponentially. To see

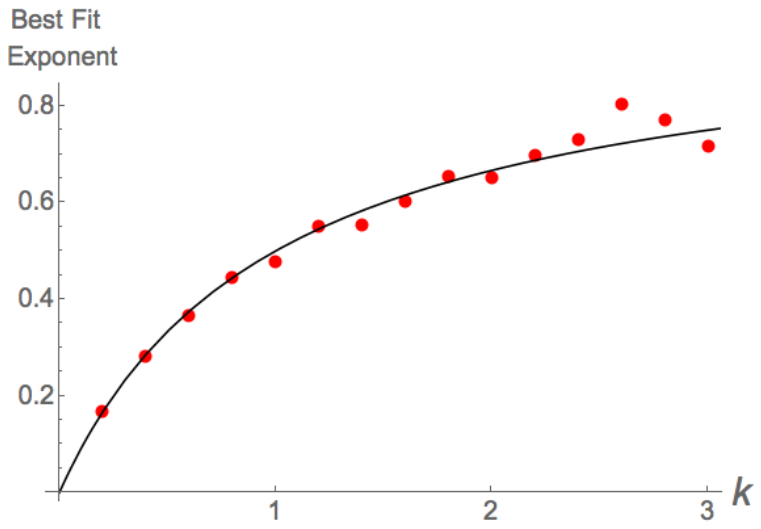

Figure 2.5: Magnitude of the decay constant in $c(t)$ for coupling scheme (a). Black curve, theoretical prediction $k/(k+1)$; red dots, simulation results for $N = 5000$ oscillators.

this, consider the Taylor expansion of $c(t)$ for large t and $k = 1$:

$$c(t) = 1/(W(1^{-1}e^{(1+t)})) \approx \frac{1}{t} + O\left(\frac{1}{t^2}\right) \tag{2.14}$$

as $t \to \infty$.

Intuitively, the physical reason for the non-exponential decay is that model (b) assumes that large clusters fire with the same strength as small ones. In contrast, the original model displayed exponential growth of synchrony (or equivalently, exponential decay of the disorder parameter $c(t)$) because it assumed that clusters fire with strength proportional to their size, which sets up a positive feedback loop in which the big clusters get bigger at the expense of smaller ones (because they fire more strongly and therefore absorb other oscillators in a snowballing fashion). That is why the level of synchrony grows exponentially fast in the original model, but not in the modified model.

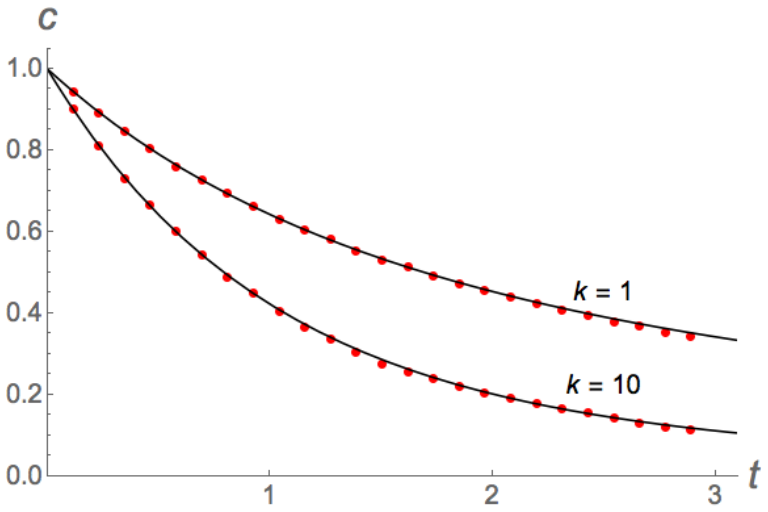

Figure 2.6: Decay of the disorder parameter $c(t)$ for coupling scheme (b), which assumes a fixed absorption distance k/N. Black curves, theoretical predictions; red dots, simulation results for $N = 5000$ oscillators. Plots for $k = 1$ and $k = 10$ are shown.

Having solved for $c(t)$ in models (a) and (b), we could now go on to solve for the individual cluster densities c_i and the moments M_i. Nothing qualitatively new happens (compared to what we saw in the main text for the original model), so we omit the details.

2.5 III. Deterministic Oscillators

The Scrambler model is a toy model. We introduced it to give the simplest possible mean-field model of pulse-coupled oscillators. Specifically, it was the random shuffling of oscillators during each firing event that simplified their analysis. It conveniently kept the voltages of all oscillators (and all clusters of oscillators) uniformly distributed on $[0, 1]$ at all times.

The extreme randomness of this resetting rule, however, is contrived. Traditional models of pulse-coupled oscillators, such as those analyzed by Peskin [23] and Mirollo and Strogatz [29], obey deterministic resetting rules, much as real biological oscillators obey deterministic phase-response curves. As we will show in this section, models with deterministic resetting can be reasonably approximated within the framework developed here.

For simplicity, we will restrict attention to an almost absurdly idealized model of pulse-coupled oscillators, even more idealized than the models discussed in Refs. [23, 29]. It consists of what we will refer to as Firefly oscillators. (Fictional Firefly oscillators would be a more apt description, given that essentially everything about the model is unrealistic for real fireflies.)

The equations of motion for the Fireflies are the same as for the Scramblers: $\dot{x}_i = 1$ (in between firing events). The initial voltages are again drawn from a uniform distribution. However, when a cluster of j synchronized Fireflies reaches the threshold value of 1, that cluster does just two things: (i) it imparts a voltage pulse of size j/N to all other oscillators. Any subsequent oscillators that reach threshold by virtue of this extra j/N, and hence get absorbed by the firing cluster, do not fire until the *next* time they reach threshold. As before, this is to avoid complications that would be caused by chain reactions of firings. (ii) The cluster resets to $x = 0$ along with any oscillators it absorbed.

Firefly oscillators of this type can be viewed as a special case, in two respects, of the oscillators studied by Peskin [23] and Mirollo and Strogatz [29]. First, the oscillators of Refs. [23, 29] have a nonlinear charging curve rather than a linear

one; their voltage dynamics are governed by $\dot{x}_i = S_0 - \gamma x_i$, where the parameters $S_0 > 0$ and $\gamma \geq 0$. Second, when an oscillator fires in the models of [23, 29], it imparts a voltage pulse of size ϵ to all other oscillators. The long-term behavior of the system then depends on the values of ϵ and γ.

Thus, the Firefly model studied here corresponds to the case $S_0 = 1$, $\gamma = 0$, and $\epsilon = 1/N$. For these parameters, the system almost always achieves global synchrony [27]. Our goal, then, is not to establish how the long-term behavior depends on parameters, since we already know that synchrony is inevitable for this simple model. Rather, the goal is to quantify how synchrony builds up over time. To put it another way, we want to predict the kinetics of cluster formation, growth, and coalescence as the system evolves toward synchrony.

2.5.1 Fireflies vs. Scramblers

Will the deterministic Fireflies behave (as hoped) like the stochastic Scramblers? There are two differences between them that complicate the analysis of the Fireflies.

The first difference has to do with how their speed $\dot{x}$ evolves as the system moves toward complete synchrony. The average speed of the Fireflies doesn't remain constant at unity, as it does for the Scramblers. This is because when a cluster of size j fires, all the other oscillators receive a pulse of size j/N which boosts them up on their voltage curve (whereas the Scramblers were just randomly reassigned on $[0, 1]$, and so keep the same speed $v_i = 1$ on average).

The speeds of the Fireflies are thus $v_i = 1 + v_{\text{pulse}}$, where v_{pulse} must be determined.

A second difference is that the firing rate of the Fireflies is piecewise constant (unlike that of the Scramblers, which as shown in the main text is a smooth function of time: $R_i(t) = c_i(t)/2$.) To see why the firing rate for the Fireflies is piecewise constant, first define Δx_i as the distance between the i^{th} and $(i+1)^{th}$ oscillators (or cluster of synchronized oscillators). Second, divide the temporal evolution into periods $\{T_n\}$, where each period is the time taken for the full population to complete a full cycle. Then ask, how does Δx_i behave during the first period, the time taken for the first wave of oscillators to complete their first cycle? Since all oscillators receive the same number of pulses and have the same speed, we see that each Δx_i won't change while the oscillators complete their virgin ascent through [0,1]. This implies a constant firing rate during this first period.

We will later show that this is not unique to the first period: the firing rates will take different, but constant, values during each period; they will be piecewise constant. This is in stark contrast to the Scramblers, where each Δx_i is constantly changing as the oscillators get reshuffled on [0, 1] during each firing event.

2.5.2 Intuitive Picture of the Dynamics

With these differences in mind, we begin with a qualitative description of the dynamics. As mentioned, the speed of each oscillator, and the effect of a pulse

on each oscillator, is the same. This means that the initial ordering of oscillators, or clusters of oscillators, will be invariant throughout the dynamics. They all march forward through $[0, 1]$ in a line with no passing.

Then we consider the *average* behavior as $N \to \infty$. In this limit, the average spacing between oscillators $\langle \Delta x_i \rangle$ approaches N^{-1}. Now, what happens when the first oscillator fires? It captures all oscillators on the interval $[1 - 1/N, 1)$. Since $\langle \Delta x_i \rangle = N^{-1}$, there will be exactly one oscillator on this interval, on average, and so one oscillator will be captured. This procedure will repeat itself for the next oscillator that fires, and the oscillator after that, such that every oscillator that fires captures the one behind itself. In this mean-field sense, then, the first wave of oscillators will be an orderly sequence of fire/capture/fire/capture, so that at the end of the first period, all oscillators will have synchronized into pairs spaced equally apart.

Of course, Δx_i will have fluctuations about the mean value of N^{-1}. For the oscillators spaced such that $\Delta x_i < N^{-1}$, no captures will take place. For $\Delta x_i > N^{-1}$, at least one capture will take place, and possibly more. So, at $t = T_1$ there will be a number of clusters of different sizes. It is not clear how these clusters are distributed on $[0, 1]$. Say, for example, that mostly clusters of size 2 and 3 formed, while a cluster of size 4 was the first capture, and a cluster of size 6 was the last capture. Then the clusters of sizes 2 and 3 will be approximately uniformly distributed in voltage, while the distribution of those of sizes 4 and 6 will be more sharply peaked.

Nevertheless we assume that the clusters of size i are uniformly distributed on

$[0, 1]$ for each i. We recognize that this won't be accurate for each i for all values of t. It will however be accurate for those values of i which contain most of the oscillator "mass" and less so for those with less of the mass. So, our assumption will be imprecise for those c_i which are small, but since they are small, the inaccuracy won't matter, to first order.

Now that we understand the first period, how will the second period proceed? We again consider the *average* behavior as $N \to \infty$. In this mean-field description, we earlier concluded that at the end of the first period, all oscillators would have synchronized into pairs spaced equally apart on $[0, 1]$. When the first pair fires, therefore, there will again be exactly one pair of synchronized oscillators behind them, and so as before, the second period will be an orderly sequence of fire/capture/fire/capture, resulting in all oscillators having synchronized into clusters of size 4.

Continuing this logic, we see the size of clusters will double during each period. Moreover, we observed that the clusters of oscillators will begin each period evenly spaced from each other. This implies the aforementioned piecewise constant firing rate. The mean-field dynamics are therefore trivial: there is a train-like progression of clusters of the same size through $[0, 1]$, with each cluster that reaches threshold doubling in size by absorbing the cluster behind it.

2.5.3 Mean-field Analysis

With this picture in mind, we analyze the rate equation for our disorder parameter: $\dot{c} = -\sum_i R_i L_i$.

To calculate L_i, we again find all oscillators in the interval $[1 - i/N, 1)$. It is tempting to write down $L_i(t) = \sum_j (i/N) N_j(t)$. This isn't strictly true however, because as we observed, Δx_i will be constant during each period, which means we must evaluate N_j at the start of said period. To make this clear, define $\tilde{x}(t) = x(t = T_{n-1})$ for $T_{n-1} < t < T_n$. The tilde notation signifies that throughout the n^{th} period the quantity x is fixed at its value at the start of that period. In terms of this tilde notation, the desired result is $L_i = \sum_j (i/N) \tilde{N}_j(t) = i\tilde{c}$.

To find the firing rate, we follow the procedure used in the main text for the Scramblers: we decompose the firing rate into two parts: $R_i = R_i^0 - R_i^a$. The rate of firing in the absence of absorption will be $R_i^0 = \tilde{c}_i v_i = \tilde{c}_i(1 + v_{\text{pulse}})$. The absorption rate will again be $R_i^a = \sum_j (j/N) \tilde{N}_i R_j = \tilde{c}_i \sum_j j R_j$.

We next determine v_{pulse}. The "pulse velocity" due to a cluster of size j will be (absolute number of pulses per sec) $\times$ (distance per pulse). Since R_j is the rate of firing of c_i, $R_j N$ will give the absolute number of fires. The distance per pulse is j/N. The total pulse velocity is thus $v_{\text{pulse}} = \sum_j (N R_j)(j/N) = \sum_j j R_j$. This gives $R_i^0 = \tilde{c}_i(1 + \sum_j j R_j)$. Putting all this together, we find

$$R_i = \tilde{c}_i(1 + \sum_j j R_j) - \tilde{c}_i \sum_j j R_j \tag{2.15}$$

which reduces to $R_i = \tilde{c}_i$.

Substituting our expressions for R_i and L_i into the rate equation for c then yields

$$\dot{c} = -\sum_i L_i R_i = -\sum_i i\tilde{c}\tilde{c}_i = -\tilde{c}\sum_i i\tilde{c}_i = -\tilde{c} \tag{2.16}$$

and hence $\dot{c} = -\tilde{c}$. Thus, we see that $\dot{c}$ will be a piecewise linear function.

To solve for this function, we need to determine the periods $\{T_n\}$. The average speed is $v = 1 + v_{\text{pulse}} = 1 + \sum_j jR_j = 1 + \sum_j jc_j = 2$. This gives $\{T_n\} = \{0, 0.5, 1, \dots\}$. For these values of $\{T_n\}$ and the initial condition $c(0) = 1$, the solution of Eq. (3.2) is the piecewise linear function

$$c(t) = \frac{p + 2 - 2t}{2^{p+1}} \quad \text{for} \quad \frac{p}{2} < t < \frac{p+1}{2}, \tag{2.17}$$

where $p = 0, 1, 2, \dots$. Hence, for the Firefly model, the disorder parameter $c(t)$ is a series of line segments of length 0.5, with slopes that are sequentially reduced by a factor of 2.

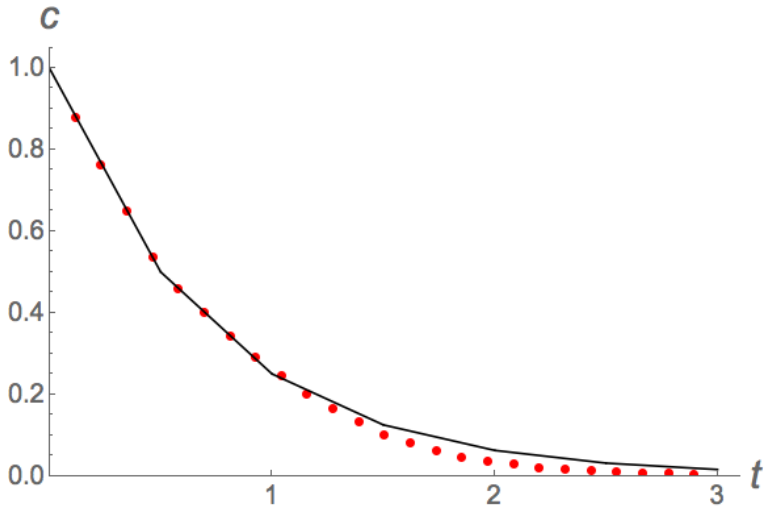

Figure 2.7: Piecewise linear behavior of the disorder parameter $c(t)$ for the Firefly model. Black curve, theoretical prediction (2.17); red dots, simulation results for $N = 5000$ oscillators.

We stress however that our results for $L_i, R_i, v_{\text{pulse}}$, and $c(t)$ are only leading-

order approximations, based on mean-field arguments. We expect fluctuations around these values. Our hope is that these will be small.

Figure 2.7 shows the simulated behavior of $c(t)$ against the mean-field prediction (2.17). As can be seen, there is reasonable agreement until late times.

Now that we have $c(t)$, the next target is $c_i(t)$. Carrying out the same analysis as for the Scramblers, we find

$$\dot{c}_i = -2\tilde{c}_i + \sum_{k=1}^{i} \tilde{c}_k \; e^{-k\tilde{c}} \sum_{\sum p a_p = i-k} \left(\prod_{p \geq 1} \frac{(k\tilde{c_p})^{a_p}}{a_p!} \right). \tag{2.18}$$

Since the quantities on the right hand side are held fixed over each period, solving for each c_i is straightforward. Figure 2.8 shows the resulting solutions along with simulated values. Reasonable agreement is evident. We restate that our results are mean-field equations, so some discrepancy is expected. The moments $M_i(t)$ will also be piecewise linear, and can be obtained in a similarly straightforward manner, following the methods shown in the main text.

2.5.4 Nonlinear charging curve

Until now in our treatment of the deterministic Firefly model, we have assumed that the oscillators rise linearly to threshold. This assumption is valid for the electronic oscillators used in sensor networks, but for neurons and cardiac pacemaker cells, a nonlinear rise to threshold is more appropriate. For this reason, and also on mathematical grounds, it is natural to ask how a nonlinear charging curve would affect the transient dynamics.

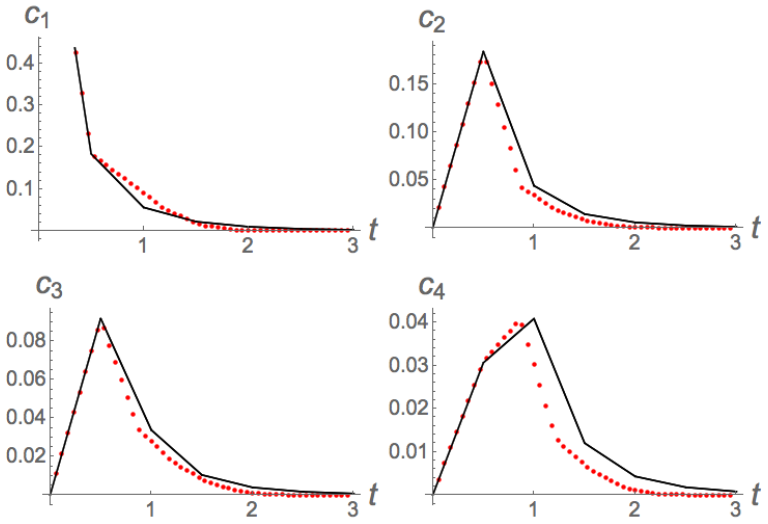

Figure 2.8: Time evolution of the cluster densities c_1 through c_4 for the Firefly model with uniformly random initial conditions. Black curves, theoretical predictions derived from Eq. (2.18); red dots; simulation results for $N = 3 \times 10^4$ oscillators. The discrepancies are due to finite-N effects.

To do so, we return to the traditional Peskin model [23, 29]. Its voltage dynamics are governed by $\dot{x}_i = S_0 - \gamma x_i$, where the parameters $S_0 > 0$ and $\gamma \geq 0$. When an oscillator fires, it kicks all other oscillators up by ϵ or up to threshold, whichever is less. For γ strictly greater than zero, the system is guaranteed to end up with all oscillators firing in unison, as proven in Ref. [29], but almost nothing is known about the model's cluster dynamics en route to synchrony. The analysis becomes much more diffult when concavity is included, for reasons discussed in the main text; in short, one can no longer assume that the oscillators are uniformly distributed in voltage at all times. Hence we numerically explore the effect of concavity. We find that for small nonlinearity $\gamma > 0$, the cluster dynamics are similar to what we have already discussed for the Scrambler and linear Firefly models.

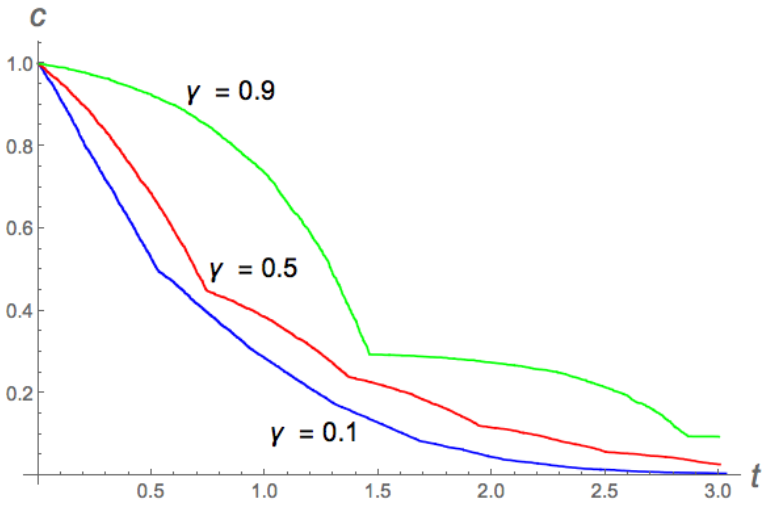

Figure 2.9: Time evolution of the disorder parameter c for the Firefly model with uniformly random initial conditions and concave-down charging curve, for three values of the concavity parameter γ. Simulation results are shown for $N = 10^4$ oscillators and pulse size $\epsilon = 1/N$.

Figure 2.9 shows that for the lowest value of γ, the decay curve for c closely resembles that shown in Figure 2.7 for the linear Firefly model. The same is true for the individual cluster densities; compare, for example, the curves for c_2 in Figure 2.10 and Figure 2.8. For higher values of γ, the cluster dynamics show new effects, not yet understood theoretically. Although these results are preliminary, they suggest that the simplified Scrambler and linear Firefly models are insensitive to small amounts of concavity in the charging curve. In this sense, these models provide reasonable approximations to systems with charging curves that are concave down, the case most often studied in the literature.

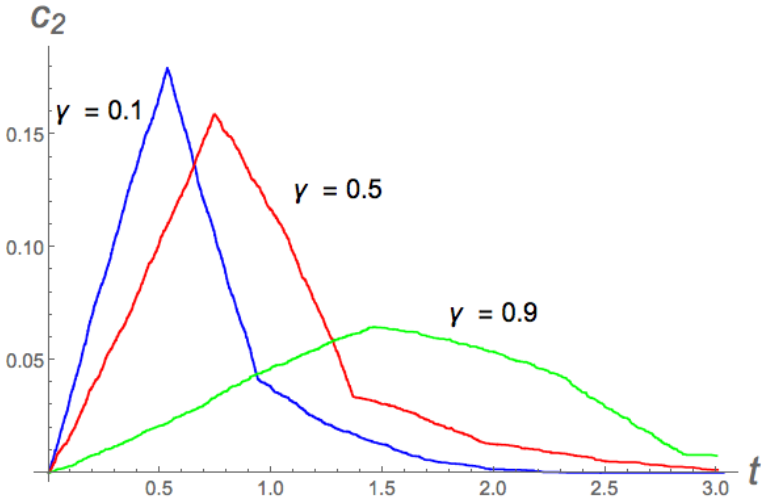

Figure 2.10: Time evolution of c_2, the density of 2-clusters, for the Firefly model with uniformly random initial conditions and concave-down charging curve, for three values of the concavity parameter γ. Simulation results are shown for $N = 10^4$ oscillators and pulse size $\epsilon = 1/N$.

2.6 IV. Breakdown of approximations

When deriving the rate equations for c_i, we made the assumption that $N_i \gg 1$ for all i (recall that N_i denotes the number of clusters of size i). This assumption let us define c_i in terms of ensemble averages, $c_i := \langle N_i \rangle / N$, which in turn let us use probabilistic arguments in our analysis. This assumption clearly cannot be satisfied for all cluster sizes i. For instance, when $t \to \infty$, we know just one giant cluster of size N remains: $N_N = 1$ and $N_i = 0$ for all other i.

It is not entirely clear how to estimate the time T at which our assumptions break down, but one can obtain an upper bound based on when $c(T) \approx 1/N$, at which time almost the whole system has coalesced into a giant synchronized cluster. Combining this criterion with our chief result that $c(t) = \exp(-t/2)$, we

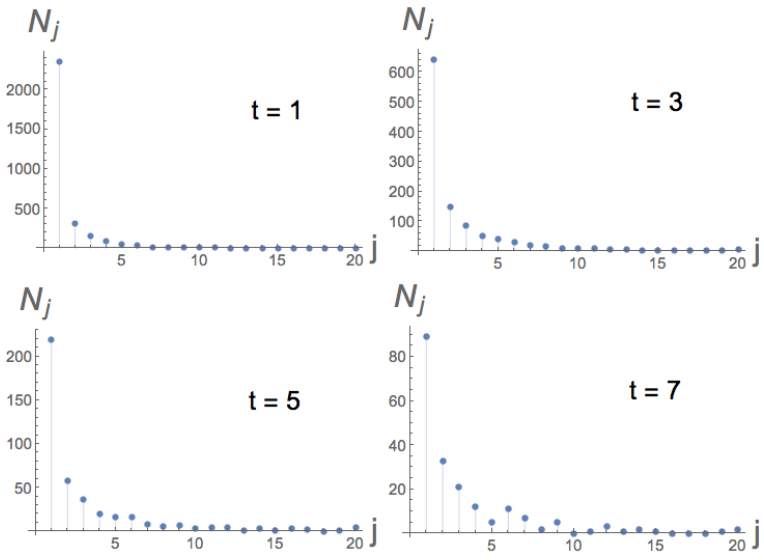

Figure 2.11: Simulated N_j versus j for $N = 5000$ oscillators at various times. N_j decreases (roughly) monotonically with increasing cluster size j at every t.

get $T < 2\ln(N)$. Thus the breakdown time scales at most logarithmically with the size of the system.

What effect does this breakdown of our assumptions have on the predicted cluster kinetics? We examine this issue by simulating the Scrambler model for $N = 5000$ oscillators, and computing N_j for $j \leq 20$ at various times. The results are shown in Figure 2.11.

The first observation is that N_j decreases (roughly) monotonically with increasing cluster size at every t. Smaller clusters are more abundant than larger ones for the times shown; $N_j \gg N_k$ for $j \ll k$. Our assumption that $N_j \gg 1$ therefore gets worse as the cluster size j increases.

The second observation is that the above is true for much of the synchronization

process. For systems of size $N = 5000$, the total time elapsed until complete synchronization occurs is typically $t \approx 10$. But $c(t)$ has typically fallen to ≈ 0.1 by $t = 5$, and to ≈ 0.01 by $t = 7$. So for $t < 7$ approx, our assumption that $N_j \gg 1$ is good for small clusters, and gets worse for larger clusters. This means our analytical results for c_j should correspondingly worsen for increasing j. As can be seen in Figure 2.12, this is indeed the case.

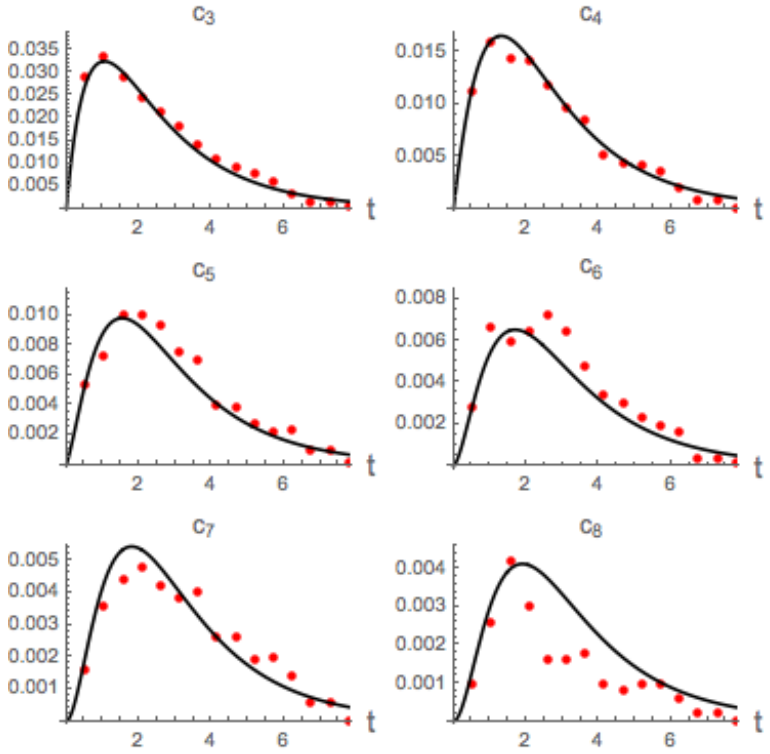

Figure 2.12: Theoretical and simulated cluster densities $c_3(t)$ through $c_8(t)$. Solid line show theoretical predictions computed from numerically integration of equation (3.29). Data points show simulation results for $N = 5000$ oscillators. As expected, the agreement between theory and simulation gets steadily worse with increasing cluster size.

These observations explain why the mean-field approximation for $c(t)$ is good for so long. By definition, $c(t) = \sum_i c_i(t)$. Because $c_j \gg c_k$ for $j \ll k$ (which follows from $N_j \gg N_k$ for $j \ll k$), $\sum_i c_i$ is dominated by small cluster sizes i, for which

the theory is most accurate.

2.6.1 Final stages of synchronization

For $t > 7$ (again, for a system of size $N = 5000$), only a few large clusters remain. Our assumption $N_j \gg 1$ is then clearly violated, and so the mean-field results no longer apply. In the final stages of the synchronization process, then, we expect large deviations from the mean-field results. To verify this, we calculated the relative errors shown in Figure 2.13 below.

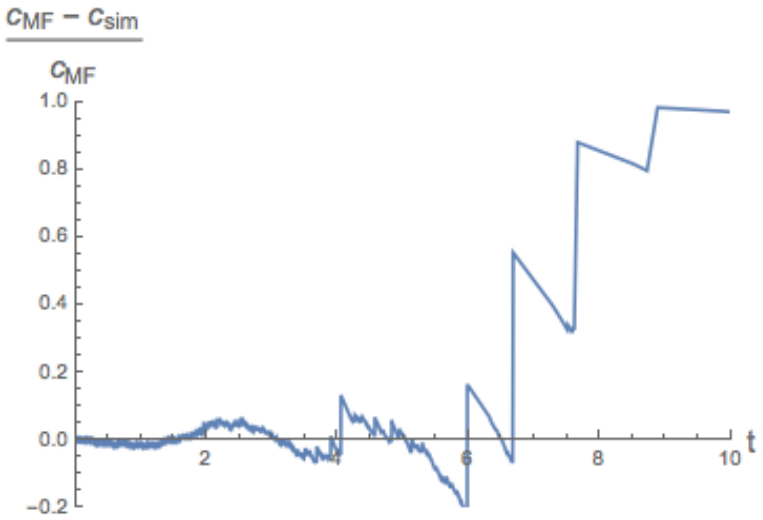

Figure 2.13: Relative error between mean-field prediction and simulated $c(t)$, for $N = 5000$ oscillators.

Around $t \approx 7$ the error starts to climb. In particular, we see $c_{MF} > c_{sim}$, meaning the synchronization of the simulated system is faster than the mean-field prediction. As $t \to \infty$ the process becomes more and more stochastic, whereby c_{sim} oscillates noisily around c, until it drops stochastically to zero, as we show in Figure 2.14.

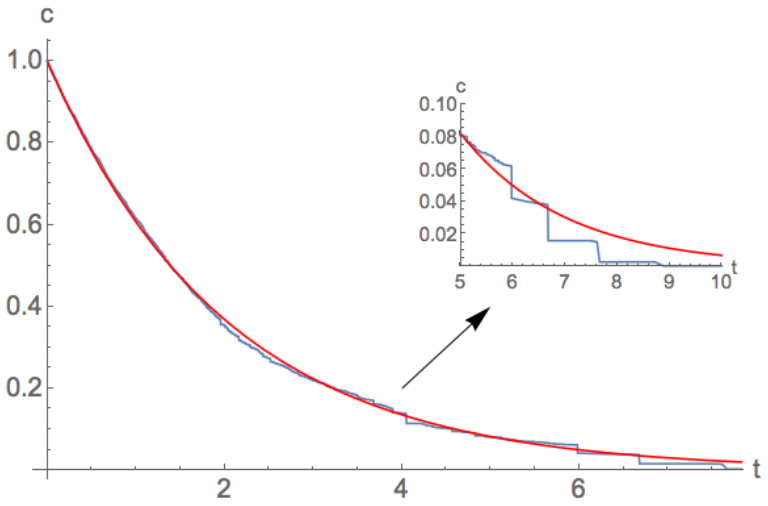

Figure 2.14: Mean-field prediction (red curve) and simulated $c(t)$ (blue curve) for $N = 5000$ oscillators. As can be seen in the inset, the simulation results get noisy for large t.

CHAPTER 3

TRANSIENT DYNAMICS OF PULSE-COUPLED OSCILLATORS WITH NONLINEAR CHARGING CURVES

Bionn gach tosach lag.

(You have to learn to crawl before learn to walk) – Irish proverb

3.1 Introduction [1]

During each heartbeat, thousands of pacemaker cells discharge in concert. This collective firing causes the contraction of cardiac muscles, which pump blood around the body. Should these firing fall out of step, heartbeats can become erratic, which inhibits blood flow. In order to maintain healthy heart function, the pacemaker cells must maintain their synchronous firing.

In 1975, Peskin gave the first mathematical analysis of the pacemaker as a self-synchronizing system [37]. He modeled the pacemaker cells as leaky 'integrate-and-fire' oscillators that communicate with each other by firing sudden impulses. He then conjectured that a population of identical leaky oscillators with all-to-all pulsatile coupling would self-organize into synchrony for all $N \geq 2$ and for almost all initial conditions. Mirollo and Strogatz [38] later proved this conjecture.

[1] This chapter is reproduced from: O'Keeffe, Kevin P. "Transient dynamics of pulse-coupled oscillators with nonlinear charging curves." Physical Review E 93.3 (2016): 032203.

Since then, pulse-coupled oscillators have been used as models in many other contexts, for example, sensor networks [13], low-powered radio transmission [14], firing neurons [39, 40], earthquakes [41], and economic booms and busts [42]. For greater realism, the associated theoretical work relaxes Peskin's original assumptions, by allowing for example local coupling in lattices [43–45] or networks [46–49]. These effects lead to new phenomena, such as traveling waves, self-organized criticality, partial synchrony, and coexistence. The inclusion of interactions with delays and different sign [50–52] have also been considered, which give rise to multi-stable clustering, transient clustering, phase-lagged synchronization.

Yet even within the simplified context of Peskin's all-to-all model, unanswered theoretical questions remain. In particular, little is known about transient dynamics: in a self-synchronizing system, what does the buildup to synchrony look like? A first step in this direction was presented in [53]. It was shown that synchrony developed through clustering; oscillators start to synchronize in small groups, which grow steadily larger over time. Using tools from aggregation theory [10], this clustering was described quantitatively. In the analysis, it was assumed that each oscillator had a linear charging curve. This idealization is appropriate for electronic oscillators such as those in sensor networks, but not for biological oscillators, like the aforementioned cardiac pacemaker cells or firing neurons. We here extend the analysis in [53] to explore the manner in which these more complicated oscillators achieve synchrony.

3.2 The model

We consider $N \gg 1$ identical oscillators coupled all-to-all. Each oscillator is characterized by a voltage-like state variable x_i, which increases from a baseline value of 0 to a threshold set to 1, according to

$$\dot{x}_i = S_0 - \gamma x_i. \tag{3.1}$$

When an oscillator reaches threshold it does two things: (i) It fires a pulse of size $1/N$. This pulse is received by all other oscillators instantaneously, causing them to discontinuously raise their voltage from x_j to $\min(x_j + 1/N, 1)$. This way, oscillators never exceed the threshold value of 1. To avoid complications with chain reactions of firing oscillators, we assume any oscillators which reach threshold by receiving a pulse, do not themselves fire until the *next* time they reach threshold. (ii) The firing oscillator then resets its voltage to 0, along with any secondary oscillators that were brought to threshold. These oscillators then begin their next cycle synchronized.

If $j > 1$ oscillators reach threshold together, each one fires, so that the pulse has total size j/N (although we later consider other types of pulse).

We note that there is some parameter redundancy, since by rescaling time we could set $S_0 = 1$ without loss of generality. For reasons that will become clear later, a different choice of S_0 is more convenient, so we leave it as a free pa-

rameter for now. We remark however that S_0 must be chosen so that $\dot{x}_i > 0$ for $0 \leq x_i \leq 1$.

3.3 Results

Assume the initial voltages of the oscillators are drawn uniformly at random. How will the dynamics unfold? At the beginning, the oscillators simply increase their voltage according to $\dot{x}_i = S_0 - \gamma x_i$. Then the first oscillator reaches threshold, fires a pulse, and perhaps brings some other oscillators to threshold. As described, these oscillators begin their next cycle in step. The primary, firing oscillator, and the secondary oscillators it incited to threshold, form a synchronous *cluster*.

As time goes on, other oscillators start firing pulses and start absorbing oscillators which are close enough to threshold. More clusters of synchronized oscillators emerge. In turn, these clusters start reaching threshold and absorbing *other* clusters, growing progressively larger. We note that clusters can only ever increase in size. They can never break apart because (a) the oscillators are identical, and therefore sync'd oscillators have the same speed, and (b) all oscillators receive the same number of pulses (thanks to the global coupling).

The picture is now clear; the system synchronizes through an aggregation phenomenon. Clusters of sync'd oscillators form and get steadily bigger by coalescing with each other. At any time t therefore, there are clusters of various sizes.

Let $N_j(t)$ denote the number of clusters of size j at time t: N_1 is the number of singletons, N_2 is the numbers doublets, and so on. These N_j are correlated random quantities. They are correlated because oscillators belonging to clusters of one size are unavailable to clusters of another size, and they are random because of the initial conditions.

To analyze the system's dynamics, we imagine $N_j \gg 1$ so that fluctuations from different realizations of the system are small. Of course, this condition cannot be satisfied for every j, at all t. For example, at the final stages of the process, there will be a small number of very large clusters. We therefore restrict our attention to the portion of the process where $N_j \gg 1$ is approximately true – the opening and middle game, as opposed to the end game.

But how does the end game play out? That is, how does the process terminate? Strogatz and Mirollo [38] showed that for $\gamma > 0$ and pulse size $> 1/N$, then full sync is guaranteed for all IC except for a set of measure zero; the clustering continues until there is one giant cluster of size N. For other values of γ and other pulse sizes, full sync is possible, but not certain to occur.

In this work, we focus only on the transient dynamics, the evolution *to* synchrony. So from now on we implicitly assume we in the early and middle stages of the process, where $N_j \gg 1$ is a valid approximation. We then use ensemble averages to define the *individual cluster densities*,

$$c_j := \langle N_j \rangle / N. \tag{3.2}$$

We then make the following strong assumptions: (i) fluctuations about the ensemble averages are small, $N^{-1}N_j = c_j + O(N^{-1/2})$, and that (ii) different cluster densities are asymptotically uncorrelated, $N^{-2}N_iN_j = c_ic_j + O(N^{-1/2})$.

We can use these c_j to define a disorder parameter for our system. This is the *total cluster density*,

$$c = \sum_{j=1}^{N} c_j \tag{3.3}$$

where the index runs over all cluster *sizes*, which range from 1 to N. This density is a measure of the total fragmentation of the system, which we interpret as a kind of disorder. To see this, consider that at $t = 0$, there are N singletons, so $c_1 = 1$, and $c_j = 0,\ \forall j \neq 1$. This means that $c(0) = 1$, correctly identifying that the system begins maximally disordered. At the other extreme as $t \to \infty$, we know there is one giant cluster of size N, so $c = 1/N \approx 0$ for large N. Hence c decreases from 1 to 0 as the system evolves from complete disorder to full synchrony.

3.3.1 Total Cluster Density

We first analyse c. It obeys the following rate equation, where R_i is the rate at which clusters of size i fire, and L_i is the number of clusters absorbed during such a firing, for $i = 1, \ldots, N$ (i.e. over all cluster *sizes*):

$$\dot{c} = -\sum_i R_i(t) L_i(t) \tag{3.4}$$

To find $L_i(t)$, we first define the 'voltage-density' $\rho_j(x,t)dx$ to be the number of j-clusters with voltage between x and $x + dx$ at time t. This has the normalization condition $\int_0^1 \rho_j(x,t)dx = N_j$. Now, when an i-cluster fires, all clusters on the interval $[1 - i/N, 1)$ will be absorbed. This means,

$$L_i(t) = \sum_j \int_{1-i/N}^{1} \rho_j(x,t)dx. \tag{3.5}$$

We digress briefly to comment on difficulties imposed by voltage density $\rho_j(x,t)$. The nonlinearities in the oscillators' charging curves make $\rho_j(x,t)$ behave non-trivially. It is this key fact which makes nonlinear charging curves much harder to analyze than linear ones. In the linear case, (i.e. when $\gamma = 0$), as considered in [53], $\rho_j(x,t)$ is well approximated by a uniform density. This substantially simplified the calculation of L_i and R_i, and in turn the remainder of the analysis. For instance, (3.5) reduces to simply $L_i = \sum_j (i/N)N_j = ic$. But when $\gamma \neq 0$, $\rho_j(x,t)$ is an unknown quantity which obeys a complicated PDE. Approximately solving this PDE is a key result of the paper.

We now return to our calculation of L_i. To proceed, we make an approximation. As stated earlier, we are only interested in transient time scales – the opening and middle game. In this regime, most clusters will be small relative to the system size: $j \ll N$. This lets us approximate the integral above, $\int_{1-i/N}^{1} \rho_j(x,t)dx \approx$

$(i/N)\rho_j(x = 1, t)$. Of course, this approximation will get worse as time goes on. We discuss this further in Section 3.4. Our expression for L_i is then

$$L_i(t) = \frac{i}{N} \sum_j \rho_j(x = 1, t). \tag{3.6}$$

To continue the analysis, we need to find $\rho_j(x, t)$. As mentioned, its behavior is complicated so we defer its calculation, and instead find the firing rate R_i. Naively, one might think that this is simply the flux of i-clusters at threshold: $N^{-1}(\rho_i v)|_{x=1}$ (where N^{-1} is required, since R_i measures the rate of firing of c_i, not N_i). However not every cluster that reaches threshold gets the chance to fire, since some will be absorbed. To account for this effect, we decompose the rate into

$$R_i = R_i^0 - R_i^a. \tag{3.7}$$

The term R_i^0 is a 'background' firing rate, where we pretend all oscillators get to fire even if they are absorbed. R_i^a is the rate at which i-clusters are *being* absorbed by other clusters of various sizes, and hence deprived of their chance to fire.

We start with R_i^0. To be clear, by background firing rate, we mean the rate i-clusters would fire at, if *every* oscillator fired a pulse when it reached threshold. That is, imagine relaxing our imposition that any secondary oscillators that reach threshold by virtue of a pulse do not fire. In that case,

$$R_i^0 = N^{-1}(\rho_i v)\Big|_{x=1}. \tag{3.8}$$

The speed v of each cluster is non-trivial. This is because in addition to its natural speed $v_0 = \dot{x} = S_0 - \gamma x$, each oscillator receives a steady stream of pulses from firing clusters which increase its voltage:

$$v(x,t) = v_0(x) + v_{pulse}(t). \tag{3.9}$$

This "pulse velocity" due to the firing of just j-clusters will be (absolute number of pulses per sec) $\times$ (distance per pulse). Since R_j is the firing rate of c_j, $R_j N$ is the absolute number of pulses, while the distance per pulse is j/N. To find the total pulse speed we then sum over all j-clusters: $\sum_j (NR_j)(j/N)$, giving

$$v_{pulse}(t) = \sum_j jR_j(t). \tag{3.10}$$

Our next target is the absorption rate R_i^a. The calculation is similar to finding L_i, and is given by $R_i^a = \sum_j R_j \int_{1-j/N}^1 \rho_i(x,t)dx$, which after approximating the integral as before gives,

$$R_i^a = \sum_j R_j(j/N)\rho_i(x=1,t). \tag{3.11}$$

Substituting R_i^0 and R_i^a into (3.7) finally gives

$$R_i = \frac{S_0 - \gamma}{N} \rho_i(x = 1, t). \tag{3.12}$$

We now analyze $\rho_j(x, t)$. In principle, it satisfies the the continuity equation with appropriate terms for the absorption of j-clusters at threshold, and the formation of j-clusters from smaller clusters:

$$\dot{\rho_j} + \partial_x(v\rho_j) + Absorption + Gain = 0 \tag{3.13}$$

$$\partial_x(v\rho_j)|_{x=0} = \partial_x(v\rho_j)|_{x=1} \tag{3.14}$$

Solving this PDE is the hardest part of the analysis. The absorption and gain terms are the main problem, because they *couple* the voltage densities; through them, $\rho_j(x, t)$ depends all the other $\rho_k(x, t)$. This is because a j-cluster can be created or absorbed by the action of various combinations of other clusters. Enumerating these combinations is by itself difficult, not to mention understanding how they affect the PDE. On top of all that, there is also the non-smoothness of the oscillators' velocity at threshold (which discontinuously jumps from $\dot{x} = S_0 - \gamma$ at $x = 1$, to $\dot{x} = S_0$ at $x = 0$) to deal with.

We can however make progress by observing that the evolution of the system naturally divides into periods $\{T_n\}$. We define a period to be the time take for the full population of oscillators to complete a voltage cycle. More carefully, T_n

is earliest time when every oscillator has completed n cycles.

We then solve the continuity equation during a given period, not worrying about what happens before or after. This lets us avoid the complication of the aforementioned non-smoothness of the oscillators' behavior at the boundaries. We also neglect the absorption term. As previously discussed, when an i-cluster fires, only oscillators on $(1 - i/N, 1]$ get absorbed. This is a small interval for the 'opening' and 'middle' game we are considering. Hence the absorption term is 0 on most of $[0, 1]$ and so we omit it.

But we still have to compute the gain term. As previously discussed, this is combinatorially intensive (we explicitly compute this term later, when calculating the individual cluster densities). However, we can neglect this cumbersome term entirely, by making the following key observation.

Looking at equations (3.6), (3.12), we see our desired quantities R_j and L_j depend only on the density of clusters at threshold: $\rho_j(x = 1, t)$. Therefore, *during each period*, R_i and L_i are only affected by j-clusters which existed *at the start of that period*, which we call 'original' j-clusters. This is because any 'new' j-clusters won't reach threshold until the *next* period. By 'new', we mean (a) j-clusters that fired during a period, didn't absorb any other clusters, and so returned to threshold, and (b) any j-clusters that were created by the firing and absorption of other smaller clusters.

So for the purposes of calculating $\rho_j(x = 1, t)$ *during a given period*, there is a 'lightcone' between original and new j-clusters. We therefore need to solve the

continuity equation for the original j-clusters only, for which the gain term is zero. The problem is then given by (3.15) below, where $v(x,t)$ is given by (3.9), $f_0(x)$ is the initial distribution of $\rho_j^{original}$, and the Heaviside functions $H(x), H(1-x)$ are included to confine the I.C. to the interval $[0,1]$.

$$\begin{aligned} &\dot{\rho_j}^{original} + \partial_x(v\rho_j^{original}) = 0 \\ &\rho_j^{original}(x,0) = f_0(x)H(x)H(1-x) \end{aligned} \tag{3.15}$$

We don't yet know the speed $v(x,t)$. However its structure, $v(x,t) = v_0(x) + v_{pulse}(t)$, lets us derive an approximate solution for $\rho_j^{original}(x,t)$ given by (3.16) below. The derivation of this key result and the definition of $\Gamma(x,t)$ are shown in the Appendix.

$$\rho_j^{original}(x,t) = e^{\gamma t} f_0\Big(\Gamma(x,t)\Big) H\Big(\Gamma(x,t)\Big) H\Big(1-\Gamma(x,t)\Big). \tag{3.16}$$

What all this means is, if we known $\rho_j^{original}(x,t)$ at the start of a period, then we know how it will evolve until that period ends. For later convenience, we introduce the following notation. Let $\tilde{x}$ denote that during a period, x is held fixed at its value at the start of that period: $\tilde{x} = x(t = T_n)$ for $T_n < t < T_{n+1}$.

We next make the strong assumption that clusters of all sizes are distributed uniformly in voltage on $[0,1)$ at the start of each period: $\rho_j(x, t = T_n) = \tilde{N}_j$. Then in (3.16), $f_0(x) = \tilde{N}_j$. We discuss the legitimacy of making this assumption in

Section 3.4.

The Heaviside functions make (3.16) look complicated. But really, they only enforce that the $\rho_j^{original}$ is zero behind the final j-cluster (cluster with smallest voltage), and ahead of the first j-cluster (cluster with largest voltage). We remark that as it stands, the solution (3.16) propagates into the unphysical $x \geq 1$ regime. But we of course restrict our attention to just $x \in [0, 1]$.

The behavior of $\rho_j^{original}(x, t)$ during each period is therefore simple. The density at each point x simply grows at rate $e^{\gamma t}$ until it drops discontinuously to 0, as the final 'original' j-cluster passes by. This behavior is shown in Figure 3.1.

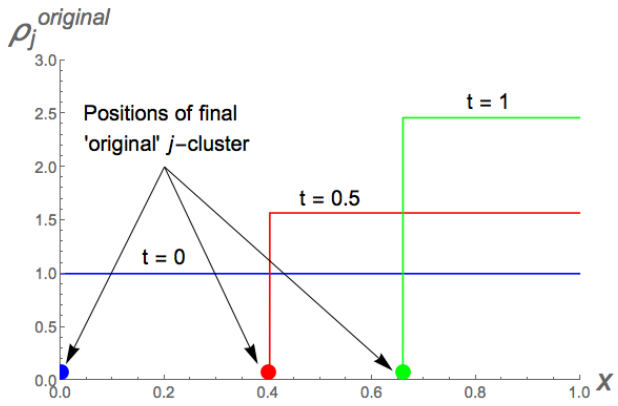

Figure 3.1: Evolution of voltage density of original j-clusters during a period, with initial condition $\rho_j^{original}(x, 0) = 1$.

Now that we have an expression for $\rho_j(x = 1, t)$, which we have argued is $\rho_j^{original}(x = 1, t)$, we can complete our expressions for L_i and R_i given by (3.6) and (3.12). We then plug the results into (3.4) to obtain our sought after rate equation for the disorder parameter $c(t)$,

$$\dot{c} = -(S_0 - \gamma)e^{2\gamma t}\tilde{c}. \tag{3.17}$$

which has solution,

$$c(t) = \frac{\tilde{c}}{2\gamma}\Big(S_0 + \gamma + e^{2\gamma t}(\gamma - S_0)\Big). \tag{3.18}$$

We restate that equations (3.17) and (3.18) are only valid during a given period. We can however use (3.18) to find $c(t)$ for all t, by stitching solutions during successive periods together.

But we still don't know the periods $\{T_n\}$ themselves. To find them, we need the speed v as per (3.9). Recalling $v_0 = S_0 - \gamma x$, and substituting R_i from (3.12), gives

$$v(x,t) = (S_0 - \gamma x) + (S_0 - \gamma)e^{\gamma t}. \tag{3.19}$$

We see that v is the same during each period (i.e. there are no 'tilde' quantities, we denote different values during different periods.). This means that the length of each period is the same: $T_n = nT_0$. We can find T_0 from $T_0(S_0, \gamma) = \int_0^1 v(x)dx$. To compare the effects of different amounts of concavity on equal footing, we want $T_0 = 1$ for every γ. We can achieve this by selecting an appropriate value for S_0, which we have strategically left as a free parameter until now. Doing the

integral, this value for S_0 is

$$S_0 = \frac{\left(e^{2\gamma} + 2e^{\gamma} - 1\right)\gamma}{(e^{\gamma} - 1)(e^{\gamma} + 3)}. \tag{3.20}$$

We must be careful when using (3.20). This is because for sufficiently negative γ, S_0 can become negative. While this choice of S_0 ensures the *total* speed $v = v_0 + v_{pulse}$ is positive, the natural speed $v_0 = \dot{x} = S_0 - \gamma x$ can become negative if S_0 is too negative. This means that the oscillators *decrease* in voltage in the absence of coupling. We avoid this unphysical regime by requiring $v_0 > 0$ for $0 \leq x \leq 1$, which leads to $\gamma_{min} \approx -0.881$.

Figure 3.2 shows the agreement between theory and simulation for $c(t)$ when $\gamma < 0$ and $\gamma > 0$. For comparison, we also show when $\gamma = 0$, which corresponds to the linear charging curve studied in [53]. In the linear case, $c(t)$ is a series of line segments whose slope decreases by a factor of 2 from period to period. But when $\gamma \neq 0$, $c(t)$ has more complicated behavior; it no longer decays linearly during each period.

As can be seen, $c(t)$ declines faster and slower when $\gamma > 0$ and $\gamma < 0$ respectively. This makes physical sense. When $\gamma > 0$, oscillators slow down as they increase in voltage, which makes them clump closer together near $x = 1$. When $\gamma < 0$, the opposite happens; clusters spread further apart closer to threshold. Now suppose a j-cluster fires. When $\gamma > 0$ the interval $[1 - j/N, 1)$ is more likely to contain oscillators than when $\gamma < 0$, thanks to the 'clumping' and 'spreading out', which in turn makes an absorption more likely. The case of zero concavity

then interpolates between these two regimes, as evidenced by Figure 3.2.

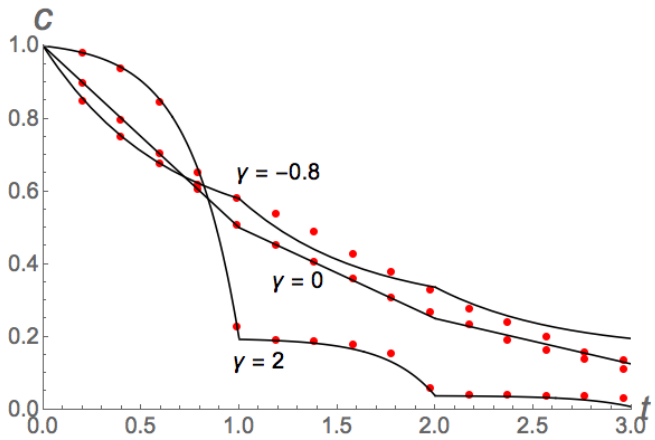

Figure 3.2: (Color online) Theoretical and simulated $c(t)$ for $\gamma = 2$, $\gamma = 0$, and $\gamma = -0.8$. Solid lines show theoretical prediction (3.18), while data points show simulated results for $N = 5 \times 10^4$ oscillators.

We can use (3.18) for c to estimate the timescale of the transient dynamics. We say transience ends when a cluster of size $\sim N$ has formed, so that $c \sim 1/N$. Our assumptions will likely break down before this, so this is best interpreted as an upper bound. Looking at (3.18), we see that after one period, c decreases by a factor of,

$$B := \frac{1}{2\gamma}\Big(S_0 + \gamma + e^{2\gamma}(\gamma - S_0)\Big) = \frac{2}{e^{\gamma} + 3}. \tag{3.21}$$

After n periods, it decreases by B^n. After some algebra, and rounding T_{trans} to the nearest period, we get,

$$T_{trans} \sim \log N / \log B^{-1} \tag{3.22}$$

3.3.2 Individual Cluster Densities

How do the individual densities c_i evolve? We begin with the 1-clusters, whose density is c_1. They are the easiest density to analyze, since they can only decrease. There are two ways this can happen: (i) the loss of a *firing* singleton, when it absorbs other clusters of any size, and (ii) the loss of *absorbed* singletons, due to the firing of another cluster:

$$\dot{c_1} = \mathcal{L}_1^{firing} + \mathcal{L}_1^{absorbed}. \tag{3.23}$$

We begin with $\mathcal{L}_1^{firing}$. From (3.12) we know singletons fire at rate $R_1 = (S_0 - \gamma)e^{\gamma t}\tilde{c_1}$. During such a firing, an absorption will take place if there is at least one cluster on $[1 - 1/N, 1)$. This interval contains on average $Nc \times 1/N = c(t) = \tilde{c}e^{\gamma t}$ clusters. Further, the probability that it contains n clusters is given by the Poisson distribution: $\Pi_n = \frac{(\tilde{c}e^{\gamma t})^n}{n!}e^{-\tilde{c}e^{\gamma t}}$. This is the mathematical statement that the clusters are distributed randomly without correlations. The probability that $[1-1/N, 1)$ is occupied by at least one cluster is therefore $1-e^{-\tilde{c}e^{\gamma t}}$. If an absorption takes place, N_1 decreases by 1, since we're only considering the loss of the *firing* oscillator here. The expected loss rate is then $(S_0-\gamma)e^{\gamma t}\tilde{c_1}[1\times(1-e^{-\tilde{c}e^{\gamma t}})+0\times e^{-\tilde{c}e^{\gamma t}}]$, leading to,

$$\mathcal{L}_1^{firing} = (S_0 - \gamma)e^{\gamma t}\tilde{c_1}(1 - e^{-\tilde{c}e^{\gamma t}}). \tag{3.24}$$

To calculate $\mathcal{L}_1^{absorbed}$, imagine a j-cluster fires and absorbs all the singletons on the interval $[1-j/N, 1)$. As before, this interval will have on average $Nc_1(t)\times j/N = jc_1e^{\gamma t}$ such singletons. Multiplying this by R_j and summing over j then gives $\sum_j(1-\gamma)\tilde{c}_j e^{\gamma t} \times j\tilde{c_1}e^{\gamma t}$, which leads to

$$\mathcal{L}_1^{absorbed} = (1-\gamma)e^{2\gamma t}\tilde{c}_i \tag{3.25}$$

Substituting $\mathcal{L}_1^{firing}$ and $\mathcal{L}_1^{absorbed}$ into (3.23) gives,

$$\dot{c_1} = -(S_0-\gamma)\tilde{c_1}\Big[(1+e^{\gamma t}) - e^{-\tilde{c}e^{\gamma t}}\Big]. \tag{3.26}$$

This looks intimidating, but since the quantities $\tilde{c}_i$ are held constant over each period, the R.H.S. is a function of only t. It therefore has an analytic solution, which we show plotted in Figure 3.3.

Will larger clusters behave similarly? They differ from the singletons in that they can be created as well as absorbed, which makes them harder to calculate. Their absorption rate is easily generalized from that of the singletons:

$$\mathcal{L}_i^{firing} + \mathcal{L}_i^{absorbed} = (S_0-\gamma t)\tilde{c}_i\Big[(1+e^{\gamma t}) - e^{-i\tilde{c}e^{\gamma t}}\Big]. \tag{3.27}$$

Their gain rate is calculated as follows. An i-cluster is created when a cluster

of size $k < i$ fires, and absorbs the right combination of other clusters. Suppose there are a_1 1-clusters, a_2 2-clusters, ..., on the interval $[1 - k/N, 1)$. If $a_1 + 2a_2 + \cdots + k = i$, then an i-cluster will be created. Such a combination occurs with probability $\frac{(kc_1)1_1^a}{a_1!}e^{-kc_1} \times \frac{(kc_2)1_2^a}{a_2!}e^{-kc_2} \times \ldots$. Summing first over all such combinations, and then over all k, gives an expected rate gain of

$$\sum_{k=1}^{i-1}(S_0 - \gamma)\tilde{c_k}e^{\gamma t}e^{-k\tilde{c}e^{\gamma t}} \sum_{a_1+2a_2+\cdots=i-k}\left(\prod_{p\geq 1}\frac{(k\tilde{c_p}e^{\gamma t})^{a_p}}{a_p!}\right) \tag{3.28}$$

After combining the loss and gain terms, and some algebraic manipulation, we finally obtain the desired rate equation for i-clusters,

$$\begin{aligned}\dot{c_i} = &-(S_0 - \gamma)e^{\gamma t}(1 + e^{\gamma t})\tilde{c_i} + \\ &\sum_{k=1}^{i}(S_0 - \gamma)\tilde{c_k}e^{\gamma t}e^{-k\tilde{c}e^{\gamma t}} \sum_{\sum pa_p=i-k}\left(\prod_{p\geq 1}\frac{(k\tilde{c_p}e^{\gamma t})^{a_p}}{a_p!}\right)\end{aligned} \tag{3.29}$$

This is a set of recursive equations, and so we can solve them successively. As with c_1, the R.H.S. is a pure function of t, so analytic solutions are findable. Figure 3.3 shows theoretical predictions versus simulation results for c_1 through c_4 when $\gamma = 0.9$. We remark that the effect of a nonlinear versus linear charging curve on the c_i is the same as that for the disorder parameter c: it causes them to no longer decay/grow linearly during each period (note we do not show c_i for the linear charging curve, $\gamma = 0$, for illustrative purpose. See Fig. 6 in the supplemental materials of [53])

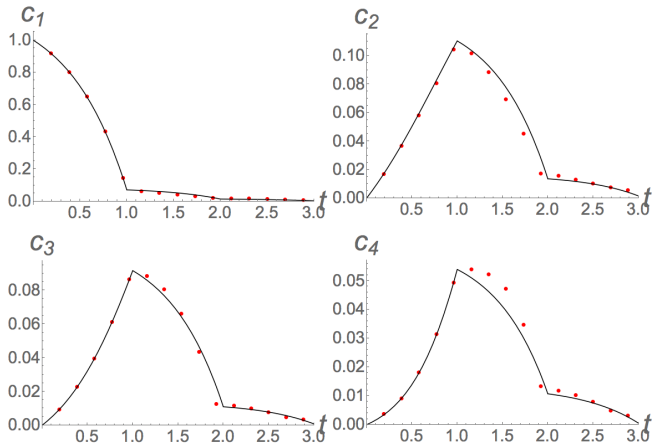

Figure 3.3: (Color online) Theoretical and simulated cluster densities c_1 though c_4 for $\gamma = 0.9$. Solid black lines show analytic solutions to (3.29). Red data points show simulation results for 5×10^4 oscillators.

3.3.3 Alternate Coupling Rules

We thus far assumed an i-cluster fired a pulse of size i/N. We now consider two alternatives. The first is simply the original pulse strength with a tunable strength K: $(Ki)/N$. The second is a fixed pulse strength of K/N regardless of the size of the firing cluster. These alterations only modestly change the analysis, so we simply list the results for L_i, R_i, v_{pulse} and c in the table below, where S_0 is given by (3.20). For illustrative purposes we do not include a formula for c_i, but its calculation is straightforward.

	Variable Pulse: Kj/N	Fixed Pulse: K/N
L_i	$Ki\tilde{c}e^{\gamma t}$	$K\tilde{c}e^{\gamma t}$
R_i	$(S_0-\gamma)\tilde{c}_i e^{\gamma t}$	$(S_0-\gamma)\tilde{c}_i e^{\gamma t}$
v_{pulse}	$K(S_0-\gamma)e^{\gamma t}$	$K(S_0-\gamma)e^{\gamma t}\tilde{c}$
$\dot{c}$	$-K(S_0-\gamma)e^{2\gamma t}\tilde{c}$	$-K(S_0-\gamma)e^{2\gamma t}\tilde{c}^2$
$c(t)$	$\frac{\tilde{c}\left(K(\gamma-S_0)(e^{2\gamma t}-1)+2\gamma\right)}{2\gamma}$	$\frac{\tilde{c}\left(\tilde{c}K(\gamma-S_0)(e^{2\gamma t}-1)+2\gamma\right)}{2\gamma}$

As can be seen, there are mostly only minor differences between the two cases.

The first thing to note is that $c(t)$ decays more slowly with a fixed pulse strength K/N. Intuitively, this is because large and small clusters now fire with the same strength, which means they absorb all clusters on the fixed interval $[1 - K/N, 1)$. In contrast, for a pulse strength Kj/N, bigger clusters fire bigger pulses, and therefore absorb clusters on an interval proportional to their size: $[1 - Kj/N, 1)$. This is mathematically manifested as $L_i^{variable\ pulse}$ being bigger than $L_i^{fixed\ pulse}$ by a factor of i, which leads to $\dot{c}^{variable\ pulse}$ being bigger than $\dot{c}^{fixed\ pulse}$ by a factor of $\tilde{c}^{-1}$ (remember, $0 \leq c \leq 1$). In turn, this means $c(t)^{fixed\ pulse}$ decays more slowly than $c(t)^{variable\ pulse}$

Also note that v_{pulse} depends on $\tilde{c}$ for the fixed pulse case. The mechanism discussed above is also at play here: since there are fewer clusters in successive periods, and the pulse per cluster is constant, the total 'current' per period will get smaller. This is in contrast to the pulse = Kj/N case, where there are fewer clusters per period also, but larger clusters fire larger pulses, keeping the total 'current' per period constant. A consequence of this decrease in v_{pulse} is that the periods won't be constant for pulse = K/N, as there are for Kj/N. They will get longer as v_{pulse} decreases from period to period.

3.4 Breakdown of Approximations

3.4.1 Uniformity Assumption

We now discuss the approximations and assumptions we made in our analysis. The first of these was that each cluster density was distributed uniformly in voltage at the start of each period, $\rho_j(x,t) = \tilde{N}_j$. From this, we derived equations (3.6) and (3.12) for L_i and R_i, which in turn led us to our disorder parameter c.

This uniformity assumption clearly cannot be satisfied for each i, at every t. For instance, consider the end of the first period. Perhaps mostly clusters of size < 5 were formed, with only a few larger clusters of size > 10. Then, $\rho_{j<5}(x, t = 1)$ will be approximately uniform, but $\rho_j(x, t = 1)$ will be more sharply peaked. So the uniformity assumption is inaccurate for large clusters, which are few in number. This explains why (3.18) approximates $c(t)$ well. Since $c(t) = \sum_j c_j$, we see that the sum will be dominated by those c_i which are large, for which the uniformity assumption is accurate.

The fact that the uniformity assumption worsens for larger clusters also means that our results for c_i should get worse for larger i. Figure 3.4 below shows that this is indeed the case.

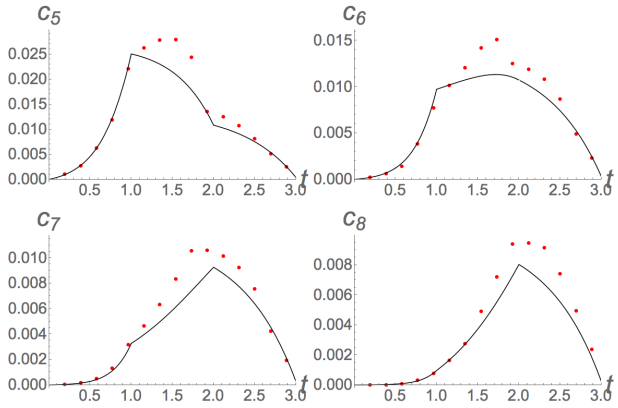

Figure 3.4: (Color online) Theoretical and simulated cluster densities c_5 though c_8 for $\gamma = 0.9$. Solid black lines show analytic solutions to (3.29). Red data points show simulations results for 5×10^4 oscillators. As can be seen, theory and simulation start to disagree

3.4.2 Final stages of process

Throughout our analysis, we assumed $N_j \gg 1$. As discussed, this cannot be true $\forall j$, at every t. This assumption is most blatantly incorrect at the end of the process, where there are a small number of macroscopic clusters. Our results should thus substantially disagree with simulation for large t, as is evidenced by Figure 3.5.

3.5 Conclusion

We have studied the transient dynamics of pulse coupled oscillators with non-linear charging curves. We derived approximations for the total cluster density

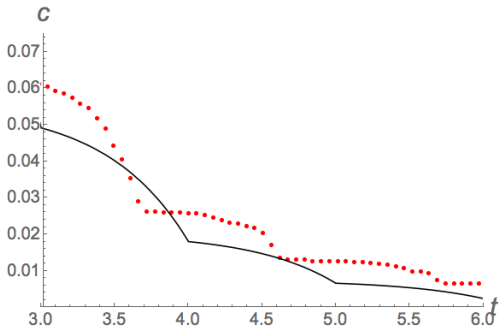

Figure 3.5: (Color online) Theoretical and simulated total cluster density $c(t)$ for $\gamma = 0.9$ and $t > 3$. Solid black lines show analytic solution (3.2). Red data points show simulation results for $N = 10^4$ oscillators. There is a significant disagreement between theory and simulation for later times, when the approximations we made in the analysis breakdown.

$c(t)$ and individual cluster densities $c_i(t)$. These approximations were good up to the final stages of the process, where the assumptions made in the analysis breakdown.

Our work could be used to understand clustering in other systems of pulse-coupled oscillators. For instance, Ernst et al [50] reported multi-stable clusters for all-to-all, inhibitory coupling with delays. They found the average number N_c of clusters obeyed $N_c \sim \tau^{-1}$, where τ is the delay. Perhaps adjustments to our analysis could analytically recover this result; the pulse velocity (3.10) could be made negative to account for the inhibitory coupling, and 'delayed' versions of equations (3.5), (3.7) for the firing and loss rates R_i, L_i could be derived.

Furthermore, Mauroya and Sepulcher [49] studied the long term behavior of the system (3.1) with $\gamma > 0$: the complement to our 'opening' and 'middle' game. They analytically determined the final number of synchronized clusters formed (we remind the reader that when $\gamma > 0$, full synchrony is not *guaranteed*

to occur, and so multiple, stable clusters are possible). Perhaps our transient analysis could be united with their steady state results to characterize the full evolution of the Peskin model.

Another possible application of our results is in network detection. Gomez-Gardenes et al [54] showed that transient clustering in the Kuramoto model can be used to approximate the underlying network structure. Could our results could be used to the same effect in networks of pulse-coupled oscillators? Local coupling would however mean that clusters could break apart as well as coalesce. One could account for this by including additional loss terms in our rate equations for c and c_i, (3.4), (3.29).

Our model has several idealizations that could be relaxed in future work. For example, local coupling, delayed coupling, and heterogeneity in oscillator speeds and pulse size could be studied. Another modification would be to allow chain reactions, by permitting any clusters that are brought to threshold by another firing cluster, to fire themselves.

3.6 Acknowledgments

This research was supported in part by the National Science Foundation through Grant No. DMS-1513179. We thank Steven Strogatz and Paul Krapivsky for helpful discussions.

3.7 Appendix

We here approximate the density $\rho_j^{original}(x,t)$. For convenience, we drop the superscript 'original'. As shown in the main body of the text, the density solves equation (3.30) below,

$$\begin{aligned} &\dot{\rho_j} + \partial_x(v\rho_j) = 0 \\ &\rho_j(x,0) = f_0(x)H(x)H(1-x) \end{aligned} \tag{3.30}$$

where, $f_0(x) = \tilde{N_j}$ (since we are assuming a initial uniform distribution), and

$$v(x,t) = v_0(x) + v_{pulse}(t). \tag{3.31}$$

While we know $v_0(x) = S_0 - \gamma x$, we don't yet have a complete expression for $v_{pulse}(t)$. In the main text, we derived $v_{pulse} = \sum_j jR_j$, which using (3.12) for R_i gives

$$v_{pulse}(t) = \sum_j \frac{S_0 - \gamma}{N} j\, \rho_j(x=1,t) \tag{3.32}$$

This is the source of our difficulty. Our PDE for $\rho_j(x,t)$ depends on v_{pulse}, which

in turn depends on the voltage density for *every other* cluster size $\rho_k(x,t)$. To overcome this difficulty, we use a technique similar to the 'leap-frog' or 'split' method used in certain numerical schemes. This involves making a series of recursive approximations for v_{pulse} and ρ_j:

$$v_{pulse} = \left(v_{pulse}^{(0)}, v_{pulse}^{(1)}, v_{pulse}^{(2)}, \dots\right) \tag{3.33}$$

$$\rho_j = \left(\rho_j^{(0)}, \rho_j^{(1)}, \rho_j^{(2)}, \dots\right) \tag{3.34}$$

Graphically, our scheme is given by the following, where we have placed the labels of equations used to make the approximations over the arrows.

$$v_{pulse}^{(0)} \xrightarrow{(3.30)} \rho_j^{(0)} \xrightarrow{(3.32)} v_{pulse}^{(1)} \xrightarrow{(3.30)} \rho_j^{(1)} \xrightarrow{(3.32)} + \dots \tag{3.35}$$

We begin by setting $v_{pulse}^{(0)} = 0$. The speed is then,

$$v(x,t)^{(0)} = v_0(x) + 0 = S_0 - \gamma x. \tag{3.36}$$

We plug this into (3.30) and solve for resulting PDE for $\rho_j^{(0)}(x,t)$. This has solution,

$$\rho_j^{(0)}(x,t) = e^{\gamma t}\tilde{N}_j H\Big(\Gamma_0(x,t)\Big)H\Big(1-\Gamma_0(x,t)\Big). \tag{3.37}$$

with $\Gamma_0(x,t) = \gamma^{-1}[S_0 - (S_0 - \gamma x)e^{\gamma t}]$. We then use $\rho_j^{(0)}$ to find $v_{pulse}^{(1)}$ using (3.32), which gives

$$v_{pulse}^{(1)} = (S_0 - \gamma)e^{\gamma t}\tilde{c}. \tag{3.38}$$

This completes the first step of our scheme. We then repeat the process to find $\rho_j^{(1)}$ and $v_{pulse}^{(2)}$. We use $v_{pulse}^{(1)}$ to update the speed,

$$\begin{aligned} v(x,t)^{(1)} &= v_0(x) + v_{pulse}^{(1)} \\ &= (S_0 - \gamma x) + (S_0 - \gamma)e^{\gamma t}\tilde{c}. \end{aligned} \tag{3.39}$$

and then plug this into (3.30) to obtain a PDE for $\rho_j^{(1)}$, which we solve to get,

$$\rho_j^{(1)}(x,t) = e^{\gamma t}\tilde{N}_j H\Big(\Gamma_1(x,t)\Big)H\Big(1 - \Gamma_1(x,t)\Big) \tag{3.40}$$

where $\Gamma_1(x,t) = \Big[\frac{3S_0-\gamma}{2\gamma} + \frac{e^{\gamma t}}{2\gamma}(2\gamma x - 2S_0) + \frac{e^{2\gamma t}}{2\gamma}(\gamma - S_0)\Big]$.

Looking at (4.78) and (3.40), we see that $\rho_j^{(0)}$ and $\rho_j^{(1)}$ have the same functional form. They only differ in the arguments of the Heaviside function: $\Gamma_0(x,t) \neq \Gamma_1(x,t)$. This remarkable similarity between $\rho_j^{(0)}$ and $\rho_j^{(1)}$ has an important consequence: it 'closes' our approximation scheme. We see this by substituting $\rho_j^{(1)}$ into (3.32) to find,

$$v^{(2)}_{pulse} = (S_0 - \gamma)e^{\gamma t}\tilde{c} = v^{(1)}_{pulse}, \tag{3.41}$$

which implies that $\rho^{(2)}_j = \rho^{(1)}_j$, which in turn implies our scheme terminates at $(v_{pulse}, \rho_j) = (v^{(2)}_{pulse}, \rho^{(1)}_j)$. Our final approximations for v_{pulse} and $\rho_j(x,t)$ are then,

$$v_{pulse}(t) \approx (S_0 - \gamma)e^{\gamma t}\tilde{c} \tag{3.42}$$

$$\rho_j(x,t) \approx e^{\gamma t}\tilde{N}_j H\Big(\Gamma(x,t)\Big)H\Big(1 - \Gamma(x,t)\Big) \tag{3.43}$$

with $\Gamma(x,t) = \Big[\frac{3S_0-\gamma}{2\gamma} + \frac{e^{\gamma t}}{2\gamma}(2\gamma x - 2S_0) + \frac{e^{2\gamma t}}{2\gamma}(\gamma - S_0)\Big]$.

This concludes our analysis. We state bluntly that our approach is not rigorously justified. Its legitimacy is supported only by the agreement between our analytic results and numerical simulation. We hope future work will elucidate the cause of its efficacy.

CHAPTER 4

SWARMALATORS: OSCILLATORS THAT SYNC AND SWARM

Truth is error burned up.

Norman O Brown

4.1 Introduction [1]

This year marks the fiftieth anniversary of a breakthrough in the study of synchronization. In 1967, Winfree proposed a coupled oscillator model for the circadian rhythms that underlie daily cycles of activity in virtually all plants and animals [55]. He discovered that above a critical coupling strength, synchronization breaks out spontaneously, in a manner reminiscent of a phase transition. Then Kuramoto simplified Winfree's model and solved it exactly [56], leading to an explosion of interest in the dynamics of coupled oscillators [12, 57, 58]. Kuramoto's model in turn has been generalized to other large systems of biological oscillators, such as chorusing frogs [59], firing neurons [60–63], and even human concert audiences clapping in unison [64]. The analyses often borrow techniques from statistical physics, such as mean-field approximations, renormalization group analyses [65, 66], and finite-size scaling [67, 68]. There has also been traffic in the other direction, from biology back to physics. For example, insights from biological synchronization have shed light on neutrino oscillations

[1] This chapter is reproduced from: O'Keeffe, Kevin P., and Steven H. Strogatz. "Swarmalators: Oscillators that sync and swarm." arXiv preprint arXiv:1701.05670 (2017).

[69], phase locking in Josephson junction arrays [70], the dynamics of power grids [71, 72], and the unexpected wobbling of London's Millennium Bridge on opening day [73].

A similarly fruitful interplay between physics and biology has occurred in the study of the coordinated movement of groups of animals. Fish schools, bird flocks, and insect swarms [74–78] have been illuminated by maximum entropy methods [79], agent-based simulations [80], and analytically tractable models based on self-propelled particles [81], and continuum limits [82–85].

Studies of swarming and synchronization have much in common. Both involve large, self-organizing groups of individuals interacting according to simple rules. Both lie at the intersection of nonlinear dynamics and statistical physics. Nevertheless the two fields have, by and large, remained disconnected. Studies of swarms focus on how animals move, while neglecting the dynamics of their internal states. Studies of synchronization do the opposite: they focus on oscillators' internal dynamics, not on their motion. In the past decade, however, a few studies of "mobile oscillators," motivated by applications in robotics and developmental biology, have brought the two fields into contact [86–90]. Even so, the assumption has been that the oscillators' locations affect their phase dynamics, but not conversely. Their motion has been modeled as a random walk or as externally determined, without feedback from the oscillators' phases.

We suspect that somewhere in nature and technology there must be mobile oscillators whose phases affect how they move. For instance, many species of frogs, crickets, and katydids call periodically, and synchronize in vast choruses

[59, 91–93]. When the individuals hop around, do they tend to move toward or away from others, depending on the relative phases of their calling rhythms? If so, what spatiotemporal patterns would we expect?

A clue comes from the physics of magnetic colloids [7, 94, 95] and microfluidic mixtures of active spinners [96, 97], both of which show rich collective behavior. In these systems, the particles or spinners attract or repel one another, depending on their orientations. Given that orientation is formally analogous to the phase of an oscillation (both being circular variables), a similarly rich phenomenology is expected for mobile oscillators whose phases affect their motion. We call these hypothetical systems *swarmalators* because they generalize swarms and oscillators.

One possible instance of a swarmalator system is a population of myxobacteria, modeled in 2001 by Igoshin and colleagues [98]. The movements of these bacteria in space are thought to be influenced by an internal, biochemical degree of freedom, which appears to vary cyclically. Igoshin et al. [98] modeled it as a phase oscillator. Experimental evidence suggests that the evolution of this phase is influenced by the spatial density of neighboring cells; thus there appears to be a bidirectional coupling between spatial and phase dynamics, as required of swarmalators.

Tanaka and colleagues also made an early contribution to the modeling of swarmalators [99, 100]. They analyzed a broad class of models in the hope of finding phenomena which were not system-specific. They considered chemotactic oscillators, whose movements in space are mediated by the diffusion of a back-

ground chemical. The oscillators' consumption of this chemical depends on their internal states, thereby completing the bidirectional space-phase coupling. Tanaka et al. [99, 100] began with a general model with these ingredients, from which they derived a simpler model by means of center manifold and phase-reduction methods.

Here we take a bottom-up approach. We propose a simple model of a swarmalator system which lets us study some of its collective states analytically. We hope our work will draw attention to this class of problems, and stimulate the discovery and characterization of natural and technological systems of swarmalators.

4.2 The model

We consider swamalators free to move in the plane. The governing equations are

$$\dot{\vec{x}}_i = \vec{v}_i + \frac{1}{N}\sum_{j=1}^{N}\left[\vec{I}_{att}(\vec{x}_j - \vec{x}_i)F(\theta_j - \theta_i) - \vec{I}_{rep}(\vec{x}_j - \vec{x}_i)\right], \tag{4.1}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} H_{att}(\theta_j - \theta_i)G(\vec{x}_j - \vec{x}_i) \tag{4.2}$$

for $i = 1, \ldots, N$, where N is the population size, $\vec{x}_i = (x_i, y_i)$ is the position of the i-th swarmalator, and θ_i, ω_i and $\vec{v}_i$ are its phase, natural frequency, and self-propulsion velocity. The functions $\vec{I}_{att}$ and $\vec{I}_{rep}$ represent the spatial attraction and repulsion between swarmalators, while the phase interaction is captured by H_{att}. The function F in equation (4.1) measures the influence of phase similarity

on spatial attraction, while G in equation (4.2) measures the influence of spatial proximity on the phase attraction.

Consider the following instance of this model:

$$\dot{\vec{x}}_i = \vec{v}_i + \frac{1}{N}\sum_{j\neq i}^{N}\left[\frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|}\Big(A + J\cos(\theta_j - \theta_i)\Big) - B\frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|^2}\right] \tag{4.3}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j\neq i}^{N}\frac{\sin(\theta_j - \theta_i)}{|\vec{x}_j - \vec{x}_i|}. \tag{4.4}$$

For simplicity, we chose power laws for I_{att}, I_{rep} and G along with analytically convenient exponents. The sine function in H_{att} was similarly motivated, in the spirit of the Kuramoto model [56]. We first consider identical swarmalators so that $\omega_i = \omega$ and $\vec{v}_i = \vec{v}$. Further, we assume propulsion with constant magnitude and direction $\vec{v} = v_0\hat{n}$ where $\hat{n}$ is a constant vector (we relax these simplifications later). Then by a choice of reference frame we can set $\omega = v_0 = 0$ without loss of generality. Finally, by rescaling time and space we set $A = B = 1$. This leaves us with a system with two parameters (J, K).

The parameter K is the phase coupling strength. For $K > 0$, the phase coupling between swarmalators tends to minimize their phase difference, while for $K < 0$, this phase difference is maximized. The parameter J measures the extent to which phase similarity enhances spatial attraction. For $J > 0$, "like attracts like": swarmalators prefer to be near other swarmalators with the same phase. When $J < 0$, we have the opposite scenario: swarmalators are preferentially attracted in space to those with *opposite* phase. And when $J = 0$, swaramalators are phase-agnostic, their spatial attraction being independent of their phase. To

keep $I_{att}(x) > 0$, we constrain J to satisfy $-1 \leq J \leq 1$.

Before stating our results, we pause to discuss our model's features. As mentioned above, the model's purpose is to study the interplay between synchronization and swarming. But what precisely do we mean by swarming? While, to our knowledge, there is no unanimous classification, elements of a swarming system typically (i) attract and repel each other, leading to aggregation, and (ii) align their orientations so as to move in the same direction. Succinctly then, a swarming system models *aggregation* and/or *alignment*.

Our model accounts for aggregation, but *not* for alignment: the spatial dynamics (4.1) model phase-dependent aggregation, while the phase dynamics (4.2) model position-dependent synchronization. There are no alignment terms. Indeed, the particles of our system do not have an orientation so there is nothing to align! We chose to neglect an orientation state variable, and thus alignment, for two reasons. The first was simply because we believe there are swarmalator systems in which orientation does not play a role, such as the Japanese tree frogs [59, 93] or chemotactic oscillators [99, 100]. The second was that modeling orientable swarmalators adds an additional layer of complexity; it gives each swarmalator an orientation β, increasing the number of state variables per swarmalator from three (a two-dimensional position (x, y) and an internal phase θ) to four.

In the interest of minimalism we wished to avoid this complication for now. Hence as it stands our model applies only to swarmalators without an orientation. However we later show that our results are robust to the inclusion of

simple alignment dynamics, indicating their potential to hold for systems of orientable swarmalators as well.

4.3 Results

We performed numerical experiments to probe the behavior of our system. Unless otherwise stated, the simulations were run using python's ODE solver 'odeint'. We initially positioned the swarmalators in a box of length 2 and drew their phases from $[-\pi, \pi]$, both uniformly at random. We found the system settles into five states (Supplementary Movies 1-5). In three of these states, the swarmalators are ultimately static in space and phase. In the remaining two, the swarmalators move. However in all states, the density of swarmalators $\rho(\vec{x}, \theta, t)$ is time-independent, where $\rho(\vec{x}, \theta, t)\, d\vec{x}\, d\theta$ gives the fraction of swarmalators with positions between $\vec{x}$ and $\vec{x}+d\vec{x}$, and phases between θ and $\theta+d\theta$ at time t. In Fig. 4.1 we show where these states occur in the (J, K) parameter plane. We next discuss these five states.

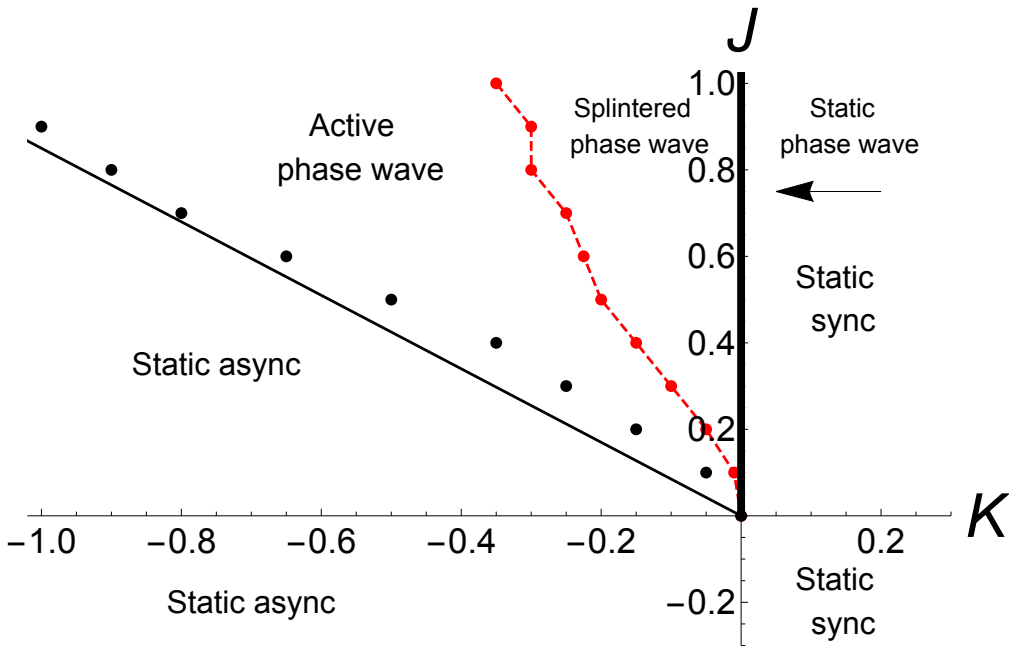

Figure 4.1: Phase diagram for the model defined by equations (4.95) and (4.96) with $A = B = 1$ and $\vec{v}_i = \omega_i = 0$. The straight line separating the static async and active phase wave states is a semi-analytic approximation given by (4.92). Black dots show simulation data. These were calculated by finding where the order parameter S bifurcates from zero, defined by where its second derivative is largest. Similarly, the red dots separating the active phase wave and splintered phase wave states were found by finding where the order parameter γ bifurcates from 0. The red dashed line simply connects these points and was included to make the boundary clearer.

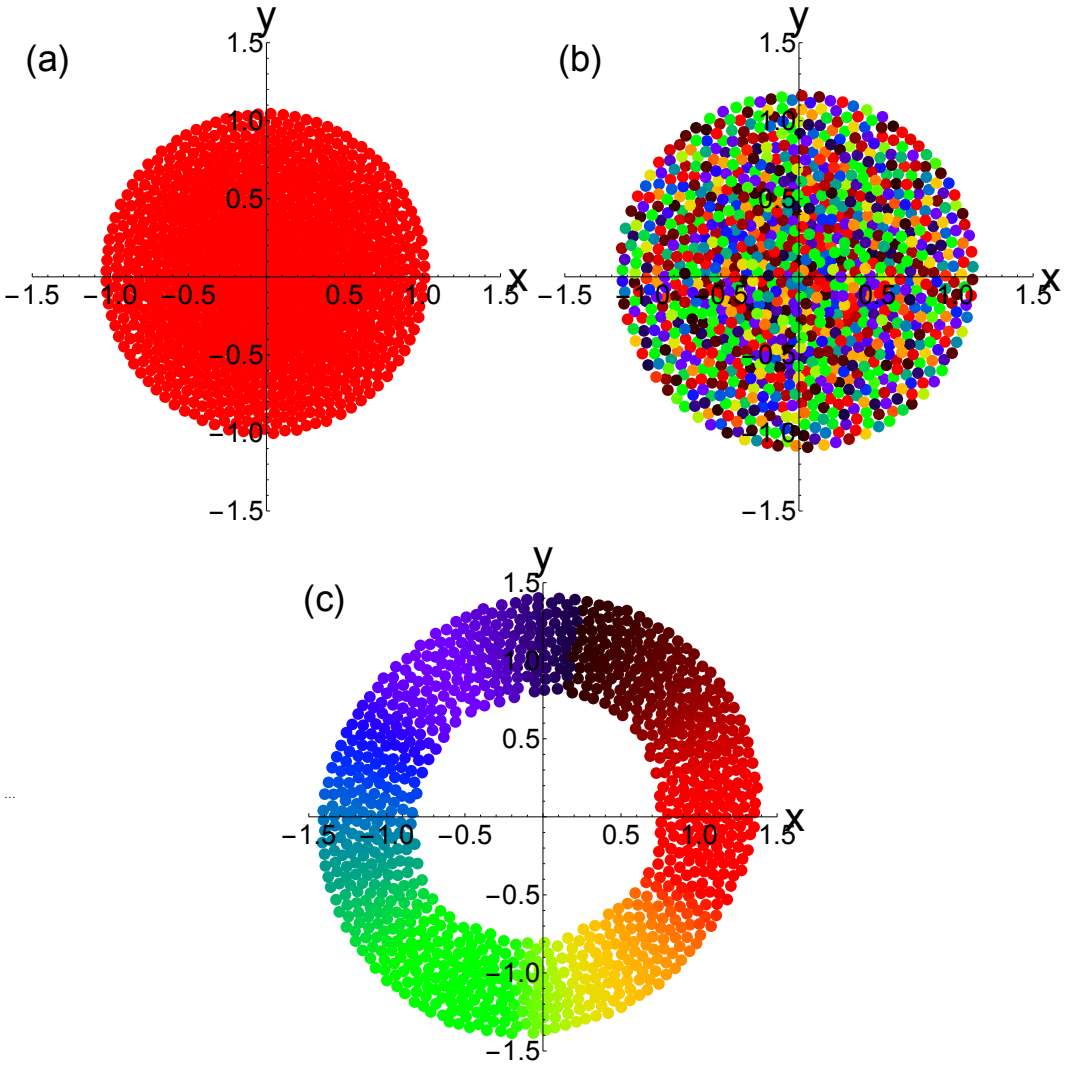

Figure 4.2: Scatter plots of three states in the (x, y) plane, where the swarmalators are colored according to their phase. Simulations were for $N = 1000$ swarmalators for $T = 100$ time units and stepsize $dt = 0.1$. Supplementary Movies 1-3 correspond to panels (a)-(c). (a) Static sync state for $(J, K) = (0.1, 1)$. (b) Static async state $(J, K) = (0.1, -1)$. (c) Static phase wave state $(J, K) = (1, 0)$

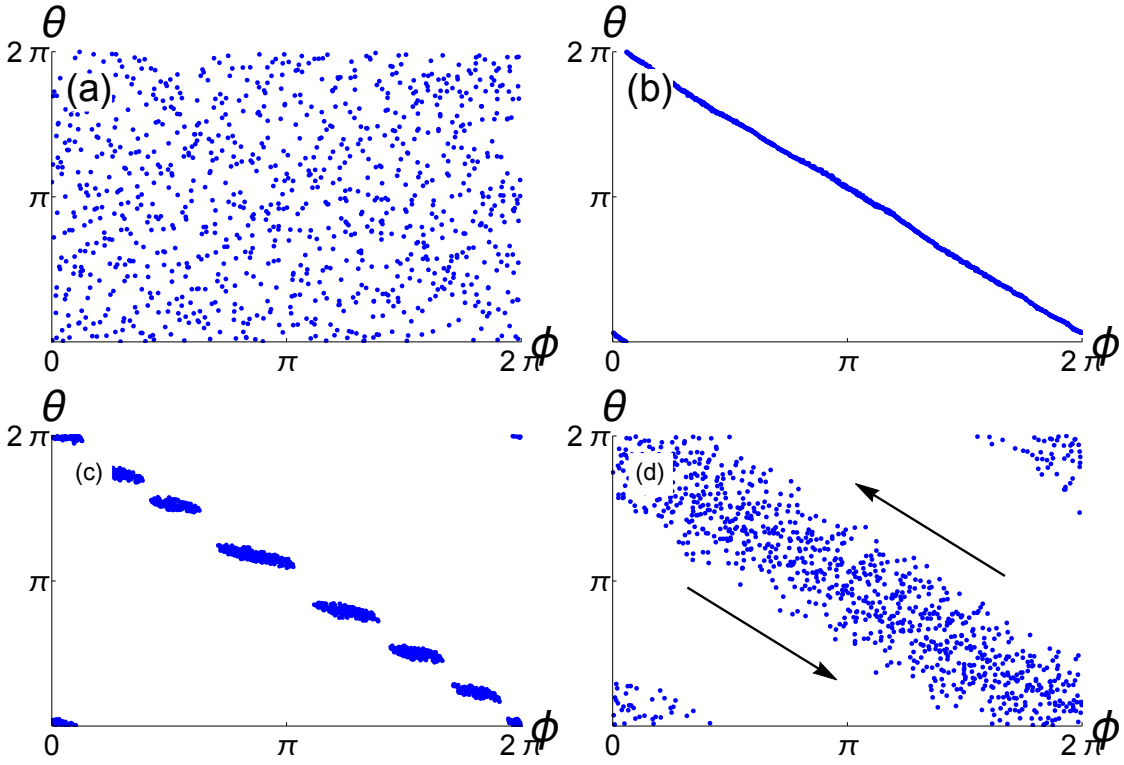

Figure 4.3: Distributions in (ϕ, θ) space corresponding to different states, where $\phi = \tan^{-1}(y/x)$. Simulations were run with $N = 1000$ swarmalators for variable numbers of time units T and stepsize $dt = 0.1$. (a) Static async state for $(J, K) = (0.1, -1)$ and $T = 100$. (b) Static phase wave state $(J, K) = (1, 0)$ and $T = 100$. (c) Splintered phase wave state $(J, K) = (1, -0.1)$ and $T = 1000$. (d) Active phase wave state $(J, K) = (1.0, -0.75)$ and $T = 1000$. Black arrows indicate the shear flow motion of swarmalators. Supplementary Movies 6 and 7 correspond to panels (c) and (d).

1. Static synchrony. The first state is shown in Fig. 4.2(a). The swarmalators form a circularly symmetric, crystal-like distribution in space, and are fully synchronized in phase, as indicated by all of them having the same color in Fig. 4.2(a). Since the swarmalators are ultimately stationary in $\vec{x}$, and they all end up at the same phase θ, we call this the *static sync* state. It occurs for $K > 0$ and for all J, as seen in Fig. 4.1.

In the continuum limit, this state is described by $\rho(r, \phi, \theta, t) = \frac{1}{2\pi} g_1(r)\delta(\theta - \theta_0)$, where ϕ is the spatial angle $\phi = \tan^{-1}(y/x)$, and the final phase θ_0 is determined from the initial conditions. In [1] we use a technique used by Kololnikov et al [101] when studying swarms to derive the following pair of integral equations for g_1:

$$\int_0^R \Big[(s-r)\mathcal{K}\left(\frac{4rs}{(r+s)^2}\right) + (r+s)\mathcal{E}\left(\frac{4rs}{(r+s)^2}\right) + \frac{\pi^2}{2J}(r-s)\Big]\frac{2Js}{r} g_1(s)\, ds = 0 \tag{4.5}$$

$$g_1(r) = \frac{2(1+J)}{\pi}\int_0^R \mathcal{K}\left(\frac{4sr}{(r+s)^2}\right)\frac{g_1(s)}{s+r} s\, ds, \tag{4.6}$$

where $\mathcal{K}, \mathcal{E}$ are the complete elliptic integral of the first and second kinds, and R is the radius of the disk in the (x, y) plane which must be determined. We were unable to solve these equations for $g_1(r)$ and R, so instead solve them numerically in [1]. Analytic progress can however be made if a linear attraction kernel $I_{att}(\vec{x}) = \vec{x}$, is used instead of the unit vector kernel we are currently considering. Then, as shown in Kolokolnikov et al. [101], the radial density becomes $g_1(r) = 1$, i.e swarmalators are uniformly distributed. In this special case we can also calculate R analytically,

$$R_{sync} = (1+J)^{-1/2}. \tag{4.7}$$

We show a full derivation in [1]. In dimensionful units, this reads $R = \sqrt{B/(A+J)}$. Thus the radius is determined by the ratio of the strengths of the attractive to the repulsive forces I_{att}, I_{rep} (in the static sync state, the effective attraction force is $A + J\cos(\theta_j - \theta_i) = A + J$, since all swarmalators have the same phase). Figure 4.4(a) shows the prediction (4.7) agrees with simulation results.

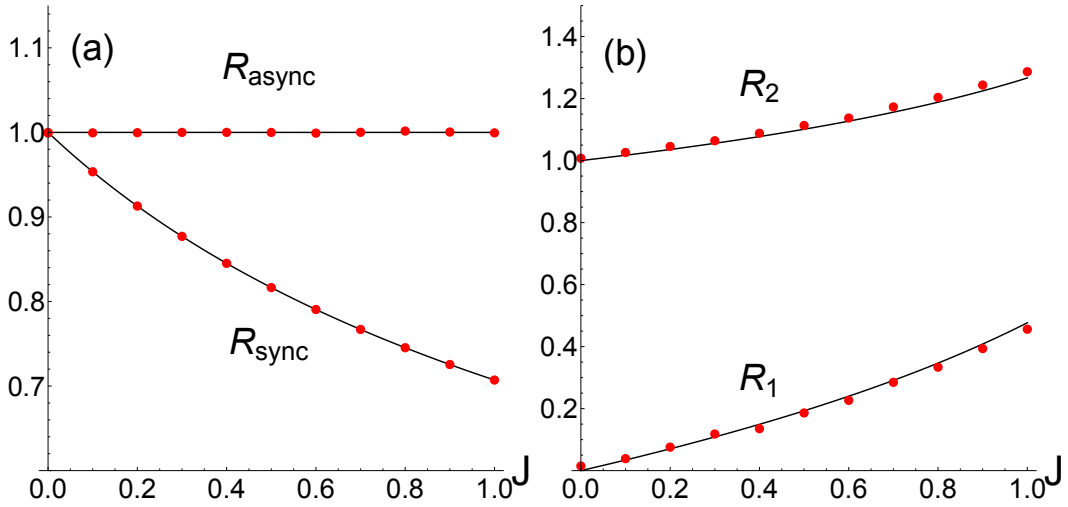

Figure 4.4: Radii of stationary states for $N = 800$ swarmalators for a linear attraction kernel $I_{att}(\vec{x}) = \vec{x}$. Red dots show simulation data, while black curves show theoretical predictions. (a): Radius of crystal formed in static sync state (for $K = 1$) and static async state (for $K = -2$) versus J. (b): Inner and outer radii of annulus in static phase wave state versus J.

2. Static asynchrony. Swarmalators can also form a *static async* state, illustrated in Fig. 4.2(b). At any given location $\vec{x}$, all phases θ can occur, and hence all colors are present everywhere in Fig. 4.2(b). This is seen more clearly in a scatter plot of the swarmalators in the (ϕ, θ) plane, depicted in Fig. 4.3(a). Notice that the swarmalators are distributed uniformly, meaning that every phase occurs everywhere. This completely asynchronous state occurs in the quadrant $J < 0$, $K < 0$, and also for $J > 0$ as long as J lies in the wedge $J < |K_c|$ shown in the

phase diagram in Fig. 4.1. As for the static sync state, we were able to calculate the radius of the circular distribution when a linear attraction kernel $\vec{I}(\vec{x})$ was used. In [1] we show this radius is given by

$$R_{async} = 1 \tag{4.8}$$

which agrees with simulation as shown in Fig. 4.4(a).

3. Static phase wave. The final stationary state occurs for the special case $K = 0$ and $J > 0$. This means the swarmalators' phases are frozen at their initial values. How, then, does the population evolve? Since $J > 0$, 'like attracts like': swarmalators want to settle near others with similar phase. The result is an annular structure where the spatial angle ϕ of each swarmalator is perfectly correlated with its phase θ, as seen in Fig. 4.2(c) and 4.3(b). Since the phases run through a full cycle as the swarmalators arrange themselves around the ring, we call this state the *static phase wave*.

In density space, this static phase wave state is described by $\rho(r, \phi, \theta) = g_2(r)\delta(\phi \pm \theta + C_1)$ where the $\pm$ and the constant C_1, are determined by the initial conditions. In [1] we again consider the linear attraction kernel, and find that $g_2(r)$ can be obtained analytically,

$$g_2(r) = 1 - \frac{\Gamma_J}{r}, \quad R_1 \leq r \leq R_2 \tag{4.9}$$

with $\Gamma_J = 2J(R_2^3 - R_1^3)\Big(3J(R_2^2 - R_1^2) + 12\Big)^{-1}$. This in turn lets us calculate the inner

and outer radii R_1, R_2 of the annulus:

$$R_1 = \Delta_J \frac{-\sqrt{3}J - 3\sqrt{12-5J}\sqrt{J+4} + 12\sqrt{3}}{12J}, \tag{4.10}$$

$$R_2 = \frac{\Delta_J}{2\sqrt{3}} \tag{4.11}$$

with $\Delta_J = \sqrt{\frac{3J-\sqrt{36-15J}\sqrt{J+4}-12}{J-2}}$. Figure 4.4(b) shows agreement between these predictions and simulation.

4. Splintered phase wave. Moving from $K = 0$ into the $K < 0$ half-plane, we encounter the first non-stationary state, shown in Fig. 4.5(a) and Fig. 4.3(c). As can be seen, the static phase wave splinters into disconnected clusters of distinct phases. Accordingly we call this state the *splintered phase wave*. It is unclear what determines the number of clusters. Fewer are found when smaller length scales for the interaction functions $\vec{I}_{att}, \vec{I}_{rep}, G$ are used. However the parameters J, K also play a role, although how precisely has not yet been determined. Within each cluster, the swarmalators "quiver," executing small amplitude oscillations in both position and phase about their mean values.

5. Active phase wave. As K is further decreased, these oscillations increase in amplitude until the swarmalators start to execute regular cycles in both spatial angle and phase. This motion is best illustrated in Fig. 4.3(d), in which shear flow about the $\phi_i = \theta_i \pm C$ axis is evident. This type of flow follows from a conserved quantity in the model: $\langle\dot{\phi}\rangle = \langle\dot{\theta}\rangle = 0$, which can be seen by averaging equations (4.95) and (4.96) over the population. There are also oscillations in the radial position, where each swarmalator travels from the inner rim to the outer

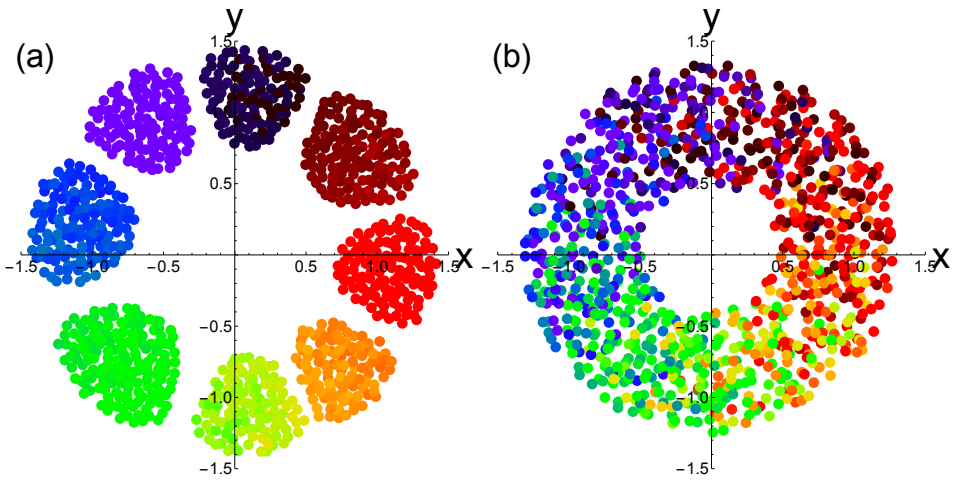

Figure 4.5: Two non-stationary states for $N = 1000$ swarmalators for $T = 1000$ time units and stepsize $dt = 0.1$. In all cases, swarmalators were initially placed in a box of length 2 uniformly at random, while their phases we drawn from $[-\pi, \pi]$. (a) Splintered phase wave $(J, K) = (1, -0.1)$. Note, there is a long transient until this state is achieved. See Supplementary Movie 4. (b) Active phase wave $(J, K) = (1, -0.75)$. See Supplementary Movie 5.

rim and back, in one orbit around the annulus.

This new, and final, state is similar to the double milling states found in biological swarms [102], where populations split into counter-rotating subgroups. It is also similar to the vortex arrays formed by groups of sperm [103], where the angular position ϕ of each sperm is correlated with the phase θ associated with the rhythmic beating of its tail.

At the density level, the state is like a blurred version of the static phase wave, insofar as the spatial angle and phase of a given swarmalator are roughly correlated, as evident in Fig. 4.3(d). However unlike the static phase wave, the swarmalators are non-stationary. To highlight this difference, we name this state the *active phase wave*.

Order parameters. Having described the five states of our system, we next discuss how to distinguish them. We define the following order parameter,

$$W_{\pm} = S_{\pm} e^{i\Psi_{\pm}} = \frac{1}{N} \sum_{j=1}^{N} e^{i(\phi_j \pm \theta_j)}, \tag{4.12}$$

where $\phi_i := \tan^{-1}(y_i/x_i)$. As shown in Fig. 4.6, the magnitude $S_{\pm}$ varies from 1 to 0 as we decrease K from 0, passing through all the states in the upper left quadrant of the (J, K) plane. (Note that all states except for static sync occur in this part of parameter space, so we hereafter confine our attention to just this region.)

To see why $S_{\pm}$ varies in this manner, recall that in the static phase wave, the spatial angle and phase of each swarmalator are perfectly correlated, $\phi_i = \pm\theta_i + C_1$ (recall that the $\pm$ and C_1 are determined by the initial conditions. This means either S_+ or S_- is non-zero). Therefore $S_{\pm} = 1$ at $K = 0$, where the static phase wave state is realized. Moving into the $K < 0$ plane we encounter the splintered phase wave. Here the correlation between ϕ_i and θ_i is not perfect, and so $S_{\pm} < 1$. As K is decreased the decay of this correlation is non-monotonic, which induces a dip in $S_{\pm}$ as seen in Fig. 4.6. Once the active phase wave is reached however this non-monotonicity disappears. As a result $S_{\pm}$ declines uniformly until it finally drops to zero when the static async state is reached, in which ϕ_i and θ_i are fully uncorrelated.

To sum up, $S_{\pm}$ is zero in the static async state, bifurcates from zero at a critical coupling strength K_c, is non-zero in the non-stationary splintered and active phase wave states, and is one in the static phase wave state.

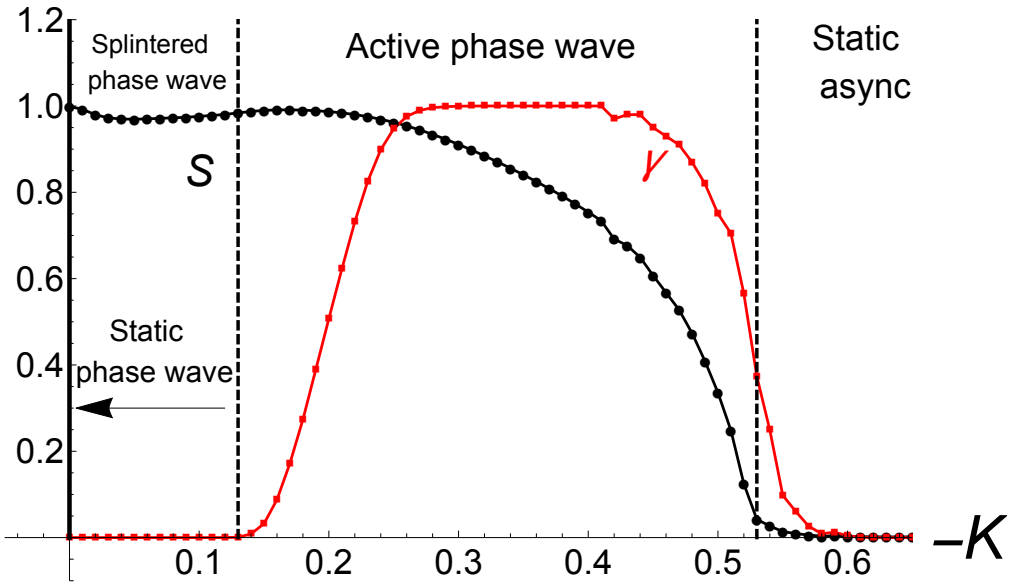

Figure 4.6: Asymptotic behavior of the order parameter $S := \max(S_+, S_-)$ (black dots) and γ (red dots) for $J = 0.5$ and $N = 800$. Note the bifurcation of S from at $K \approx -0.53$ near the approximation (4.92) $K_c = -1.2J = -0.6$. Data were collected using Heun's method for $T = 1000$ time units with stepsize $dt = 0.01$, of which the first half were discarded as transients. Each data point represents the average of one hundred realizations. Swarmalators were initially placed in a box of length 2 uniformly at random for all values of K with a common seed, while their phases we drawn from $[-\pi, \pi]$.

Notice however that since $S_\pm$ is non-zero for both the splintered and active phase wave, it cannot distinguish between these states. To do this, we use another order parameter γ. We define this to be the fraction of swarmalators that have executed at least one full cycle in phase and position, after transients have been discarded. Then γ is zero for the splintered phase wave, and non-zero for the active phase wave. Using γ in concert with $S_\pm$ then allows us to discern all the macroscopic states of our system as illustrated in Fig. 4.6.

Stability analysis. To calculate the critical coupling strength K_c at which the static async state loses stability, we consider perturbations η in density space defined by

$$\rho(\vec{x}, \theta, t) = \rho_0(\vec{x}, \theta) + \epsilon\eta(\vec{x}, \theta, t) \tag{4.13}$$

where $\rho_0(\vec{x}, \theta, t) = (4\pi^2)^{-1} g_1(r)$ is density in the static async state. In [1], we substitute this ansatz into the continuity equation, expand η in a Fourier series, $\eta(\vec{x}, \theta, t) = \sum_{n=0} b_n(\vec{x}, t)e^{in\theta} + c.c.$, and derive the following expressions for the harmonics:

$$\begin{aligned} \dot{b}_1(\vec{x}, t) = -\frac{J}{2}\, \vec{\nabla}\rho_0(\vec{x}).\int \frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|} b_1(\vec{\tilde{x}}, t)\, d\vec{\tilde{x}} + \\ \frac{(J+K)}{2}\rho_0(\vec{x}) \int \frac{1}{|\vec{\tilde{x}} - \vec{x}|} b_1(\vec{\tilde{x}}, t) d\vec{\tilde{x}}, \end{aligned} \tag{4.14}$$

$$\dot{b}_n(\vec{x}, t) = 0, \quad n \neq 1. \tag{4.15}$$

Similar equations are obtained for the complex conjugates $\bar{b}_n(\vec{x}, t)$. We see the first Fourier harmonic $b_1(\vec{x}, t)$ is distinguished. To study its stability we expand it in a Fourier series $b_1(\vec{x}, t) = \sum_{m=0}^{\infty} f_m(r, t)e^{in\phi} + c.c.$. Substituting this ansatz into (4.88) leads to a evolution equation for each mode $f_m(r, t)$. We then set $f_m(r, t) = e^{\lambda_m t} c_m(r)$ and derive the following eigenvalue equation:

$$\lambda_m c_m(r) = \int_0^R H_m(r, s) c_m(s) s\, ds \tag{4.16}$$

where R is the radius of the support of the density in the static async state. We focus first on the zeroth mode f_0 for which we can compute $H_0(r, s)$ analytically:

$$\lambda_0 c_0(r) = \int_0^R H_0(r, s) c_0(s) s\, ds, \tag{4.17}$$

$$\begin{aligned} H_0(r, s) = \frac{J(r^2 - s^2)g'(r) + 2rg(r)(J + K)}{4\pi^2 r(r + s)} \mathcal{K}\left(\frac{4rs}{(r + s)^2}\right) \\ + \frac{J(r + s)g'(r)}{4\pi^2 r} \mathcal{E}\left(\frac{4rs}{(r + s)^2}\right) \end{aligned} \tag{4.18}$$

where $\mathcal{K}, \mathcal{E}$ are the complete elliptic integral of the first and second kinds. We were unable to solve (4.17) for λ_0 analytically. Instead, we found it numerically by approximating the integral using gaussian quadrature. This reduces (4.17) to the form $\lambda' c_i = M_{ij} c_j$ where $M_{ij} = H_0(r_i, r_j) w_j$, w_j are gaussian quadrature weights, $r_i = i * (R/N')$ and $i = 1 \ldots N'$. The eigenvalues $\lambda_0'(N')$ of M_{ij}, which depend on the number of grid points N' used in the quadrature, then approximate λ_0.

The eigenvalues $\lambda_0'(N')$ have unexpected properties. The real part of the most unstable eigenvalue, denoted $\lambda_0^{*\prime}(N')$, is positive for all J, K. This tells us that f_0 is always unstable, which in turn tells us that the static async state is always unstable! In Fig. 4.13 we plot $\lambda_0^{*\prime}(N')$ versus K for $J = 0.5$ and $N' = 200$ grid points. As can be seen it is small but positive for sufficiently negative K. Note however that there is a transition-like point $K_0^* \approx -0.5$ beyond which $\lambda_0^{*\prime}(N')$ increases sharply. Figure 4.13 also shows $\lambda_m^{*\prime}(N')$ for $m = 1, 2, 3, 4$, which have the same behavior as $\lambda_0^{*\prime}(N')$: they are small but positive for $K < K_m^*$, and grow sharply for $K > K_m^*$.

Small but positive eigenvalues for $K < K_m^*$ were a surprise. We were expecting them to be negative, since simulations show the static async state is stable. We were thus suspicious of these results, and doubted the accuracy of the approximation $\lambda_0^{*\prime}(N')$ to the true λ_0^*. We therefore repeated the calculation for different values of N' up to $N' = 1600$ in [1]. Contrary to our expectations, we found that while the $\lambda_m^*(N')$ got smaller, they consistently remained positive for $K < K_m^*$.

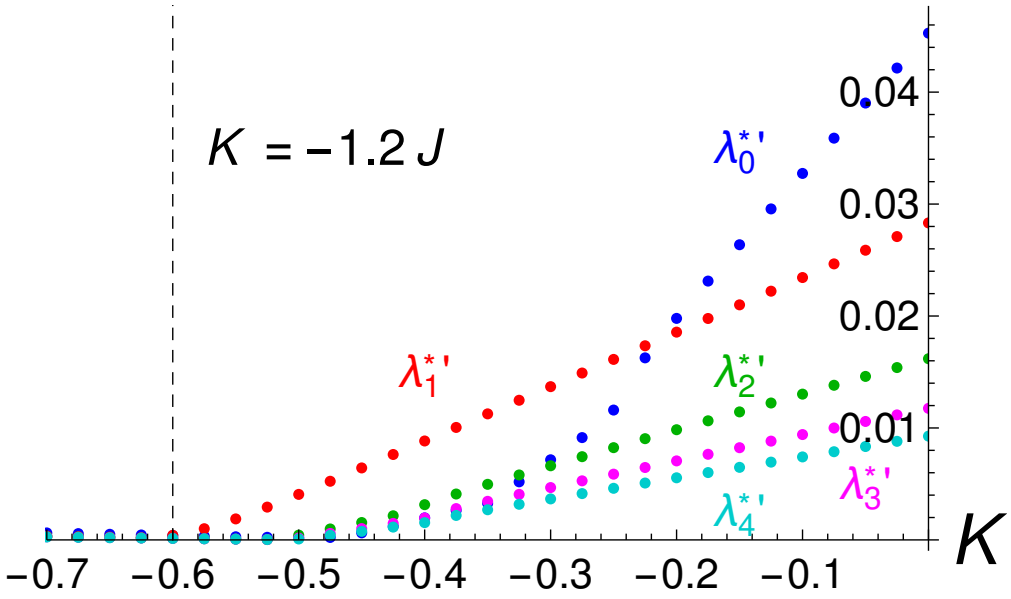

Figure 4.7: The real part of the most unstable eigenvalue $\lambda_m^{*\prime}$ of the first five modes f_m calculated from equation (4.90) for $J = 0.5$. Notice that they are all is positive for all K. Each $\lambda_m^{*\prime}$ was calculated by approximating the integral of the R.H.S. of (4.90) using gaussian quadrature with $N' = 200$ grid points and diagonalizing the resulting matrix. The upper limit of integration $R = 1.15$ was measured from simulations. The radial density $g(r)$ was determined numerically as discussed in [1]. The kernels H_m in equation (4.90) for $m > 1$ were calculated numerically. The dashed line marks the approximation to the critical coupling strength (4.92).

We also crudely investigated the $N' \to \infty$ limit by (i) fitting our data to curves of the form $a + b(N')^c$ and (ii) using Richardson extrapolation. Due to the small magnitudes of the $\lambda_m^{*\prime}(N')$ however, the results were rather unconvincing. Typical values for the best fit parameter a, which represents the limiting behavior of $\lambda_m^{*\prime}$, were $a \sim 10^{-6}$. The confidence interval for this parameter also contained positive and negative values. On top of that the approximations from methods (i) and (ii) gave inconsistent results. Hence we were unable to reliably determine the sign of $\lambda_m^{*\prime}$ when $K < K_m^*$ and $N' \to \infty$, which preventing us from accurately ascertaining the stability of the static async state. We restate however that the

fact that $\lambda_m^{*\prime} > 0$ for the large but finite value of N' we used is significant evidence that the unanticipated *instability* of the static async state is genuine.

While a rigorous determination of the sign of $\lambda_m^{*\prime}$ when $K < K_m^*$ remains elusive, our analysis certifiably shows its magnitude is very small. Hence, whatever the stability or instability of the m-th mode f_m turns out to be, it must be *weak*. In turn, then, the static async state has weak stability properties for $K < K_c$, where $K_c = \min_m K_m^*$ (i.e., at the point the most unstable f_m loses stability). How can we find this K_c? In Fig. 4.13 we see the f_1 becomes unstable first. There are of course an infinite number of modes, but as can be seen, λ_m^* appears to decrease with increasing m. Thus we assume $\min_m K_m^* = 1$. In [1] we approximate $K_1^* = \frac{d^2\lambda_1^{*\prime}}{dK^2}$, calculate it for different J, and find the following linear relation:

$$K_c \approx -1.2J. \tag{4.19}$$

Summarizing our main result: in the continuum limit $N \to \infty$, the static async state is unstable for $K > K_c$, and either (i) weakly stable, (ii) neutrally stable, or (iii) weakly unstable for $K < K_c$. Further, numerical evidence suggests that (iii) is the most likely. While this result is perhaps unsatisfying from a technical perspective, in practice it has utility. For example as shown in Fig. 4.1, the approximation (4.92) for K_c agrees reasonably well with finite N simulations.

Genericity. Our analysis so far has been for the instance (4.95), (4.96), of the model defined by (4.1), (4.2). This begs the question: are the phenomena we found generic to the model? Or specific to this instance of the model? To answer this question, we ran simulations for different choices of the functions $\vec{I}_{rep}, \vec{I}_{att}$

and G; see [1].

In all but one case, we found the same phenomena. The exception is when a linear attraction kernel $\vec{I}_{att}(\vec{x}) = \vec{x}$ is used. Here we found new states, which we call *non-stationary phase waves*. They are similar to the active phase wave, except now the phase $\Psi_\pm$ of the order parameter $W_\pm$ begins to rotate, reminiscent of the traveling wave states found in the Kuramoto model with distributed coupling strengths [104, 105]. We further discuss this and other properties in [1].

4.4 Extensions to the model

Noise and disordered natural frequencies. The swarmalators previously considered were identical and noiseless. We now relax these idealizations. Then the governing equations are

$$\dot{\vec{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N}\left[\frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|}\Big(1 + J\cos(\theta_j - \theta_i)\Big) - \frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|^2}\right] + \vec{\xi}_i^{\vec{x}}(t), \tag{4.20}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j\neq i}^{N}\frac{\sin(\theta_j - \theta_i)}{|\vec{x}_j - \vec{x}_i|} + \eta_i(t), \tag{4.21}$$

where ω_i are random variables drawn from a Lorentzian $g(\omega) = (\sigma/\pi)\Big[(\omega-\mu)^2 + \sigma^2\Big]^{-1}$. By a change of frame we set $\mu = 0$ leaving just σ which quantifies the strength of the disorder. We choose white noise variables $\eta_i(t)$ and $\vec{\xi}_i^{\vec{x}}(t)$ with zero mean and strengths D_x, D_y, D_θ characterized by $\langle \xi_i^x(t)\xi_j^x(t')\rangle = 2D_x\delta_{ij}\delta(t-t')$, etc.

Simulations show that when just phase noise D_θ is turned on, noisy versions of all the states are realized. The splintered phase wave however degenerates into the active phase wave for all but the smallest noise $D_\theta \gtrsim 10^{-3}$. In the remaining states, the spatial densities remain compact supported with the same radii, except now the swarmalators have noisy phase motion (this induces some spatial movement, which disappears when $N \to \infty$ as we show in [1]). Hence the following states, where we have swapped the descriptor 'static' with 'noisy', are robustly realized when $D_\theta > 0$: (i) *noisy phase wave*, (ii) *active phase wave*, (iii) *noisy async*.

Frequency disorder $\sigma > 0$ has a more serious effect. Since $g(\omega)$ is symmetric about zero, there are equal numbers of swarmalators with oppositely signed natural frequencies. This turns the static/noisy phase wave into the active phase wave, in the sense that counter-rotating groups develop. This is not seen in the async state. Here, there are noisy spatial movements which vanish as $N \to \infty$, as in the noisy async state. In contrast however, the swarmalators execute noisy, but full, phase cycles. To highlight this distinction, we rename the state the active async state. The states realized are then: (i) *active phase wave*, (ii) *active async*.

Finally spatial noise $D_x, D_y > 0$ simply blurs the spatial densities of the states. No other phenomena are induced. Hence when $D_\theta, \sigma, D_x, D_y > 0$, we again get the (i) *active phase wave*, and (ii) *active async* states.

In Fig. 4.8 we plot the order parameter $S(K)$ for different amounts of noise and frequency disorder. As for the original model, S simply declines to zero as K is

decreased, with the noise and disordered frequencies changing just the shape of the curves and the value of K_c. Note the disappearance of the dip in S for small K, which indicates the absence of the splintered phase wave state. Note also we do not plot the second order parameter γ which discerns the splintered phase wave since this state does not robustly exist when $\sigma, D_\theta \neq 0$.

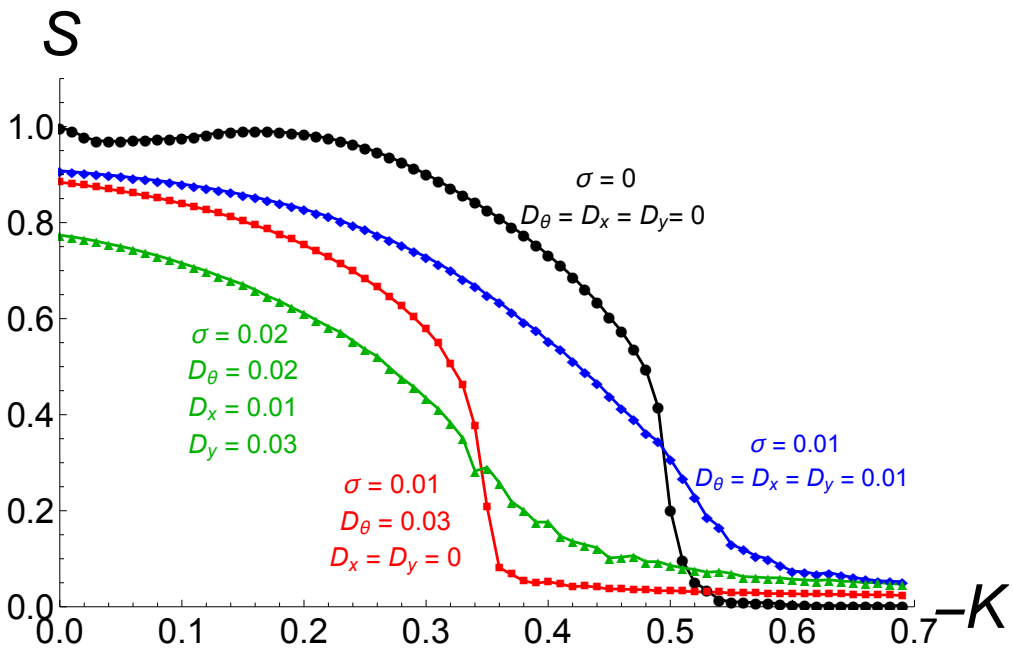

Figure 4.8: Order parameter $S = \max(S_+, S_-)$ versus K for $J = 0.5$ and different amounts of disorder as quantified by the width of the distribution of natural frequencies σ and the noise strengths, D_θ, D_x, and D_y. As can be seen, greater amount of disorder stabilize the async state, as indicated by $-K_c$ becoming smaller and smaller. Note also the disappearance of the dip in the $S(K)$ curve, which tells us the splintered phase wave state does not exist in the presence of noise of this strength. Simulations were run for $N = 500$ swarmalators using Heun's method for $T = 1000$ time units with stepsize $dt = 0.01$, the first half of which were discarded. Each data point represents the average of 10 realizations.

Swarmalators in 3D. So far we have considered swarmalators moving in two di-

mensions. While there are physical systems where this approximation is valid, such as certain active colloids [106] or sperm, which are often attracted to two dimensional surfaces [107], this restriction was mostly for mathematical convenience. Here we explore the more physically realistic case of motion in three spatial dimensions (in [1] we also explore motion in one dimension). For simplicity we consider the case of identical swarmalators with no noise, although we relax these idealizations in [1]. Our system is then

$$\dot{\vec{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N}\left[\frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|}\Big(1 + J\cos(\theta_j - \theta_i)\Big) - \frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|^3}\right], \tag{4.22}$$

$$\dot{\theta}_i = \frac{K}{N}\sum_{j\neq i}^{N}\frac{\sin(\theta_j - \theta_i)}{|\vec{x}_j - \vec{x}_i|}, \tag{4.23}$$

where $\vec{x}_i = (x_i, y_i, z_i)$. These are the same as equation (4.95) and (4.96), except the exponent of the hard shell repulsion is now 3 (we choose this because it yields simple formulas for the radii of certain states).

Simulations show that analogues of the states found in 2D are realized. We show these as scatter plots in the (x, y, z) plane in Fig. 4.9. We also provide movies of the evolution to these states in [1]. The static sync and async states become spheres (note we do not plot the static sync state due to space limitations) as seen in panel (a). As in the 2D case, we can calculate their radii when a linear attraction kernel is used,

$$R_{sync} = (1 + J)^{-1/3}, \tag{4.24}$$

$$R_{async} = 1, \tag{4.25}$$

which agree with simulation as shown in [1].

In panel (b) we show the static phase wave becomes a sphere with a cylindrical hole through its center. The orientation of this cylinder is determined by the initial conditions. The phase and azimuthal angle $\phi = \tan^{-1}(y/x)$ are correlated in the same way for each value of the polar angle $\alpha = \cos^{-1}(z/\sqrt{x^2+y^2+z^2})$ (when the azimuthal and polar angles are measured relative to the axis of the cylindrical hole). We show this more clearly in a scatter plot in the (θ, ϕ) plane in [1].

As in the 2D model, this correlation between ϕ and θ persists for the splintered phase waves and active phase wave states as can be seen in panels (c) and (d) of Fig. 4.9. The motion of the swarmalators in these states are as before: in the splintered phase wave they 'quiver', executing small oscillations in space and phase, while in the active phase wave they execute full rotations (note the spatial component of these rotations are in the azimuthal direction $\hat{\phi}$ only, not in the polar direction $\hat{\alpha}$). In [1] we show how the order parameters $S_{\pm}, \gamma$ can also be used to differentiate these 3D states.

Alignment and self-propulsion. Up to now we have considered the trivial case of swarmalators that propel themselves with constant magnitude and direction, in a manner uninfluenced by their neighbors. This allowed us to set this term to zero via a change of reference. In many real systems, however, such behavior is unrealistic: individuals often adjust the direction of their motion to align with that of their neighbors. Vicsek studied this alignment effect in a seminal work [81].

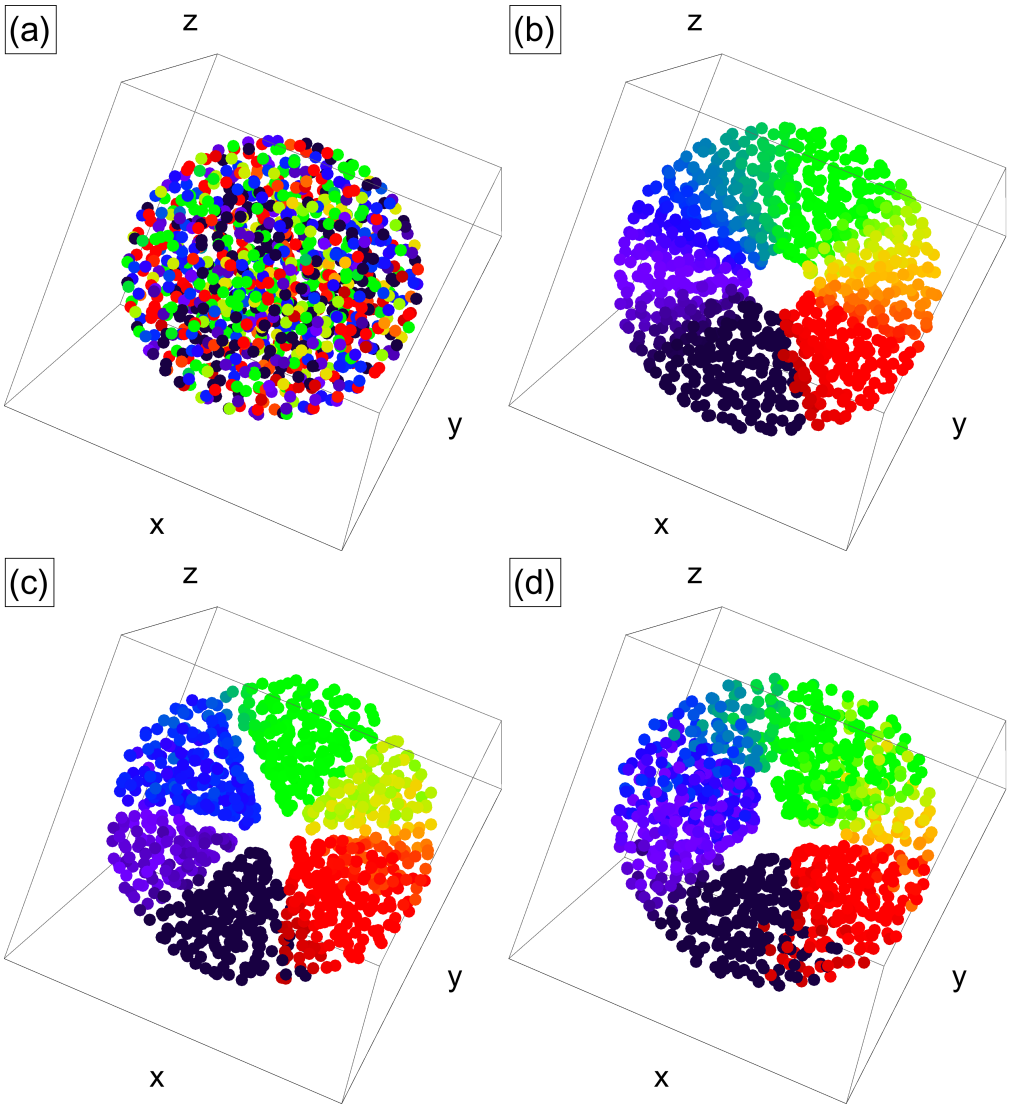

Figure 4.9: Scatter plots of four states in the (x, y, z) plane, where the swarmalators are colored according to their phase. Data were collected for $J = 0.5$ and $N = 1000$ swarmalators for $T = 5000$ time units with stepsize $dt = 0.001$ using Heun's method. (a) Static async state for $K = -1$. (b) Static phase wave for $K = 0$ (c) Splintered phase wave for $K = -0.05$. (d) Active phase wave state for $K = -0.6$. Supplementary movies 9-12 correspond to panels (a)-(d).

We here partially explore the effect of alignment on swarmalator systems. Accordingly we endow each swarmalator with an orientation β, which characterizes the direction of its self-propulsion. The inclusion of alignment makes our

model complicated; there are now four state variables (x, y, θ, β) per swarmalator, which could interact with each other in potentially many ways. Furthermore, there are six parameters $(J, K, \sigma, D_\theta, D_x, D_y)$, not to mention any additional parameters governing the evolution of β. An exhaustive study of orientable swarmalators is thus beyond the scope of the present work. Hence, we restrict ourselves to answering a simple question: are the states of our swarmalator system robust to the inclusion of simple alignment dynamics?

To this end, we study the simplest possible extension to our current model: we choose Vicsek type interactions between $\vec{x}$ and β, and leave β and the phase θ uncoupled (although they are indirectly coupled through the position $\vec{x}$). Our system then reads

$$\dot{\vec{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N}\left[\frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|}\Big(1 + J\cos(\theta_j - \theta_i)\Big) - \frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|^2}\right] + \vec{\xi_i^x}(t) + v_0\hat{n}, \tag{4.26}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j\neq i}^{N}\frac{\sin(\theta_j - \theta_i)}{|\vec{x}_j - \vec{x}_i|} + \eta_i(t), \tag{4.27}$$

$$\dot{\beta}_i = -\beta_i + \frac{1}{|\Lambda_i|}\sum_{j\in\Lambda_i}\beta_j + \zeta_i(t), \tag{4.28}$$

where $\hat{n} = (\cos\beta, \sin\beta)$, Λ_i is the set of swarmalators within a distance δ of the i-th swarmalator, and $|\Lambda_i|$ is the number of such neighbors. The $\zeta_i(t)$ is a white noise variable with zero mean and strength D_β characterized by $\langle\zeta_i(t)\zeta_j(t')\rangle = 2D_\beta\delta_{ij}\delta(t - t')$.

Simulations show that for certain parameter values aligned versions of all our states persist. We plot two of these in panels (a) and (b) of Fig. 4.10, where each swarmalator is depicted as a colored arrow, oriented according to β, and colored

according to phase. As can be seen the swarmalators are aligned, with their space-phase distributions being the same as before. In contrast to the original model, however, the center of mass of each distribution now moves (in a direction determined by the initial conditions). In this sense the states are mobile. They are however equivalent to their static versions via a change of reference frame, $\vec{x} \to \vec{x} + \vec{v}_0 t$. For larger D_β, the unaligned versions of the same states are realized, as illustrated in panels (c) and (d) of Fig. 4.10.

We have demonstrated that the phenomena of our system are insensitive to the inclusion of simple alignment dynamics. We restate however that we have not comprehensively explored the space defined by the other parameters $(J, K, \sigma, v_0, D_x, D_y)$ given its large size. Thus it remains to be seen if new states will be found.

4.5 Discussion

We have examined the collective dynamics of swarmalators. These are mobile particles or agents with both phase and spatial degrees of freedom, which lets them sync and swarm. Furthermore, their phase and spatial dynamics are coupled. By studying simple models, we found this coupling leads to rich spatiotemporal patterns which we explored analytically and numerically. These patterns were robust to modifications to the model, namely motion in one, two, and three spatial dimensions, distributed natural frequencies, noisy interactions, and alignment dynamics. We thus believe they could be realized in nature

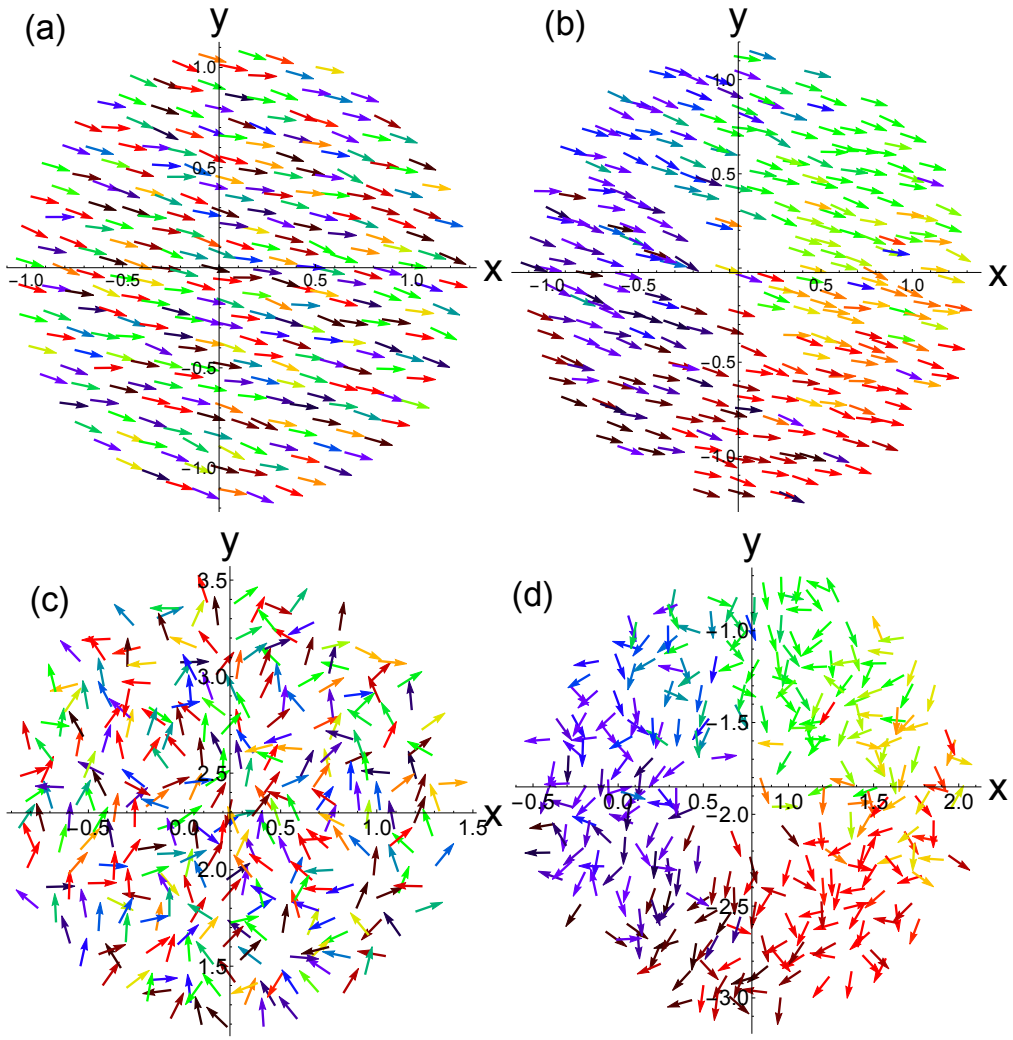

Figure 4.10: Scatter plots of four states in the (x, y) plane where the swarmalators are depicted as colored arrows, whose orientation represents β, and whose color represents the phase θ. Data were collected for $N = 300$ swarmalators for $T = 5000$ time units with stepsize $dt = 0.01$ using Heun's method. In each panel, parameter values were $J = \delta = 0.5$, $\sigma = D_\theta = D_\beta = 0.01$, $D_x = D_y = 0$ and $v_0 = 0.001$. (a) Aligned active async for $(K, D_\beta) = (-1.0, 0.01)$. (b) Aligned noisy phase wave for $K = (-0.1, 0.01)$. (c) Unaligned active async for $(K, D_\beta) = (-1.0, 1.0)$. (d) Unaligned noisy phase wave for $(K, D_\beta) = (-0.1, 1.0)$.

or technology.

A pertinent future goal, then, is to investigate the behavior of real-world systems of swarmalators. As mentioned in the introduction, colloidal suspensions of magnetic particles [7, 94, 95] or active spinners [96, 97] are promising candidates. For example, structures equivalent to the static phase wave state have been experimentally realized by Snezhko and Aranson, when studying the behavior of ferromagnetic colloids at liquid-liquid interfaces [94] (the particles comprising the colloids can be considered swarmalators if we interpret the angle subtended by their magnetic dipole vectors as their phase). As shown in Fig. 4 of [94], the colloids can form asters. These are structures composed of radial chains of magnetically ordered particles, which "decorate slopes of a self-induced circular standing wave" [94], analogous to the annular pattern of correlated phases and positions of the static phase wave shown in Fig. 4.2(c).

Could colloidal equivalents of the splintered and active phase wave states also be realized? Aside from being theoretically interesting, the ability to engineer these states could have practical application. For instance, Snezhko and Aranson also show that asters can be manipulated to capture and transport target particles. Perhaps the non-stationary behavior of the splintered and active wave states could also have locomotive utility. Tentative evidence for this claim is provided by populations of cilia, whose collective metachronal waves, similar to the motion of swarmalators in the aforementioned states, are known to facilitate biological transport [108–110].

Other plausible systems of real-world swarmalators are biological microswimmers, self-propelled micro-organisms capable of collective behavior [111]. One such contender is populations of spermatoza, which exhibit rich swarming be-

havior such as trains [112, 113] and vortex arrays [103], the latter of which is reminiscent of the active phase wave state, as mentioned in the Results section. The phase variable for each sperm is associated with the rhythmic beating of the sperm's tail, which can synchronize with that of a neighboring sperm [114, 115]. It has been theorized that this can induce spatial attraction [116], leading to clusters of synchronized sperm, consistent with experimentally observed behavior [117].

There are also theoretical avenues to explore within our proposed model of swarmalators. For instance the curious stability properties of the static async state deserve further study. Another route would be to include more realism by including heterogeneity in the coupling parameters K, J, or by choosing more complicated interaction functions I_{att}, I_{rep}, G, H. For example we chose $H(\theta) = \sin(\theta)$ to mimic the Kuramoto model, but as we saw, it led to just the trivial static sync state when $K > 0$. Perhaps choosing the more realistic Winfree model for the phase dynamics, which gives rise to richer collective behavior, would lead to more interesting swarmalator phenomena in this parameter regime.

Perhaps the most important direction for future work is to more fully explore the interplay among aggregation, alignment, and synchronization—or put another way, to explore the collective behavior of particles with a position $\vec{x}$, an orientation β, and an internal phase θ. The primary goal of our work is to draw attention to this class of problems, which we believe define a wide landscape of new emergent behavior. In this work, we have started to map out this landscape by studying a simple model that contains a subset of these three effects, namely

aggregation and synchronization.

Others have considered the remaining subsets. For example, Leon and Liverpool have explored the interaction between alignment and synchronization [118]. They introduced a new class of soft active fluids whose units have an orientation and phase. They found this mixture can either enhance or inhibit the transition from disordered states to states with polar and/or phase order. The latter states are roughly similar to the (un)aligned static (a)sync states. Yet counterparts of the static, splintered, and active phase waves were not reported.

The final combination, aggregation and alignment, is perhaps the most well studied, in both new models and old. For instance, Starnini et al. [119] recently introduced a model of mobile particles capable of aggregating and aligning their opinions, and found the emergence of echo chambers. Even in the classic Vicsek model and its numerous extensions, new phenomena are still being found. For instance, Kruk et al. found that delayed alignment in the Vicsek model produces self-propelled chimeras [120]; perhaps delayed phase interactions could lead to similar states for swarmalators. Liebchen and Levis [121] considered units with an intrinsic rotation, and found *phase separated droplets*: clusters of rotation-synchronized particles surrounded by a sea of incoherent particles (multiple droplets are also possible). These droplets are similar to our static sync states, but they differ in the crucial respect that the entire population is synchronized in our static sync state. Here too, the counterparts of our static, splintered, and active phase waves were not seen.

Thus, to the best of our knowledge, no other models display states analogous to the splintered phase waves and active phase waves found in our swarmalator model. In that sense, those two states are unprecedented.

4.6 Acknowledgments

Research supported by United States NSF Grant Nos. DMS-1513179 and CCF-1522054 (S.H.S), and by South Korean NRF Grant No. NRF-2015R1D1A3A01016345 (H.H).

4.7 Supplemental Materials

4.8 Properties of static sync and async states

We here use techniques used by Fetecau et al. [101] when studying swarm dynamics to study the static sync and static async states. We start with the async state whose density is

$$\rho(r, \phi, \theta, t) = \frac{1}{4\pi^2} g(r), \quad 0 \leq r \leq R. \tag{4.29}$$

We wish to solve for the radial density $g(r)$ and the radius R of its support. In this state the swarmalators are at rest and their phases are unchanging, so $\underline{v} \equiv \underline{0}$, where $\underline{v} = (v_x, v_y, v_\theta)$. As we will show, it is also useful to consider the divergence

of the velocity, which must also be zero (from the continuity equation for the conservation of swarmalators, and by applying the assumptions that the density for the static async state is stationary and the velocity is zero). This gives us a pair of simultaneous equations,

$$\underline{v} \equiv \underline{0}, \tag{4.30}$$

$$\nabla.\underline{v} \equiv 0. \tag{4.31}$$

We begin with divergence term (4.31). In cartesian coordinates the velocity reads

$$\vec{v}_{\vec{x}}(\vec{x},\theta,t) = \int \Bigg(\Big(\vec{\tilde{x}} - \vec{x}\Big)\Big(1 + J\cos(\tilde{\theta}-\theta)\Big) - \frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|^2} \Bigg) \rho(\vec{\tilde{x}},\ \tilde{\theta}, t)\ d\vec{\tilde{x}}\, d\tilde{\theta}, \tag{4.32}$$

$$v_\theta(\vec{x},\theta,t) = \int \frac{\sin(\tilde{\theta}-\theta)}{|\vec{\tilde{x}} - \vec{x}|} \rho(\vec{\tilde{x}},\tilde{\theta},t) d\vec{\tilde{x}}\, d\tilde{\theta}. \tag{4.33}$$

The divergence has a spatial and phase component: $\nabla.\underline{v} = \nabla_{\vec{x}}.\vec{v}_{\vec{x}} + \partial_\theta v_\theta$. The phase component $\partial_\theta v_\theta$ is trivially zero, since the swarmalators' phases are uniformly distributed in phase in the static async state. We find the spatial component by applying $\nabla_{\vec{x}}$ to (4.32):

$$\nabla_{\vec{x}}.\vec{v}_{\vec{x}} = \int -2\Big(1 + J\cos(\tilde{\theta}-\theta)\Big)\rho(\vec{\tilde{x}},\ \tilde{\theta}, t) d\vec{\tilde{x}}\, d\tilde{\theta} \tag{4.34}$$

$$+ 2\pi\delta(\vec{\tilde{x}} - \vec{x})\rho(\vec{\tilde{x}},\ \tilde{\theta}, t)\ d\vec{\tilde{x}}\, d\tilde{\theta}. \tag{4.35}$$

Here we have used the identity (expressed most cleanly in cartesian coordinates)

$$\nabla_{\vec{x}}.\frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|^2} = -2\pi\delta(\vec{\tilde{x}} - \vec{x}). \tag{4.36}$$

Simplifying this, and substituting $\partial_\theta v_\theta = 0$ gives the full divergence

$$\nabla.\underline{v} = 2\pi\rho(\vec{x},\theta) - 2\int \Big(1 + J\cos(\tilde{\theta}-\theta)\Big)\rho(\vec{\tilde{x}},\tilde{\theta},t)\ d\vec{\tilde{x}}\, d\tilde{\theta}. \tag{4.37}$$

By (4.31) we require this to be zero, which gives a self-consistent equation for ρ:

$$\rho(\vec{x},\theta,t) = \frac{1}{\pi}\int \Big(1 + J\cos(\tilde{\theta}-\theta)\Big)\rho(\vec{\tilde{x}},\tilde{\theta},t)\,d\vec{\tilde{x}}\,d\tilde{\theta}. \tag{4.38}$$

Finally substituting the ansatz (4.29) into this and performing the integration over ϕ gives

$$g(r) = 2\int_0^R g(\tilde{r})\tilde{r}\;d\tilde{r} = M = \text{const.} \tag{4.39}$$

This tells us ρ is constant inside a disc of radius R. The radius R can be determined via self-consistency: $M = \int_0^R rg(r)dr = \int_0^R rMdr \Rightarrow R = 1$. By normalizing ρ as per (4.29) we find $M = 1$ which means $g(r) = 2$. Putting this all together gives

$$\rho_{async}(r,\phi,\theta,t) = \frac{1}{2\pi^2}, \quad 0 \le r \le R_{async} \tag{4.40}$$

$$R_{async} = 1. \tag{4.41}$$

We must now check if the solutions (4.40), (4.41) imply $\underline{v} \equiv \underline{0}$ as required by (4.30). We do this in cartesian coordinates, in which

$$\rho(\vec{x},\theta,t) = \frac{1}{\pi R^2}\delta(\theta-\theta_0), \quad |\vec{x}| \le R, \tag{4.42}$$

where θ_0 is the final, common phase of each swarmalator. Substituting this into equations (4.32) and (4.33) for $v_{\vec{x}}, v_\theta$ and performing the integration gives

$$\vec{v}_{\vec{x}}(\vec{x},\theta,t) = \frac{1}{4\pi R^2}\Big(R^2 - [1 + J\cos(\theta-\theta_0)]\Big)\vec{x}, \tag{4.43}$$

$$v_\theta(\vec{x},\theta,t) = 0, \tag{4.44}$$

where we have used the identity

$$\int_{|\vec{\tilde{x}}|<R} \frac{\vec{\tilde{x}}-\vec{x}}{|\vec{\tilde{x}}-\vec{x}|^2} = \pi\vec{x}, \qquad |\vec{x}| < R. \tag{4.45}$$

We see that $\vec{v}_{\vec{x}} = 0$ at $\theta = \theta_0$ if $R = 1$, as required. Hence we have shown that the solutions (4.40), (4.41) satisfy equations (4.30) and (4.31).

Carrying out the same analysis for the static sync state leads to

$$\rho_{sync}(r,\phi,\theta,t) = \frac{1}{\pi^2}\delta(\theta-\theta_0), \qquad 0 \le r \le R_{sync} \tag{4.46}$$

$$R_{sync} = (1+J)^{-1/2}, \tag{4.47}$$

where θ_0 is the final common phase of the swarmalators in the static sync state.

Unit vector attraction kernel. We now carry out the same analysis for static async state for the unit vector attraction kernel $\vec{I}_{att}(\vec{x})$ used in the main text. This amounts to solving the pair of equations

$$\underline{v} \equiv \underline{0}, \tag{4.48}$$

$$\nabla.\underline{v} \equiv 0, \tag{4.49}$$

where the velocity is

$$\vec{v}_{\vec{x}}(\vec{x},\theta,t) = \int \left(\frac{\vec{\tilde{x}}-\vec{x}}{|\vec{\tilde{x}}-\vec{x}|}\Big(1+J\cos(\tilde{\theta}-\theta)\Big) - \frac{\vec{\tilde{x}}-\vec{x}}{|\vec{\tilde{x}}-\vec{x}|^2}\right)\rho(\vec{\tilde{x}},\,\tilde{\theta},t)\,d\vec{\tilde{x}}\,d\tilde{\theta}, \tag{4.50}$$

$$v_\theta(\vec{x},\theta,t) = \int \frac{\sin(\tilde{\theta}-\theta)}{|\vec{\tilde{x}}-\vec{x}|}\rho(\vec{\tilde{x}},\tilde{\theta},t)d\vec{\tilde{x}}\,d\tilde{\theta}. \tag{4.51}$$

And the density ansatz is the same as (4.29). After calculations we get the following pair of simultaneous equations for the unknown radial density $g(r)$ and radius R,

$$\int_0^R \Big[(s-r)\mathcal{K}\left(\frac{4rs}{(r+s)^2}\right) + (r+s)\mathcal{E}\left(\frac{4rs}{(r+s)^2}\right) + \frac{\pi^2}{2J}(r-s)\Big]\frac{2Js}{r}g_1(s)\,ds = 0, \tag{4.52}$$

$$g_1(r) = \frac{2(1+J)}{\pi}\int_0^R \mathcal{K}\left(\frac{4sr}{(r+s)^2}\right)\frac{g_1(s)}{s+r}s\,ds, \tag{4.53}$$

where $\mathcal{K}, \mathcal{E}$ are the complete elliptic integrals of the first and second kinds, and R is the radius of the disk in the (x, y) plane which must be determined. We were unable to solve these equations for $g_1(r)$ and R.

Equation (4.53) is however easily solved numerically by discretizing the interval $[0, R]$, which leads to

$$g_i = M_{ij} g_j, \tag{4.54}$$

$$M_{ij} = \frac{2(1+J)}{\pi} \mathcal{K}\left(\frac{4 r_j r_i}{(r_i + r_j)^2}\right) \frac{r_j}{r_j + r_i} w_j. \tag{4.55}$$

where $i = 1, \ldots N_{grid}$, $r_i = i * h = i * (R/N_{grid})$, $g_i = g(r_i)$ and w_j are gaussian quadrature weights. We see that g_i is simply an eigenvector of the matrix M_{ij} with eigenvalue 1. This lets us find the unknown radius R by computing the eigenvalues of M_{ij} for a selection of trial radii, and finding the R which corresponds to the eigenvalue 1. In practice this is the largest eigenvalue. In Fig. 4.11 we do this for a variety of grid sizes N_{grid}. As can be seen, the value of $R = 1.16$ produces an eigenvalue closest to 1. This is in reasonable agreement with radius of $R_{sim} \approx 1.15$ as measured from simulation data (for simulation parameters, see the caption of Fig. 4.11). Also shown in Fig. 4.11 is eigenvector g_i which approximates the radial density $g(r)$.

Determining R and $g(r)$ this way satisfies Eq. (4.53). Equation (4.52) must also be satisfied, which we checked by substitution.

Radii with unit vector attraction kernel. While we were unable to calculate the density of the static sync and async states analytically, we were able to partially

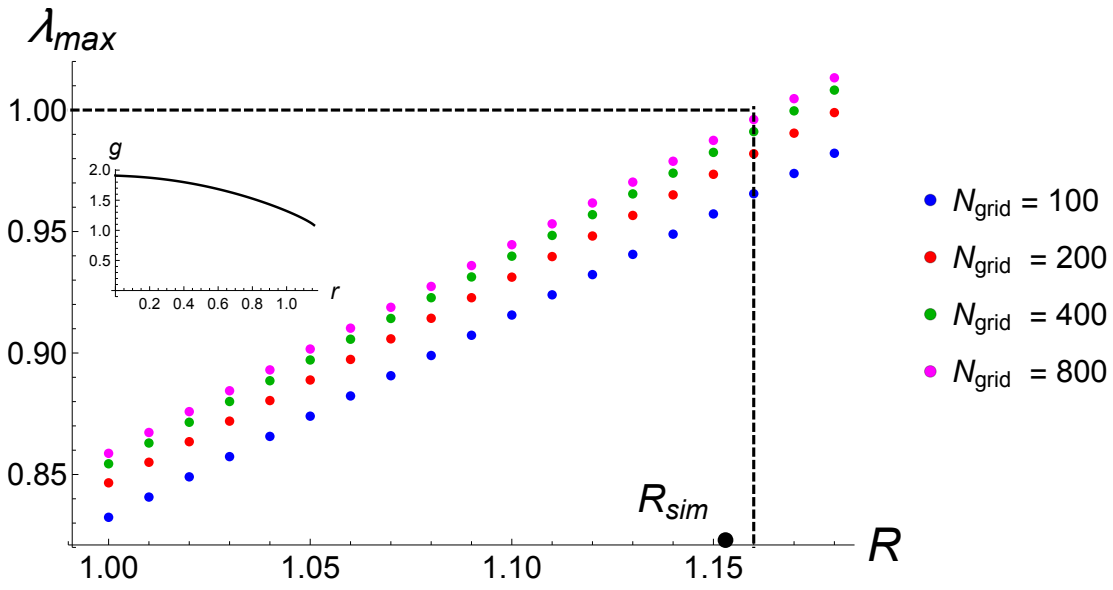

Figure 4.11: Determination of radius R of spatial density of static async state. The largest eigenvalue of the matrix M_{ij} as defined by Eq. (4.55) is plotted versus R for different N_{grid}. The black, dashed lines show that the eigenvalue closest to 1 is achieved for a radius of $R = 1.16$. This is in reasonable with the radius R_{sim} measured from simulation data for $N = 1000$ swarmalators for $(J, K) = (0, -2)$ using python's solver 'odeint' for $T = 500$ time units and a stepsize of $dt = 0.1$. The inset shows the eigenvector g_i for $N_{grid} = 800$.

calculate the radii of its support using dimensional analysis. We assume that the radii are determined when the magnitudes of attractive and repulsive forces balance. For the sync state

$$I_{att}(R_{sync})F(\theta) \propto I_{rep}(R), \tag{4.56}$$

$$(1)(1+J) \propto 1/R_{sync},$$

$$R_{sync} = C(1+J)^{-1},$$

where C is an unknown constant. Note the effective spatial attraction force is $I_{att}(\vec{x})F(\theta)$, and $F(\theta) = 1 + J\cos(0) = 1 + J$ in the static sync state (since all swarmalators have the same phase). In the static async state the calculation is the

same with $F(\theta) = 1$ giving

$$R_{async} = C. \tag{4.57}$$

We cannot find the radii directly because of the unknown constant C. Their ratio is however given by

$$\frac{R_{sync}}{R_{async}} = \frac{1}{1+J} \tag{4.58}$$

which agrees with simulation as shown in Fig. 4.12.

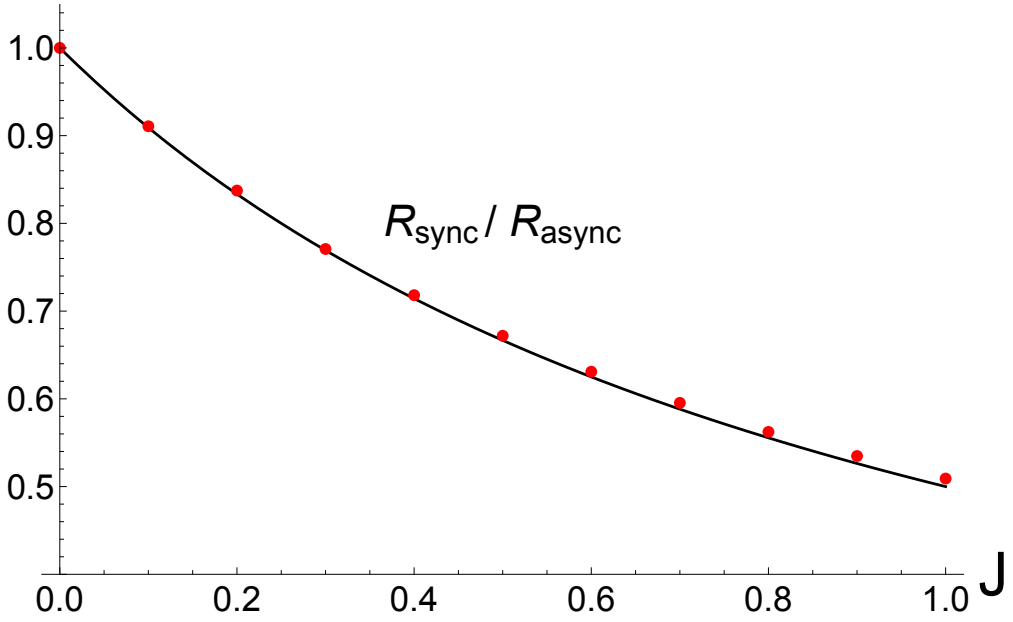

Figure 4.12: Ratio of radius of static sync (for $K = 1$) and async states (for $K = -2$) versus J for $N = 800$ swarmalators for a unit vector attraction kernel $I_{att}(\vec{x}) = \vec{x}/|\vec{x}|$.

4.9 Static phase wave state

Here we calculate the density of swarmalators, and inner and outer radii R_1, R_2 of the annulus, in the static phase wave state, when a linear attraction kernel is

used.

Density. The calculation is the same as for the static sync and async states: we assert

$$\underline{v} \equiv \underline{0}, \tag{4.59}$$

$$\nabla.\underline{v} \equiv 0. \tag{4.60}$$

The density of the static phase wave state is

$$\rho(r, \phi, \theta, t) = (2\pi)^{-1} g(r)\delta(\phi - \theta), \quad R_1 < r < R_2. \tag{4.61}$$

We first calculate the divergence, which in polar coordinates is given by

$$\nabla.\underline{v} = \frac{1}{r}\frac{\partial(rv_r)}{\partial r} + \frac{\partial(rv_\phi)}{\partial \phi} + \frac{\partial v_\theta}{\partial \theta}. \tag{4.62}$$

The velocity $\underline{v} = (v_r, v_\phi, v_\theta)$ is given by

$$\begin{aligned} v_r = \int \Big(\tilde{r}\cos(\tilde{\phi} - \phi) - r\Big)\Big(1 + J\cos(\tilde{\theta} - \theta) \\ - \frac{1}{\tilde{r}^2 - 2r\tilde{r}\cos(\tilde{\phi} - \phi) + r^2}\Big)\rho(\tilde{r}, \tilde{\phi}, \tilde{\theta})\tilde{r}\, d\tilde{r}\, d\tilde{\phi}\, d\tilde{\theta}, \end{aligned} \tag{4.63}$$

$$\begin{aligned} v_\phi = \int \frac{\tilde{r}}{r}\sin(\tilde{\phi} - \phi)\Big(1 + J\cos(\tilde{\theta} - \theta) \\ - \frac{1}{\tilde{r}^2 - 2r\tilde{r}\cos(\tilde{\phi} - \phi) + r^2}\Big)\rho(\tilde{r}, \tilde{\phi}, \tilde{\theta})\tilde{r}\, d\tilde{r}\, d\tilde{\phi}\, d\tilde{\theta}, \end{aligned} \tag{4.64}$$

$$v_\theta = \int \frac{\sin(\tilde{\theta} - \theta)}{\tilde{r}^2 - 2r\tilde{r}\cos(\tilde{\phi} - \phi) + r^2}\rho(\tilde{r}, \tilde{\phi}, \tilde{\theta})\tilde{r}\, d\tilde{r}\, d\tilde{\phi}\, d\tilde{\theta}. \tag{4.65}$$

Taking the derivatives on these, plugging in Eq. (4.61) for ρ, and substituting the result into (4.62), gives

$$\underline{\nabla}.\underline{v} = -2 + g(r) + \frac{J}{2r}\int_{R_1}^{R_2} \tilde{r}^2 g(\tilde{r}) d\tilde{r}. \tag{4.66}$$

Setting this to zero, we see $g(r)$ satisfies

$$g(r) = 2 - \frac{J}{2r}\int_{R_1}^{R_2} \tilde{r}^2 g(\tilde{r}) d\tilde{r}, \tag{4.67}$$

which means it can be determined self-consistently in terms of R_1 and R_2. The result is

$$g(r) = 1 - \frac{\Gamma_J}{r}, \quad R_1 \leq r \leq R_2 \tag{4.68}$$

with $\Gamma_J = 2J(R_2^3 - R_1^3)\Big(3J(R_2^2 - R_1^2) + 12\Big)^{-1}$.

Inner and outer radii. Next we use the result (4.68) in $\underline{v} = \underline{0}$ to compute the inner and outer radii R_1, R_2. We first evaluate v_r by substituting (4.68) into (4.63). Performing the integration we get

$$v_r(r) = Cr + \frac{D}{r} \tag{4.69}$$

with

$$C = -R_1^2 + \frac{4JR_1(R_2^3 - R_1^3)}{3J(R_2^2 - R_1^2) + 12}, \tag{4.70}$$

$$D = 1 + \frac{R_2 - R_1}{6}\Big[-6(R_2 - R_1) + \frac{8JR_1(R_2^3 - R_1^3)}{3J(R_2^2 - R_1^2) + 4}\Big]. \tag{4.71}$$

Since $\underline{v} \equiv \underline{0}$, the coefficients C, D must be zero. This yields two equations for R_1, R_2, with solutions

$$R_1 = \Delta_J \frac{-\sqrt{3}J - 3\sqrt{12 - 5J}\sqrt{J + 4} + 12\sqrt{3}}{12J}, \tag{4.72}$$

$$R_2 = \frac{\Delta_J}{2\sqrt{3}}, \tag{4.73}$$

with

$$\Delta_J = \sqrt{\frac{3J - \sqrt{36 - 15J}\sqrt{J + 4} - 12}{J - 2}} \tag{4.74}$$

and small-J expansion given by

$$R_1 = \frac{J}{3} + O\left(J^2\right), \tag{4.75}$$

$$R_2 = 1 + \frac{J}{6} + O\left(J^2\right). \tag{4.76}$$

4.10 Stability of static async state

Analytic derivation. We calculate the stability of the static async state by considering a linear perturbation η in density space,

$$\rho(\underline{x}, t) = \rho_0(\underline{x}) + \epsilon\eta(\underline{x}, t), \tag{4.77}$$

where

$$\rho_0(\underline{x}) = (4\pi^2)^{-1} g(r), \quad |\vec{x}| < R \tag{4.78}$$

is the unperturbed density. The decomposition (4.77) of the density induces a decomposition in $\underline{v}$,

$$\underline{v} = \underline{v}^{(0)} + \epsilon\underline{v}^{(1)} = \epsilon\underline{v}^{(1)}, \tag{4.79}$$

where $\underline{v}^{(0)} \equiv \underline{0}$ since swarmalators are motionless in the static async state. By normalizing (4.77) we require

$$\int \eta(\underline{x}, t) d\underline{x} = 0. \tag{4.80}$$

We substitute the ansatz (4.77) into the continuity equation for $\rho(\underline{x}, t)$ to derive an evolution equation for η. Collecting terms at $O(\epsilon)$ gives

$$\dot{\eta} + \rho_0 \nabla . \underline{v}^{(1)} + \underline{v}^{(1)} . \underline{\nabla} \rho_0 = 0. \tag{4.81}$$

We first calculate the divergence $\nabla.\underline{v}^{(1)} = \partial_{\vec{x}}\vec{v}_{\vec{x}} + \partial_\theta v_\theta$. Recalling that

$$\nabla_{\vec{x}}.\frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|} = -\frac{1}{|\vec{\tilde{x}} - \vec{x}|}, \tag{4.82}$$

we see the terms on the LHS are given by

$$\begin{aligned}\partial_{\vec{x}}\vec{v}_{\vec{x}}(\vec{x},\theta,t) = -\int\Bigg(&\frac{1}{|\vec{\tilde{x}} - \vec{x}|}\Big(1 + J\cos(\tilde{\theta} - \theta)\Big) \\ &- 2\pi\delta(\vec{\tilde{x}} - \vec{x})\Bigg)\eta(\vec{\tilde{x}},\ \tilde{\theta},t)\ d\vec{\tilde{x}}\,d\tilde{\theta},\end{aligned} \tag{4.83}$$

$$\partial_\theta v_\theta(\vec{x},\theta,t) = -K\int\frac{\cos(\tilde{\theta} - \theta)}{|\vec{\tilde{x}} - \vec{x}|}\eta(\vec{\tilde{x}},\tilde{\theta},t)d\vec{\tilde{x}}\,d\tilde{\theta}. \tag{4.84}$$

We simplify these by expanding $\eta(\vec{x},t)$ in a Fourier series over θ,

$$\eta(\vec{x},\theta,t) = \sum_{n=1}^{\infty} b_n(\vec{x},t)e^{in\theta} + c.c., \tag{4.85}$$

where 'c.c.' denotes the complex conjugate. Notice the zeroth harmonic $b_0 = 0$, which follows from the normalization condition (4.80). Plugging this ansatz into (4.83) and (4.84) and performing the integration over θ leads to

$$\begin{aligned}\underline{\nabla}.\underline{v}^{(1)} = &-e^{i\theta}\frac{J}{2}\int\frac{1}{|\vec{\tilde{x}} - \vec{x}|}b_1(\vec{\tilde{x}},t)\,d\vec{\tilde{x}} \\ &- e^{i\theta}\frac{K}{2}\int\frac{1}{|\vec{\tilde{x}} - \vec{x}|}b_1(\vec{\tilde{x}},t)\,d\vec{\tilde{x}}.\end{aligned} \tag{4.86}$$

Note the appearance of the $1/|\vec{\tilde{x}} - \vec{x}|$ in the first term of Eq. (4.83) and in Eq. (4.84). This tells us that the divergence of the spatial attraction and phase attraction have the same length scale. This is a convenient property, which results from our careful choosing of the unit vector attraction $I_{att}(\vec{x}) = \vec{x}/|\vec{x}|$, so that $\nabla.I_{att}(\vec{x}) = H(\vec{x}) = 1/|\vec{x}|$.

We next calculate the third term, $\underline{v}^{(1)}.\underline{\nabla\rho_0}$, in the ODE for η given by Eq. (4.81). Note that since $\rho_0 = (4\pi)^{-1}g(r)$ for $r < R$, then the gradient is purely in the spatial direction $\underline{\nabla\rho_0} = \vec{\nabla}\rho_0 + \partial_\theta\rho_0 = \vec{\nabla}\rho_0$, which we leave in cartesian coordinates. Thus

we only need to calculate the spatial components of $\underline{v}^{(1)} = \vec{v}^{(1)}_{\vec{x}} + v_\theta$. Plugging the Fourier ansatz into the expression for $\vec{v}^{(1)}_{\vec{x}}$ and simplifying gives

$$\vec{v}^{(1)}_{\vec{x}} = e^{i\theta} \int \frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|} b_1(\vec{\tilde{x}}, t)\, d\vec{\tilde{x}}. \tag{4.87}$$

Putting our results (4.86) and (4.87) into the ODE (4.81) for η and collecting first harmonic terms $e^{i\theta}$ leads to

$$\dot{b}_1(\vec{x}, t) = -\frac{J}{2}\, \vec{\nabla}\rho_0(\vec{x}). \int \frac{\vec{\tilde{x}} - \vec{x}}{|\vec{\tilde{x}} - \vec{x}|} b_1(\vec{\tilde{x}}, t)\, d\vec{\tilde{x}} + \frac{(J+K)}{2}\rho_0(\vec{x}) \int \frac{1}{|\vec{\tilde{x}} - \vec{x}|} b_1(\vec{\tilde{x}}, t) d\vec{\tilde{x}}, \tag{4.88}$$

$$\dot{b}_n(\vec{x}, t) = 0, \quad n \neq 1. \tag{4.89}$$

To study the stability of $b_1(\vec{x}, t)$ we expand it in a Fourier series $b_1(r, \phi, t) = \sum_{m=0}^{\infty} f_m(r, t) e^{in\phi} + c.c.$. Substituting this ansatz into Eq. (4.88) and doing some algebra leads to an evolution equation for each mode $f_m(r, t)$. We then set $f_m(r, t) = e^{\lambda_m t} c_m(r)$ which leads to the following eigenvalue equation for each mode:

$$\lambda_m c_m(r) = \int_0^R H_m(r, s) c_m(s) s\, ds, \tag{4.90}$$

where the R is the radius of the support of the density in the static async state. This is the equation that appears in the main text. There, we also plot the real parts of the eigenvalues $\lambda_m^*{}'(K)$ for the five most unstable modes, for $J = 0.5$. For convenience we replot this in Fig. 4.13 below.

Finite-size scaling of eigenvalues. As can be seen, $\lambda_m^*{}'(K)$ are small but positive for $K < K_c$. But are these positive values genuine? They could be artifacts of the discretization scheme used to approximate $\lambda_m^*(K)$, which has a parameter

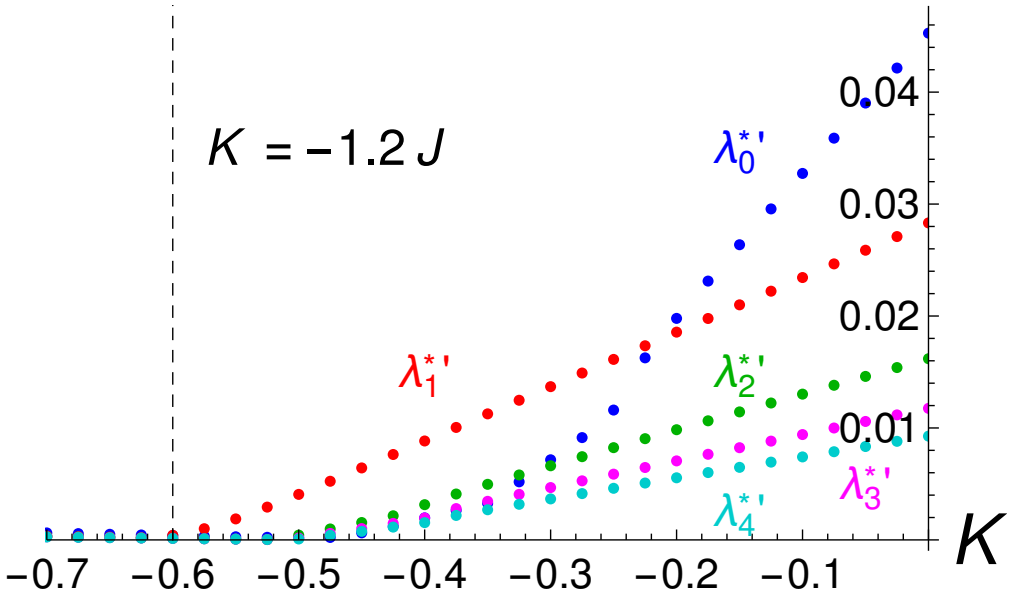

Figure 4.13: The real part of the most unstable eigenvalue $\lambda_m^{*\prime}$ of the first five of the modes f_m calculated from Eq. (4.90) for $J = 0.5$. They are all small but positive for all K, indicating the static async state is unstable as $N \to \infty$. The dashed line marks the approximation to the critical coupling strength as per Eq. (4.92). The spectra for each mode were calculated by approximating the integral of the RHS of (4.90) using a Gaussian quadrature scheme with $N' = 200$ grid points and diagonalizing the resulting matrix. The upper limit of integration $R = 1.15$ was measured from simulations. The radial density $g(r)$ was determined numerically, as described in the opening section. The kernels H_m in Eq. (4.90) for $m > 1$ were calculated numerically.

N' representing the number of grid points used to approximate the integral in (4.90). There is no reason to believe that $\lambda_m^{*\prime}$ remain positive as $N' \to \infty$; they could become zero, or negative, giving rise to different stability properties.

Hence we perform a finite-size scaling analysis of $\lambda_m^{*\prime}(K)$. In Fig. 4.14 we show $\lambda_0^{*\prime}(K)$ for $K < K_c$ for different N'. As N' is increased, $\lambda_0^{*\prime}$ gets progressively smaller while remaining positive, supporting the claim that the modes f_m are

unstable. We next used two crude approaches to probe the $N' \to \infty$ limit. First, we fit the data to a function of the form $a + bh^c$, where $h = 1/N'$. In Fig. 4.15 we show the results of this procedure for $K = -0.5$ and $K = -0.8$. As shown $\lambda_0^*(h \to 0)$ attains small but *negative* values. Second, we used Richardson extrapolation, a method used to estimate the limiting value of a converging sequence. Using the values of $\lambda_0^{*\prime}$ at $h = 1/400, 1/800, 1/1600$ it gave the approximation $\lambda_0^* = +2 \times 10^{-8}$ for $K = -0.5$ and $\lambda_0^* = -4 \times 10^{-5}$ for $K = -0.8$.

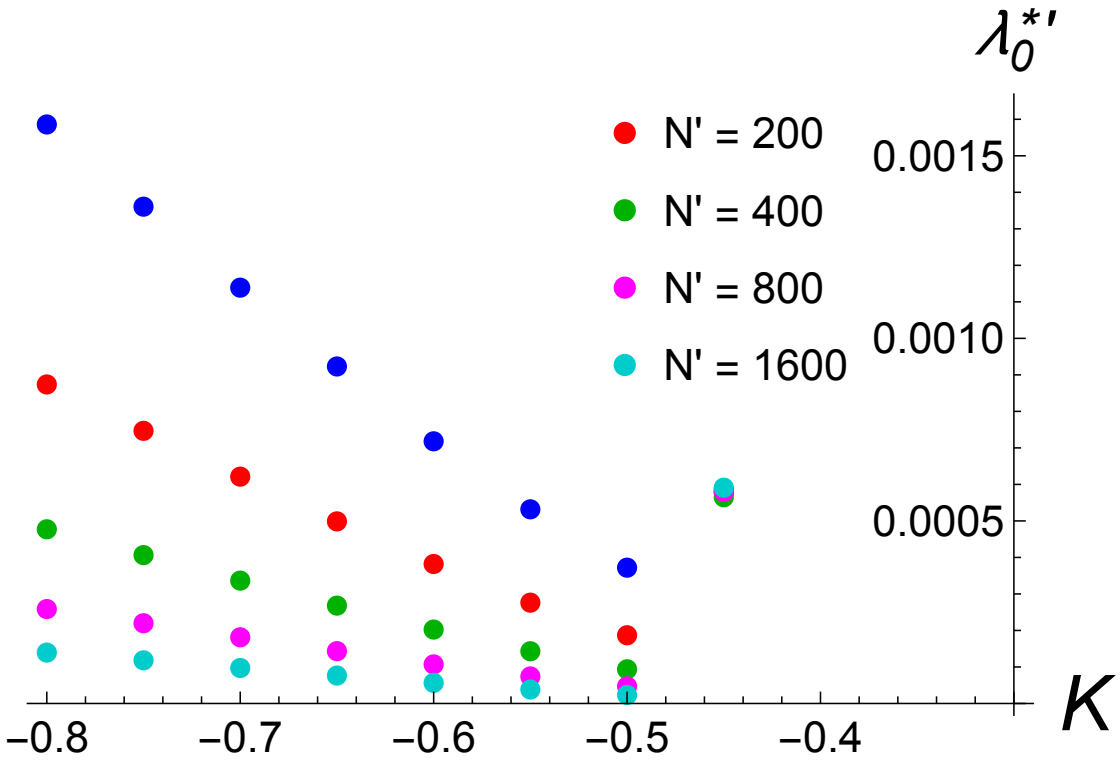

Figure 4.14: The real part of the most unstable eigenvalue of the first mode $\lambda_0^{*\prime}$ calculated from Eq. (4.90) for $J = 0.5$ for various N'. As can be seen, $\lambda_0^{*\prime}$ diminishes in magnitude for increasing N', but remains positive.

Confusingly, the two methods give approximations with different signs for $K = -0.5$. Furthermore, the magnitudes of the estimates are very small. These two facts make the results rather unconvincing. We thus declare the finite-size

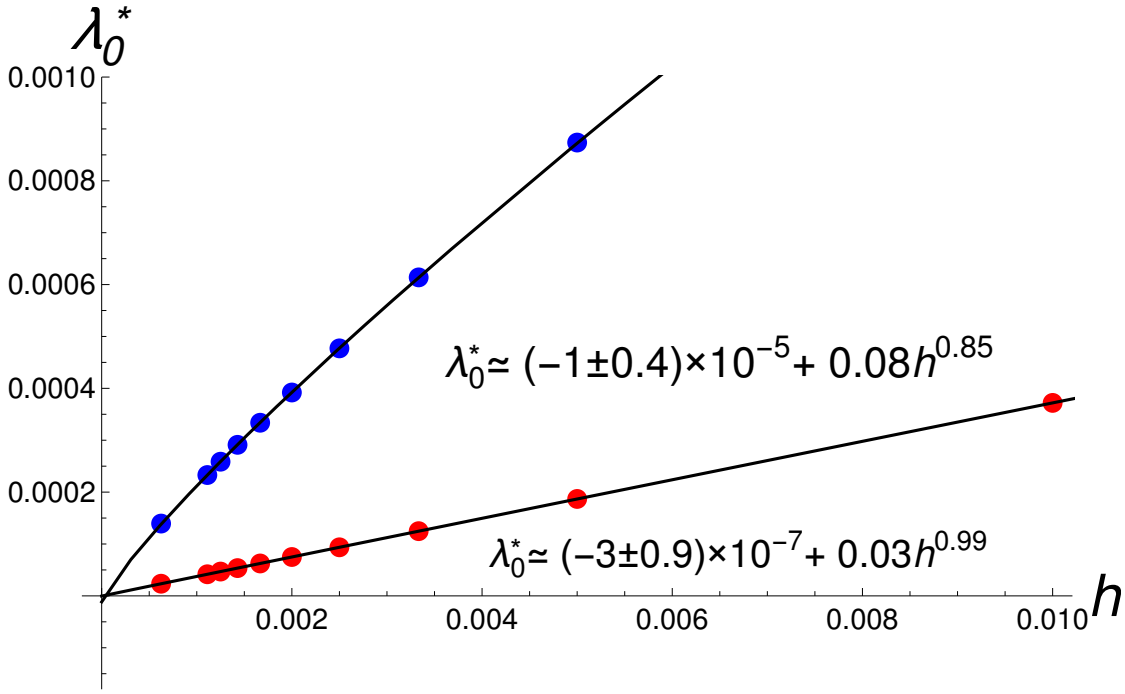

Figure 4.15: Finite scaling behavior of $\lambda_0^{*\prime}$ with respect to the $h = 1/N'$ for $J = 0.5$. Red dots are for $K = -0.5$ and blue dots are for $K = -0.8$. To obtain the $h \to 0$ limit the data were fit to curves of the form $a + bh^c$ using Mathematica. The best fit parameters a, b, c are illustrated in the plot. As can be seen $\lambda_0^{*\prime}$ approaches small but negative values as $h \to 0$. Note also the different exponents of h for $K - 0.5, -0.8$.

scaling analysis inconclusive; we have not been able to satisfactorily determine the sign of $\lambda_m^{*\prime}$ as $N' \to \infty$. This prevents us from determining the stability properties of the static async state. The only thing we can say with confidence is that there is a parameter regime $K < K_c$ where the state has *weak* stability properties, since here, their signs notwithstanding, the magnitudes of the eigenvalues are very small. For $K > K_c$, however, the eigenvalues become unambiguously positive, explicitly indicating the instability of static async state in this regime.

Critical coupling K_c. In the main text we approximated the critical coupling separating these two regimes as $K_c = \min_m K_m^* = K_2^*$, where K_m^* marks the point where $\lambda_m^{*\prime}$ starts to increase significantly. For example $K_1^* \approx -0.6$ as seen in Fig. 4.13. We use the following definition:

$$K_m^* = \frac{d^2 \lambda_m^{*\prime}}{dK^2}. \tag{4.91}$$

We calculated K_c using Eq. (4.91) for various values of J and plot the results in Fig. 4.16. As can be seen the results fall on the straight line

$$K_c = -1.2\ J. \tag{4.92}$$

To test this prediction we calculated K_c from simulation data by finding the

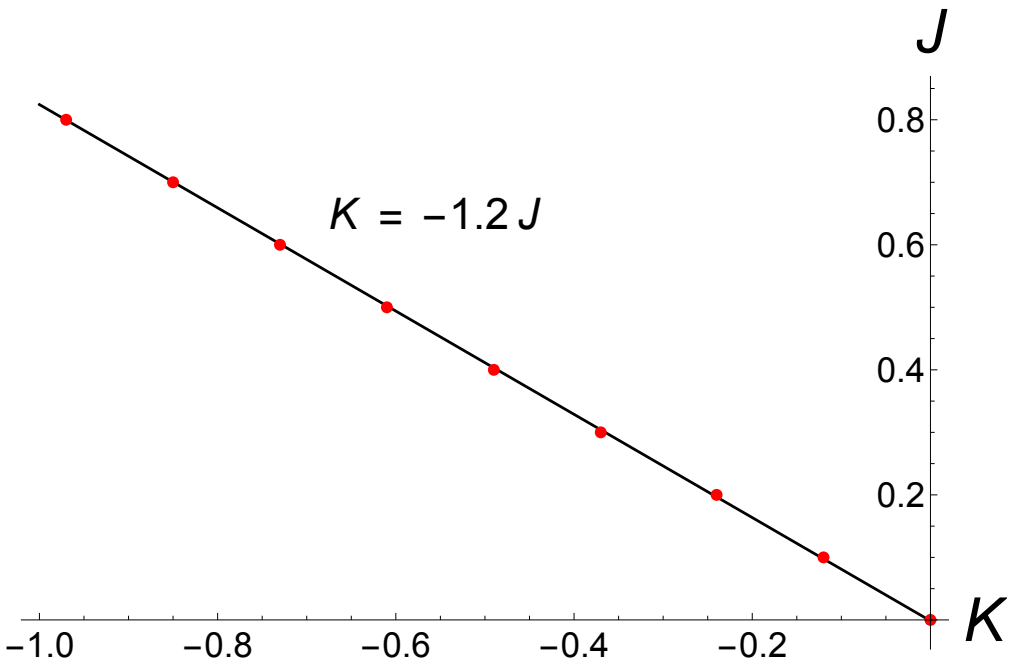

Figure 4.16: Critical coupling $K_c = K_1^*$ defined by (4.91) (red dots) versus J. The line of best fit is also shown.

point at which the order parameter $S = \max(S_+, S_-)$ bifurcates from zero. We define this K_c to be the midpoint of the interval (K_i, K_{i+1}) over which the derivative of S with respect to K is largest. We did this over a range of J, the results of

which are the black dots in the phase diagram in the main text Fig. 1. As can be seen, the straight line approximation (4.92) is reasonably accurate.

Finite-size scaling of S. Measuring K_c this way is dependent on the population size N. To get a feel for the scale of this dependence, and to see if any scaling laws are present, we here perform a finite-size scaling analysis of S and K_c. Given the computational cost of each simulation, and the necessity of collecting many realizations at each parameter value, the quality of the data we collected was limited. Hence, our analysis is preliminary.

In Fig. 4.17 we show $S(K)$ for different N. The qualitative shape of the curves do not change, but the bifurcation point K_c increases with N. In Fig. 4.18 we show $K_c(N)$. No trend is evident. We tried to fit the data to a curve of the form $a + bh^c$, but the fitting algorithms failed. This is not too surprising given the sparsity and quality of the data. Next we tried $a + bh^{1/2}$, which gave the black curve shown in the Figure. The parameter of best fit a, representing $K_c = N \to \infty$, has confidence interval $(-0.47, -0.59)$.

4.11 Noisy async state

As discussed in the main text, when white noise is added to the phase dynamics, the static async state is no longer static in the sense that the swarmalators are no longer motionless. They oscillate in space and phase, and thus we call the

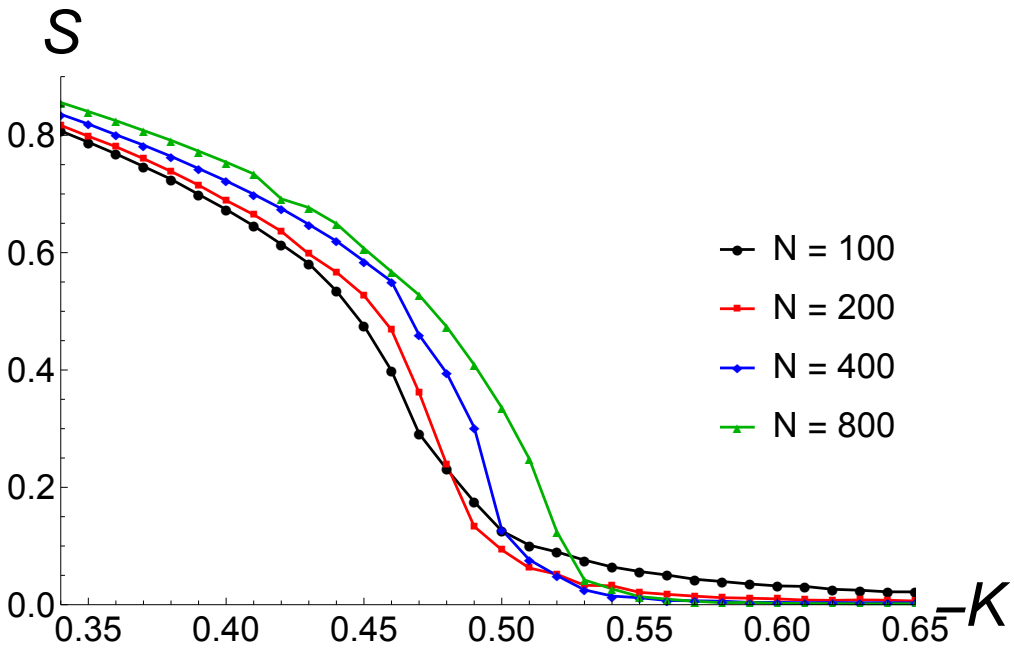

Figure 4.17: Behavior of order parameter $S = \max(S_+, S_-)$ versus K for $J = 0.5$ for increasing population sizes N. The critical coupling strength K_c at which S bifurcates from zero decreases with increasing N. Data were collected for $T = 1000$ time units with stepsize $dt = 0.01$. Each data point represents the average of 200 realizations.

new state the noisy async state. In the continuum limit, the spatial movements should vanish. In Fig. 4.19 we demonstrate this by plotting the time-averaged, mean population spatial velocity, $\langle v_x \rangle = \langle \frac{1}{N} \sum_j \sqrt{\dot{x}_j^2 + \dot{y}_j^2} \rangle_t$ where $\langle . \rangle_t$ is the time average, as a function of K for $J = 0.5$ and $D_\theta = 0.06$. As can be seen, there is a sharp drop in the velocity at a critical value of K indicating the transition to the active phase wave state. In this state, the spatial velocity decays to zero with increasing population size like $N^{-1/2}$, as indicated by the downward arrow. We show $\langle v_x \rangle$ versus N for a fixed value of K in Fig. 4.20, in which the $N^{-1/2}$ dependence is clear. We have fitted the data to a curve of form $a + bN^{-c}$, plotted along with the data.

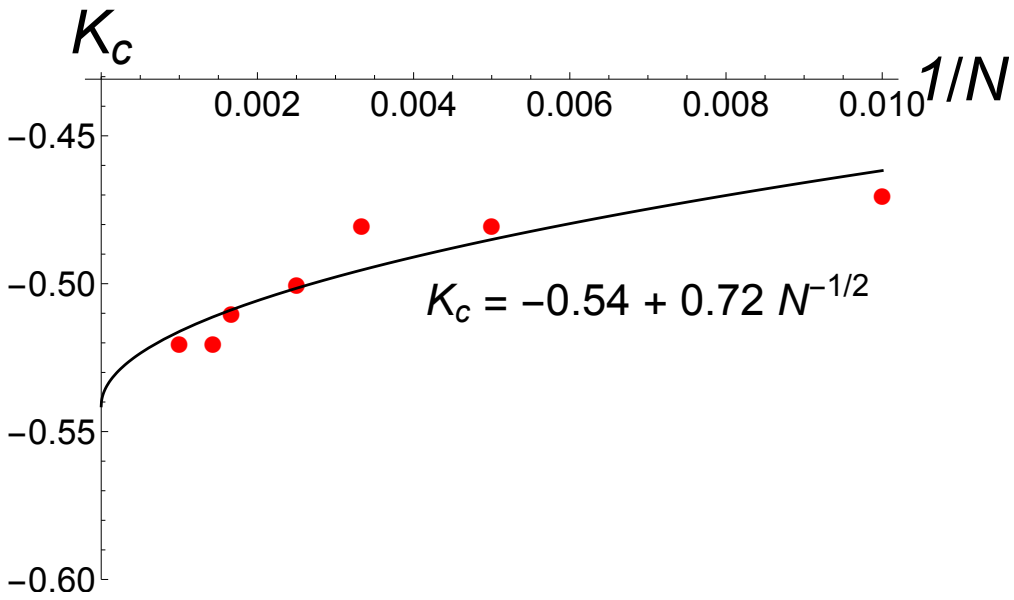

Figure 4.18: Behavior of the critical coupling strength K_c at which the order parameter S bifurcates from 0 for increasing population sizes. Red dots show simulation results. The black curve has form $K_c \approx a + bN^{-1/2}$, where the parameters of best fit a, b were determined using Mathematica.

4.12 Genericity of phenomena

The model we introduced in this work is

$$\dot{\vec{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N} \vec{I}_{att}(\vec{x}_j - \vec{x}_i)F(\theta_j - \theta_i) - \vec{I}_{rep}(\vec{x}_j - \vec{x}_i), \tag{4.93}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j\neq i}^{N} H(\theta_j - \theta_i)G(\vec{x}_j - \vec{x}_i). \tag{4.94}$$

We studied the following instance of this model:

$$\dot{\vec{x}}_i = \frac{1}{N}\sum_{j\neq i}^{N} \frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|}\Big(1 + J\cos(\theta_j - \theta_i)\Big) - \frac{\vec{x}_j - \vec{x}_i}{|\vec{x}_j - \vec{x}_i|^2}, \tag{4.95}$$

$$\dot{\theta}_i = \frac{K}{N}\sum_{j\neq i}^{N} \frac{\sin(\theta_j - \theta_i)}{|\vec{x}_j - \vec{x}_i|}. \tag{4.96}$$

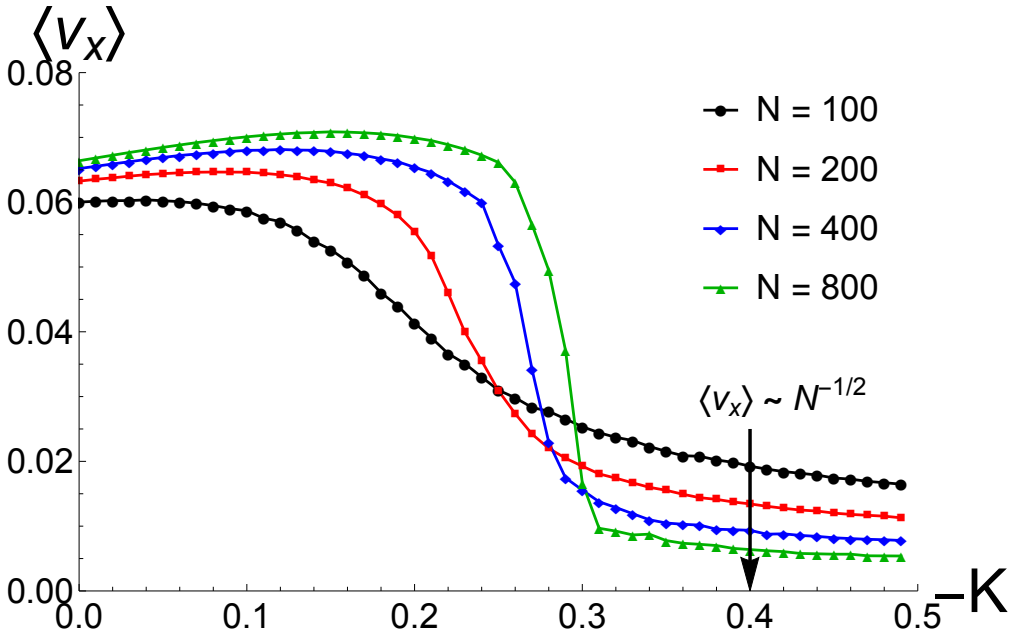

Figure 4.19: Finite-size scaling of the asymptotic spatial velocity $\langle v_x \rangle$ versus K for $J = 0.5$ and $D_\theta = 0.06$ for increasing population sizes N. Simulations were run for 10^5 timesteps of size 0.01 using Heun's method. For sufficiently negative K, there is a drop in $\langle v_x \rangle$ indicating the transition to the active async state. In this state, the velocity decays as $\langle v_x \rangle \sim N^{-1/2}$.

To check if the phenomena we found are generic for the model, as opposed to being specific to the *instance* of the model, we ran simulations for the different choices of the $I_{att}(\vec{x}), I_{rep}(\vec{x}), G(\vec{x})$ in equations (4.97)-(4.101):

$$I_{att}(\vec{x}),\ I_{rep}(\vec{x}),\ G(\vec{x}) = \frac{\vec{x}}{|\vec{x}|^2},\ \frac{\vec{x}}{|\vec{x}|^4},\ \frac{1}{|\vec{x}|^3} \tag{4.97}$$

$$= \frac{\vec{x}}{|\vec{x}|^2},\ \frac{\vec{x}}{|\vec{x}|^4},\ \frac{1}{|\vec{x}|} \tag{4.98}$$

$$= \frac{\vec{x}}{|\vec{x}|^2},\ \frac{\vec{x}}{|\vec{x}|^4},\ \exp(-|\vec{x}|^2/\nu) \tag{4.99}$$

$$= \vec{x}\exp(-|\vec{x}|^2/\nu),\ \frac{\vec{x}}{|\vec{x}|^2},\ \frac{1}{|\vec{x}|} \tag{4.100}$$

$$= \vec{x},\ \frac{\vec{x}}{|\vec{x}|^2},\ \frac{1}{|\vec{x}|} \tag{4.101}$$

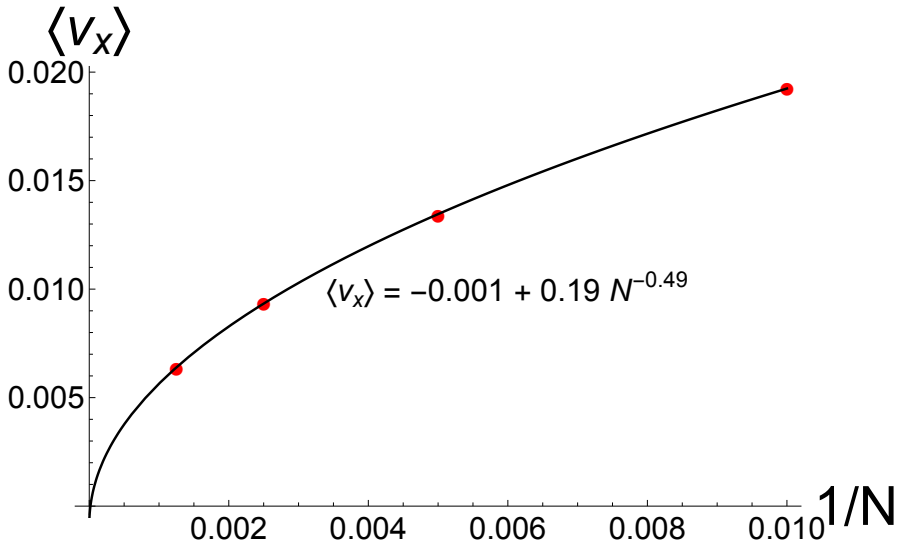

Figure 4.20: Finite-size scaling of the asymptotic spatial velocity $\langle v_x \rangle$ versus population size N for $J = 0.5$, $D_\theta = 0.06$ and $K = -0.4$. Red dots shown simulation results for runs with 10^5 timesteps of size 0.01 using Heun's method. The solid black curve has form $a + bN^{-c}$, where a, b, c are the best fit parameters to the data. The $K_c \sim N^{-1/2}$ behavior is evident.

In all these cases, we found the same phenomena, as evidenced by the behavior of the order parameters for these new instances of the model, shown in Fig. 4.21.

The first two changes were to check if our phenomena were dependent on length scale. In modification (4.97) we replaced the physically unrealistic unit vector attraction kernel with one that spatially decays. We chose a power law with exponent -1. To keep the length scale of $I_{rep}(\vec{x})$ shorter than that of $I_{att}(\vec{x})$, we changed the latter's exponent to -3. Further, to keep the length scale of the phase dynamics (captured by the function $G(\vec{x})$) the same as that of $I_{rep}(\vec{x})$, as it is in the original model, we changed the exponent of $G(\vec{x})$ to -3. In the modification (4.98), we matched the exponents of $I_{att}(\vec{x})$ and $G(\vec{x})$, so that the length scale

of the phase dynamics was the same as that of spatial attraction. As displayed in panels (a) and (b) of Fig. 4.21, these changes led to no new behavior.

The purpose of the next two modifications (4.99) and (4.100) was to check if the phenomena were independent of specific functional forms. To this end, we replaced the power-law choices for $I_{att}(\vec{x})$ and $G(\vec{x})$ with exponentials with tunable length scales ν. For modification (4.99) we varied ν from 0 to 1 and found the same phenomena. For the modification (4.100), we found that for $\sigma \gtrsim 1$, the same phenomena were realized. However when $\sigma \lesssim 1$, only the static sync and static async states were found. The theoretical reasons for this are not yet understood, and are left for future work.

Non-stationary phase waves. The final modification (4.101) corresponds to the original model with the unit vector attraction kernel replaced by a linear kernel. While a linear function is somewhat unrealistic, physically speaking, it has the advantage of being much easier to analyze. In particular, in combination with the newtonian repulsion, it leads to swarmalators being confined to disks of uniform density in the static sync and static async states. As we showed in the preceding sections, this let us solve for the radii of the support disks in these states, as well as letting us solve for the density of the static phase wave state, along with its inner and outer radii.

The disadvantage of the linear attraction kernel is that it has non-generic behavior, which we why we chose not to present it in the main text. Its non-genericity is such that it leads to *extra* states not found in other instances of the swarmalator

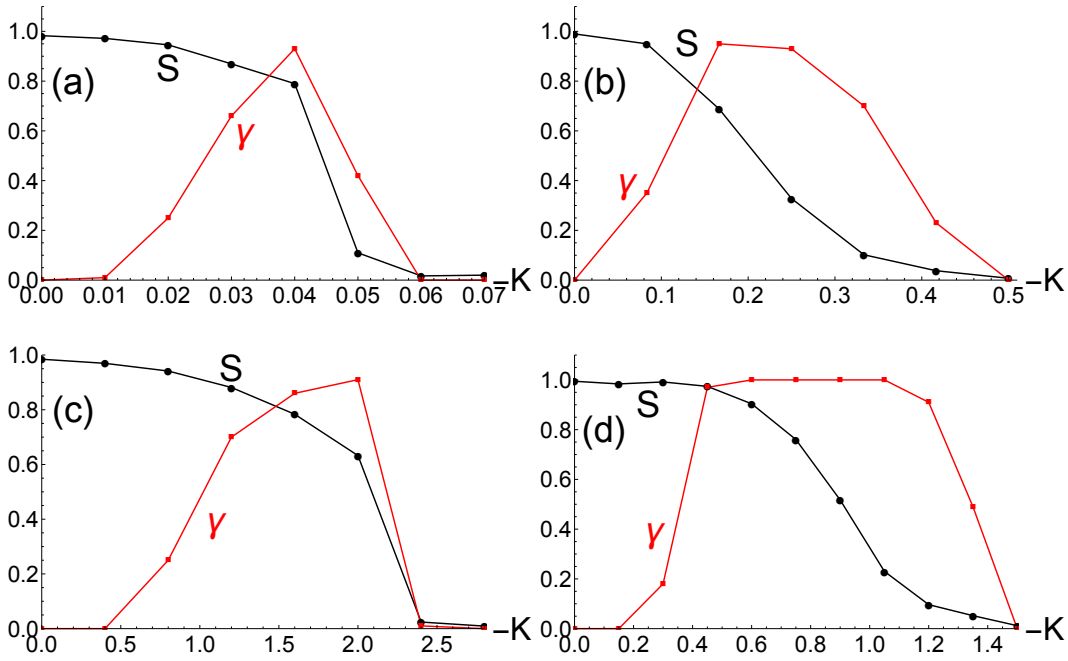

Figure 4.21: Behavior of the order parameters $S = \max(S_+, S_-)$ and γ for different choices of interactions functions. Unless otherwise stated, simulations were run for $N = 100$ swarmalators with $J = 1.0$ for $T = 1000$ time units with a stepsize of $dt = 0.5$ using python's 'odeint' solver. In all cases, the same qualitatively behavior was found, indicating the same states as the original model were realized. Panels (a) through (d) correspond to equations (4.97), through (4.100). In panel (c), $\sigma = 0.1$ and in panel (d) $\sigma = 3.0$. In panel (a) a longer time of $T = 2000$ was used.

model. We call all these new states *non-stationary phase waves*, since (in contrast to states we have studied so far) they correspond to non-stationary densities of the swarmalators. As K is decreased, the non-stationary phase waves bifurcate from the active phase wave, before ultimately morphing into the static async state, as shown in a plot of the order parameters in Fig. 4.24. During this transition, the phase $\Psi_\pm$ of the order parameter $W_\pm = S_\pm e^{i\Psi_\pm} = N^{-1}\sum_j e^{i(\phi_j \pm \theta_j)}$ changes from being constant to being time-dependent. For values of K near the active phase wave, $W_\pm$ rotates uniformly, so that $\Psi = \Omega t$, as shown in Fig. 4.22(b).

At the swarmalator level, this means the correlation between the spatial angle ϕ and phase θ becomes periodic, namely $\theta \approx \pm\phi + C(t)$, as illustrated by the moving bands in (ϕ, θ) space shown in Fig. 4.23 and in Supplementary Movie 8.

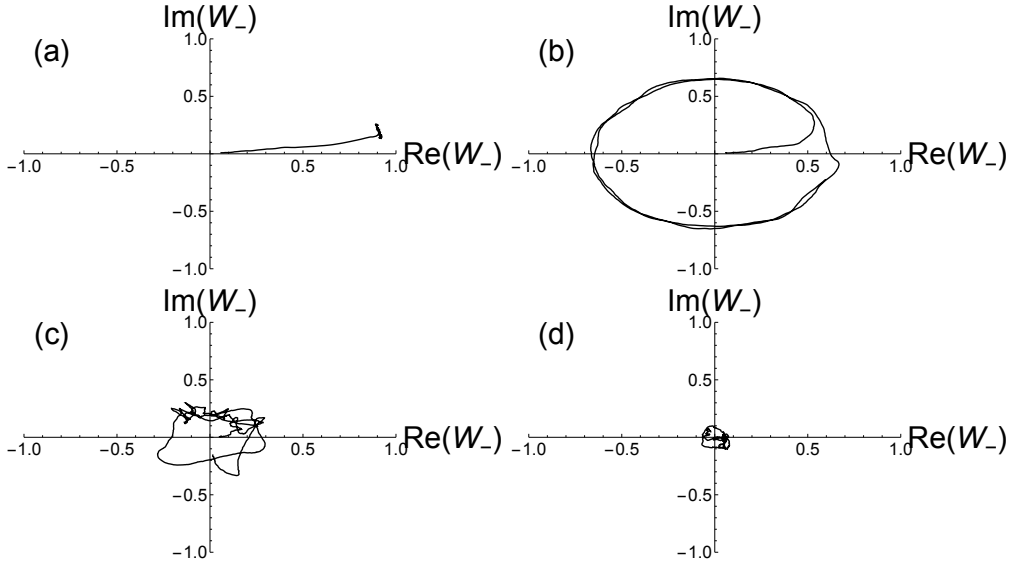

Figure 4.22: Time series of the order parameters W_- in the complex plane for different values of K. In all panels, simulations were run for $N = 1000$ swarmalators with $(dt, T) = (0.5, 500)$ and $J = 0.5$. (a) $K = -0.3$. Both the phase and amplitude of W_- approach constant values, indicating the active phase wave state. (b) $K = -0.4$. W_- rotates with constant velocity and amplitude, indicating the non-stationary phase wave state. (c) $K = -0.5$. Both the amplitude and phase of W_- oscillate irregularly, indicating another version of the non-stationary phase wave state. (d) $K = -0.8$. Both the amplitude and phase of W_- are zero (up to finite effects), indicating the static async state.

For more negative K, wilder versions of the traveling phase wave are encountered. Here, both $S_\pm$ and $\Psi_\pm$ start to oscillate irregularly, as shown in Fig. 4.22(c). The precise nature of this irregularity is not yet understood theoretically, and is left for future work. As K is further decreased, the average amplitude of S smoothly decreases to zero, at which point the static async state is achieved, as shown in Fig. 4.22(d).

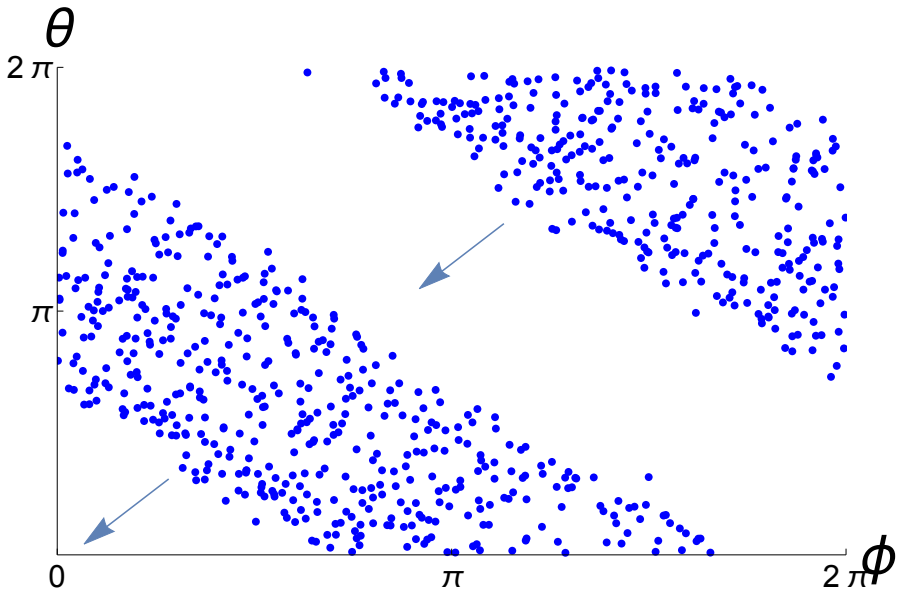

Figure 4.23: Scatter plots of the swarmalators' spatial angles ϕ and phases θ in the non-stationary phase wave state. Simulations were performed for $N = 800$ swarmalators with $dt = 0.01$ and $(J, K) = (0.9, -0.11)$. As illustrated by the arrows, the offset of the correlation between ϕ and θ changes uniformly between 0 and 2π.

To distinguish between the active, and non-stationary, phase wave states, we define the order parameter Γ. This is 1 if $\Psi_{\pm}$ has executed at least one cycle, after transients have been discarded. In conjunction with the order parameters $S_{\pm}$ and γ, this lets us discern all the macroscopic states as K is varied, as shown in Fig. 4.24.

4.13 Swarmalators in 3D

In the main text we explored the behavior of our model in three dimensions for identical swarmalators with no noise. We plotted four of the five states in the

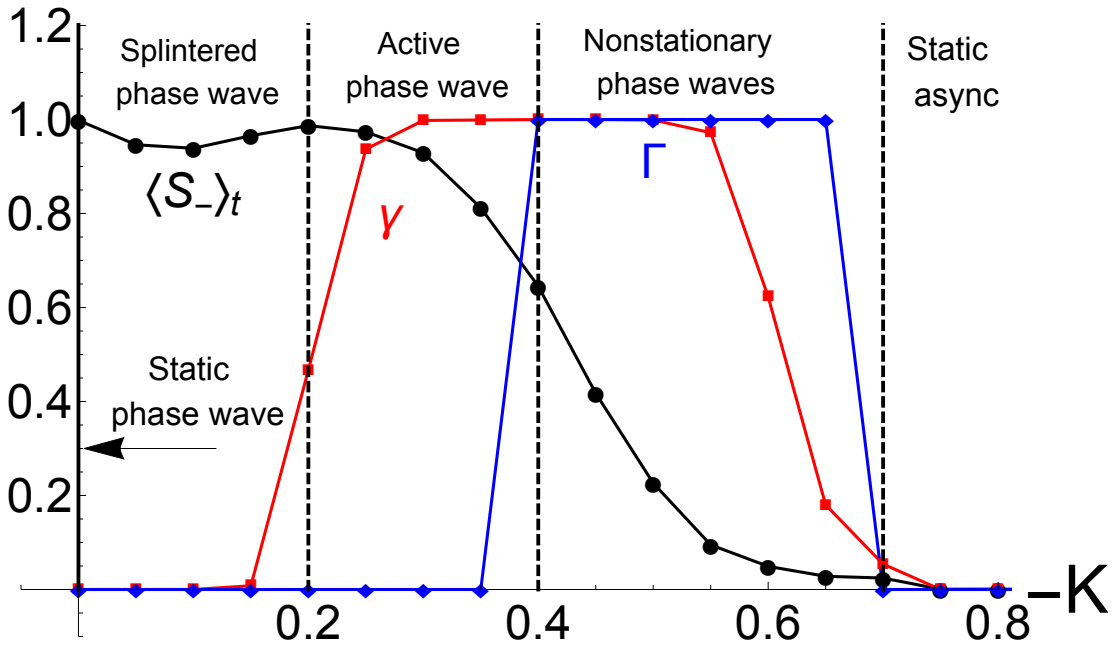

Figure 4.24: Order parameters for $N = 1000$ swarmalators with $(dt, T) = (0.5, 1000)$ and $J = 0.5$. The first 50% of data were discarded as transients. Angled brackets $\langle . \rangle$ denote time average. Plots are qualitatively similar for other values of J.

(x, y, z) plane, where each swarmalator was colored according to its phase. Here we show these same four states as scatter plots in (ϕ, θ) plane, where ϕ is the azimuthal angle of swarmalators $\phi = \tan^{-1}(y/x)$. Fig. 4.25 shows the results, where the points have been colored according to their polar angle $\alpha = \cos^{-1}(\frac{z}{\sqrt{x^2+y^2+z^2}})$. As can be seen, the plots look qualitatively the same as those obtained for the 2D model. We also show the states in the (θ, α) plane, and color particles according to their polar angle, depicted in Fig. 4.26.

The 3D analogues of the order parameters S and γ also have the same qualitative behavior as their 2D versions, as seen in Fig. 4.27.

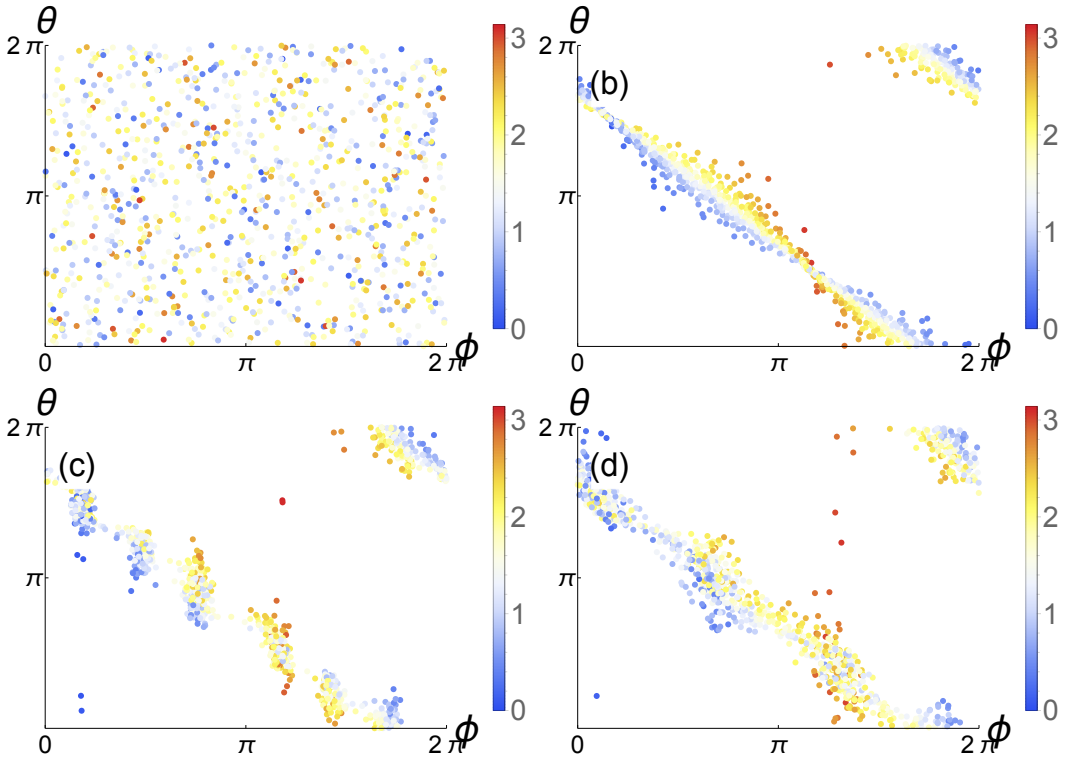

Figure 4.25: Scatter plots of four states in the $(\phi.\theta)$ plane, where the swarmalators are colored according to their polar angle α. Data were collected for $N = 1000$ swarmalators for 5×10^5 timesteps of width $dt = 0.001$ using Heun's method. The first 50% of data were discarded as transients. Parameter values were $(J, \sigma, D_\theta, D_x, D_y, D_z) = (0.5, 0, 0, 0, 0, 0)$. (a) Static async state for $K = -1$. (b) Static phase wave for $K = 0$. (c) Splintered phase wave for $K = -0.05$. (d) Active phase wave state for $K = -0.6$.

The same is true of the radii of the spherical densities of the static async and static async states, which in the main text we showed were $R_{async} = 1$ and $R_{sync} = (1 + J)^{-1/3}$. Prohibitively large population sizes would be needed to confirm these predictions in detail. To get around this, we instead investigated the finite-size scaling behavior of the radii. In Fig. 4.28(a) we show R_{async} for progressively larger population sizes N. As can be seen, the data appear to converge to the theoretical value of 1. We confirmed this convergence by fitting the data to a curve of the form $a + bN^{-c}$ using Mathematica, which returned a best fit param-

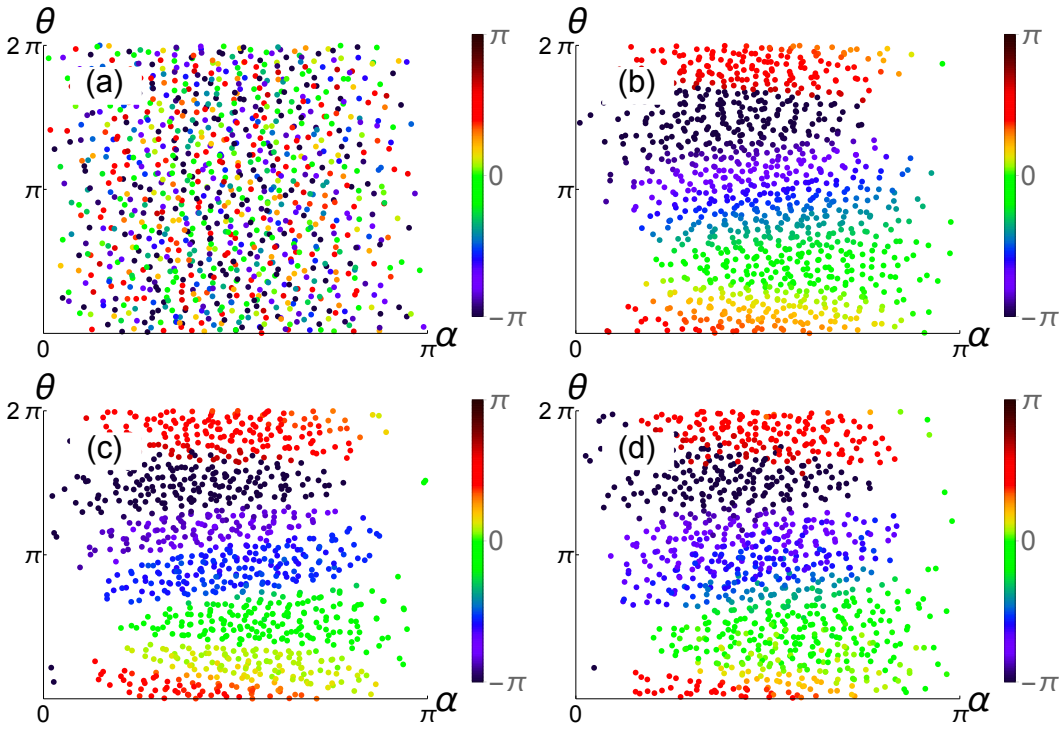

Figure 4.26: Scatter plots of four states in the $(\alpha.\theta)$ plane, where the swarmalators are colored according to their polar azimuthal angle ϕ. Parameter and simulations details are the same as for Figure 4.25

eter of $a = 0.998$. In Fig. 4.28(b), we plot R_{sync} versus J. As before, for each each value of J, the radius $R_{sync}(N)$ was calculated for increasing population sizes N, from which the $N \to \infty$ limit was found by fitting the data to a curve $a + bN^{-0.5}$, and identifying $a = R_{sync}$. As can be seen there is good agreement between simulation and theory.

We next checked robustness to noise. As in two dimensions, all but the splintered phase wave state persist in the presence of noise and distributed natural frequencies. We show these states in Fig. 4.29 below.

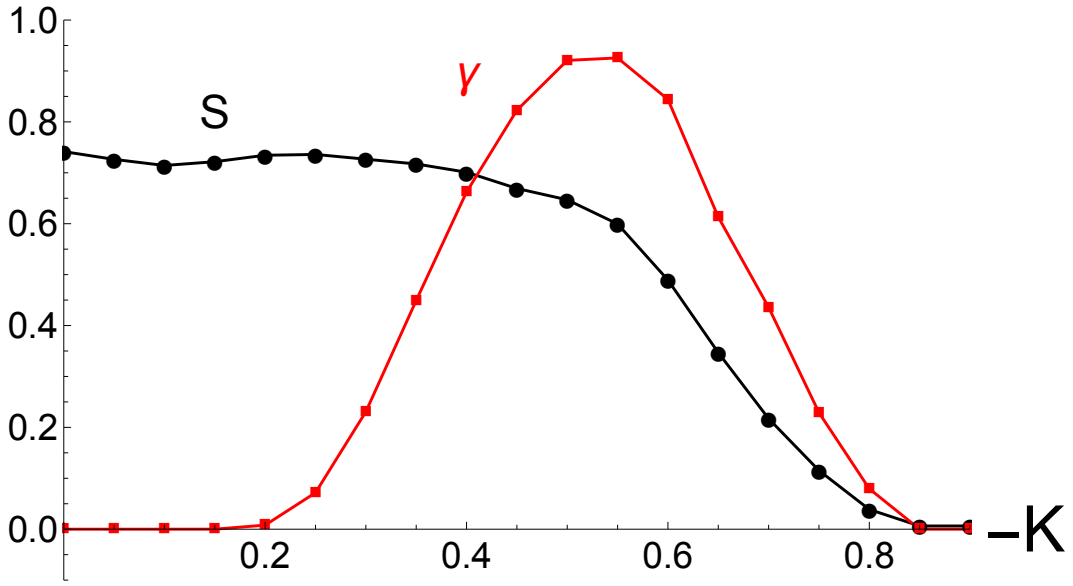

Figure 4.27: Behavior of 3D versions of the order parameters $W_{\pm} = N^{-1}\sum_j e^{i(\phi_j \pm \theta_j)}$ and order parameter γ when $J = 0.5$. The same qualitative behavior of the as the 2D case is evident. Simulations have been run for $N = 500$ swarmalators using Heun's method for 5×10^5 time steps of width 0.001. One hundred realizations were computed for each value of K, the average of which is plotted.

4.14 Swarmalators in 1D

In our work so far, swarmalators have been confined to the (x, y) plane. Our motivation for this was physical realism. However, as we discovered, the steady states of our system were often one dimensional, displaying circular symmetry. This raises the question: could we recover the same phenomena by confining swarmalators to move on a circle? In this section, we show that we can.

The 1D governing equations are

$$\dot{x}_i = \frac{1}{N}\sum_{j=1}^{N} I_{att}(x_j - x_i)F(\theta_j - \theta_i) - I_{rep}(x_j - x_i), \tag{4.102}$$

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} G_{att}(\theta_j - \theta_i)H(x_j - x_i), \tag{4.103}$$

where x_i is an angle parameterizing the position of swarmalator i on the unit circle, and the various functions are as described in the main text. We consider the following instance of this model:

$$\dot{x}_i = \frac{A}{N}\sum_{j=1}^{N} \sin(x_j - x_i)\Big[B + J\cos(\theta_j - \theta_i) - C\Big(\frac{1+\cos(x_j - x_i)}{2}\Big)^n\Big], \tag{4.104}$$

$$\dot{\theta}_i = \omega + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i)\Big(\frac{1+\cos(x_j - x_i)}{2}\Big)^n. \tag{4.105}$$

The parameters J, K are the same as those in the main text. We choose the sine function for G_{att}, I_{att} for simplicity, again in the spirit of the Kuramoto model. The function $H(x) = 2^{-n}(1+\cos x)^n$ represents a smooth pulse of unit strength. It contains the parameter n, a positive integer that controls the width of the pulse. Increasing n decreases the width, as shown in Fig. 4.30. The function $F(\theta)$ is also shown for convenience.

The remaining parameters are A, B, C and the natural frequency ω. By rescaling time and going to a co-rotating frame with $\theta_i \to \theta_i + \omega t$, we can set $A = 1$ and $\omega = 0$ respectively, without a loss in generality. To keep our model minimal, we wish to fix B, C, which control the strength of the attractive and repulsive spatial forces. We want I_{rep} to mimic hard-shell repulsion. To this end, we need $C > B$. By experimenting numerically, we find $B = 1$ and $C = 10$ produced the desired behavior. For smaller values of C, swarmalators collapse to a single point x.

Lastly, our simulations also indicate that the phenomena do not depend on the spatial length scale n. We therefore also fix $n = 1$, unless otherwise stated.

Phenomena. To investigate the behavior of our system, we again performed numerical experiments using python's ODE solver 'odeint'. We selected the initial positions and phases of the swarmalators uniformly at random from $[-\pi, \pi]$. We find the system settles into 1D versions of the five states found in 2D. Figure 4.31 shows the phase diagram, which is qualitatively the same as the Fig. 1 in the main text. The 1D stationary states are shown in Fig. 4.32 and panels (a) and (b) of Fig. 4.33, and are direct analogues of their 2D versions.

The non-stationary states are shown in Fig. 4.34 and panels (c) and (d) of Fig. 4.33. These are different from their 2D analogues, in that the density $\rho(x, \theta, t)$ is now non-stationary: in the splintered phase wave state, the clusters of similarly-phased swarmalators gently expand and contract in time. These contractions are also present in the active phase wave state, but are now more violent, as indicated in Fig. 4.33(d).

Analysis. As before, the order parameters $S_{\pm}$ and γ together let us discriminate between all the states of our systems, as illustrated in Fig. 4.35. At $K = K_1(J)$, we see a sharp transition in γ, identifying the change from the splintered phase wave to the active phase wave. At $K = K_2(J)$, we see both γ and $S_{\pm}$ drop to zero, signaling the transition to the static async state. After the following subsection, we again perform a linear stability analysis in density space to find K_2

analytically,

$$K_2(n, J) = -\frac{\sqrt{\pi}}{2}\frac{\Gamma(n)}{\Gamma(n+1/2)}(n+1)\ J \tag{4.106}$$

which for $n = 1$ reduces to the line $K_2 = -2J$ drawn in Fig. 4.31.

Genericity. To check if our phenomena are generic, we ran simulations for different choices of the functions $I_{rep}(x), F(\theta), H(x)$, summarized below:

$$I_{rep}(x) = \frac{1}{C - \cos(x)}, \tag{4.107}$$

$$F(\theta) = \exp(-\frac{\sin(\theta/2)^2}{2\sigma}), \tag{4.108}$$

$$H(x) = \exp(-\frac{\sin(x/2)^2}{2\sigma}). \tag{4.109}$$

In all cases, the phenomena we found were the same. In contrast to the 2D swarmalators, we could not find a 1D realization of the traveling phase wave states. Figure 4.36 below shows the order parameters S, γ for the $H(x)$ given by (4.109). As can be seen, their behavior is qualitatively the same as that in Fig. 4.35 above.

Stability of static async state in 1D. We investigate the stability of the static async state by performing a linear stability analysis in density space. The analysis is the same as for the 2D case. Let

$$\rho(x, \theta, t) = \rho_0(x, \theta, t) + \epsilon\eta(x, \theta, t) \tag{4.110}$$

as before, except now ρ_0 is given by

$$\rho_0(x, \theta) = \frac{1}{4\pi^2}. \tag{4.111}$$

Substituting the ansatz (4.110) into the continuity equation and collecting terms at $O(\epsilon)$ gives

$$\dot{\eta} + \rho_0 \underline{\nabla}.\underline{v}^{(1)} = 0. \tag{4.112}$$

We expand $\eta(x, \theta, t)$ in a Fourier series,

$$\begin{aligned} \eta(x,\theta,t) = &\sum_{m=0,l=0} \alpha_{m,l}(t)\cos(mx)\cos(l\theta) \\ &+ \beta_{m,l}(t)\sin(mx)\sin(l\theta) \\ &+ \gamma_{m,l}(t)\cos(mx)\sin(l\theta) \\ &+ \delta_{m,l}(t)\sin(mx)\cos(l\theta). \end{aligned} \tag{4.113}$$

Note the zero modes $\alpha_{0,0}, \beta_{0,0} \dots$ are zero via the normalization condition $\int \eta(x,\theta,t) dx\, d\theta = 0$ (but the $\alpha_{0,1}, \alpha_{1,0} \dots$ aren't necessarily). Computing the divergence, and plugging the result into (4.112), and projecting onto the various basis vectors $\cos(mx)\cos(l\theta), \dots$ leads to the ODEs for the Fourier components $\alpha_{m,l}, \beta_{m,l}, \gamma_{m,l}, \delta_{m,l}$. The calculation is difficult to achieve for a general n, so we instead tabulate the results for $n = 1, 2, \dots$). The results for $n = 1$ are

$$
\begin{aligned}
\dot{\alpha}_{0,1} + \tfrac{1}{2}(B-2)\alpha_{0,1} = 0 \qquad & \qquad \dot{\delta}_{1,0} - \tfrac{1}{2}K\delta_{1,0} = 0 \\
\dot{\gamma}_{0,1} + \tfrac{1}{2}(B-2)\gamma_{0,1} = 0 \qquad & \qquad 8\dot{\alpha}_{1,1} - \alpha_{1,1}(2J+K) = 0 \\
\dot{\alpha}_{0,2} + \tfrac{1}{2}B\alpha_{0,2} = 0 \qquad & \qquad 8\dot{\beta}_{1,1} - \beta_{1,1}(2J+K) = 0 \\
\dot{\gamma}_{0,2} + \tfrac{1}{2}B\gamma_{0,2} = 0 \qquad & \qquad 8\dot{\gamma}_{1,1} - \gamma_{1,1}(2J+K) = 0 \\
\dot{\alpha}_{1,0} - \tfrac{1}{2}K\alpha_{1,0} = 0 \qquad & \qquad 8\dot{\delta}_{1,1} - \delta_{1,1}(2J+K) = 0
\end{aligned}
\tag{4.114}
$$

where we have omitted all equations of the form $\dot{\alpha}_{l,m} = 0$, $\dot{\beta}_{l,m} = 0$, etc . Looking at the $\alpha_{0,1}, \gamma_{0,1}$ modes, we see the static async is stable when

$$B > 2. \tag{4.115}$$

We remind the reader that B is the strength of the spatial repulsion. So this condition is telling us that B needs to be sufficiently large for the static async state to exist. From the $\alpha_{1,0}, \delta_{1,0}$ and $\alpha_{0,2}, \gamma_{0,2}$ modes, we see we also require $K < 0$, and $B > 0$ respectively for the static async to be stable. These conditions are trivially satisfied in the parameter regime of interest.

Looking at the $\alpha_{1,1}, \beta_{1,1}, \ldots$ modes finally gives us the desired critical coupling strength:

$$K_2(n = 1) = -2J. \tag{4.116}$$

We next repeated the calculation for $n = 2, 3 \ldots$, and then using Mathematica's

"FindSequenceFunction", found the critical coupling K_2 for a general n:

$$K_2(n) = -\frac{\sqrt{\pi}}{2}\frac{\Gamma(n)}{\Gamma(n+1/2)}(n+1)\,J. \tag{4.117}$$

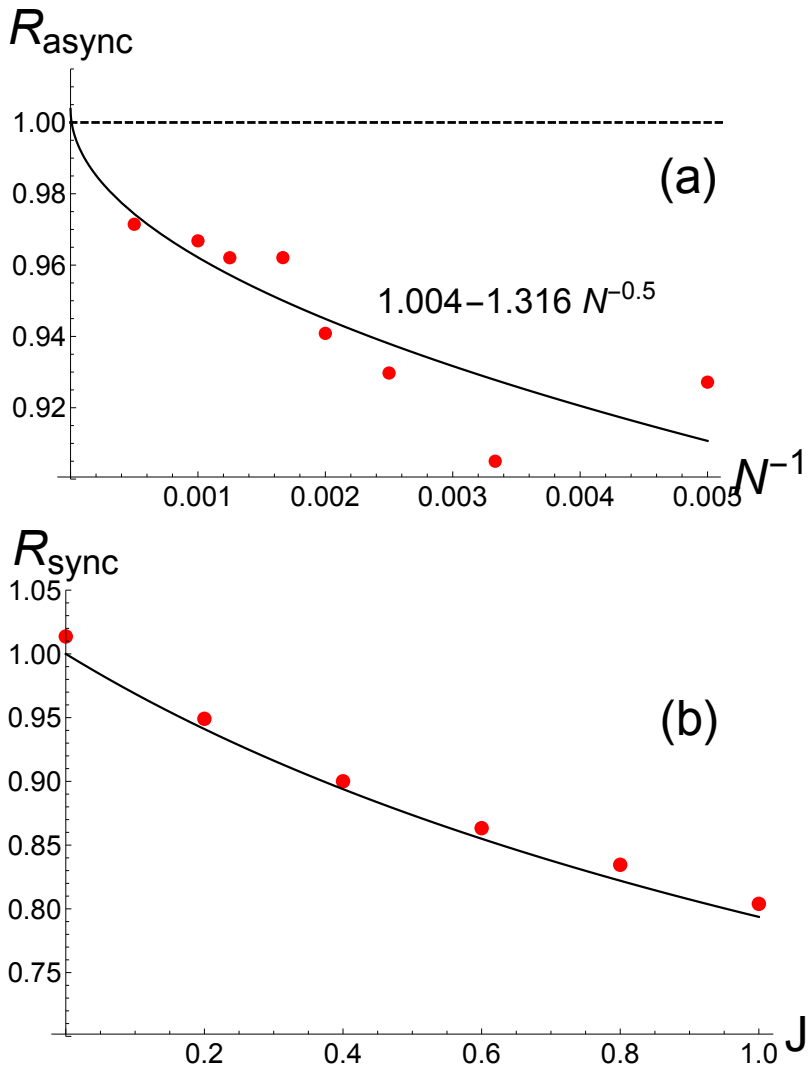

Figure 4.28: Testing predictions for the support radii. (a) Finite-size scaling of the radius of static async state for $J = K = 0$. Red dots show simulation results, the dotted black line shows theoretical prediction $R_{async} = 1$, and the solid black line shows the curve $a + bN^{-c}$ where the parameters of best fit have been found using Mathematica. As can be seen, the data approach the theoretical prediction as $N^{-1} \to 0$, as confirmed by the best fit parameter $a = 0.998$. (b) Radius of static sync state as a function of J for $K = 1$. The solid black line shows the theoretical prediction $R_{sync} = (1 + J)^{-1/3}$. For each each value of J, the radius $R_{sync}(N)$ was calculated for $N = 100, 200, 500, 1000$ swarmalators, from which the $N \to \infty$ limit was found by fitting the data to a curve $a + bN^{-0.5}$, and identifying $a = R_{sync}$. In both panels, simulations were run for $T = 500$ time units with a stepsize of 0.5 using python's solver 'odeint'.

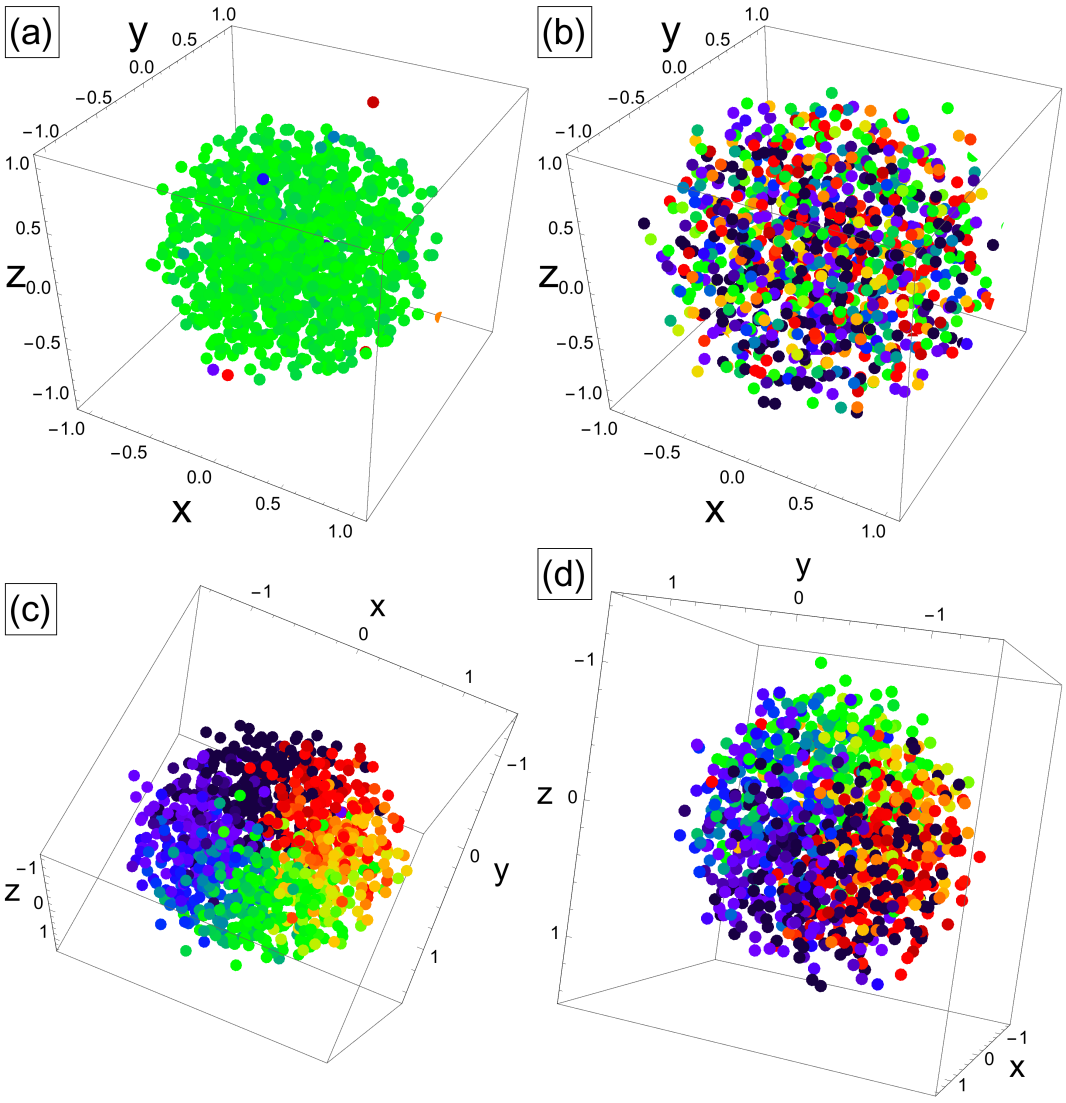

Figure 4.29: Scatter plots of four states in the (x, y, z) plane, where the swarmalators are colored according to their phase. Data were collected for $N = 1000$ swarmalators for 5×10^5 timesteps of width $dt = 0.001$ using Heun's method. Parameter values were $(J, \sigma, D_\theta, D_x, D_y, D_z) = (0.5, 0.01, 0.01, 0.01, 0.01, 0.01)$. (a) Active sync state for $K = 1$. (b) Active async state $K = -1$. (c) Active phase wave state for $K = 0$. (d) Active phase wave state for $K = -0.6$.

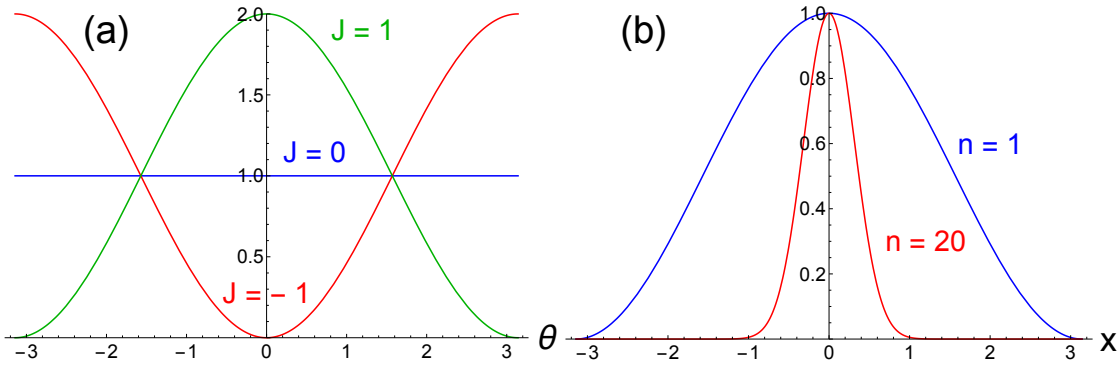

Figure 4.30: Pulse functions. (a) $F(\theta) = 1 + J\cos(\theta)$ for $J = 0, -1, 1$. (b) The pulse function $2^{-n}(1 + \cos x)^n$ for $n = 1, 20$, showing a decrease in width for increasing n.

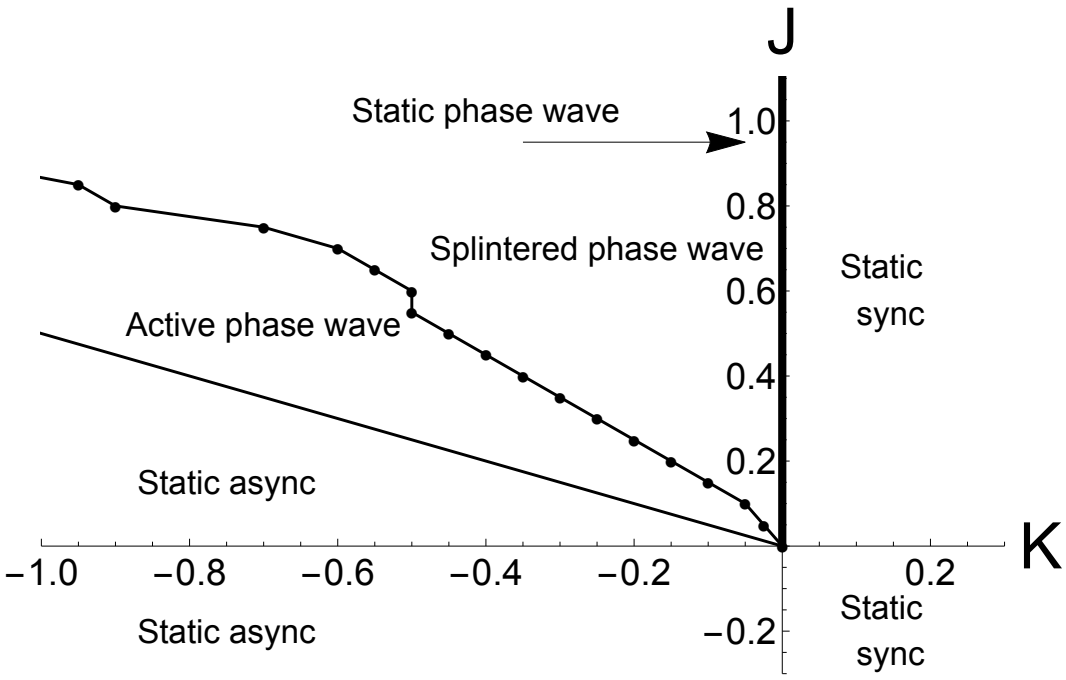

Figure 4.31: Phase diagram for the model on a one-dimensional ring. The line separating the static async and active phase wave states was calculated analytically and is given by Eq. (4.106). The line separating the active phase wave and splintered phase wave states was calculated numerically, where black dots show simulation results.

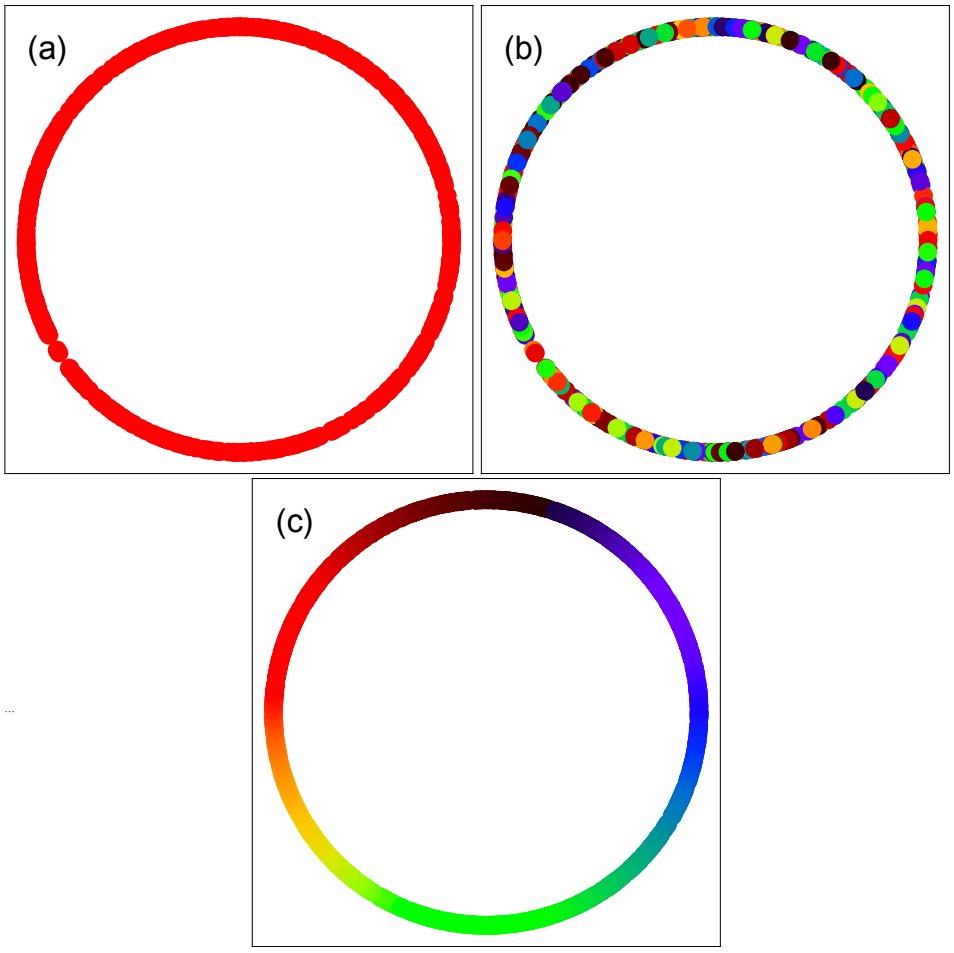

Figure 4.32: Three steady states for $N = 1000$ swarmalators with a time step of 0.1. Swarmalators' initial phases and positions were drawn uniformly at random from $[-\pi, \pi]$. Swarmalators are positioned on the unit circle and are colored according to their phase. (a) Static sync state for $(J, K) = (1, 1)$. (b) Static async state for $(J, K) = (1, -1.2)$. (c) Static phase wave state $(J, K) = (1, 0)$.

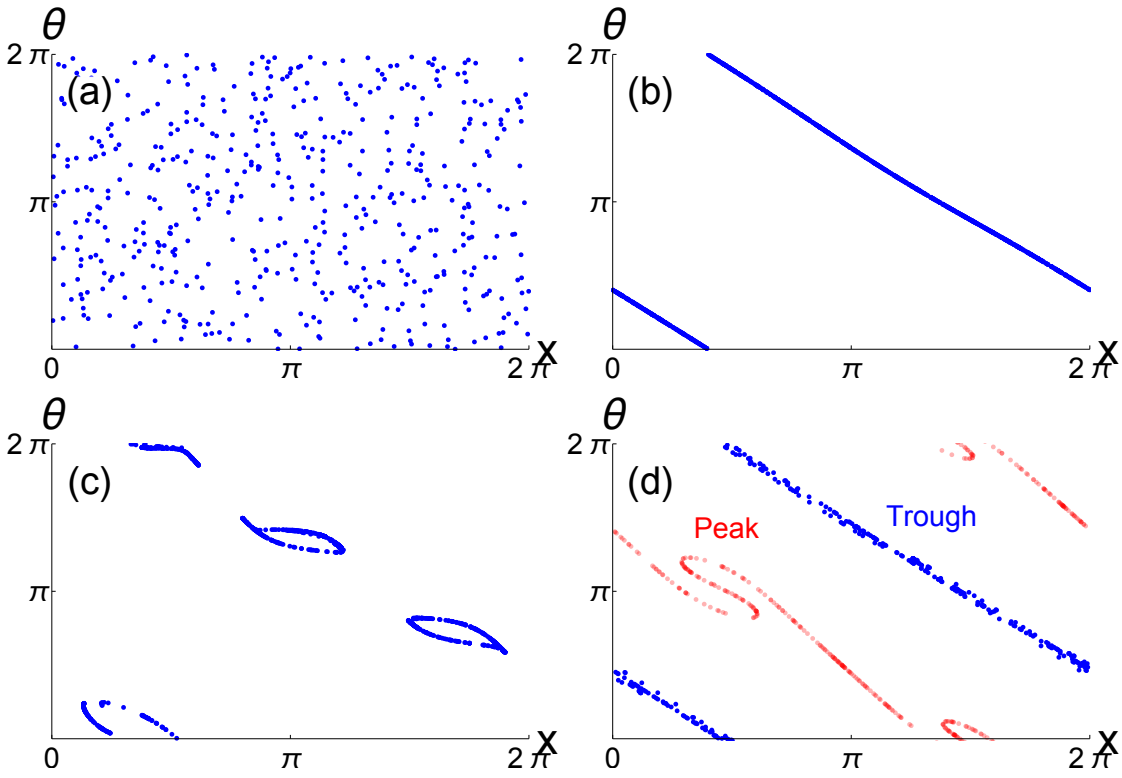

Figure 4.33: Distributions of different states in (x, θ) space. In all panels, simulations were run with $N = 1000$ swarmalators for 500 time units with a step of $dt = 0.1$. (a) Static async state for $(J, K) = (1, 1)$. (b) Static phase wave state $(J, K) = (1, 0)$. (c) Splintered phase wave state $(J, K) = (1, -0.25)$. Note the clusters gently pulsate. (d) Active phase wave state $(J, K) = (0.1, -1)$. Blue dots, distribution during a trough (when the mean population speed $\langle v \rangle$ is minimum). Red dots, distribution during a peak (when $\langle v \rangle$ is a maximum). The peak distribution has been shifted by $-\pi$ in the x direction for clarity.

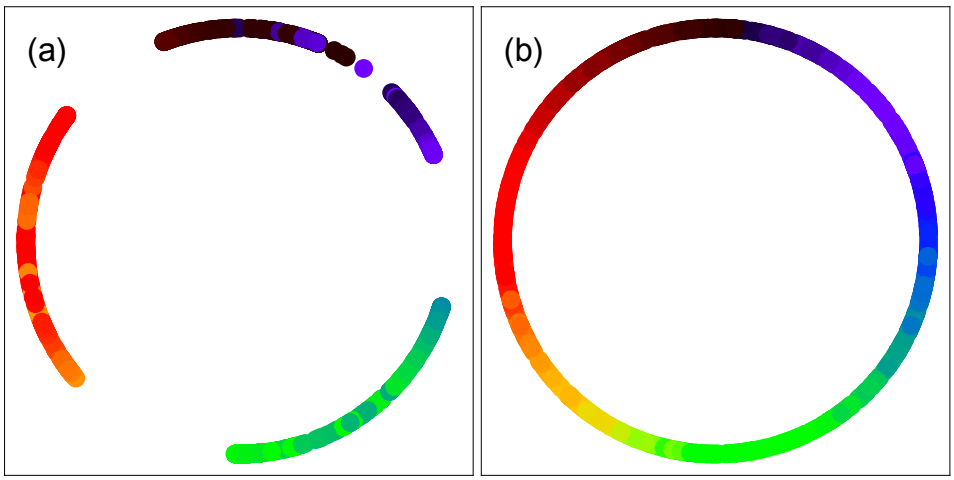

Figure 4.34: Two non-steady states for $N = 1000$ swarmalators with a time step of 0.1. In all cases, swarmalators' initial phases and positions were drawn uniformly at random from $[-\pi, \pi]$. (a) Splintered phase wave for $(J, K) = (1, -0.25)$ and $n = 3$ (as opposed to $n = 1$). (b) Active phase wave for $(J, K) = (1, -1.5)$.

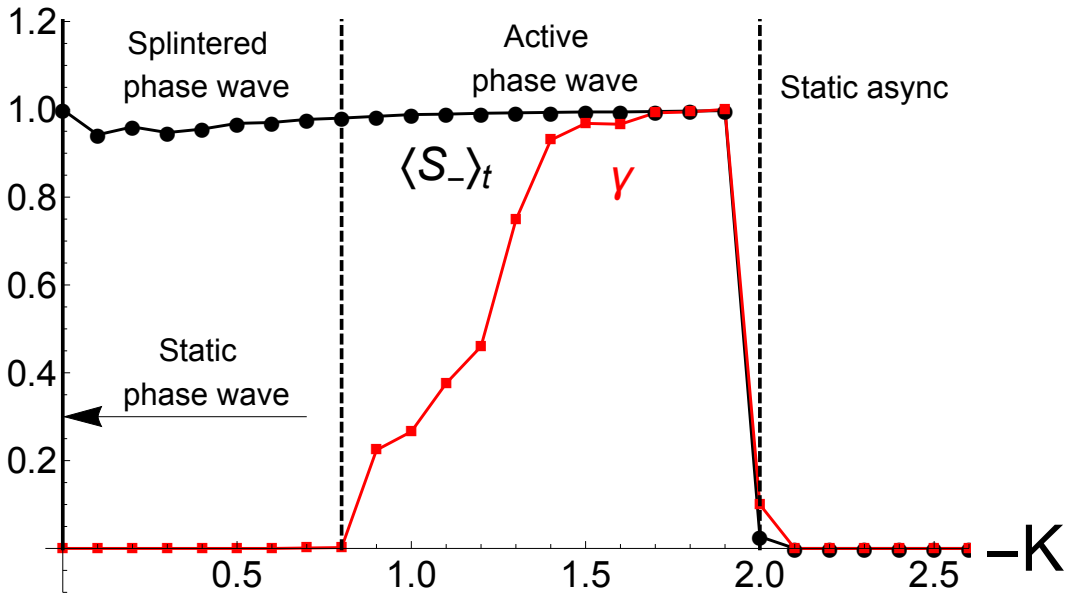

Figure 4.35: Time average of order parameter S, and γ for $J = 1.0$ and $N = 500$ swarmalators. Data was collected for $T = 3000$ time units, of which the first 50% of data were discarded as transients. As can be seen γ bifurcates from 0 at $K \approx -0.8$, signaling the transition from the splintered, to the active, phase wave states. At $K_c = -2J = -2$, S_- bifurcates to zero indicating the transition to the static async state, as predicted by equation (4.106).

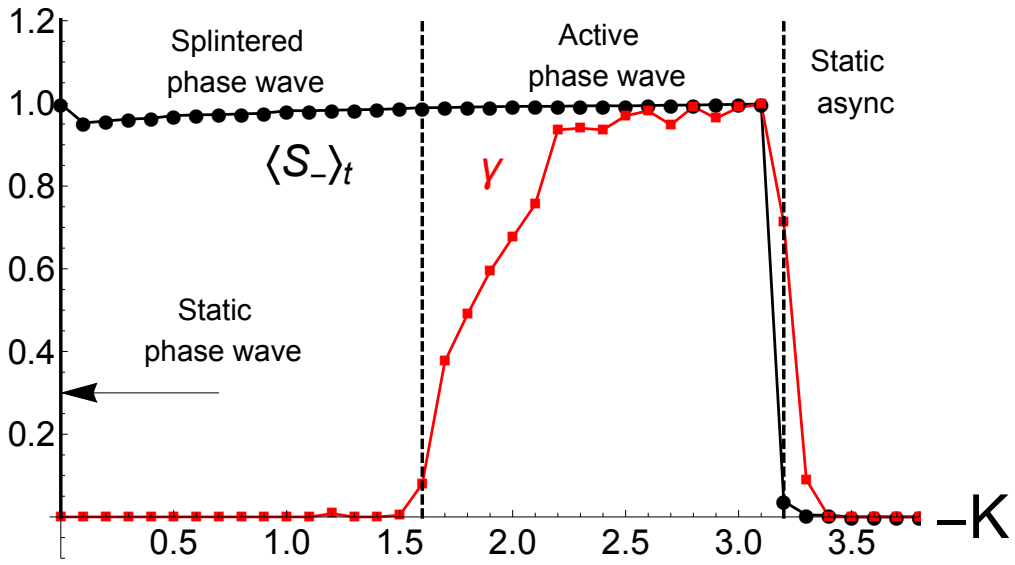

Figure 4.36: Time averaged of order parameter S, and γ for $J = 1.0$ and $N = 500$ swarmalators for $H(x)$ given by (4.109). Data was collected for $T = 3000$ time units, of which the first 50% of data were discarded as transients. Notice the order parameters for this choice of $H(x)$ have the same qualitative behavior as that for the original model (4.104), (4.105) as seen in Fig. 4.35.

BIBLIOGRAPHY

[1] *Supplemental Materials*. 2017.

[2] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.

[3] Martin Niss. History of the lenz-ising model 1920–1950: from ferromagnetic to cooperative phenomena. *Archive for history of exact sciences*, 59(3): 267–318, 2005.

[4] Martin Niss. History of the lenz–ising model 1950–1965: from irrelevance to relevance. *Archive for history of exact sciences*, 63(3):243–287, 2009.

[5] Arthur T Winfree. *The Geometry of Biological Time*. Springer, New York, 2001.

[6] Y Kuramoto. Kuramoto talks about the kuramoto model. URL `https://www.youtube.com/watch?v=lac4TxWyBOgt=21s`.

[7] Jing Yan, Moses Bloom, Sung Chul Bae, Erik Luijten, and Steve Granick. Linking synchronization to self-assembly using magnetic janus colloids. *Nature*, 491(7425):578–581, 2012.

[8] John Buck and Elisabeth Buck. Mechanism of rhythmic synchronous flashing of fireflies. *Science*, 159(3821):1319–1327, 1968.

[9] François Leyvraz. Scaling theory and exactly solved models in the kinetics of irreversible aggregation. *Physics Reports*, 383(2):95–212, 2003.

[10] Pavel L Krapivsky, Sidney Redner, and Eli Ben-Naim. *A Kinetic View of Statistical Physics*. Cambridge University Press, Cambridge, UK, 2010.

[11] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization*. Cambridge University Press, Cambridge, UK, 2003.

[12] Steven Strogatz. *Sync*. Hyperion, New York, 2003.

[13] Yao-Win Hong and Anna Scaglione. Time synchronization and reach-back communications with pulse-coupled oscillators for uwb wireless ad hoc networks. In *2003 IEEE Conference on Ultra Wideband Systems and Technologies*, pages 190–194. IEEE, 2003.

[14] Xiao Y Wang and Alyssa B Apsel. Pulse coupled oscillator synchronization for low power uwb wireless transceivers. In *50th Midwest Symposium on Circuits and Systems, 2007*, pages 1524–1527. IEEE, 2007.

[15] Geoffrey Werner-Allen, Geetika Tewari, Ankit Patel, Matt Welsh, and Radhika Nagpal. Firefly-inspired sensor network synchronicity with realistic radio effects. In *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, pages 142–153. ACM, 2005.

[16] Joel Nishimura and Eric Friedman. Robust convergence in pulse-coupled oscillators with delays. *Phys. Rev. Lett.*, 106:194101, May 2011. doi: 10.1103/PhysRevLett.106.194101.

[17] Joel Nishimura and Eric Friedman. Probabilistic convergence guarantees for type-ii pulse-coupled oscillators. *Phys. Rev. E*, 86:025201, Aug 2012. doi: 10.1103/PhysRevE.86.025201.

[18] A V M Herz and J J Hopfield. Earthquake cycles and neural reverberations: Collective oscillations in systems with pulse-coupled threshold elements. *Phys. Rev. Lett.*, 75:1222–1225, Aug 1995. doi: 10.1103/PhysRevLett.75.1222.

[19] Samuele Bottani and Bertrand Delamotte. Self-organized-criticality and synchronization in pulse coupled relaxation oscillator systems; the olami, feder and christensen and the feder and feder model. *Physica D*, 103(1): 430–441, 1997.

[20] S. Gualdi, J.-P. Bouchaud, G. Cencetti, M. Tarzia, and F. Zamponi. Endogenous crisis waves: a stochastic model with synchronized collective behavior. *ArXiv e-prints*, September 2014.

[21] Christoph Kirst, Theo Geisel, and Marc Timme. Sequential desynchronization in networks of spiking neurons with partial reset. *Phys. Rev. Lett.*, 102:068101, Feb 2009. doi: 10.1103/PhysRevLett.102.068101.

[22] J J Hopfield and A V M Herz. Rapid local synchronization of action potentials: Toward computation with coupled integrate-and-fire neurons. *Proceedings of the National Academy of Sciences*, 92(15):6655–6662, 1995.

[23] Charles S Peskin. *Mathematical Aspects of Heart Physiology*, pages 268–278. Courant Institute of Mathematical Sciences, New York, 1975.

[24] P. Bressloff, S. Coombes, and B. de Souza. Dynamics of a ring of pulse-coupled oscillators: Group-theoretic approach. *Phys. Rev. Lett.*, 79:2791–2794, Oct 1997. doi: 10.1103/PhysRevLett.79.2791.

[25] P. Bressloff and S. Coombes. Traveling waves in a chain of pulse-coupled oscillators. *Phys. Rev. Lett.*, 80:4815–4818, May 1998. doi: 10.1103/PhysRevLett.80.4815.

[26] Alvaro Corral, Conrad Pérez, Albert Diaz-Guilera, and Alex Arenas. Self-organized criticality and synchronization in a lattice model of

integrate-and-fire oscillators. *Phys. Rev. Lett.*, 74:118–121, Jan 1995. doi: 10.1103/PhysRevLett.74.118.

[27] S. Bottani. Pulse-coupled relaxation oscillators: From biological synchronization to self-organized criticality. *Phys. Rev. Lett.*, 74:4189–4192, May 1995. doi: 10.1103/PhysRevLett.74.4189.

[28] Gregory A Worrell, Stephen D Cranstoun, Javier Echauz, and Brian Litt. Evidence for self-organized criticality in human epileptic hippocampus. *NeuroReport*, 13(16):2017–2021, 2002.

[29] Renato E Mirollo and Steven H Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50 (6):1645–1662, 1990.

[30] Wulfram Gerstner and J L van Hemmen. Coherence and incoherence in a globally coupled ensemble of pulse-emitting units. *Phys. Rev. Lett.*, 71: 312–315, Jul 1993. doi: 10.1103/PhysRevLett.71.312.

[31] U. Ernst, K. Pawelzik, and T. Geisel. Synchronization induced by temporal delays in pulse-coupled oscillators. *Phys. Rev. Lett.*, 74:1570–1573, Feb 1995. doi: 10.1103/PhysRevLett.74.1570.

[32] Samuele Bottani. Synchronization of integrate and fire oscillators with global coupling. *Phys. Rev. E*, 54:2334–2350, Sep 1996. doi: 10.1103/PhysRevE.54.2334.

[33] See Supplemental Material at http://link.aps.org/supplemental/http://link.aps.org/su for results for other coupling rules.

[34] Johannes Klinglmayr, Christoph Kirst, Christian Bettstetter, and Marc

Timme. Guaranteeing global synchronization in networks with stochastic interactions. *New Journal of Physics*, 14(7):073031, 2012.

[35] Jesús Gómez-Gardenes, Yamir Moreno, and Alex Arenas. Paths to synchronization on complex networks. *Physical review letters*, 98(3):034101, 2007.

[36] Alex Arenas and Albert Diaz-Guilera. Synchronization and modularity in complex networks. *The European Physical Journal Special Topics*, 143(1): 19–25, 2007.

[37] Charles S Peskin. *Mathematical aspects of heart physiology*. Courant Institute of Mathematical Sciences, New York University, 1975.

[38] Renato E Mirollo and Steven H Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50 (6):1645–1662, 1990.

[39] Christoph Kirst, Theo Geisel, and Marc Timme. Sequential desynchronization in networks of spiking neurons with partial reset. *Physical review letters*, 102(6):068101, 2009.

[40] G Bard Ermentrout and David H Terman. *Mathematical foundations of neuroscience*, volume 35. Springer Science & Business Media, 2010.

[41] Christopher H Scholz. Large earthquake triggering, clustering, and the synchronization of faults. *Bulletin of the Seismological Society of America*, 100(3):901–909, 2010.

[42] Stanislao Gualdi, Jean-Philippe Bouchaud, Giulia Cencetti, Marco Tarzia, and Francesco Zamponi. Endogenous crisis waves: Stochastic model with

synchronized collective behavior. *Physical review letters*, 114(8):088701, 2015.

[43] P. C. Bressloff, S. Coombes, and B. de Souza. Dynamics of a ring of pulse-coupled oscillators: Group-theoretic approach. *Phys. Rev. Lett.*, 79:2791–2794, Oct 1997. doi: 10.1103/PhysRevLett.79.2791. URL http://link.aps.org/doi/10.1103/PhysRevLett.79.2791.

[44] P. C. Bressloff and S. Coombes. Traveling waves in a chain of pulse-coupled oscillators. *Phys. Rev. Lett.*, 80:4815–4818, May 1998. doi: 10.1103/PhysRevLett.80.4815. URL http://link.aps.org/doi/10.1103/PhysRevLett.80.4815.

[45] Álvaro Corral, Conrad J. Pérez, Albert Díaz-Guilera, and Alex Arenas. Self-organized criticality and synchronization in a lattice model of integrate-and-fire oscillators. *Phys. Rev. Lett.*, 74: 118–121, Jan 1995. doi: 10.1103/PhysRevLett.74.118. URL http://link.aps.org/doi/10.1103/PhysRevLett.74.118.

[46] Marc Timme, Fred Wolf, and Theo Geisel. Coexistence of regular and irregular dynamics in complex networks of pulse-coupled oscillators. *Phys. Rev. Lett.*, 89:258701, Nov 2002. doi: 10.1103/PhysRevLett.89.258701. URL http://link.aps.org/doi/10.1103/PhysRevLett.89.258701.

[47] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3): 183–208, 2000.

[48] Alex Arenas, Albert Díaz-Guilera, Jurgen Kurths, Yamir Moreno, and

Changsong Zhou. Synchronization in complex networks. *Physics Reports*, 469(3):93–153, 2008.

[49] Alexandre Mauroy and Rodolphe Sepulchre. Clustering behaviors in networks of integrate-and-fire oscillators. *Chaos: An interdisciplinary journal of nonlinear science*, 18(3):037122, 2008.

[50] U. Ernst, K. Pawelzik, and T. Geisel. Synchronization induced by temporal delays in pulse-coupled oscillators. *Phys. Rev. Lett.*, 74:1570–1573, Feb 1995. doi: 10.1103/PhysRevLett.74.1570. URL http://link.aps.org/doi/10.1103/PhysRevLett.74.1570.

[51] Wulfram Gerstner. Rapid phase locking in systems of pulse-coupled oscillators with delays. *Physical review letters*, 76(10):1755, 1996.

[52] Carl Van Vreeswijk, LF Abbott, and G Bard Ermentrout. When inhibition not excitation synchronizes neural firing. *Journal of computational neuroscience*, 1(4):313–321, 1994.

[53] Kevin P. O'Keeffe, P. L. Krapivsky, and Steven H. Strogatz. Synchronization as aggregation: Cluster kinetics of pulse-coupled oscillators. *Phys. Rev. Lett.*, 115:064101, Aug 2015. doi: 10.1103/PhysRevLett.115.064101. URL http://link.aps.org/doi/10.1103/PhysRevLett.115.064101.

[54] Jesús Gómez-Gardeñes, Yamir Moreno, and Alex Arenas. Paths to synchronization on complex networks. *Phys. Rev. Lett.*, 98: 034101, Jan 2007. doi: 10.1103/PhysRevLett.98.034101. URL http://link.aps.org/doi/10.1103/PhysRevLett.98.034101.

[55] Arthur T Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, 16(1):15–42, 1967.

[56] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In H. Araki, editor, *International Symposium on Mathematical Problems in Theoretical Physics*, pages 420–422. Springer, 1975.

[57] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, 2003.

[58] Juan A Acebrón, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137, 2005.

[59] Ikkyu Aihara, Hiroyuki Kitahata, Kenichi Yoshikawa, and Kazuyuki Aihara. Mathematical modeling of frogs calling behavior and its possible application to artificial life and robotics. *Artificial Life and Robotics*, 12(1-2): 29–32, 2008.

[60] Ernest Montbrió, Diego Pazó, and Alex Roxin. Macroscopic description for networks of spiking neurons. *Physical Review X*, 5(2):021028, 2015.

[61] Diego Pazó and Ernest Montbrió. Low-dimensional dynamics of populations of pulse-coupled oscillators. *Physical Review X*, 4(1):011009, 2014.

[62] Kevin P O'Keeffe and Steven H Strogatz. Dynamics of a population of oscillatory and excitable elements. *Physical Review E*, 93(6):062203, 2016.

[63] Carlo R Laing. Derivation of a neural field model from a network of theta neurons. *Physical Review E*, 90(1):010901, 2014.

[64] Z. Néda, E. Ravasz, T. Vicsek, Y. Brechet, and A. L. Barabási. Physics of the rhythmic applause. *Phys. Rev. E*, 61:6987–6992, Jun 2000. doi: 10.1103/PhysRevE.61.6987.

[65] Hiroaki Daido. Lower critical dimension for populations of oscillators with randomly distributed frequencies: a renormalization-group analysis. *Physical Review Letters*, 61(2):231, 1988.

[66] Per Östborn. Renormalization of oscillator lattices with disorder. *Physical Review E*, 79(5):051114, 2009.

[67] Hyunsuk Hong, Hugues Chaté, Hyunggyu Park, and Lei-Han Tang. Entrainment transition in populations of random frequency oscillators. *Physical Review Letters*, 99(18):184101, 2007.

[68] Hyunsuk Hong, Hugues Chaté, Lei-Han Tang, and Hyunggyu Park. Finite-size scaling, dynamic fluctuations, and hyperscaling relation in the kuramoto model. *Physical Review E*, 92(2):022122, 2015.

[69] J. Pantaleone. Stability of incoherence in an isotropic gas of oscillating neutrinos. *Phys. Rev. D*, 58:073002, Aug 1998. doi: 10.1103/PhysRevD.58.073002. URL `http://link.aps.org/doi/10.1103/PhysRevD.58.073002`.

[70] Kurt Wiesenfeld, Pere Colet, and Steven H Strogatz. Synchronization transitions in a disordered josephson series array. *Physical review letters*, 76(3):404, 1996.

[71] Adilson E Motter, Seth A Myers, Marian Anghel, and Takashi Nishikawa. Spontaneous synchrony in power-grid networks. *Nature Physics*, 9(3):191–197, 2013.

[72] Florian Dorfler and Francesco Bullo. Synchronization and transient stability in power networks and nonuniform kuramoto oscillators. *SIAM Journal on Control and Optimization*, 50(3):1616–1642, 2012.

[73] Steven H Strogatz, Daniel M Abrams, Allan McRobie, Bruno Eckhardt, and Edward Ott. Theoretical mechanics: Crowd synchrony on the millennium bridge. *Nature*, 438(7064):43–44, 2005.

[74] Iain Couzin. Collective minds. *Nature*, 445(7129):715–715, 2007.

[75] Jerome Buhl, David JT Sumpter, Iain D Couzin, Joe J Hale, Emma Despland, ER Miller, and Steve J Simpson. From disorder to order in marching locusts. *Science*, 312(5778):1402–1406, 2006.

[76] David JT Sumpter. *Collective Animal Behavior*. Princeton University Press, 2010.

[77] James E Herbert-Read. Understanding how animal groups achieve coordinated movement. *Journal of Experimental Biology*, 219(19):2971–2983, 2016.

[78] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008. doi: 10.1073/pnas.0711437105.

[79] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale, and Aleksandra M. Walczak. Statistical mechanics for natural flocks of birds. *Proceed-*

ings of the National Academy of Sciences, 109(13):4786–4791, 2012. doi: 10.1073/pnas.1118633109.

[80] Craig W Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH computer graphics*, 21(4):25–34, 1987.

[81] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226, 1995.

[82] Andrew J Bernoff and Chad M Topaz. Nonlocal aggregation models: A primer of swarm equilibria. *SIAM Review*, 55(4):709–747, 2013.

[83] Chad M Topaz and Andrea L Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM Journal on Applied Mathematics*, 65(1):152–174, 2004.

[84] Chad M Topaz, Andrea L Bertozzi, and Mark A Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of mathematical biology*, 68(7):1601–1623, 2006.

[85] Theodore Kolokolnikov, Hui Sun, David Uminsky, and Andrea L Bertozzi. Stability of ring patterns arising from two-dimensional particle interactions. *Physical Review E*, 84(1):015203, 2011.

[86] Koichiro Uriu, Saúl Ares, Andrew C. Oates, and Luis G. Morelli. Dynamics of mobile coupled phase oscillators. *Phys. Rev. E*, 87:032911, Mar 2013.

[87] Igor V Belykh, Vladimir N Belykh, and Martin Hasler. Blinking model and synchronization in small-world networks with a time-varying coupling. *Physica D: Nonlinear Phenomena*, 195(1):188–206, 2004.

[88] Daniel J Stilwell, Erik M Bollt, and D Gray Roberson. Sufficient conditions for fast switching synchronization in time-varying network topologies. *SIAM Journal on Applied Dynamical Systems*, 5(1):140–156, 2006.

[89] Mattia Frasca, Arturo Buscarino, Alessandro Rizzo, Luigi Fortuna, and Stefano Boccaletti. Synchronization of moving chaotic agents. *Physical Review Letters*, 100(4):044102, 2008.

[90] Naoya Fujiwara, Jürgen Kurths, and Albert Díaz-Guilera. Synchronization in networks of mobile oscillators. *Physical Review E*, 83(2):025101, 2011.

[91] Thomas J Walker. Acoustic synchrony: two mechanisms in the snowy tree cricket. *Science*, 166(3907):891–894, 1969.

[92] Michael D Greenfield. Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions. *American Zoologist*, 34(6):605–615, 1994.

[93] Ikkyu Aihara, Takeshi Mizumoto, Takuma Otsuka, Hiromitsu Awano, Kohei Nagira, Hiroshi G Okuno, and Kazuyuki Aihara. Spatio-temporal dynamics in collective frog choruses examined by mathematical modeling and field observations. *Scientific Reports*, 4:3891, 2014.

[94] Alexey Snezhko and Igor S Aranson. Magnetic manipulation of self-assembled colloidal asters. *Nature materials*, 10(9):698–703, 2011.

[95] James E Martin and Alexey Snezhko. Driving self-assembly and emergent dynamics in colloidal suspensions by time-dependent magnetic fields. *Reports on Progress in Physics*, 76(12):126601, 2013.

[96] Nguyen HP Nguyen, Daphne Klotsa, Michael Engel, and Sharon C Glotzer. Emergent collective phenomena in a mixture of hard shapes through active rotation. *Physical Review Letters*, 112(7):075701, 2014.

[97] Benjamin C van Zuiden, Jayson Paulose, William TM Irvine, Denis Bartolo, and Vincenzo Vitelli. Spatiotemporal order and emergent edge currents in active spinner materials. *Proceedings of the National Academy of Sciences*, page 201609572, 2016.

[98] Oleg A Igoshin, Alex Mogilner, Roy D Welch, Dale Kaiser, and George Oster. Pattern formation and traveling waves in myxobacteria: theory and modeling. *Proceedings of the National Academy of Sciences*, 98(26):14913–14918, 2001.

[99] Dan Tanaka. General chemotactic model of oscillators. *Physical Review Letters*, 99(13):134103, 2007.

[100] Masatomo Iwasa, Kazuki Iida, and Dan Tanaka. Hierarchical cluster structures in a one-dimensional swarm oscillator model. *Physical Review E*, 81(4):046220, 2010.

[101] Razvan C Fetecau, Yanghong Huang, and Theodore Kolokolnikov. Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity*, 24(10):2681, 2011.

[102] JA Carrillo, MR Dorsogna, and V Panferov. Double milling in self-propelled swarms from kinetic theory. *Kinetic and Related Models*, 2(2): 363–378, 2009.

[103] Ingmar H Riedel, Karsten Kruse, and Jonathon Howard. A self-organized

vortex array of hydrodynamically entrained sperm cells. *Science*, 309 (5732):300–303, 2005.

[104] Hyunsuk Hong and Steven H Strogatz. Kuramoto model of coupled oscillators with positive and negative coupling parameters: An example of conformist and contrarian oscillators. *Physical Review Letters*, 106(5): 054102, 2011.

[105] Hyunsuk Hong, Kevin P O'Keeffe, and Steven H Strogatz. Phase coherence induced by correlated disorder. *Physical Review E*, 93(2):022219, 2016.

[106] Oliver Pohl and Holger Stark. Dynamic clustering and chemotactic collapse of self-phoretic active particles. *Physical Review Letters*, 112(23): 238303, 2014.

[107] AD Maude. Non-random distribution of bull spermatozoa in a drop of sperm suspension. *Nature*, 200:381, 1963.

[108] Jens Elgeti and Gerhard Gompper. Emergence of metachronal waves in cilia arrays. *Proceedings of the National Academy of Sciences*, 110(12):4470–4475, 2013.

[109] Yasushi Okada, Shigenori Nonaka, Yosuke Tanaka, Yukio Saijoh, Hiroshi Hamada, and Nobutaka Hirokawa. Abnormal nodal flow precedes situs inversus in iv and inv mice. *Molecular cell*, 4(4):459–468, 1999.

[110] LB Wong, IRVING F Miller, and DONOVAN B Yeates. Nature of the mammalian ciliary metachronal wave. *Journal of Applied Physiology*, 75(1):458–467, 1993.

[111] Eric Lauga and Thomas R Powers. The hydrodynamics of swimming microorganisms. *Reports on Progress in Physics*, 72(9):096601, 2009.

[112] Simone Immler, Harry DM Moore, William G Breed, and Tim R Birkhead. By hook or by crook? morphometry, competition and cooperation in rodent sperm. *PloS one*, 2(1):e170, 2007.

[113] Harry Moore, Katerina Dvoráková, Nicholas Jenkins, and William Breed. Exceptional sperm cooperation in the wood mouse. *Nature*, 418(6894): 174–177, 2002.

[114] Geoffrey Taylor. Analysis of the swimming of microscopic organisms. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 209, pages 447–461. The Royal Society, 1951.

[115] Lisa J Fauci and Amy McDonald. Sperm motility in the presence of boundaries. *Bulletin of mathematical biology*, 57(5):679–699, 1995.

[116] Yingzi Yang, Jens Elgeti, and Gerhard Gompper. Cooperation of sperm in two dimensions: synchronization, attraction, and aggregation through hydrodynamic interactions. *Physical Review E*, 78(6):061903, 2008.

[117] Fumio Hayashi. Insemination through an externally attached spermatophore: bundled sperm and post-copulatory mate guarding by male fishflies (megaloptera: Corydalidae). *Journal of insect physiology*, 42(9):859–866, 1996.

[118] M Leoni and TB Liverpool. Synchronization and liquid crystalline order in soft active fluids. *Physical Review Letters*, 112(14):148104, 2014.

[119] Michele Starnini, Mattia Frasca, and Andrea Baronchelli. Emergence of metapopulations and echo chambers in mobile agents. *Scientific Reports*, 6, 2016.

[120] Nikita Kruk, Yuri Maistrenko, Nicolas Wenzel, and Heinz Koeppl. Self-propelled chimeras. *arXiv preprint arXiv:1511.04738*, 2015.

[121] Benno Liebchen and Demian Levis. The rotating viczek model: Pattern formation and enhanced flocking in chiral active matter. *arXiv preprint arXiv:1701.00091*, 2016.