EURANDOM PREPRINT SERIES

2017-003

March 4, 2017

Synchronization of phase oscillators on the hierarchical lattice

D. Garlaschelli, F. den Hollander, J. Meylahn, B. Zeegers
ISSN 1389-2355

Synchronization of phase oscillators on the hierarchical lattice

D. Garlaschelli [1]
F. den Hollander [2]
J. Meylahn [2]
B. Zeegers [2]

March 4, 2017

Abstract

Synchronization of neurons forming a network with a hierarchical structure is essential for the brain to be able to function optimally. In this paper we study synchronization of phase oscillators on the most basic example of such a network, namely, the hierarchical lattice. Each oscillator has a natural frequency, drawn independently from a common probability distribution. In addition, pairs of oscillators interact with each other at a strength that depends on their hierarchical distance, modulated by a sequence of interaction parameters. We look at block averages of the oscillators on successive hierarchical scales, which we think of as block communities. Also these block communities are given a natural frequency, drawn independently from a common probability distribution that depends on their hierarchical scale. In the limit as the number of oscillators per community tends to infinity, referred to as the hierarchical mean-field limit, we find a separation of time scales, i.e., each block community behaves like a single oscillator evolving on its own time scale. We show that the evolution of the block communities is given by a renormalized mean-field noisy Kuramoto equation, with a synchronization level that depends on the hierarchical scale of the block community. We identify three universality classes for the synchronization levels on successive hierarchical scales, with explicit characterizations in terms of the sequence of interaction parameters and the sequence of natural frequency probability distributions. We show that disorder reduces synchronization when the natural frequency probability distributions are symmetric and unimodal, with the reduction gradually vanishing as the hierarchical scale goes up.

Mathematics Subject Classification 2010. 60K35, 60K37, 82B20, 82C27, 82C28.

Key words and phrases. Hierarchical lattice, phase oscillators, noisy Kuramoto model, block communities, renormalization, universality classes.

Acknowledgment. DG is supported by EU-project 317532-MULTIPLEX. FdH and JM are supported by NWO Gravitation Grant 024.002.003–NETWORKS.

[1] Lorentz Institute for Theoretical Physics, Leiden University, P.O. Box 9504, 2300 RA Leiden, The Netherlands.
[2] Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands.

1 Introduction

The concept of *spontaneous synchronization* is ubiquitous in nature. Single oscillators in a population (like flashing fireflies, chirping crickets or spiking brain cells) may rotate incoherently, at their own natural frequency, when they are isolated from the population, but within the population they adapt their rhythm to that of the other oscillators, acting as a system of coupled oscillators. There is no central driving mechanism, yet the population reaches a globally synchronized state via mutual local interactions.

The omnipresence of spontaneous synchronization triggered scientists to search for a mathematical approach in order to understand the underlying principles. The first steps were taken by Winfree [16], [17], who recognized that spontaneous synchronization should be understood as a threshold phenomenon: if the coupling between the oscillators is sufficiently strong, then a macroscopic part of the population freezes into synchrony. Although the model that Winfree proposed was too difficult to solve analytically, it inspired Kuramoto [6], [7] to suggest a more mathematically tractable model that captures the same phenomenon. The Kuramoto model has since been used successfully to study synchronization in a variety of different contexts. By now there is an extended literature, covering aspects like phase transition, stability, and effect of disorder (for a review, see Acébron *et al.* [1]).

Mathematically, the Kuramoto model still poses many challenges. As long as the interaction is *mean-field* (meaning that every oscillator interacts equally strongly with every other oscillator), a fairly complete theory has been developed. However, as soon as the interaction has a *non-trivial geometry*, computations become cumbersome. There is a large literature for the Kuramoto model on complex networks, where the population is viewed as a random graph whose vertices carry the oscillators and whose edges represent the interaction. Numerical and heuristic results have been obtained for networks with a small-world, scale-free and/or community structure, showing a range of interesting phenomena (for a review, see Arenas *et al.* [2]). In the present paper we focus on one particular network with a community structure, namely, the *hierarchical lattice*.

The remainder of this paper is organised as follows. Sections 1.1–1.3 are devoted to the mean-field noisy Kuramoto model. In Section 1.1 we recall definitions and basic properties. In Section 1.2 we recall the McKean-Vlasov equation, which describes the evolution of the probability density for the phase oscillators and their natural frequencies in the *mean-field limit*. In Section 1.3 we take a closer look at the scaling properties of the order parameters towards the mean-field limit. In Section 1.4 we define the hierarchical lattice and in Section 1.5 introduce the noisy Kuramoto model on the hierarchical lattice, which involves a sequence of interaction strengths $(K_k)_{k\in\mathbb{N}}$ and a sequence of disorder distributions $(\mu_k)_{k\in\mathbb{N}}$ acting on successive hierarchical levels. Section 2 contains our main results, presented in the form of three theorems for the non-disordered system and their analogues for the disordered system. These theorems are valid in the *hierarchical mean-field limit* and show that, for each $k \in \mathbb{N}$, the block communities at hierarchical level k behave like the mean-field noisy Kuramoto model, with an interaction strength and a noise that depend on k and are obtained via a *renormalization transformation* connecting successive hierarchical levels. There are three universality classes for $(K_k)_{k\in\mathbb{N}}$ and $(\mu_k)_{k\in\mathbb{N}}$, corresponding to loss of synchronization at a finite hierarchical level, loss of synchronization as the hierarchical level tends to infinity, and no loss of synchronization. The renormalization transformation allows us to describe these classes in some detail. In Section 3 we prove the renormalization scheme, in Section 4 we prove the criteria for the universality classes *subject to a technical inequality that is verified numerically for three different types of disorder.* Appendix A contains the proof of certain inequalities linking the system with disorder to the system without disorder. Appendix B provides numerical examples and computations.

1.1 Mean-field Kuramoto model

We begin by reviewing the mean-field Kuramoto model. Consider a population of $N \in \mathbb{N}$ oscillators, and suppose that the i^{th} oscillator has a natural frequency ω_i, such that

$$\blacktriangleright \quad \begin{array}{l} \omega_i,\ i = 1, \ldots, N, \text{ are i.i.d. and are drawn from} \\ \text{a common probability distribution } \mu \text{ on } \mathbb{R}. \end{array} \tag{1.1}$$

Let the phase of the i^{th} oscillator at time t be $\theta_i(t) \in \mathbb{R}$. If the oscillators were not interacting, then we would have the system of uncoupled differential equations

$$\frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t} = \omega_i, \quad i = 1, \ldots, N. \tag{1.2}$$

Kuramoto [6], [7] realized that the easiest way to allow for synchronization was to let every oscillator interact with every other oscillator according to the sine of their phase difference, i.e., to replace (1.2) by:

$$\frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t} = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin\left[\theta_j(t) - \theta_i(t)\right], \quad i = 1, \ldots, N. \tag{1.3}$$

Here, $K \in (0, \infty)$ is the *interaction strength*, and the factor $\frac{1}{N}$ is included to make sure that the total interaction per oscillator stays finite in the thermodynamic limit $N \to \infty$. The coupled evolution equations in (1.3) are referred to as the *mean-field Kuramoto model.* An illustration of the interaction in this model is given in Fig. 1.

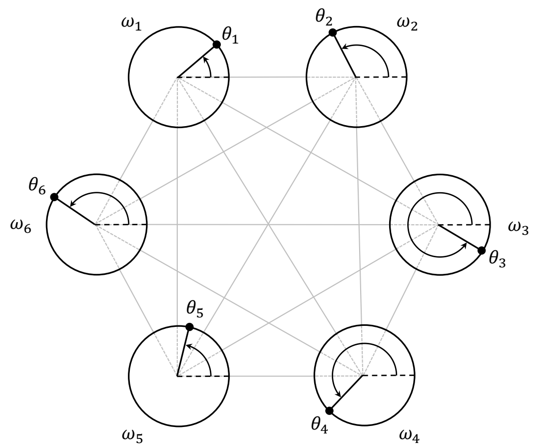


Figure 1: Mean-field interaction of $N = 6$ oscillators with natural frequencies ω_i and phases $\theta_i = \theta_i(t)$, evolving according to (1.3).

If noise is added, then (1.3) turns into the *mean-field noisy Kuramoto model*, given by

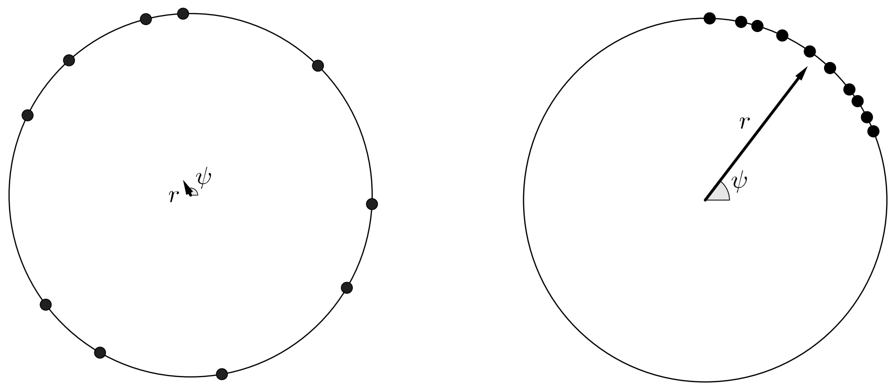
$$\mathrm{d}\theta_i(t) = \omega_i \,\mathrm{d}t + \frac{K}{N}\sum_{j=1}^{N} \sin\left[\theta_j(t) - \theta_i(t)\right] \mathrm{d}t + D\,\mathrm{d}W_i(t), \quad i = 1, \dots, N. \tag{1.4}$$

Here, $D \in (0, \infty)$ is the *noise strength*, and $(W_i(t))_{t\geq 0}$, $i = 1, \dots, N$, are independent standard Brownian motions on $\mathbb{R}$. The coupled evolution equations in (1.4) are stochastic differential equations in the sense of Itô (see e.g. Karatzas and Shreve [5]).

In order to exploit the mean-field nature of (1.4), the complex-valued order parameter

$$r_N(t)\,\mathrm{e}^{i\psi_N(t)} = \frac{1}{N}\sum_{j=1}^{N} \mathrm{e}^{i\theta_j(t)} \tag{1.5}$$

is introduced. In (1.5), $r_N(t)$ is the *synchronization level* at time t and takes values in $[0,1]$, while $\psi_N(t)$ is the *average phase* at time t and takes values in $[0, 2\pi]$. (Note that $\psi_N(t)$ is properly defined only when $r_N(t) > 0$.) The order parameter (r, ψ) is illustrated in Fig. 2 ($r = 0$ corresponds to the oscillators being completely unsynchronized, $r = 1$ to the oscillators being completely synchronized).



(a) $r = 0.095$.

(b) $r = 0.929$.

Figure 2: Phase distribution of oscillators for two different values of r. The arrow represents the complex number $re^{i\psi}$.

By rewriting (1.4) in terms of (1.5) as

$$\mathrm{d}\theta_i(t) = \omega_i \,\mathrm{d}t + K r_N(t) \sin\left[\psi_N(t) - \theta_i(t)\right] \mathrm{d}t + D\,\mathrm{d}W_i(t), \quad i = 1, \dots, N, \tag{1.6}$$

we see that *the oscillators are coupled via the order parameter*, i.e., the phases θ_i are pulled towards ψ_N with a strength proportional to r_N.

In the *mean-field limit* $N \to \infty$, the system in (1.6) exhibits what is called "propagation of chaos", i.e., the evolution of single oscillators becomes *autonomous*. Indeed, suppose that

- ► $\theta_i(0)$, $i = 1, \dots, N$, are i.i.d. and are drawn from a common probability distribution ρ on $[0, 2\pi]$. (1.7)

The order parameter associated with ρ is the pair (R, Φ) defined by

$$R\,\mathrm{e}^{i\Phi} = \int_0^{2\pi} \rho(\mathrm{d}\theta)\,\mathrm{e}^{i\theta}. \tag{1.8}$$

Suppose that $R > 0$, so that Φ is properly defined. Suppose further that

- ▶ the disorder distribution μ is symmetric. (1.9)

Then, as we will see in Section 1.2–1.3 and 3.1, the limit as $N \to \infty$ of the evolution of a single oscillator, say θ_1, is given by

$$\mathrm{d}\theta_1(t) = \omega_1\,\mathrm{d}t + Kr(t)\sin\left[\Phi - \theta_1(t)\right]\mathrm{d}t + D\,\mathrm{d}W_1(t), \tag{1.10}$$

where $(W_1(t))_{t\geq 0}$ is a standard Brownian motion, and $r(t)$ is driven by a *deterministic relaxation equation* such that

$$r(0) = R, \qquad \lim_{t\to\infty} r(t) = r \text{ for some } r \in [0,1). \tag{1.11}$$

The parameter $r = r(\mu, D, K)$ will be identified in (1.21) below (and the convergence holds at least when R is close to r; see Remark 1.1 below). The evolution in (1.10) is *not closed* because of the presence of $r(t)$, but after a *transient period* it converges to the *closed* evolution equation

$$\mathrm{d}\theta_1(t) = \omega_1\,\mathrm{d}t + Kr\sin\left[\Phi - \theta_1(t)\right]\mathrm{d}t + D\,\mathrm{d}W_1(t). \tag{1.12}$$

Without loss of generality, we may *calibrate* $\Phi = 0$ by rotating the circle $[0, 2\pi)$ over $-\Phi$. After that the parameters R, Φ associated the initial distribution ρ are gone, and only r remains as the relevant parameter. It is known (see e.g. (1.22) below) that there exists a critical threshold $K_c = K(\mu, D) \in (0, \infty)$ separating two regimes:

(I) For $K \in (0, K_c)$ there is relaxation to an *unsynchronized state* ($r = 0$).

(II) For $K \in (K_c, \infty)$ there is relaxation to a *partially synchronized state* ($r \in (0,1)$).

See Strogatz [13] and Luçon [9] for overviews.

1.2 McKean-Vlasov equation

For the system in (1.4), Sakaguchi [10] showed that in the limit as $N \to \infty$, the probability density for the phase oscillators and their natural frequencies (with respect to $\lambda \times \mu$, with λ the Lebesgue measure on $[0, 2\pi]$ and μ the disorder measure on $\mathbb{R}$) evolves according to the *McKean-Vlasov equation*

$$\frac{\partial}{\partial t}\,p(t;\theta,\omega) = -\frac{\partial}{\partial\theta}\left[p(t;\theta,\omega)\left\{\omega + Kr(t)\sin\left[\psi(t) - \theta\right]\right\}\right] + \frac{D}{2}\frac{\partial^2}{\partial\theta^2}\,p(t;\theta,\omega), \tag{1.13}$$

where

$$r(t)\,\mathrm{e}^{i\psi(t)} = \int_{\mathbb{R}} \mu(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta\,\mathrm{e}^{i\theta}\,p(t;\theta,\omega), \tag{1.14}$$

is the continuous counterpart of (1.5). If ρ has a density, say $\rho(\theta)$, then $p(0;\theta,\omega) = \rho(\theta)$ for all $\omega \in \mathbb{R}$.

By (1.9), we can again *calibrate* the average phase to be zero, i.e., $\psi(t) = \psi(0) = \Phi = 0$, $t \geq 0$, in which case the stationary solutions of (1.13) satisfy

$$0 = -\frac{\partial}{\partial\theta}\left[p(\theta,\omega)\,(\omega - Kr\sin\theta)\right] + \frac{D}{2}\frac{\partial^2}{\partial\theta^2}\,p(\theta,\omega) \tag{1.15}$$

and therefore are of the form

$$p_\lambda(\theta,\omega) = \frac{A_\lambda(\theta,\omega)}{\int_0^{2\pi} \mathrm{d}\phi\, A_\lambda(\phi,\omega)}, \qquad \lambda = 2Kr/D, \tag{1.16}$$

with

$$\begin{aligned} A_\lambda(\theta,\omega) &= B_\lambda(\theta,\omega)\left(\mathrm{e}^{4\pi\omega}\int_0^{2\pi}\frac{\mathrm{d}\phi}{B_\lambda(\phi,\omega)} + (1-\mathrm{e}^{4\pi\omega})\int_0^{\theta}\frac{\mathrm{d}\phi}{B_\lambda(\phi,\omega)}\right), \\ B_\lambda(\theta,\omega) &= \mathrm{e}^{\lambda\cos\theta + 2\theta\omega}. \end{aligned} \tag{1.17}$$

After rewriting

$$A_\lambda(\theta,\omega) = B_\lambda(\theta,\omega)\left(\int_{\theta-2\pi}^{0}\frac{\mathrm{d}\phi}{B_\lambda(-\phi,-\omega)} + \int_0^{\theta}\frac{\mathrm{d}\phi}{B_\lambda(\phi,\omega)}\right) \tag{1.18}$$

and noting that $B_\lambda(\phi,\omega) = B_\lambda(-\phi,-\omega)$, we easily see that

$$p_\lambda(\theta,\omega) = p_\lambda(-\theta,-\omega), \tag{1.19}$$

a property we will need later. In particular, in view of (1.9), we have

$$\int_{\mathbb{R}} \mu(\mathrm{d}\omega)\int_0^{2\pi}\mathrm{d}\theta\, p_\lambda(\theta,\omega)\,\sin\theta = 0. \tag{1.20}$$

Since $\psi(t) = \psi(0) = \Phi = 0$, we see from (1.14) that $p_\lambda(\theta,\omega)$ in (1.16) is a solution if and only if r satisfies

$$\int_{\mathbb{R}} \mu(\mathrm{d}\omega)\int_0^{2\pi}\mathrm{d}\theta\, p_\lambda(\theta,\omega)\,\cos\theta = r, \qquad \lambda = 2Kr/D. \tag{1.21}$$

This gives us a *self-consistency relation* for r, which can in principle be solved (and possibly has more than one solution). The equation in (1.21) always has a solution with $r = 0$: the unsynchronized state (I) corresponding to $p_0(\theta,\omega) = \frac{1}{2\pi}$ for all θ,ω. A (not necessarily unique) solution with $r \in (0,1)$ exists when the coupling strength K exceeds a critical threshold $K_c = K_c(\mu, D)$. When this occurs, we say that the oscillators are in a *partially synchronized state* (II). As K increases also r increases. Moreover, $r \uparrow 1$ as $K \to \infty$ and we say that the oscillators convergence to a *fully synchronized state*. When K crosses K_c, the system exhibits a second-order phase transition.

For the case where the frequency distribution μ is *symmetric and unimodal*, an explicit expression is known for K_c:

$$\frac{1}{K_c} = \int_{\mathbb{R}} \mu(\mathrm{d}\omega)\,\frac{D}{D^2 + 4\omega^2}. \tag{1.22}$$

Thus, when the spread of μ is large compared to K, the oscillators are not able to synchronize and they rotate near their own frequencies. As K increases, this remains the case until K reaches K_c. After that a small fraction of synchronized oscillators starts to emerge, which becomes of macroscopic size when K moves beyond K_c. For μ symmetric and unimodal it is *conjectured* that for $K > K_c$ there is a *unique* synchronized solution $p_\lambda(\cdot,\cdot)$ with $r \in (0,1)$ solving (1.21) (Luçon [9, Conjecture 3.12]). This conjecture has been proved when μ is narrow, i.e., the disorder is small (Luçon [9, Proposition 3.13 3.12]).

Remark 1.1. Stability of stationary solutions has been studied by Strogatz and Morillo [14], Strogatz, Morillo and Matthews [15], Luçon [9, Section 3.4]. For symmetric unimodal disorder, the unsynchronized state is linearly stable for $K < K_c$ and linearly unstable for $K > K_c$, while the synchronized state for $K > K_c$ is linearly stable at least for small disorder. Not much is known about stability for general disorder. □

There is no closed form expression for K_c beyond symmetric unimodal disorder, except for special cases, e.g. symmetric binary disorder. We refer to Luçon [9] for an overview.

1.3 Scaling towards mean-field

In Section 3.1 below we will prove the following key property.

Theorem 1.2. *Assume that $r > 0$. Then*

$$\begin{aligned} &\lim_{N\to\infty} \psi_N(Nt) = \psi_*(t), && \mathrm{d}\psi_*(t) = D_*\,\mathrm{d}W_*(t), && \psi_*(0) = \Phi, \\ &\lim_{N\to\infty} r_N(t) = r(t), && \lim_{t\to\infty} r(t) = r, && r(0) = R, \end{aligned} \tag{1.23}$$

where $(W_*(t))_{t\geq 0}$ is a standard Brownian motion and $D_* \in (0,\infty)$ is a diffusion constant. □

Thus, for large N and in a partially synchronized state, the average phase evolves on time scale Nt. In particular, $\lim_{N\to\infty} \psi_N(t) = \lim_{N\to\infty} \psi_N(0) = \Phi = 0$ for fixed $t \geq 0$, i.e., the average phase does not evolve on time scale t. A similar result was proved by Ha and Slemrod [4] for the Kuramoto model with disorder and without noise, while an approximate solution was obtained by Sonnenschein and Schimansky-Geier [12] for the Kuramoto model without disorder and with noise.

Theorems 2.2 and 2.6 below will expand on the above observations. In the next section we take a closer look at the hierarchical mean-field limit.

1.4 Hierarchical lattice

The hierarchical lattice of order N consist of countable many vertices that form communities of sizes N, N^2, etc. For example, the hierarchical lattice of order $N = 3$ consists of vertices that are grouped into 1-block communities of 3 vertices, which in turn are grouped into 2-block communities of 9 vertices, and so on. Each vertex is assigned a label that defines its location at each block level (see Fig. 3).

Formally, the hierarchical group Ω_N of order $N \in \mathbb{N}\backslash\{1\}$ is the set

$$\Omega_N = \left\{ \eta = (\eta^\ell)_{\ell\in\mathbb{N}_0} \in \{0,1,\dots,N-1\}^{\mathbb{N}_0} \colon \sum_{\ell\in\mathbb{N}_0} \eta^\ell < \infty \right\} \tag{1.24}$$

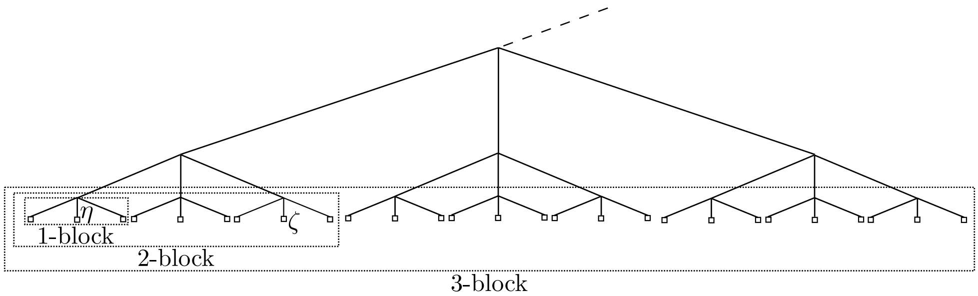


Figure 3: The hierarchical lattice of order $N = 3$. The vertices live at the lowest level. The tree visualizes their distance: the distance between two vertices η, ζ is the height of their lowest common branching point in the tree: $d(\eta, \zeta) = 2$ in the picture.

with addition modulo N, i.e., $(\eta + \zeta)^\ell = \eta^\ell + \zeta^\ell \,(\mathrm{mod}\, N)$, $\ell \in \mathbb{N}_0$. The distance on Ω_N is defined as

$$d\colon\ \Omega_N \times \Omega_N \to \mathbb{N}_0, \quad (\eta, \zeta) \mapsto \min\big\{k \in \mathbb{N}_0\colon\ \eta^\ell = \zeta^\ell \ \forall\, \ell \geq k\big\}, \tag{1.25}$$

i.e., the distance between two vertices is the smallest index from which onwards the sequences of hierarchical labels of the two vertices agree. This distance is ultrametric:

$$d(\eta, \zeta) \leq \min\{d(\eta, \xi), d(\zeta, \xi)\} \qquad \forall\, \eta, \zeta, \xi \in \Omega_N. \tag{1.26}$$

For $\eta \in \Omega_N$ and $k \in \mathbb{N}_0$, the k-block around η is defined as

$$B_k(\eta) = \{\zeta \in \Omega_N\colon\ d(\eta, \zeta) \leq k\}. \tag{1.27}$$

In what follows, for each $\ell \in \mathbb{N}_0$ we write $\Omega_N^{[\ell]}$ to denote the set of vertices in the tree at height ℓ, and $\eta^{[\ell]}$ to denote the ancestor of η in $\Omega_N^{[\ell]}$.

1.5 Hierarchical Kuramoto model

We are now ready to define the model that will be our object of study. Each vertex $\eta \in \Omega_N$ carries a phase oscillator, whose phase at time t is denoted by $\theta_\eta(t)$. Oscillators interact in pairs, but at a strength that depends on their hierarchical distance. To modulate this interaction, we introduce a sequence of interaction strengths

$$(K_k)_{k\in\mathbb{N}} \in (0, \infty)^{\mathbb{N}}, \tag{1.28}$$

and we let each pair of oscillators $\eta, \zeta \in \Omega_N$ at distance $d(\eta, \zeta) = d$ interact as in the mean-field Kuramoto model with K/N replaced by K_d/N^{2d-1}, where the scaling factor is chosen to ensure that the model remains well behaved in the limit as $N \to \infty$. In addition, we subject each oscillator to a sequence of random natural frequencies at each hierarchical level. Thus, our coupled evolution equations read

$$\mathrm{d}\theta_\eta(t) = \sum_{k\in\mathbb{N}_0} \frac{1}{N^k}\, \omega_{\eta^{[k]}}\, \mathrm{d}t + \sum_{\zeta\in\Omega_N} \frac{K_{d(\eta,\zeta)}}{N^{2d(\eta,\zeta)-1}} \sin\big[\theta_\zeta(t) - \theta_\eta(t)\big]\, \mathrm{d}t + D\, \mathrm{d}W_\eta(t),\ \ \eta \in \Omega_N, t \geq 0, \tag{1.29}$$

where $(W_\eta(t))_{t\geq 0}$, $\eta \in \Omega_N$, are i.i.d. standard Brownian motions, and

$$\blacktriangleright \quad \begin{array}{l} (\omega_\zeta)_{\zeta\in\Omega_N^{[k]}} \text{ are i.i.d. and are drawn from a common probability} \\ \text{distribution } \mu^{[k]}(\mathrm{d}\omega) \text{ at each height } k \in \mathbb{N}_0 \text{ in the tree.} \end{array} \tag{1.30}$$

As initial condition we take, as in (1.7),

$$\blacktriangleright \quad \begin{array}{l} \theta_\eta(0),\ \eta \in \Omega_N, \text{ are i.i.d. and are drawn from} \\ \text{a common probability distribution } \rho(\mathrm{d}\theta) \text{ on } [0, 2\pi]. \end{array} \tag{1.31}$$

We will be interested in understanding the evolution of the order parameter associated with the N^k oscillators in the k-block around η at time $N^{k-1}t$, defined by

$$R_\eta^{[k]}(t)\,\mathrm{e}^{i\Phi_\eta^{[k]}(N^{-1}t)} = \frac{1}{N^k} \sum_{\zeta\in B_k(\eta)} \mathrm{e}^{i\theta_\zeta(N^{k-1}t)}, \qquad \eta \in \Omega_N, t \geq 0, \tag{1.32}$$

where $R_\eta^{[k]}(t)$ is the synchronization level and $\Phi_\eta^{[k]}(N^{-1}t)$ is the average phase. The time scales in (1.32) are a natural choice in view of (1.23) and the scaling factors in (1.29). Our goal will be to pass to the limit $N \to \infty$, look at the limiting synchronization levels around a given vertex, say $\eta = 0$, and classify the scaling behavior of these synchronization levels as $k \to \infty$ into universality classes according to the choice of $(K_k)_{k\in\mathbb{N}}$ in (1.28).

Note that, for every $\eta \in \Omega_N$, we can telescope to write

$$\begin{aligned} \sum_{\zeta\in\Omega_N} \frac{K_{d(\zeta,\eta)}}{N^{2d(\eta,\zeta)-1}} \sin\big[\theta_\zeta(t) - \theta_\eta(t)\big] &= \sum_{k\in\mathbb{N}} \frac{K_k}{N^{2k-1}} \sum_{\zeta\in B_k(\eta)/B_{k-1}(\eta)} \sin\big[\theta_\zeta(t) - \theta_\eta(t)\big] \\ &= \sum_{k\in\mathbb{N}} \Big(\frac{K_k}{N^{2k-1}} - \frac{K_{k+1}}{N^{2(k+1)-1}}\Big) \sum_{\zeta\in B_k(\eta)} \sin\big[\theta_\zeta(t) - \theta_\eta(t)\big]. \end{aligned} \tag{1.33}$$

Inserting (1.33) into (1.29) and using (1.32), we get

$$\begin{aligned} \mathrm{d}\theta_\eta(t) &= \sum_{k\in\mathbb{N}_0} \frac{1}{N^k}\, \omega_{\eta^{[k]}}\, \mathrm{d}t \\ &\quad + \sum_{k\in\mathbb{N}} \frac{1}{N^{k-1}} \Big(K_k - \frac{K_{k+1}}{N^2}\Big)\, R_\eta^{[k]}(N^{1-k}t)\, \sin\Big[\Phi_\eta^{[k]}(N^{-k}t) - \theta_\eta(t)\Big]\,\mathrm{d}t + D\,\mathrm{d}W_\eta(t). \end{aligned} \tag{1.34}$$

This shows that, like in (1.6), the oscillators are coupled via the order parameters associated with the k-blocks for all $k \in \mathbb{N}$, suitably weighted.

When we pass to the limit $N \to \infty$ in (1.34), in the first sum only the term with $k = 0$ survives, while in the second sum only the term with $k = 1$ survives, so that we end up with an *autonomous* evolution equation similar to (1.10). The goal of the present paper is to show that when (compare with (1.9))

$$\blacktriangleright \quad \text{the disorder distributions } \mu^{[k]},\ k \in \mathbb{N}_0, \text{ are symmetric,} \tag{1.35}$$

a similar decoupling occurs *at all block levels*. Indeed, we expect the successive time scales at which synchronization occurs to separate. If there is synchronization at scale k, then we expect the average of the k-blocks around the origin forming the $(k+1)$-blocks (of which there are N in total) to behave *as if they were single oscillators* at scale $k+1$.

2 Main theorems

In Sections 2.1 and 2.2 we state our main theorems for the multi-scaling of the system without disorder (Theorems 2.2–2.4 below), respectively, the system with disorder (Theorems 2.6–2.8 below). The proofs of these theorems are given in Sections 3–4. In Section 2.2 we further show that spreading out the disorder lowers the synchronization level (Theorem 2.9 below). The proof of this fact is given in Appendix A.1. The system without disorder already exhibits interesting and intricate scaling behaviour, which is why we treat it separately and do not present it as a special case of the system with disorder.

Without loss of generality we set $D = 1$ in (1.29). *Throughout this section we assume that* (1.30)–(1.31) *and* (1.35) *are in force.* We further assume that the disorder is *small*, although we believe this assumption to be redundant (see Remark 4.1 below).

2.1 Non-disordered system

Define

$$Z(\lambda) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi\ \mathrm{e}^{\lambda\cos\phi}, \qquad \lambda > 0. \tag{2.1}$$

After normalization, the integrand becomes what is called the von Mises probability density function on the unit circle with parameter λ, which is $\phi \mapsto p_\lambda(\phi, 0)$ (recall (1.16)–(1.17)).

Definition 2.1. (Renormalization map) For $K \in (0,\infty)$, let $\mathcal{T}_K\colon [0,1]\times[\frac{1}{2},1] \to [0,1]\times[\frac{1}{2},1]$ be the map

$$(\bar{R},\bar{Q}) = \mathcal{T}_K(R,Q) \tag{2.2}$$

defined by

$$\begin{aligned} \bar{R} &= R\,\frac{Z'}{Z}\big(2K\bar{R}\sqrt{Q}\,\big),\\ \bar{Q}-\tfrac{1}{2} &= (Q-\tfrac{1}{2})\left[2\frac{Z''}{Z}\big(2K\bar{R}\sqrt{Q}\,\big)-1\right]. \end{aligned} \tag{2.3}$$

The first equation is a *consistency relation* and the second equation is a *recursion relation.* They must be used in that order to find the image point $(\bar{R},\bar{Q})$ of the original point (R,Q) under the map $\mathcal{T}_K$. □

Our first theorem shows that the average phase of the k-blocks behaves like that of the noisy mean-field Kuramato model described in Theorem 1.2.

Theorem 2.2. (Multi-scaling for the block average phases) *Fix $k \in \mathbb{N}$ and assume that $R^{[k]} > 0$. Then, in the limit as $N \to \infty$, $(\Phi_0^{[k]}(t))_{t\geq 0}$ evolves according to the Itô-equation*

$$d\Phi_0^{[k]}(t) = K_{k+1}\frac{Q^{[k]}}{R^{[k]}}\,R_0^{[k+1]}(t)\,\sin\big[\Phi-\Phi_0^{[k]}(t)\big]\,\mathrm{d}t + \frac{\sqrt{Q^{[k]}}}{R^{[k]}}\,\mathrm{d}W_0^{[k]}(t), \quad t\geq 0, \tag{2.4}$$

where $(W_0^{[k]}(t))_{t\geq 0}$ is a standard Brownian motion, $\Phi = 0$ by calibration, and

$$(R^{[k]},Q^{[k]}) = \big(\mathcal{T}_{K_k}\circ\cdots\circ\mathcal{T}_{K_1}\big)(R^{[0]},Q^{[0]}), \qquad k\in\mathbb{N}, \tag{2.5}$$

with $(R^{[0]},Q^{[0]}) = (1,1)$. □

The evolution in (2.4) is that of a mean-field noisy Kuramoto model with *renormalized coefficients* (compared with (1.6)). The composition in (2.5) is to be viewed as the result of the iteration of *renormalization maps* $\mathcal{T}_{K_1}, \ldots, \mathcal{T}_{K_k}$ acting on block communities at levels $1, \ldots, k$ successively, starting from the initial value $(R^{[0]}, Q^{[0]}) = (1,1)$. This initial value comes from the fact that single oscillators are completely synchronized by definition.

The evolution in (2.4) is *not closed* because of the presence of the term $R_0^{[k+1]}(t)$, which comes from the $(k+1)$-st block community one hierarchical level up from k. Similarly as in (1.11), $R_0^{[k+1]}(t)$ is driven by a *deterministic relaxation equation* such that

$$R_0^{[k+1]}(0) = R, \qquad \lim_{t\to\infty} R_0^{[k+1]}(t) = R^{[k+1]}. \tag{2.6}$$

This relaxation equation will be of no concern to us here. Convergence holds at least for R close to $R^{[k+1]}$ (recall Remark 1.1). Thus, after a *transient period*, (2.4) converges to the *closed* evolution equation

$$d\Phi_0^{[k]}(t) = K_{k+1} \frac{Q^{[k]}}{R^{[k]}} R^{[k+1]} \sin\left[\Phi - \Phi_0^{[k]}(t)\right] \mathrm{d}t + \frac{\sqrt{Q^{[k]}}}{R^{[k]}} \mathrm{d}W_0^{[k]}(t), \quad t \geq 0. \tag{2.7}$$

We will see that $k \mapsto (R^{[k]}, Q^{[k]})$ is non-increasing in both components. In particular, synchronization cannot increase when the hierarchical level goes up.

There are *three universality classes* depending on the choice of $(K_k)_{k\in\mathbb{N}}$ in (1.28), illustrated in Fig. 4:

(1) Synchronization is lost at a finite level:
$R^{[k]} > 0$, $0 \leq k < k_*$, $R^{[k]} = 0$, $k \geq k_*$ for some $k_* \in \mathbb{N}$.

(2) Synchronization is lost asymptotically:
$R^{[k]} > 0$, $k \in \mathbb{N}_0$, $\lim_{k\to\infty} R^{[k]} = 0$.

(3) Synchronization is not lost asymptotically:
$R^{[k]} > 0$, $k \in \mathbb{N}_0$, $\lim_{k\to\infty} R^{[k]} > 0$.

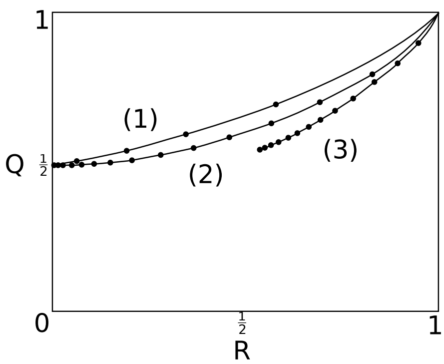


Figure 4: The dots represent the map $k \mapsto (R^{[k]}, Q^{[k]})$ for the three universality classes, starting from $(R^{[0]}, Q^{[0]})(1,1)$. The dots move left and down as k increases.

Our second theorem provides sufficient conditions for universality classes (1) and (3) in terms of the sum $\sum_{k\in\mathbb{N}} K_k^{-1}$.

Theorem 2.3. (Criteria for the universality classes)

- $\sum_{k\in\mathbb{N}} K_k^{-1} \geq 4 \Longrightarrow$ *universality class* (1).
- $\sum_{k\in\mathbb{N}} K_k^{-1} \leq \frac{1}{\sqrt{2}} \Longrightarrow$ *universality class* (3). □

Two examples are: (1) $K_k = \frac{3}{2\log 2}\log(k+1)$; (3) $K_k = 4e^k$. The scaling behaviour for these examples can be seen from the numerical analysis in Appendix B (see, in particular, Fig. 8 and Fig. 9 below).

The criteria in Theorem 2.3 are not sharp. Universality class (2) corresponds to a *critical surface* in the space of parameters $(K_k)_{k\in N}$ that appears to be rather complicated and certainly is not (!) of the type $\sum_{k\in\mathbb{N}} K_k^{-1} = c$ for some $\frac{1}{\sqrt{2}} < c < 4$ (see Fig. 5). Note that the full sequence $(K_k)_{k\in\mathbb{N}}$ determines in which universality class the system is.

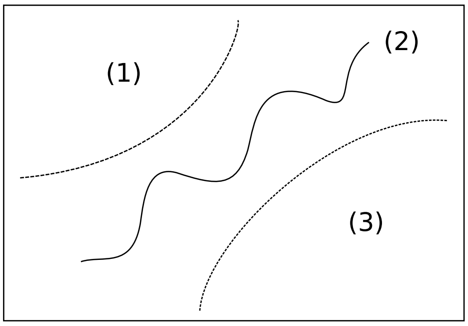


Figure 5: Caricature showing the critical surface in the parameter space and the bounds provided by Theorem 2.3.

The behaviour of K_k as $k \to \infty$ determines the speed at which $R^{[k]} \to R^{[\infty]}$ in universality classes (2) and (3). Our third theorem provides upper and lower bounds.

Theorem 2.4. (Bounds for the block synchronization levels)

- *In universality classes* (2) *and* (3),
$$\tfrac{1}{4}\sigma_k \leq R^{[k]} - R^{[\infty]} \leq \sqrt{2}\,\sigma_k, \qquad k \in \mathbb{N}_0, \tag{2.8}$$
with $\sigma_k = \sum_{\ell>k} K_\ell^{-1}$.
- *In universality class* (1), *the upper bound holds for* $k \in \mathbb{N}_0$, *while the lower bound is replaced by*
$$R^{[k]} - R^{[k_*-1]} \geq \tfrac{1}{4}\sum_{\ell=k+1}^{k_*-1} K_\ell^{-1}, \qquad 0 \leq k \leq k_* - 2, \tag{2.9}$$
which implies that
$$k_* \leq \max\left\{k \in \mathbb{N}\colon \sum_{\ell=1}^{k-1} K_\ell^{-1} < 4\right\} \tag{2.10}$$
because $R^{[0]} = 1$ *and* $R^{[k_*-1]} > 0$. □

In universality classes (2) and (3) we have $\lim_{k\to\infty} \sigma_k = 0$. Depending on how fast $k \mapsto K_k$ grows, various speeds of convergence are possible: logarithmic, polynomial, exponential, superexponential.

2.2 Disordered system

In this section we show how Definition 2.1 and Theorems 2.2–2.4 generalize in the presence of disorder.

Once the stationary state has been reached by the k-blocks in a given $(k+1)$-block with a current synchronization level $R^{[k+1]}$, the k-blocks have an equilibrium distribution that is given by

$$p^{[k]}(\theta,\omega) = \frac{A_\lambda(\theta,\omega)}{\int_0^{2\pi} \mathrm{d}\phi\, A_\lambda(\phi,\omega)}, \qquad \lambda = 2K_{k+1}R^{[k+1]}\sqrt{Q^{[k]}}. \tag{2.11}$$

Indeed, these are the expressions in (1.16)–(1.17) with the replacements

$$K \to K_{k+1}Q^{[k]}/R^{[k]}, \quad r \to R^{[k+1]}, \quad D \to \sqrt{Q^{[k]}}/R^{[k]}, \tag{2.12}$$

which are the renormalized coefficients in (2.4).

Definition 2.5. (Renormalization map) *For $k \in \mathbb{N}_0$, let $\mathcal{T}_{K_{k+1}}\colon [0,1]\times[\frac{1}{2},1] \to [0,1]\times[\frac{1}{2},1]$ be the map*

$$(R^{[k+1]}, Q^{[k+1]}) = \mathcal{T}_{K_{k+1}}(R^{[k]}, Q^{[k]}) \tag{2.13}$$

with

$$\begin{aligned} R^{[k+1]} &= R^{[k]}\left(\int_0^{2\pi} \mathrm{d}\theta\, \cos\theta \int_{\mathbb{R}} \mu^{[k]}(\mathrm{d}\omega)\, p^{[k]}(\theta,\omega)\right), &(2.14)\\ Q^{[k+1]} - \tfrac{1}{2} &= (Q^{[k]} - \tfrac{1}{2})\left[\left(2\int_0^{2\pi} \mathrm{d}\theta\, \cos^2\theta \int_{\mathbb{R}} \mu^{[k]}(\mathrm{d}\omega)\, p^{[k]}(\theta,\omega)\right) - 1\right]. &(2.15) \end{aligned}$$

The derivation of (2.14)–(2.15) is given in Section 3.3. It is easy to see that $k \mapsto (R^{[k]}, Q^{[k]})$ is non-increasing in both components, and that $(R^{[k]}, Q^{[k]}) \in [0,1]\times[\frac{1}{2},1]$.

Theorem 2.6. (Multi-scaling for the block average phases) *Fix $k \in \mathbb{N}$ and assume that $R^{[k]} > 0$. Then, in the limit as $N \to \infty$, the same evolution equation as in (2.4) holds, but with an extra term $\omega_{0^{[k]}}\,\mathrm{d}t$ in the right-hand side, and with $\mathcal{T}_{K_k}$ replaced by the map defined in (2.13)–(2.15).* □

The same observations apply as for the non-disordered case: there are three universality classes, illustrated by Fig. 4–5. The following two theorems are the analogues of Theorems 2.3–2.4, and require that, in addition to (1.35),

- ► the disorder distributions $\mu^{[k]}$, $k \in \mathbb{N}_0$, are unimodal. (2.16)

Let

$$I_k = \int_{\mathbb{R}} \mu^{[k]}(d\omega)\, \frac{1}{1+4\omega^2}. \tag{2.17}$$

The following two theorems are subject to a *technical inequality*, stated in Lemma 4.9 below, which is verified *numerically* in Appendix B.2 for three different types of disorder.

Theorem 2.7. (Criteria for the universality classes)

- $\sum_{k\in\mathbb{N}} K_k^{-1} \geq 4 \Longrightarrow$ *universality class* (1).
- $\sum_{k\in\mathbb{N}} (I_k K_k)^{-1} \leq \frac{1}{\sqrt{2}} \Longrightarrow$ *universality class* (3). $\square$

Theorem 2.8. (Bounds for the block synchronization levels)

- *In universality classes* (2) *and* (3),
$$\tfrac{1}{4}\sigma_k \leq R^{[k]} - R^{[\infty]} \leq \sqrt{2}\,\hat{\sigma}_k, \qquad k \in \mathbb{N}_0, \tag{2.18}$$
with $\sigma_k = \sum_{\ell>k} K_\ell^{-1}$ *and* $\hat{\sigma}_k = \sum_{\ell>k} (I_\ell K_\ell)^{-1}$.
- *In universality class* (1), *the upper bound holds for* $k \in \mathbb{N}_0$, *while the same lower bound holds as in* (2.9), *which again implies* (2.10). $\square$

Theorems 2.7–2.8 indicate how the disorder affects the presence of synchronization. Indeed, the integral in (2.17) is < 1 unless $\mu^{[k]} = \delta_0$. Hence, with disorder synchronization is harder: universality class (3) comes with a *more stringent criterion* in Theorem 2.7 than in Theorem 2.3. We expect that, similarly, universality class (1) should come with a *less stringent criterion* in Theorem 2.7 than in Theorem 2.3. However, our estimates are not good enough to make this apparent. The reason is that, as we move up the hierarchy and look at larger and larger blocks, the disorder has a tendency to wash out (because μ is symmetric), so that for large k the evolution of the k-th block average gets close to that without disorder. See Section 4 for more details.

In Appendix A.1 we show that *synchronization is monotone in the spread of the disorder.* More precisely, recall (1.35) and let $\bar{\mu}^{[k]}(\mathrm{d}\omega) = \mu^{[k]}(\mathrm{d}\omega) + \mu^{[k]}(-\mathrm{d}\omega)$ be the projection of $\mu^{[k]}$ on $[0,\infty)$ obtained by reflection.

Theorem 2.9. (Less synchronization for more spread-out disorder) *For all* $k \in N_0$, *if* $\bar{\mu}^{[k],*}$ *is stochastically strictly larger than* $\bar{\mu}^{[k]}$, *then the synchronization level under* $\bar{\mu}^{[k],*}$ *is strictly smaller than the synchronization level under* $\bar{\mu}^{[k]}$. $\square$

3 Proof of the multi-scaling for the block average phases

In Section 3.1 we prove Theorem 1.2. The proof of the diffusive scaling of the average phase in the mean-field noisy Kuramato model, as shown in the first line of (1.23), serves as a prelude to the proof of the multi-scaling of the block average phases in the hierarchical noisy Kuramoto model, stated in Theorems 2.2 and 2.6. The proof of the latter is given in Section 3.2. In Section 3.3 we derive (2.14)–(2.15), the formulas for the renormalization map.

3.1 Scaling of average phase for mean-field Kuramato

Proof. For the derivation of the second line of (1.23) we combine (1.13)–(1.14), to obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} r(t) = &\int_{\mathbb{R}} \mu(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta\, \cos\theta \\ &\times \left\{ -\frac{\partial}{\partial\theta}\Big[p_\lambda(t;\theta,\omega)\left\{\omega + Kr(t)\sin[\psi(t)-\theta]\right\}\Big] + \frac{D}{2}\frac{\partial^2}{\partial\theta^2} p_\lambda(t;\theta,\omega)\right\}\end{aligned} \tag{3.1}$$

with $\lambda = 2Kr/D$. After partial integration with respect to θ this becomes (use that $\theta \mapsto p_\lambda(t;\theta,\omega)$ is periodic)

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t) = \int_{\mathbb{R}} \mu(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta \, p_\lambda(t;\theta,\omega) \left\{ (-\sin\theta)\big\{\omega + Kr(t)\sin(-\theta)\big\} + (-\cos\theta)\,\frac{D}{2} \right\}, \tag{3.2}$$

where we use that $\psi(t) = \psi(0) = 0$. Define

$$\begin{aligned} C(t) &= \int_{\mathbb{R}} \mu(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta \, p_\lambda(t;\theta,\omega)\, \omega \sin\theta, \\ q(t) &= \int_{\mathbb{R}} \mu(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta \, p_\lambda(t;\theta,\omega)\, \cos^2\theta. \end{aligned} \tag{3.3}$$

Then (3.2) reads

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t) = -C(t) + \left[K(1-q(t)) - \frac{D}{2} \right] r(t). \tag{3.4}$$

We know that $\lim_{t\to\infty} r(t) = r$, $\lim_{t\to\infty} C(t) = C$ and $\lim_{t\to\infty} q(t) = q$, with $0 = -C + [K(1-q) - \frac{D}{2}]\, r$. The details of the relaxation are delicate and depend on the full solution of the McKean-Vlasov equation in (1.13).

For the derivation of the first line of (1.23) we use the symmetry of the equilibrium distribution (recall (1.16)–(1.17)), i.e.,

$$p_\lambda(\theta,\omega) = p_\lambda(-\theta,-\omega) \tag{3.5}$$

and the symmetry of the disorder distribution (recall (1.9)), i.e.,

$$\mu(\mathrm{d}\omega) = \mu(-\mathrm{d}\omega), \tag{3.6}$$

together with the fact that $x \mapsto \cos x$ is a symmetric function and $x \mapsto \sin x$ is an asymmetric function.

Itô's rule applied to (1.5) gives

$$\mathrm{d}\psi_N(t) = \sum_{i=1}^N \frac{\partial \psi_N}{\partial \theta_i}(t)\, \mathrm{d}\theta_i(t) + \frac{1}{2} \sum_{i=1}^N \frac{\partial^2 \psi_N}{\partial \theta_i^2}(t) \big(\mathrm{d}\theta_i(t)\big)^2 \tag{3.7}$$

with

$$\frac{\partial \psi_N}{\partial \theta_i} = \frac{1}{Nr} \cos(\psi - \theta_i), \tag{3.8}$$

$$\frac{\partial^2 \psi_N}{\partial \theta_i^2} = -\frac{1}{(Nr)^2} \sin\big(2(\psi - \theta_i)\big) + \frac{1}{Nr} \sin(\psi - \theta_i). \tag{3.9}$$

Inserting (1.6), we get

$$\mathrm{d}\psi_N(t) = \big[I_1(N;t) + I_2(N;t) + I_3(N;t)\big]\, \mathrm{d}t + \mathrm{d}J(N;t) \tag{3.10}$$

with

$$
\begin{aligned}
I_1(N;t) &= \frac{1}{Nr_N(t)} \sum_{i=1}^{N} \cos\big[\psi_N(t) - \theta_i(t)\big]\,\omega_i(t),\\
I_2(N;t) &= \frac{K}{N} \sum_{i=1}^{N} \cos\big[\psi_N(t) - \theta_i(t)\big] \sin\big[\psi_N(t) - \theta_i(t)\big],\\
I_3(N;t) &= -\frac{D^2}{2\big(Nr_N(t)\big)^2} \sum_{i=1}^{N} \sin\big(2\big[\psi_N(t) - \theta_i(t)\big]\big),\\
\mathrm{d}J(N;t) &= \frac{D}{Nr_N(t)} \sum_{i=1}^{N} \cos\big[\psi_N(t) - \theta_i(t)\big]\,\mathrm{d}W_i(t),
\end{aligned}
\tag{3.11}
$$

where we use that $\sum_{i=1}^{N} \sin[\psi_N(t) - \theta_i(t)] = 0$ by (1.5). We assume that the system converges to a partially synchronized state, i.e.,

$$
\lim_{N\to\infty} r_N(t) = r > 0. \tag{3.12}
$$

Consequently, as $N \to \infty$ the first term in (3.11) vanishes in distribution because of (3.5)–(3.6), $\lim_{N\to\infty} \psi_N(t) = \psi(t) = \psi(0) = \Phi = 0$ for all $t \geq 0$ (recall Section 1.2), and the fact that the oscillators decouple from the average phase. The second term vanishes in distribution for the same reasons, while the third term vanishes trivially. The fourth term is equal in distribution to $\bar{J}(N;t)\,\mathrm{d}W_*(t/N)$ with

$$
\bar{J}(N;t) = \frac{D}{r_N(t)} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \cos^2\big[\psi_N(t) - \theta_i(t)\big]}. \tag{3.13}
$$

Replacing t by Nt, we see that

$$
\lim_{N\to\infty} \psi_N(Nt) = \psi_*(t) \quad \text{with} \quad \psi_*(t) = D_*\,W_*(t), \quad \psi_*(0) = \Phi = 0, \tag{3.14}
$$

where

$$
D_* = \lim_{N\to\infty} \bar{J}(N;Nt) = \lim_{N\to\infty} \frac{D}{r_N(Nt)} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \cos^2\big[\psi_N(Nt) - \theta_i(Nt)\big]}. \tag{3.15}
$$

But $\lim_{N\to\infty} r_N(Nt) = r$, while the term under the square root converges to q defined in (3.3). The latter holds because $\theta_i(Nt)$, $i = 1, \ldots, N$, are asymptotically independent and $\theta_i(Nt)$ converges in distribution to $\theta \mapsto p_\lambda(\theta, \omega_i)$ *relative* to the value of $\psi_N(Nt)$, which itself evolves but only *slowly* on time scale Nt. Hence we get the claim in the first line of (1.23) with $D_* = D\sqrt{q}/r$. □

3.2 Multi-scaling of block average phases for hierarchical Kuramoto

We give only the main idea behind the proof of Theorems 2.2 and 2.6. The argument runs along the same line as in Section 3.1, but is more involved because of the hierarchical interaction. The details can be filled in with more effort, based on the techniques developed for the McKean-Vlasov equation in (1.13)–(1.14) (see e.g. Dai Pra and den Hollander [3]).

What is crucial for the argument is the *separation of space-time scales*:

- Each k-block consists of N disjoint $(k-1)$-blocks, and evolves on a time scale that is N times larger than the time scale on which the constituent blocks evolve.
- In the limit as $N \to \infty$, the constituent $(k-1)$-blocks rapidly achieve equilibrium subject to the *current* value of the k-block, which allows us to treat them as a noisy mean-field Kuramoto model with coefficients that depend on their synchronization level and their average phase, with an effective interaction that depends on the current value of the synchronization level of the k-block (recall (2.12)).
- The k-block itself interacts with the $(k+1)$-block it is part of, with interaction strength K_{k+1}, while the interaction with the even larger blocks it is part of is negligible as $N \to \infty$.

Recall that in (1.32) we already reduced the times scale of the average phase by a factor N^{-1}, in accordance with the scaling found in the first line of (1.23). Thus, we have that $\Phi_0^{[k]}(t)$ is the average phase of the k-block around 0 at time $N^k t$.

Proof. Itô's rule applied to (1.32) gives

$$\mathrm{d}\Phi_0^{[k]}(t) = \sum_{\zeta \in B_k(0)} \frac{\partial \Phi_0^{[k]}}{\partial \theta_\zeta}(t)\, \mathrm{d}\theta_\zeta(N^k t) + \frac{1}{2} \sum_{\zeta \in B_k(0)} \frac{\partial^2 \Phi_0^{[k]}}{\partial \theta_\zeta^2}(t) \left(\mathrm{d}\theta_\zeta(N^k t)\right)^2 \tag{3.16}$$

with

$$\frac{\partial \Phi_0^{[k]}}{\partial \theta_\zeta} = \frac{1}{N^k R_0^{[k]}} \cos\left[\Phi_0^{[k]} - \theta_\zeta\right], \tag{3.17}$$

$$\frac{\partial^2 \Phi_0^{[k]}}{\partial \theta_\zeta^2} = -\frac{1}{\left[N^{2k} R_0^{[k]}\right]^2} \sin\left(2\left[\Phi_0^{[k]} - \theta_\zeta\right]\right) + \frac{1}{N^k R_0^{[k]}} \sin\left[\Phi_0^{[k]} - \theta_\zeta\right]. \tag{3.18}$$

Inserting (1.34), we find

$$d\Phi_0^{[k]}(t) = \left[I_1(k,N;t) + I_2(k,N;t) + I_3(k,N;t)\right] \mathrm{d}t + \mathrm{d}J(k,N;t) \tag{3.19}$$

with

$$\begin{aligned}
I_1(k,N;t) &= \frac{1}{R_0^{[k]}(Nt)} \sum_{\ell \in \mathbb{N}_0} \frac{1}{N^\ell} \sum_{\zeta \in B_k(0)} \cos\left[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\right] \omega_{\zeta^{[\ell]}},\\
I_2(k,N;t) &= \frac{1}{R_0^{[k]}(Nt)} \sum_{\ell \in \mathbb{N}} \left(K_\ell - \frac{K_{\ell+1}}{N^2}\right)\\
&\quad \times \frac{1}{N^{\ell-1}} \sum_{\zeta \in B_k(0)} R_\zeta^{[\ell]}(N^{k-\ell+1}t) \sin\left[\Phi_\zeta^{[\ell]}(N^{k-\ell}t) - \theta_\zeta(N^k t)\right] \cos\left[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\right],\\
I_3(k,N;t) &= -\frac{1}{2N^k\left[R_0^{[k]}(Nt)\right]^2} \sum_{\zeta \in B_k(0)} \sin\left(2\left[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\right]\right),\\
\mathrm{d}J(k,N;t) &= \frac{1}{N^{k/2} R_0^{[k]}(Nt)} \sum_{\zeta \in B_k(0)} \cos\left[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\right] \mathrm{d}W_\zeta(t).
\end{aligned} \tag{3.20}$$

We assume that the k-block converges to a partially synchronized state, i.e.,

$$\lim_{N\to\infty} R_0^{[k]}(Nt) = R^{[k]} > 0. \tag{3.21}$$

We must take the limit $N \to \infty$ and show that (3.19)–(3.20) reduces to (2.4) with $\omega_{0^{[k]}}\mathrm{d}t$ added on the right-hand side. To do so we will use the symmetry of the equilibrium distribution at all levels (recall (2.11)), i.e.,

$$p^{[k]}(\theta,\omega) = p^{[k]}(-\theta,-\omega), \qquad k \in \mathbb{N}_0, \tag{3.22}$$

and the symmetry of the disorder distribution at all levels (recall (1.35)), i.e.,

$$\mu^{[k]}(\mathrm{d}\omega) = \mu^{[k]}(-\mathrm{d}\omega), \qquad k \in \mathbb{N}_0. \tag{3.23}$$

The key idea is that, in the limit as $N \to \infty$, the average phases of the k-blocks around ζ *decouple* and converge in distribution to $\theta \mapsto p^{[k]}(\theta, \omega_\zeta^{[k]})$ for all $k \in \mathbb{N}_0$, just as for the noisy mean-field Kuramoto model discussed in Section 3.1. This is the reason why a *recursive structure* is in place, captured by the renormalization maps $\mathcal{T}_{K_k}$, $k \in \mathbb{N}$.

Along the way we need the quantities

$$\begin{aligned} R_0^{[k]}(Nt) &= \frac{1}{N^k} \sum_{\zeta\in B_k(0)} \cos\big[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\big], \\ Q_0^{[k]}(Nt) &= \frac{1}{N^k} \sum_{\zeta\in B_k(0)} \cos^2\big[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\big]. \end{aligned} \tag{3.24}$$

$\underline{I_1(k,N;t)}$. The terms with $\ell \geq k+1$ vanish trivially. The terms with $\ell \leq k-1$ vanish in distribution because of (3.23) and the fact that the k-block average *decouples* from the disorder at levels $\ell \leq k-1$. Hence only the term with $\ell = k$ survives, which converges to $\omega_{0^{[k]}}$ because $\zeta^{[k]} = 0^{[k]}$ for all $\zeta \in B_k(0)$.

$\underline{I_2(k,N;t)}$. The terms with $\ell \geq k+2$ vanish trivially. The terms with $\ell \leq k-1$ vanish in distribution, as can be shown with the help of a *telescoping argument*. Indeed, we may write

$$\Phi_\zeta^{[\ell]}(N^{k-\ell}t) - \theta_\zeta(N^k t) = \sum_{m=1}^{\ell} \Big[\Phi_\zeta^{[m]}(N^{k-m}t) - \Phi_\zeta^{[m-1]}(N^{k-(m-1)}t)\Big]. \tag{3.25}$$

Since

$$\begin{aligned} \sin\left(\sum_{m=1}^{\ell} x_m\right) &= \sum_{\substack{J\subset\{1,\dots,\ell\}\\ |J|\text{ odd}}} (-1)^{(|I|-1)/2} \prod_{m\notin J}\cos x_m \prod_{i\in J}\sin x_m, \\ \cos\left(\sum_{m=1}^{\ell} x_m\right) &= \sum_{\substack{J\subset\{1,\dots,\ell\}\\ |J|\text{ even}}} (-1)^{|J|/2} \prod_{m\notin J}\cos x_m \prod_{i\in J}\sin x_m, \\ & x_1,\dots,x_\ell \in \mathbb{R},\ \ell\in\mathbb{N}, \end{aligned} \tag{3.26}$$

we see that $\sin[\Phi_\zeta^{[\ell]}(N^{k-\ell}t)-\theta_\zeta(N^k t)]$ is a sum of terms containing an odd number of factors of the form

$$\sin\left[\Phi_\zeta^{[m]}(N^{k-m}t)-\Phi_\zeta^{[m-1]}(N^{k-(m-1)}t)\right], \qquad 1\le m\le \ell, \tag{3.27}$$

while $\cos[\Phi_0^{[k]}(t)-\theta_\zeta(N^k t)]$ is a sum of terms containing an even number of factors of the form

$$\sin\left[\Phi_\zeta^{[m]}(N^{k-m}t)-\Phi_\zeta^{[m-1]}(N^{k-(m-1)}t)\right], \qquad 1\le m\le k, \tag{3.28}$$

where we note that $\Phi_0^{[k]}(t)=\Phi_\zeta^{[k]}(t)$ for all $\zeta\in B_k(0)$. However, for every $1\le m\le \ell$,

- $\Phi_\zeta^{[m-1]}(N^{k-(m-1)}t)$ *decouples* from $R_\zeta^{[\ell]}(N^{k-\ell+1}t)$ and $\Phi_\zeta^{[m]}(N^{k-m}t)$ and converges in distribution to $\theta\mapsto p^{[m-1]}(\theta,\omega_{\zeta^{[m-1]}})$ *relative* to $\Phi_\zeta^{[m]}(N^{k-m}t)$, which itself evolves slowly compared to $\Phi_\zeta^{[m-1]}(N^{k-(m-1)}t)$.

Therefore the factor in (3.27) *vanishes on average* because of (3.22). By the same *telescoping argument* also the term with $\ell=k$ vanishes in distribution, where we note that $R_\zeta^{[k]}(Nt)=R_0^{[k]}(Nt)$ for all $\zeta\in B_k(0)$. Hence only the term with $\ell=k+1$ survives.

$I_3(k,N;t)$. This term vanishes trivially.

$\mathrm{d}J(k,N;t)$. This term is equal in distribution to

$$\frac{\sqrt{Q_0^{[k]}(Nt)}}{R_0^{[k]}(Nt)}\,\mathrm{d}W_0^{[k]}(t) \tag{3.29}$$

and converges in distribution to

$$\frac{\sqrt{Q^{[k]}}}{R^{[k]}}\,\mathrm{d}W_0^{[k]}(t) \tag{3.30}$$

with

$$\begin{aligned} R^{[k]} &= \lim_{N\to\infty} R_0^{[k]}(Nt),\\ Q^{[k]} &= \lim_{N\to\infty} Q_0^{[k]}(Nt). \end{aligned} \tag{3.31}$$

The quantities defined in (3.31) are relaxation limits and do not depend on t (recall the remark leading up to (2.6)).

Combining the above observations we see that, in the limit as $N\to\infty$, (3.19)–(3.20) reduce to

$$\mathrm{d}\Phi_0^{[k]}(t)=\omega_{0^{[k]}}\,\mathrm{d}t+K_{k+1}\frac{R_0^{[k+1]}(t)}{R^{[k]}}\,\bar I_2(k,N;t)\,\mathrm{d}t+\frac{\sqrt{Q^{[k]}}}{R^{[k]}}\,\mathrm{d}W_0^{[k]}(t) \tag{3.32}$$

with

$$\bar I_2(k,N;t)=\frac{1}{N^k}\sum_{\zeta\in B_k(0)}\sin\left[\Phi_0^{[k+1]}(N^{-1}t)-\theta_\zeta(N^k t)\right]\cos\left[\Phi_0^{[k]}(t)-\theta_\zeta(N^k t)\right], \tag{3.33}$$

where we use that $\Phi_\zeta^{[k+1]}(N^{-1}t) = \Phi_0^{[k+1]}(N^{-1}t)$ for all $\zeta \in B_k(0)$. With the help of the trigonometric identity $\sin(x-y+z)\cos z = \sin(x-y)\cos z - \frac{1}{2}\cos(x-y)\sin(2z)$, $x,y,z \in \mathbb{R}$, we can rewrite (pick $x = \Phi_0^{[k+1]}(N^{-1}t)$, $y = \Phi_0^{[k]}(t)$, $z = \Phi_0^{[k]}(t) - \theta_\zeta(N^k t)$)

$$\begin{aligned}\bar{I}_2(k,N;t) &= \sin\left[\Phi_0^{[k+1]}(N^{-1}t) - \Phi_0^{[k]}(t)\right] Q_0^{[k]}(Nt) \\ &\quad - \tfrac{1}{2}\sin\left[\Phi_0^{[k+1]}(N^{-1}t) - \Phi_0^{[k]}(t)\right] \frac{1}{N^k} \sum_{\zeta\in B_k(0)} \sin\left(2\left[\Phi_0^{[k]}(t) - \theta_\zeta(N^k t)\right]\right).\end{aligned} \tag{3.34}$$

Because of (3.22), the last term vanishes in distribution, where we again use a *telescoping argument*. Combining (3.31)–(3.34) and using that $\lim_{N\to\infty} \Phi_0^{[k+1]}(N^{-1}t) = \Phi = 0$, we get the evolution equation in Theorem 2.6, which is (2.4) with an extra term $\omega_{0^{[k]}}\mathrm{d}t$ in the right-hand side.

The proof for Theorem 2.2 is identical, except that the term $I_1(k,N;t)$ is absent in the non-disordered system, and hence so is the term $\omega_{0^{[k]}}$ to which it converges. □

3.3 Renormalization map for disordered system

Proof. To derive (2.15), we return to the definition of $Q_0^{[k+1]}(Nt)$ in the second line of (3.24). With the help of the trigonometric identity $\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos(2x)$, $x \in \mathbb{R}$, we can rewrite (pick $x = \Phi_0^{[k+1]}(t) - \theta_\zeta(N^{k+1}t)$)

$$Q_0^{[k+1]}(Nt) - \tfrac{1}{2} = \tfrac{1}{2}\left[\frac{1}{N^{k+1}} \sum_{\zeta\in B_{k+1}(0)} \cos\left(2\left[\Phi_0^{[k+1]}(t) - \theta_\zeta(N^{k+1}t)\right]\right)\right]. \tag{3.35}$$

Split the sum over $\zeta \in B_{k+1}(0)$ into a sum over the constituent k-blocks B_k^i, $i = 1,\dots,N$, and a sum over $\zeta \in B_k^i$. With the help of the trigonometric identity $\cos(y+z) = \cos y\cos z - \sin y \sin z$, $y,z \in \mathbb{R}$, we can rewrite the term between square brackets in the right-hand side as (pick $y = 2[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)]$, $z = 2[\Phi_{B_k^i}^{[k]}(Nt) - \theta_\zeta(N^{k+1}t)]$)

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^N \cos\left(2\left[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)\right]\right) \frac{1}{N^k} \sum_{\zeta\in B_k^i} \cos\left(2\left[\Phi_{B_k^i}^{[k]}(Nt) - \theta_\zeta(N^{k+1}t)\right]\right) \\ &\qquad - \frac{1}{N}\sum_{i=1}^N \sin\left(2\left[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)\right]\right) \frac{1}{N^k} \sum_{\zeta\in B_k^i} \sin\left(2\left[\Phi_{B_k^i}^{[k]}(Nt) - \theta_\zeta(N^{k+1}t)\right]\right),\end{aligned} \tag{3.36}$$

where $\Phi_{B_k^i}^{[k]}(Nt)$ denotes the average phase of the block B_k^i at time $N^{k+1}t$. Because of (3.22), the fourth sum vanishes in distribution as $N \to \infty$, where we again use a *telescoping argument*. Moreover, the second sum equals $2(Q_{B_k^i}^{[k]}(N^2t) - \frac{1}{2})$, with $Q_{B_k^i}^{[k]}(N^2t)$ the analogue of $Q_0^{[k]}(N^2t)$ that is obtained after replacing $B_k(0)$ by B_k^i in the second line of (3.24). Since $Q_{B_k^i}^{[k]}(N^2t)$ is equal in distribution to $Q_0^{[k]}(N^2t)$ for every B_k^i, we can combine (3.35)–(3.36) with the second lines of (3.24) and (3.31), to get

$$
\begin{aligned}
Q^{[k+1]} - \tfrac{1}{2} &= \big(Q^{[k]} - \tfrac{1}{2}\big) \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \cos\big(2\big[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)\big]\big) \\
&= \big(Q^{[k]} - \tfrac{1}{2}\big) \left(\left[2 \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \cos^2 \big[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)\big] \right] - 1 \right).
\end{aligned}
\tag{3.37}
$$

However,

- $\Phi_{B_k^i}^{[k]}(Nt)$, $i = 1, \dots, N$, are asymptotically independent and converge in distribution to $\theta \mapsto p^{[k]}(\theta, \omega^{[k]})$ *relative* to $\Phi_0^{[k+1]}(t)$, which itself evolves slowly compared to $\Phi_{B_k^i}^{[k]}(Nt)$.

Hence we have

$$
\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \cos^2 \big[\Phi_0^{[k+1]}(t) - \Phi_{B_k^i}^{[k]}(Nt)\big] = \int_{\mathbb{R}} \mu^{[k]}(\mathrm{d}\omega) \int_0^{2\pi} \mathrm{d}\theta\, p^{[k]}(\theta, \omega)\, \cos^2 \theta. \tag{3.38}
$$

Combine (3.37)–(3.38) to get (2.15).

The proof of (2.14) is similar and is based on the first line of (3.24). □

4 Universality classes and synchronization levels

In Section 4.1 we derive some basic properties of the renormalization map (Lemmas 4.2–4.4 below). In Section 4.2 we prove Theorems 2.3 and 2.7 for the non-disordered system. The proof relies on convexity and sandwich estimates (Lemmas 4.5–4.7 below). In Section 4.3 we prove Theorems 2.4 and 2.8 for the disordered system. The convexity and sandwich estimates are more delicate in the presence of disorder (Remark 4.1 and Lemma 4.9 below). The latter are proved in Appendix A.

4.1 Properties of the renormalization map

For $\lambda \in [0, \infty)$, define

$$
\begin{aligned}
V_\mu(\lambda) &= \int_0^{2\pi} \mathrm{d}\theta\, \cos\theta \int_{\mathbb{R}} \mu(\mathrm{d}\omega)\; p_\lambda(\theta, \omega), \\
W_\mu(\lambda) &= \int_0^{2\pi} \mathrm{d}\theta\, \cos^2\theta \int_{\mathbb{R}} \mu(\mathrm{d}\omega)\; p_\lambda(\theta, \omega),
\end{aligned}
\tag{4.1}
$$

where the probability distribution $p_\lambda(\theta, \omega)$ is given by (1.16) with $D = 1$. The renormalization map $\mathcal{T}_K$ in (2.2) can be written as $(\bar{R}, \bar{Q}) = \mathcal{T}_K(R, Q)$ with

$$
\begin{aligned}
\bar{R} &= R V_\mu(\lambda), \\
\bar{Q} - \tfrac{1}{2} &= (Q - \tfrac{1}{2})\big[2W_\mu(\lambda) - 1\big],
\end{aligned}
\tag{4.2}
$$

and $\lambda = 2K\bar{R}\sqrt{Q}$. In the non-disordered case, we have $V_{\delta_0} = V = Z'/Z$, $W_{\delta_0} = W = Z''/Z$ with $Z(\lambda)$ the integral defined in (2.1). In the disordered case, V_μ and W_μ are more difficult. Nevertheless it is possible to prove some general properties.

Remark 4.1. It is believed that $\lambda \mapsto V_\mu(\lambda)$ is *strictly concave* on $[0,\infty)$ when μ is symmetric and unimodal (see Luçon [9, Conjecture 3.12]). However, this conjecture has only been proved when μ is also narrow, i.e., the disorder is small (see Luçon [9, Proposition 3.13]). $\square$

Lemma 4.2. *The map $K \mapsto \bar{R}(R,K)$ is strictly increasing.* $\square$

Proof. The derivative of $\bar{R}$ w.r.t. K exists by the implicit function theorem, so that

$$\begin{aligned}\frac{d\bar{R}}{dK} &= 2RV'(2K\bar{R})\left[\bar{R} + K\frac{d\bar{R}}{dK}\right],\\ \frac{d\bar{R}}{dK}\left[1 - 2KRV'(2K\bar{R})\right] &= 2R\bar{R}V'(2K\bar{R}).\end{aligned} \tag{4.3}$$

Note that $\bar{R}$ is the solution to $\bar{R} = RV(2K\bar{R})$, which is non-trivial only when $1 < 2RKV'(2K\bar{R})$ due to the concavity of the map $R \mapsto RV(2K\bar{R})$. This implies that $2KRV'(2K\bar{R}) < 1$ at the solution, which makes the term in (4.3) between square brackets positive. The claim follows since we proved previously that $R, \bar{R} \in [0,1)$ and $V'(2K\bar{R}) > 0$. $\square$

Lemma 4.3. *The map $K \mapsto \bar{Q}(\bar{R},K,Q)$ is strictly increasing.* $\square$

Proof. The derivative of $\bar{Q}$ w.r.t. K exists by the implicit function theorem, so that

$$\frac{d\bar{Q}}{dK} = (Q - \tfrac{1}{2})\,4\sqrt{Q}\,W'\big(2\sqrt{Q}K\bar{R}\big)\left[\bar{R} + K\frac{d\bar{R}}{dK}\right]. \tag{4.4}$$

We have that $(Q-\frac{1}{2})\sqrt{Q} \geq 0$ because $Q \in [\frac{1}{2},1)$, $W'(2\sqrt{Q}K\bar{R}) > 0$ as proven before, and $[\bar{R} + K\frac{d\bar{R}}{dK}] > 0$ as in the proof of Lemma 4.2. The claim therefore follows. $\square$

Lemma 4.4. *The map $(R,Q) \mapsto (\bar{R},\bar{Q})$ is non-increasing in both components, i.e.,*

(i) $R \mapsto \bar{R}(K,R)$ *is non-increasing.*

(ii) $Q \mapsto \bar{Q}(K,\bar{R},Q)$ *is non-increasing.* $\square$

Proof. (i) We have

$$\bar{R} = R\,V\big(2\sqrt{Q}K\bar{R}\big). \tag{4.5}$$

But $V(\sqrt{Q}K\bar{R}) \in [0,1)$, and so $\bar{R} \leq R$.
(ii) We have

$$\bar{Q} - \tfrac{1}{2} = (Q - \tfrac{1}{2})\left[2W\big(2\sqrt{Q}K\bar{R}\big) - 1\right]. \tag{4.6}$$

But $W(2\sqrt{Q}K\bar{R}) \in [\frac{1}{2},1)$, and so $\bar{Q} \leq Q$. In fact, since both $V(2\sqrt{Q}K\bar{R})$ and $W(2\sqrt{Q}K\bar{R})$ are < 1, both maps are strictly decreasing until $R=0$ and $Q=\frac{1}{2}$ are hit, respectively. $\square$

4.2 Non-disordered system

Recall (2.1). To prove Theorems 2.3 and 2.7 we need the following lemma.

Lemma 4.5. *The map $\lambda \mapsto \log Z(\lambda)$ is analytic, strictly increasing and strictly convex on $(0,\infty)$, with*

$$Z(\lambda) = 1 + \tfrac{1}{4}\lambda^2\,[1+O(\lambda^2)], \quad \lambda \downarrow 0, \qquad Z(\lambda) = \frac{\mathrm{e}^{\lambda}}{\sqrt{2\pi\lambda}}\,[1+O(\lambda^{-1})], \quad \lambda \to \infty. \tag{4.7}$$

□

Proof. Analyticity is immediate from (2.1). Strict convexity follows because the numerator of $[\log Z(\lambda)]''$ equals

$$\begin{aligned} Z''(\lambda)Z(\lambda) - Z'(\lambda)Z'(\lambda) &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\psi\,[\cos^2\phi - \cos\phi\cos\psi]\,\mathrm{e}^{\lambda(\cos\phi+\cos\psi)} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\psi\,[\cos\phi - \cos\psi]^2\,\mathrm{e}^{\lambda(\cos\phi+\cos\psi)} > 0, \end{aligned} \tag{4.8}$$

where we symmetrize the integrand. Since $\log Z(0) = 0$, $\log Z(\lambda) > 0$ for $\lambda > 0$ and $\lim_{\lambda\to\infty} \log Z(\lambda) = \infty$, the strict monotonicity follows. The asymptotics in (4.7) is easily deduced from (2.1) via a saddle point computation. □

Since $V = Z'/Z = [\log Z]'$, we obtain from (4.7) and convexity that

$$V(\lambda) \sim \tfrac{1}{2}\lambda, \qquad \lambda \downarrow 0, \tag{4.9}$$

$$1 - V(\lambda) \sim \frac{1}{2\lambda}, \qquad \lambda \to \infty. \tag{4.10}$$

This limiting behaviour of $V(\lambda)$ inspires the choice of bounding functions in the next lemma.

Lemma 4.6. *$V^+(\lambda) \geq V(\lambda) \geq V^-(\lambda)$ for all $\lambda \in (0,\infty)$ with (see Fig. 6)*

$$\begin{aligned} V^+(\lambda) &= \frac{2\lambda}{1+2\lambda}, \\ V^-(\lambda) &= \frac{\frac{1}{2}\lambda}{1+\frac{1}{2}\lambda}. \end{aligned} \tag{4.11}$$

□

Proof. Recall that $V = Z'/Z$, with $Z = I_0$ and $Z' = I_1$ modified Bessel functions of the first kind. Segura [11, Theorem 1] shows that

$$V(\lambda) < V_*^+(\lambda) = \frac{\lambda}{\frac{1}{2} + \sqrt{(\frac{1}{2})^2 + \lambda^2}}, \qquad \lambda > 0. \tag{4.12}$$

Since $\lambda < \sqrt{(\frac{1}{2})^2 + \lambda^2}$, it follows that $V_*^+(\lambda) < V^+(\lambda)$. Laforgia and Natalini [8, Theorem 1.1] show that

$$V(\lambda) > V_*^-(\lambda) = \frac{-1 + \sqrt{\lambda^2+1}}{\lambda}, \qquad \lambda > 0. \tag{4.13}$$

Abbreviate $\eta = \sqrt{\lambda^2+1}$. Then $\lambda = \sqrt{(\eta-1)(\eta+1)}$, and we can write

$$V_*^-(\lambda) = \sqrt{\frac{\eta-1}{\eta+1}} = \frac{\lambda}{\eta+1} = \frac{\lambda}{2+(\eta-1)}. \tag{4.14}$$

Since $\lambda > \eta - 1$, it follows that $V_*^-(\lambda) > V^-(\lambda)$. $\square$

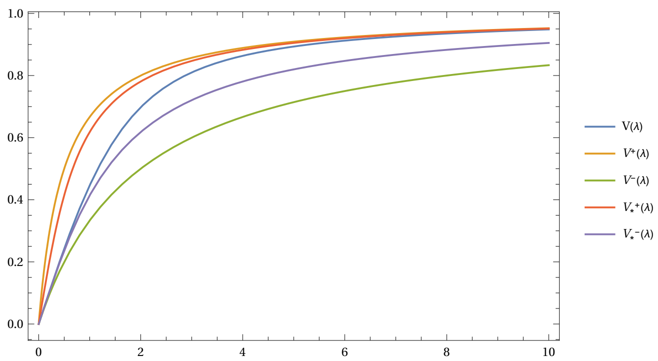


Figure 6: Plots of the tighter bounds in the proof of Lemma 4.6 and the looser bounds needed for the proof of Theorem 2.3.

Note that both V^+ and V^- are strictly increasing and concave on $(0,\infty)$, which guarantees the uniqueness and non-triviality of the solution to the consistency relation in the first line of (4.2) when we replace $V(\lambda)$ by either $V^+(\lambda)$ or $V^-(\lambda)$.

In the sequel we write V, W, R_k, Q_k instead of $V_{\delta_0}, W_{\delta_0}, R^{[k]}, Q^{[k]}$ to lighten the notation. We know that $(R_k)_{k\in\mathbb{N}_0}$ is the solution of the sequence of consistency relations

$$R_{k+1} = R_k V\big(2\sqrt{Q_k}K_{k+1}R_{k+1}\big), \qquad k \in \mathbb{N}_0. \tag{4.15}$$

This requires as input the sequence $(Q_k)_{k\in\mathbb{N}_0}$, which is obtained from the sequence of recursion relations

$$Q_{k+1} - \tfrac{1}{2} = (Q_k - \tfrac{1}{2})\big[2W\big(2\sqrt{Q_k}K_{k+1}R_{k+1}\big) - 1\big]. \tag{4.16}$$

By using that $Q_k \in [\frac{1}{2}, 1]$ for all $k \in \mathbb{N}_0$, we can remove Q_k from (4.15) at the cost of doing estimates. Namely, let $(R_k^+)_{k\in\mathbb{N}_0}$ and $(R_k^-)_{k\in\mathbb{N}_0}$ denote the solutions of the sequence of consistency relations

$$\begin{aligned} R_{k+1}^+ &= R_k V^+\big(2K_{k+1}R_{k+1}^+\big), && k \in \mathbb{N}_0,\\ R_{k+1}^- &= R_k V^-\big(2\sqrt{\tfrac{1}{2}}K_{k+1}R_{k+1}^-\big), && k \in \mathbb{N}_0. \end{aligned} \tag{4.17}$$

Lemma 4.7. *$R_k^+ \ge R_k \ge R_k^-$ for all $k \in \mathbb{N}$.* $\square$

Proof. If we replace $V(\lambda)$ by $V^+(\lambda)$ (or $V^-(\lambda)$) in the consistency relation for R_{k+1} given by (4.15), then the new solution R_{k+1}^+ (or R_{k+1}^-) is larger (or smaller) than R_{k+1}. Indeed, we have

$$R_{k+1} = R_k V(2K_{k+1}R_{k+1}\sqrt{Q_k}) \le R_k V^+(2K_{k+1}R_{k+1}). \tag{4.18}$$

Because V^+ is concave, it follows from (4.18) and the first line of (4.17) that $R_{k+1} \le R_{k+1}^+$. $\square$

We are now ready to prove Theorems 2.3–2.4.

Proof. From the first lines of (4.11) and (4.17) we deduce

$$R_k > \frac{1}{4K_{k+1}} \quad \Longleftrightarrow \quad R_{k+1}^+ > 0 \quad \Longrightarrow \quad R_k - R_{k+1}^+ = \frac{1}{4K_{k+1}}. \tag{4.19}$$

Hence, with the help of Lemma 4.7, we get

$$R_k > \frac{1}{4K_{k+1}} \quad \Longrightarrow \quad R_k - R_{k+1} \geq \frac{1}{4K_{k+1}}. \tag{4.20}$$

Iteration gives (recall that $R_0 = 1$)

$$1 - R_k \geq \min\left\{1, \sum_{\ell=1}^{k} \frac{1}{4K_\ell}\right\}. \tag{4.21}$$

As soon as the sum in the right-hand side is ≥ 1, we know that $R_k = 0$. This gives us the criterion for universality class (1) in Theorem 2.3. Similarly, from the second lines of (4.11) and (4.17) we deduce

$$R_k > \frac{2\sqrt{2}}{K_{k+1}} \quad \Longleftrightarrow \quad R_{k+1}^- > 0 \quad \Longrightarrow \quad R_k - R_{k+1}^- = \frac{\sqrt{2}}{K_{k+1}}. \tag{4.22}$$

Hence, with the help of Lemma 4.7, we get

$$R_k > \frac{\sqrt{2}}{K_{k+1}} \quad \Longrightarrow \quad R_k - R_{k+1} \leq \frac{\sqrt{2}}{K_{k+1}}. \tag{4.23}$$

Iteration gives

$$1 - R_k \leq \max\left\{1, \sum_{\ell=1}^{k} \frac{\sqrt{2}}{K_\ell}\right\}. \tag{4.24}$$

As soon as the sum in the right-hand side is < 1, we know that $R_k > 0$. This gives us the criterion for universality class (3) in Theorem 2.3.

In universality classes (2) and (3) we have $R_k^+ \geq R_k > 0$ for $k \in \mathbb{N}$, and hence

$$R_k - R_\infty = \sum_{\ell \geq k} (R_\ell - R_{\ell+1}) \geq \sum_{\ell \geq k} (R_\ell - R_{\ell+1}^+) = \sum_{\ell \geq k} \frac{1}{4K_{\ell+1}}, \qquad k \in \mathbb{N}_0. \tag{4.25}$$

In universality class (1), on the other hand, we have $R_k^+ \geq R_k > 0$ for $1 \leq k < k_*$ and $R_k = 0$ for $k \geq k_*$, and hence

$$R_k - R_{k_*-1} = \sum_{\ell=k}^{k_*-2} (R_\ell - R_{\ell+1}) \geq \sum_{\ell=k}^{k_*-2} (R_\ell - R_{\ell+1}^+) = \sum_{\ell=k}^{k_*-2} \frac{1}{4K_{\ell+1}}, \qquad 0 \leq k \leq k_* - 2. \tag{4.26}$$

Finally, with no assumption on $(R_k)_{k\in\mathbb{N}}$, we have

$$R_k - R_\infty = \sum_{\ell \geq k} (R_\ell - R_{\ell+1}) \leq \sum_{\ell \geq k} (R_\ell - R_{\ell+1}^-) \leq \sum_{\ell \geq k} \frac{\sqrt{2}}{K_{\ell+1}}, \tag{4.27}$$

where the last inequality follows from (4.22). The bounds in (4.25)–(4.27) yields the sandwich in Theorem 2.4. □

Remark 4.8. In the proof of Theorem 2.3–2.4 we exploited the fact that $Q_k \in [\frac{1}{2}, 1]$ to get estimates that involve a consistency relation in only R_k. In principle we can improve these estimates by exploring what effect Q_k has on R_k. Namely, in analogy with Lemma 4.6, we have $W^+(\lambda) \geq W(\lambda) \geq W^-(\lambda)$ for all $\lambda \in (0, \infty)$ with (see Fig. 7)

$$W^+(\lambda) = \frac{1+\lambda}{2+\lambda}, \qquad W^-(\lambda) = \frac{1-\lambda+\lambda^2}{2+\lambda^2}. \tag{4.28}$$

This allows for better control on Q_k and hence better control on R_k. However, the formulas are cumbersome to work with and do not lead to a sharp condition anyway. □

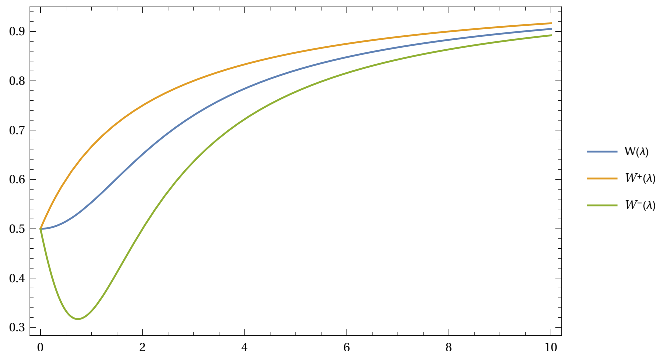


Figure 7: Bounding functions for $W(\lambda)$.

4.3 Disordered system

In this section we analyze what has to be modified in Section 4.2 to prove the analogues of Theorem 2.3–2.4. Clearly, we need to find simple bounds for $V_\mu(\lambda)$.

The analogues of (4.9)–(4.10), proven in Appendix A.2, read

$$\begin{aligned} V_\mu(\lambda) &\sim \tfrac{1}{2} I_\mu \lambda, & \lambda \downarrow 0, \quad (4.29)\\ 1 - V_\mu(\lambda) &\sim \frac{1}{2\lambda}, & \lambda \to \infty, \quad (4.30)\end{aligned}$$

where I_μ is given by (2.17). Note that the scaling for small λ is modified by the disorder, while the scaling for large λ is not. The key is the following analogue of Lemma 4.6.

Lemma 4.9. *$V_{\delta_0}(\lambda) \geq V_\mu(\lambda) \geq V_{\delta_0}(I_\mu \lambda)$ for all $\lambda \in (0, \infty)$.* □

Proof. The upper bound is proven in Appendix A.1. The lower bound is verified *numerically* in Appendix B.2 for three different types of disorder. □

With the help of the bounds in Lemma 4.9 we can repeat the proof in Section 4.2. Indeed, since the bounds on V_μ are in terms of $V_{\delta_0} = V$, we can use Lemma 4.6 to obtain

$$V_+(\lambda) \geq V(\lambda) \geq V_\mu(\lambda) \geq V(I_\mu \lambda) \geq V_-(I_\mu \lambda). \tag{4.31}$$

With the help of this sandwich we can prove Theorems 2.7–2.8 in the same way as Theorems 2.3–2.4.

A Comparison of the systems with and without disorder

In Appendix A.1 we prove Theorem 2.9, which states that the synchronization level strictly decreases when the spread of the disorder strictly increases. In Appendix A.2 we prove the asymptotics in (4.29)–(4.30).

A.1 Monotonicity in the spread of the disorder

Proof. It follows from (1.16)–(1.17) and (4.1) that

$$V_\mu(\lambda) = \int_0^\infty \bar{\mu}(\mathrm{d}\omega)\, F(\omega, \lambda), \tag{A.1}$$

where $\bar{\mu}$ is the projection of μ on $[0, \infty)$ obtained by reflection (recall that μ is assumed to be symmetric), and

$$F(\omega, \lambda) = \frac{\int_0^{2\pi} \mathrm{d}\theta\, (\cos\theta)\, S(\theta, \omega, \lambda)}{\int_0^{2\pi} \mathrm{d}\theta\, S(\theta, \omega, \lambda)} = F(-\omega, \lambda). \tag{A.2}$$

with

$$S(\theta, \omega, \lambda) = \int_0^\theta \mathrm{d}u\, \mathrm{e}^{G_{\lambda,\omega}(\theta,u)} + \mathrm{e}^{4\pi\omega} \int_\theta^{2\pi} \mathrm{d}u\, \mathrm{e}^{G_{\lambda,\omega}(\theta,u)} \tag{A.3}$$

and

$$G_{\lambda,\omega}(\theta, u) = 2\omega(\theta - u) + \lambda(\cos\theta - \cos u). \tag{A.4}$$

Noting that

$$4\pi\omega + G_{\lambda,\omega}(\theta, u) = G_{\lambda,\omega}(\theta, u - 2\pi), \tag{A.5}$$

we can apply the transformation $u \to u - 2\pi$ in the second integral of (A.3) to get the expression

$$S(\theta, \omega, \lambda) = \int_{\theta-2\pi}^\theta \mathrm{d}u\, \mathrm{e}^{G_{\lambda,\omega}(\theta,u)}. \tag{A.6}$$

The following lemma implies that $V_\mu(\lambda) \le V_{\delta_0}(\lambda)$ for all $\lambda \in (0, \infty)$, as claimed in Lemma 4.9.

Lemma A.1. *For all $\lambda \in (0, \infty)$, the map $\omega \mapsto F(\omega, \lambda)$ is strictly decreasing on $[0, \infty)$.* □

Proof. We show that $\frac{\partial}{\partial\omega} F(\omega, \lambda) < 0$. By (A.2), this amounts to showing that $\Delta < 0$ with

$$\Delta = \int_0^{2\pi} \mathrm{d}\theta\, (\cos\theta) \int_0^{2\pi} \mathrm{d}\phi\, \left[\left(\frac{\partial}{\partial\omega} S(\theta, \omega, \lambda)\right) S(\phi, \omega, \lambda) - S(\theta, \omega, \lambda) \left(\frac{\partial}{\partial\omega} S(\phi, \omega, \lambda)\right)\right]. \tag{A.7}$$

It follows from (A.6) that

$$\frac{\partial}{\partial\omega} S(\theta, \omega, \lambda) = \int_{\theta-2\pi}^\theta \mathrm{d}u\, \left(\frac{\partial}{\partial\omega} G_{\lambda,\omega}(\theta, u)\right) \mathrm{e}^{G_{\lambda,\omega}(\theta,u)}. \tag{A.8}$$

Since $\frac{\partial}{\partial\omega} G_{\lambda,\omega}(\theta, u) = 2(\theta - u)$, substitution into (A.7) yields the expression

$$\Delta = \int_0^{2\pi} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \int_{\theta-2\pi}^\theta \mathrm{d}u \int_{\phi-2\pi}^\phi \mathrm{d}v\ \mathrm{e}^{G_{\lambda,\omega}(\theta,u)}\, \mathrm{e}^{G_{\lambda,\omega}(\phi,v)} \big[(\cos\theta) 2(\theta - u) - 2(\phi - v)\big]. \tag{A.9}$$

By the symmetry of the integrals, the term between square brackets may be replaced by

$$(\cos\theta)(\theta-u)+(\cos\phi)(\phi-v)-(\theta-u)-(\phi-v) = -(\theta-u)(1-\cos\theta)-(\phi-v)(1-\cos\phi). \quad \text{(A.10)}$$

Since this is manifestly ≤ 0, the claim follows. □

Lemma A.1 implies Theorem 2.9. Indeed, if $\bar{\mu}'$ is stochastically strictly larger than $\bar{\mu}$, then (A.1) and the fact that $\omega \mapsto F(\omega,\lambda)$ is strictly decreasing imply that $V_{\mu'}(\lambda) < V_{\mu}(\lambda)$ for all $\lambda \in (0,\infty)$, which via (4.2) implies the claim. □

A.2 Asymptotics

Proof. The asymptotics for $\lambda \downarrow 0$ in (4.29) is derived in Luçon [9]. The asymptotics for $\lambda \to \infty$ in (4.30) is obtained as follows.

Return to (A.6). The map $u \mapsto G_{\lambda,\omega}(\theta,u)$ achieves its maximum at u solving $\sin u = 2\omega/\lambda$ and $\cos u < 0$. For λ large this means that $\pi - u$ is close to $2\omega/\lambda$. Writing $u = \pi - \epsilon$ with $\epsilon = -2\omega/\lambda$ and expanding up to order ϵ^2, we get

$$G_{\lambda,\omega}(\theta,u) = 2\omega(\theta-\pi+\epsilon) + \lambda\cos\theta + \lambda - \frac{\lambda}{2}\epsilon^2 + O(\lambda\epsilon^4). \quad \text{(A.11)}$$

Writing $\theta = \pi - \delta$, we get

$$S(\theta,\omega,\lambda) = \begin{cases} \mathrm{e}^{4\pi\omega}\mathrm{e}^{-2\omega\delta+\lambda[1+\cos(\pi-\delta)]}\int_{-\pi}^{\delta}\mathrm{e}^{2\omega\epsilon-\frac{\lambda}{2}\epsilon^2+O(\lambda\epsilon^4)}\mathrm{d}\epsilon, & \delta>0,\\ \mathrm{e}^{-2\omega\delta+\lambda[1+\cos(\pi-\delta)]}\int_{\delta}^{\pi}\mathrm{e}^{2\omega\epsilon-\frac{\lambda}{2}\epsilon^2+O(\lambda\epsilon^4)}\mathrm{d}\epsilon, & \delta<0. \end{cases} \quad \text{(A.12)}$$

To perform the integrals over ϵ, we write

$$2\omega\epsilon - \frac{\lambda}{2}\epsilon^2 = -\frac{\lambda}{2}\Big(\epsilon - \frac{2\omega}{\lambda}\Big)^2 + \frac{2\omega^2}{\lambda}, \quad \text{(A.13)}$$

put $v = \epsilon - \frac{2\omega}{\lambda}$ and integrate over v, to obtain

$$S(\theta,\omega,\lambda) = \Big[\text{(i)}\,\mathrm{e}^{4\pi\omega}\,\mathbf{1}_{\{\theta<\pi\}} + \text{(ii)}\,\mathbf{1}_{\{\theta>\pi\}}\Big]\,\mathrm{e}^{-2\omega(\pi-\theta)+\lambda(1+\cos\theta)}. \quad \text{(A.14)}$$

with

$$\text{(i)} = \int_{-\pi}^{\delta}\mathrm{e}^{2\omega\epsilon-\frac{\lambda}{2}\epsilon^2+O(\lambda\epsilon^4)}\mathrm{d}\epsilon = \mathrm{e}^{\frac{2\omega^2}{\lambda}}\sqrt{\frac{\pi}{2\lambda}}\Big(\mathrm{erf}\Big[\frac{\lambda\delta-2\omega}{\sqrt{2\lambda}}\Big] + \mathrm{erf}\Big[\frac{\pi\lambda+2\omega}{\sqrt{2\lambda}}\Big]\Big)\big[1+O(1/\lambda)\big], \quad \text{(A.15)}$$

$$\text{(ii)} = \int_{\delta}^{\pi}\mathrm{e}^{2\omega\epsilon-\frac{\lambda}{2}\epsilon^2+O(\lambda\epsilon^4)}\mathrm{d}\epsilon = \mathrm{e}^{\frac{2\omega^2}{\lambda}}\sqrt{\frac{\pi}{2\lambda}}\Big(\mathrm{erf}\Big[\frac{\pi\lambda-2\omega}{\sqrt{2\lambda}}\Big] - \mathrm{erf}\Big[\frac{\lambda\delta-2\omega}{\sqrt{2\lambda}}\Big]\Big)\big[1+O(1/\lambda)\big]. \quad \text{(A.16)}$$

We next integrate $(\cos\theta)\,S(\theta,\omega,\lambda)$ over θ. The map $\theta \mapsto -2\omega(\pi-\theta)+\lambda(1+\cos\theta)$ achieves its maximum when $\sin\theta = 2\omega/\lambda$ and $\cos\theta > 0$. For large λ this means that θ is close to $2\omega/\lambda$. Expanding around $\theta = 0$, we get

$$\frac{\int_0^{2\pi}\mathrm{d}\theta\,(\cos\theta)\,S(\theta,\omega,\lambda)}{\int_0^{2\pi}\mathrm{d}\theta\,S(\theta,\omega,\lambda)} = 1 + \frac{-\frac{1}{2}\int_0^{2\pi}\mathrm{d}\theta\,[\theta^2+O(\theta^4)]\,S(\theta,\omega,\lambda)}{\int_0^{2\pi}\mathrm{d}\theta\,S(\theta,\omega,\lambda)}. \quad \text{(A.17)}$$

To find the term of order $1/\lambda$, the asymptotics we are after, we calculate the numerator and denominator separately. The numerator equals

$$-\tfrac{1}{2}\int_0^{2\pi} \mathrm{d}\theta\, \theta^2\, S(\theta,\omega,\lambda) = -\tfrac{1}{2}\int_0^{2\pi} \mathrm{d}\theta\, \theta^2\, (\mathrm{i})\, \mathrm{e}^{4\pi\omega}\, \mathrm{e}^{-2\omega(\pi-\theta)+\lambda(1+\cos\theta)} \tag{A.18}$$

$$= -[1+O(\mathrm{e}^{-2\lambda})]\, \tfrac{1}{2}(\mathrm{i})\, \mathrm{e}^{2\pi\omega+2\lambda} \int_0^{2\pi} \mathrm{d}\theta\, \theta^2\, \mathrm{e}^{2\omega\theta-\frac{\lambda}{2}\theta^2+O(\lambda\theta^4)}, \tag{A.19}$$

where we may pull (i) in front because the main contribution comes from $\theta = \pi - \delta$ small, namely, order $1/\sqrt{\lambda}$. Similarly, the denominator equals

$$[1+O(\mathrm{e}^{-2\lambda})]\, (\mathrm{i})\, \mathrm{e}^{2\pi\omega+2\lambda} \int_0^{2\pi} \mathrm{d}\theta\, \mathrm{e}^{2\omega\theta-\frac{\lambda}{2}\theta^2+O(\lambda\theta^4)}. \tag{A.20}$$

The prefactor cancels in the quotient, which leaves us with

$$\frac{-\frac{1}{2}\int_0^{2\pi} \mathrm{d}\theta\, \theta^2\, \mathrm{e}^{2\omega\theta-\frac{\lambda}{2}\theta^2}}{\int_0^{2\pi} \mathrm{d}\theta\, \mathrm{e}^{2\omega\theta-\frac{\lambda}{2}\theta^2}}\, [1+O(1/\lambda)]\,. \tag{A.21}$$

The quotient equals

$$-\frac{1}{2}\left(\frac{\frac{2\omega}{\lambda^2}-2\lambda^{-2}\,\mathrm{e}^{-2\pi^2\lambda+4\pi\omega}(\pi\lambda+\omega)+\frac{1}{2\lambda^{5/2}}\,\mathrm{e}^{\frac{2\omega^2}{\lambda}}\sqrt{2\pi}(\lambda+4\omega^2)\Big(\mathrm{erf}\Big[\frac{\sqrt{2}(\pi\lambda-\omega)}{\sqrt{\lambda}}\Big]+\mathrm{erf}\Big[\frac{\sqrt{2}\omega}{\sqrt{\lambda}}\Big]\Big)}{\mathrm{e}^{\frac{2\omega^2}{\lambda}}\sqrt{\frac{\pi}{2\lambda}}\Big(\mathrm{erf}\Big[\frac{\sqrt{2}(\pi\lambda-\omega)}{\sqrt{\lambda}}\Big]+\mathrm{erf}\Big[\frac{\sqrt{2}\omega}{\sqrt{\lambda}}\Big]\Big)}\right), \tag{A.22}$$

which equals $-\frac{1}{2\lambda}+O(1/\lambda^2)$. □

B Numerical analysis

In Appendix B.1 we numerically compute the iterates of the renormalization map for the non-disordered system for two specific choices of $(K_k)_{k\in\mathbb{N}}$, belonging to universality classes (1) and (3), respectively. In Appendix B.2 we numerically compare the bounds in Lemma 4.9 for three types of disorder and show that they are asymptotically sharp.

B.1 Non-disordered system: renormalization

In Fig. 8 we show an example in universality class (1): $K_k = \frac{3}{2\log 2}\log(k+1)$. Synchronization is lost at level $k=4$. When we calculate the sum that appears in our sufficient criterion for universality class (1), stated in Theorem 2.3, up to level $k=4$, we find that

$$\sum_{k=1}^{4} \frac{2\log 2}{3\log(k+1)} = 1.70774. \tag{B.1}$$

This does not exceed 4, which shows that our sufficient criterion is not tight. It only gives us an upper bound for the level above which synchronization is lost for sure (recall (2.10)), although it may be lost earlier.

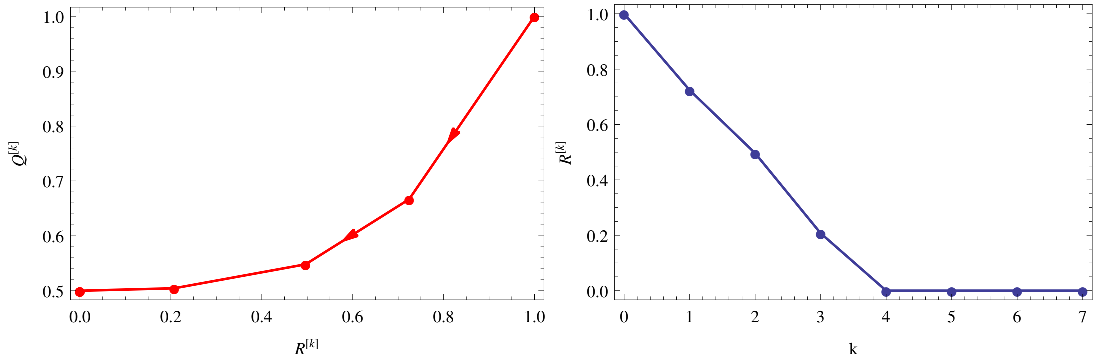


Figure 8: A plot of the renormalization map $(R^{[k]}, Q^{[k]})$ for $k = 0, \ldots, 7$ (left) and the corresponding values of $R^{[k]}$ (right) for the choice $K_k = \frac{3}{2\log 2}\log(k+1)$.

In Fig. 9 we show an example of universality class (3), where $K_k = 4\,\mathrm{e}^k$. There is synchronization at all levels. To check our sufficient criterion we calculate the sum

$$\sum_{k\in\mathbb{N}} \frac{1}{4\,\mathrm{e}^k} \approx 0.145494 < \frac{1}{\sqrt{2}} \approx 0.7071. \tag{B.2}$$

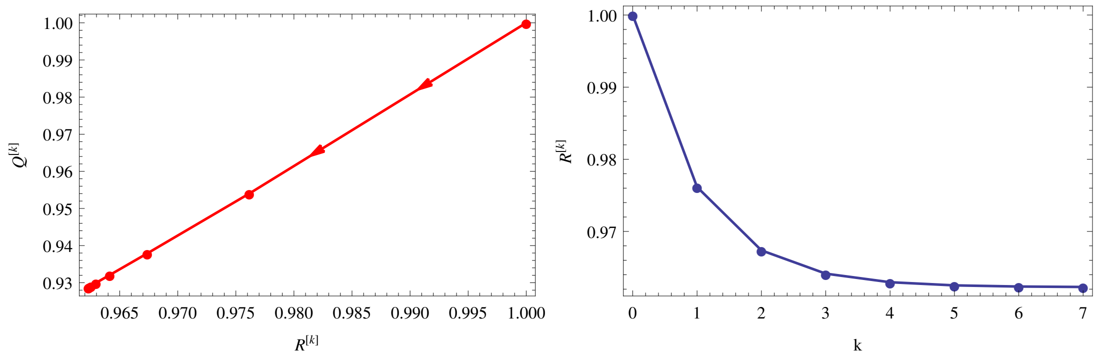


Figure 9: A plot of the renormalization map $(R^{[k]}, Q^{[k]})$ for $k = 0, \ldots, 7$ (left) and the corresponding values of $R^{[k]}$ (right) for the choice $K_k = 4\,\mathrm{e}^k$.

To find a sequence $(K_k)_{k\in\mathbb{N}}$ for universality class (2) is difficult because we do not know the precise criterion for criticality. An artificial way of producing such a sequence is to calculate the critical interaction strength at each hierarchical level and taking the next interaction strength to be 1 larger.

B.2 Disordered system: bounds

In this section we show that the bounds for $V_\mu(\lambda)$ in Lemma 4.9 are satisfied for three types of disorder and are sharp in the limit as $\lambda \downarrow 0$ and $\lambda \to \infty$, respectively,

Gaussian disorder

In Fig. 10 we take the disorder to be Gaussian with mean zero and standard deviation $\sigma = \frac{1}{2}$ and $\sigma = 3$. The numerics show that both the upper and the lower bound are correct.

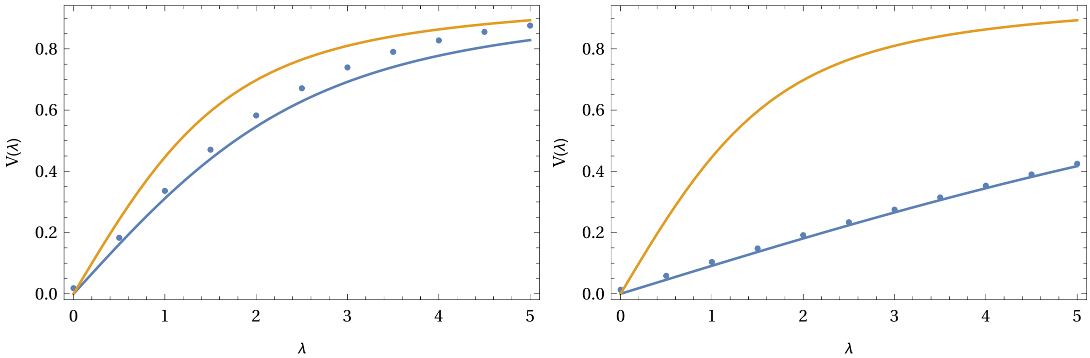


Figure 10: $V_\mu(\lambda)$ (dotted), $V_{\delta_0}(\lambda)$ (yellow) and $V_{\delta_0}(I_\mu \lambda)$ (blue) for Gaussian disorder with $\sigma = \frac{1}{2}$ (left) and $\sigma = 3$ (right).

In Fig. 11 the second plot in Fig. 10 is continued for large λ.

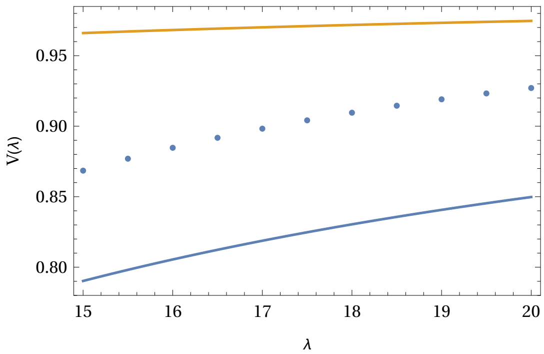


Figure 11: $V_\mu(\lambda)$ (dotted), $V_{\delta_0}(\lambda)$ (yellow) and $V_{\delta_0}(I_\mu \lambda)$ (blue) for Gaussian disorder with $\sigma = 3$.

Tent disorder

Here we take the disorder to be

$$\tilde{\mu}_{(}\omega) = 2\left(-\frac{\omega}{A^2} + \frac{1}{A}\right), \qquad \omega \in [0, A]. \tag{B.3}$$

Fig. 12 provides numerics for $A = 1$.

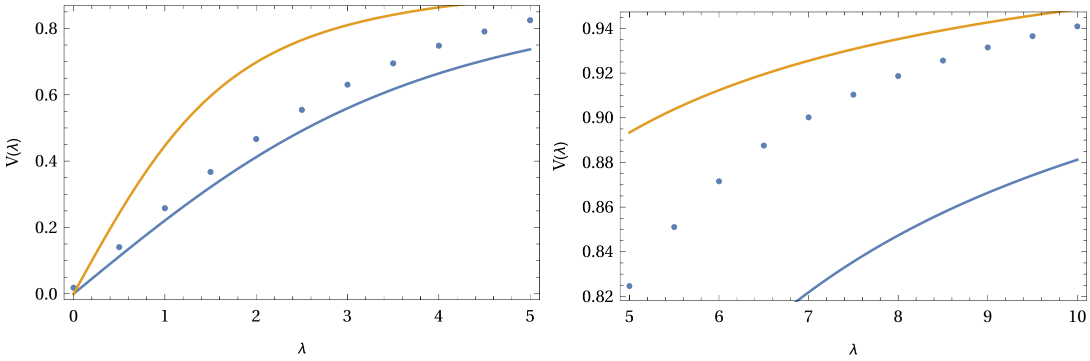


Figure 12: $V_\mu(\lambda)$ (dotted), $V_{\delta_0}(\lambda)$ (yellow) and $V_{\delta_0}(I_\mu\lambda)$ (blue) for tent disorder with $A = 1$.

Wigner semicircle disorder

Here we take the disorder to be

$$\tilde{\mu}(\omega) = \frac{4}{\pi B^2}\sqrt{B^2 - \omega^2}, \qquad \omega \in [0, B]. \tag{B.4}$$

Fig. 13 provides numerics for $B = 1$.

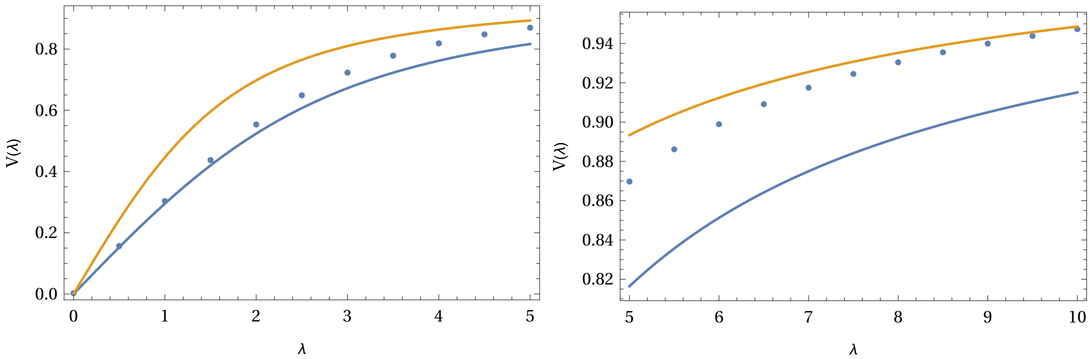


Figure 13: $V_\mu(\lambda)$ (dotted), $V_{\delta_0}(\lambda)$ (yellow) and $V_{\delta_0}(I_\mu\lambda)$ (blue) for Wigner semicircle disorder with $B = 1$.

References

[1] J.A. Acebrón, L.L. Bonilla, C.J. Pérez Vicente, F. Ritort and R. Spigler, The Kuramoto model: A simple paradigm for synchronization phenomena, Rev. Mod. Phys. 77 (2005) 137–185.

[2] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno and C. Zhou, Synchronization in complex networks, Physics Reports 469 (2008) 93–153.

[3] P. Dai Pra and F. den Hollander, McKean-Vlasov limit for interacting random processes in random media, J. Stat. Phys. 84 (1996) 735–772.

[4] S. Ha and M. Slemrod, A fast-slow dynamical systems theory for the Kuramoto type phase model, J. Differ. Equ. 251 (2011) 2685–2695.

[5] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics 113 (2nd. ed.), Springer, New York, 1998.

[6] Y. Kuramoto, Self-entrainment of a population of coupled non-linear oscillators. In: International Symposium on Mathematical Problems in Theoretical Physics, pp. 420–422. Lecture Notes in Phys. 39, Springer, Berlin, 1975.

[7] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer, New York, 1984.

[8] A. Laforgia and P. Natalini, Some inequalities for modified bessel functions, J. Ineq. App. 2010, article 253035 (2010) 1–10.

[9] E. Lucon, *Oscillateurs couplés, désordre et renormalization*, PhD thesis, Université Pierre et Marie Curie-Paris VI, 2012.

[10] H. Sakaguchi, Cooperative phenomena in coupled oscillators systems under external fields, Prog. Theor. Phys. 79 (1988) 39–46.

[11] J. Segura, Bounds for ratios of modified Bessel functions and associated Turin-type inequalities, J. Math. An. App. 372 (2011) 516–528.

[12] B. Sonnenschein and L. Schimansky-Geier, Approximate solution to the stochastic Kuramoto model, Phys. Rev. E 88, 052111 (2013) 1–5.

[13] S.H. Strogatz, From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators, Phys. D 143 (2000) 1–20.

[14] S.H. Strogatz and R.E. Morillo, Stability of incoherence in a population of coupled oscillators, J. Stat. Phys. 63 (1991) 613–635.

[15] S.H. Strogatz, R.E. Morillo and P.C. Matthews, Coupled nonlinear oscillators below the synchronization threshold: relaxation by generalized Landau damping, Phys. Rev. Lett. 68 (1992) 2730–2733.

[16] A.T. Winfree, Biological rhythms and the behavior of populations of coupled oscillators, J. Theor. Biol. 16 (1967)15–42.

[17] A.T. Winfree, *The Geometry of Biological Time*, Springer, New York, 1980.