Heterogeneous nucleation in finite size adaptive dynamical networks

Jan Fialkowski[1],[*] Serhiy Yanchuk[2],[3], Igor M. Sokolov[4], Eckehard Schöll[1],[2],[5], Georg A. Gottwald[6], and Rico Berner[1],[4][†]

[1]*Institute for Theoretical Physics, Technische Universität Berlin, Hardenbergstr. 36, 10623 Berlin, Germany*
[2]*Potsdam Institute for Climate Impact Research, Telegrafenberg A 31, 14473 Potsdam, Germany*
[3]*Institute of Mathematics, Humboldt University Berlin, 12489 Berlin, Germany*
[4]*Department of Physics, Humboldt Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany*
[5]*Bernstein Center for Computational Neuroscience Berlin,*
Humboldt Universität, 10115 Berlin, Germany and
[6]*School of Mathematics and Statistics, University of Sydney, Camperdown NSW 2006, Australia*
(Dated: July 6, 2022)

Phase transitions in equilibrium and nonequilibrium systems play a major role in the natural sciences. In dynamical networks, phase transitions organize qualitative changes in the collective behavior of coupled dynamical units. Adaptive dynamical networks feature a connectivity structure that changes over time, co-evolving with the nodes' dynamical state. In this Letter, we show the emergence of two distinct first-order nonequilibrium phase transitions in a finite size adaptive network of heterogeneous phase oscillators. Depending on the nature of defects in the internal frequency distribution, we observe either an abrupt single-step transition to full synchronization or a more gradual multi-step transition. This observation has a striking resemblance to heterogeneous nucleation. We develop a mean-field approach to study the interplay between adaptivity and nodal heterogeneity and describe the dynamics of multicluster states and their role in determining the character of the phase transition. Our work provides a theoretical framework for studying the interplay between adaptivity and nodal heterogeneity.

Phase transitions play an important role in the natural sciences [1]. Complex dynamical networks [2, 3] exhibit a plethora of nonequilibrium phase transitions organizing their collective dynamics in response to variations in control parameters such as interaction strength or noise [4, 5]. In particular, transitions between coherence and incoherence have attracted significant attention in static [6] and temporally evolving complex networks [7]. The Kuramoto model [8] has served as a testbed to study phase transitions in networks of coupled oscillators. It exhibits either first or second-order phase transitions from incoherence to full synchronization, depending on the natural frequency distribution [8–13]. Similarly, the network structure [14] and the weight distribution [15] can have strong effects on the nature of the synchronization transitions leading to first-order transitions and hysteresis.

To better describe real-world phenomena, the original Kuramoto model has been extended and modified. Beyond the classical Kuramoto model, generalizations to static and time-evolving networks have been developed [7, 16–21]. The inclusion of additional dynamical degrees of freedom, e.g. to describe power grids [22–26] in terms of the Kuramoto model with inertia, has introduced much richer synchronization transitions with regimes of coexisting cluster states. Recently, adaptive dynamical network models were introduced which are capable of describing chemical [27, 28], epidemic [29], biological [30], neurological [31–33], transport [34], and social systems [35, 36]. Adaptive dynamical networks are characterized by the coevolution of network structure and functionality. A paradigmatic example of adaptively coupled phase oscillators has recently attracted much atten-

tion [37–42], and has shown promise in predicting and describing phenomena in more realistic and complex physical systems such as neuronal and biological systems [43–46], as well as power grid models [47]. However, the type and nature of phase transition in this important class of models remains unclear.

This work contributes to understanding nonequilibrium phase transitions in adaptive networks and their characterisation. We study phase transitions in a finite-size Kuramoto model equipped with adaptive coupling weights. The natural frequencies are considered to be uniformly distributed, i.e., all possible frequencies are equally probable, and therefore disorders induced by finite size realizations of this distribution have a direct impact on the synchronization behavior. We find two qualitatively distinct types of first-order transitions to synchrony akin to the first-order transition phenomenon of heterogeneous nucleation [48]. The first, *multi-step* type of synchronization transition is characterized by the nucleation and growth of a dominant cluster, similar to Ostwald ripening in equilibrium and nonequilibrium systems [49], until the system reaches synchrony. The second, *single-step* transition type features multiple stable synchronization nuclei and the transitions to full synchrony is caused by an abrupt merging of large clusters of similar size. These two paths to synchrony exhibit a high degree of multistability. We identify the location of fluctuations in the realization of the natural frequency distribution as the cause for the two different scenarios. Methodologically, we present a framework to reduce high dimensional adaptive networks to a few mesoscopic variables and show that a collection of partially synchronized clusters of approximately equal size is more stable with

respect to changes in the coupling strength. With this, we extend the scope of mean-field approaches, which have been so far used only for static networks, to adaptive networks.

We consider the adaptively coupled phase oscillator model

$$\frac{d\phi_i}{dt} = \omega_i - \frac{\sigma}{N} \sum_{j=1}^{N} \kappa_{ij} \sin(\phi_i - \phi_j), \tag{1}$$

$$\frac{d\kappa_{ij}}{dt} = -\epsilon \left(\kappa_{ij} + \sin(\phi_i - \phi_j + \beta) \right), \tag{2}$$

for N phase oscillators $i = 1, ..., N$ with phases $\phi_i(t) \in [0, 2\pi)$ coevolving with adaptive coupling weights $\kappa_{ij}(t)$. The natural frequency ω_i of the ith oscillator is drawn randomly from a uniform distribution $\omega_i \in [-\hat{\omega}, \hat{\omega}]$ [73]. The overall coupling strength is σ, the time-scale separation between the fast oscillator dynamics and the slower adaptation of the coupling weights is denoted by ϵ, and β is a tuning parameter for the adaptivity function [51, 52].

Coherence can be quantified by the synchronization index S which measures the fraction of frequency-synchronized oscillator pairs $S := \frac{1}{N^2} \sum_{i,j=1}^{N} s_{ij}$, where $s_{ij} = 1$ for equal mean phase velocities of the oscillators i and j, $\langle \dot{\phi}_i \rangle = \langle \dot{\phi}_j \rangle$, and $s_{ij} = 0$ otherwise [74]. Here $\langle x \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{T_0}^{T_0+T} x(t) dt$ with sufficiently large transient time T_0. For $S = 1$ the system is fully frequency-synchronized, whereas for $S = 0$ the system is asynchronous.

With increasing coupling strength σ, system (6)–(7) undergoes a transition from asynchrony to full synchronization, see Fig. 1(a) [75]. The routes to synchrony with increasing coupling strength σ in Fig. 1(a) follow two paths of first-order transitions: a more gradual multi-step (upper, light yellow) and a more abrupt single-step (lower, dark green) path. In what follows, we describe these paths and determine the finite size features in the realization of the natural frequency distribution that lead to either a multi-step or single-step transition to synchrony.

Figures 1(b) and 1(c) show the multi-step and single-step transitions, respectively, for two representative realizations of the natural frequencies (displayed as insets) and 30 different initial conditions of the phases $\phi_i(0)$. Depending on initial conditions and natural frequencies the system can develop a large number of coexisting states. The multi-step path in Fig. 1(b) exhibits a higher degree of multistability even for the fixed realization of frequencies and a large number of transitions between coexisting states. Figs. 2(a)–(d) show snapshots of the coupling matrix corresponding to the multi-step transition. It is seen that a cluster nucleus emerges and, upon increasing the coupling strength, entrains more and more oscillators leading to the fully frequency synchronized state in Fig. 2(d), similarly as observed in [22]. Depending on the initial conditions the exact structure of the system's
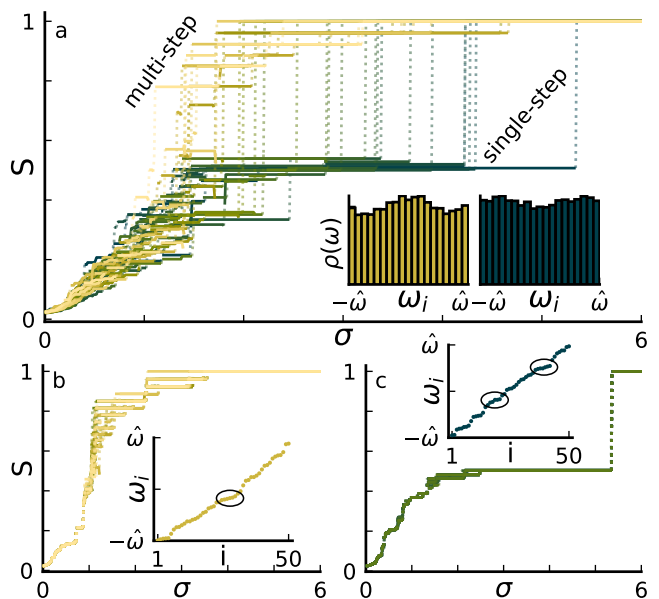

FIG. 1. Paths to synchrony for system (6)–(7). (a) Synchronization index S as a function of the coupling strength σ for 100 simulations with $N = 50$ oscillators. Each run was initiated with random initial conditions and σ was increased in steps of $\Delta\sigma = 0.01$. For further details on the up-sweep protocol, we refer to [56]. For each run, natural frequencies are drawn independently from a uniform distribution $\omega_i \in [\hat{\omega}, \hat{\omega}]$ with $\hat{\omega} = 0.25$. The dotted lines indicate jumps in the synchronization index during the transition. The inset shows realizations of the finite-size frequency distributions for the multi-step (left) and single-step (right) paths generated from 1000 simulations. (b) and (c): Synchronization index S for fixed realization of the natural frequencies and 30 different initial conditions, consistently leading to the multi-step (b) or single-step transition (c). Insets show the natural frequencies used in the simulations. The circles highlight areas of higher frequency densities. The synchronization index S is determined with an averaging time window $T = 7 \cdot 10^3$ and transient time $T_0 = 3 \cdot 10^3$. Other parameters: $\beta = -0.53\pi$, $\epsilon = 0.01$.

state can vary, e.g. the size of the nucleus for different σ. This gives rise to a multitude of stable states for most values of σ.

The single-step transition is shown in Fig. 1(c). Regardless of the initial condition, the dynamics follows very similar paths of the synchronization index. The phase oscillators organize into a small number of clusters within each of which all oscillators move with the same mean phase velocity for small values of σ. Notably, there is no single cluster at any time that is significantly larger than the others, in contrast to the multi-step branch and prior findings for multicluster states [39]. Further, for the single-step phase transition path an intermediate state with $S \approx 0.5$ emerges, which is stable for a wide range of coupling strengths σ. The transition to full frequency synchronization occurs discontinuously at $\sigma \approx 5.35$. The formation of small initial clusters and the intermediate

state are shown in Figs. 2(e)–(h). The intermediate state (g) consists of two almost evenly sized clusters (nuclei) that are formed simultaneously.

Whether the system undergoes a multi-step or a single-step transition path is determined by the particular realization of the natural frequency distribution, see the insets in Fig 1(a). Frequency distributions corresponding to multi-step phase transitions are characterized by a higher density around the average frequency $\bar{\omega} = 0$, leading to an initial cluster with average cluster frequency close to the overall average frequency of the network. On the other hand, frequency distributions corresponding to single-step transitions are characterized by deviations which are concentrated away from the average frequency $\omega = 0$. This leads to local clusters emerging at low coupling strength around those non-zero seeds. These clusters survive a further increase in coupling strength entraining further oscillators before for larger coupling strength collapsing into the fully synchronized state with mean phase velocity close to 0.

The distinction between these two qualitatively very different scenarios of phase transitions is hence caused by finite size induced inhomogeneities of the natural frequency distribution. The fluctuations in the realization of the natural frequencies dominates the influence of the choice of initial conditions, see Fig 1(b) and (c). To further probe that the choice of transition path is indeed a finite-size effect, we study equidistant draws from the frequency distribution, to mimic the thermodynamic limit of infinite network size. In this case the transition path which is taken by the system is determined instead by the initial condition of the phases, see [56].

Performing a down-sweep from $\sigma = 6$ to $\sigma = 0$ for the given frequency distributions in Fig. 1(b) and (c), hysteresis and bistability between full and partial cluster synchrony is observed (see [56] for details).

During the transition to synchrony, oscillators group into frequency locked clusters. The fluctuations in realization of the natural frequencies determine the shapes of the emerging cluster states, which in turn govern the type of the transition. In systems with adaptive coupling weights there are many ways in which an arbitrary number of oscillators can form a multicluster structure [39, 57]. In the following, we study the coexistence of multicluster states and show that the states with equally sized clusters are stable for a larger interval in σ than the states with strongly different cluster sizes.

We develop an approach that reduces the problem of multicluster states from $N + N^2$ dimensions to $2M - 1 + M^2$ dimensions, where M is the number of clusters. Our approach uses the collective coordinate method first introduced in [58]. This method successfully describes the synchronization of the Kuramoto system for general, in particular multimodal [59], frequency distributions, with complex coupling topology [60] or chaotic cluster dynamics. In contrast to other well known reduction meth-
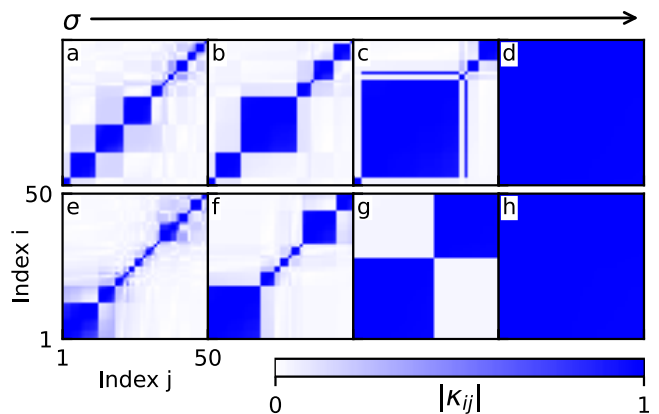

FIG. 2. Coupling matrices for specific values of σ, corresponding to two different types of synchronization transition. (a)–(d): multi-step, and (e)–(h) single-step transition. The values of σ are (a) 0.9, (b) 1.05, (c) 1.75, (d) 6.0 and (e) 0.55, (f) 1.75 (g) 2.75 and (h) 6.0. The snapshots are taken after 10^4 time units. Other parameters: $\hat{\omega} = 0.25$, $\beta = -0.53\pi$ and $\epsilon = 0.01$. Due to the $\phi \to \phi + \pi$ symmetry, κ_{ij} and $-\kappa_{ij}$ are indistinguishable, therefore the absolute values $|\kappa_{ij}|$ are plotted, see [56].

ods such as Watanabe-Strogatz reduction [61] and Ott-Antonsen-Ansatz [62], this approach is able to capture finite-size effects while still reproducing the findings of mean-field theories in the $N \to \infty$ limit for a well chosen ansatz [63].

Motivated by [59], we consider the following cluster ansatz:

$$\phi_i \approx \hat{\phi}_i^\mu = \vartheta_\mu(\omega_i - \Omega_\mu) + f_\mu, \qquad (3)$$

$$\kappa_{ij} \approx \hat{\kappa}_{ij}^{\mu\nu} = \varkappa_{\mu\nu}. \qquad (4)$$

Here, we split up the dynamical description for each phase oscillator into two coordinates. The coordinates $\vartheta_\mu(t)$ and f_μ describe the spread of oscillators within the μth cluster with relative frequencies $(\omega_i - \Omega_\mu)$ and the collective phase of each cluster, respectively. Similarly, we reduce the coupling weights κ_{ij} to an inter-cluster coupling $\varkappa_{\mu\nu}$ between clusters μ and ν. We assign the index i of each oscillator to a cluster μ of size $N_\mu = n_\mu N$ and with the set of indices $\mathcal{C}_\mu$; $\Omega_\mu = N_\mu^{-1} \sum_{i \in \mathcal{C}_\mu} \omega_i$ is the mean natural frequency of cluster μ. This approach changes the microscopic description of each oscillator to a mesoscopic description for the clusters with the new collective coordinates ϑ_μ, f_μ, and $\varkappa_{\mu\nu}$ [76].

The error of the collective coordinate approximation (8)–(9) is minimized by ensuring that the error is orthogonal to the space spanned by the ansatz functions, see [58, 63]. This Galerkin approximation determines the equations of motion for each of the collective coordinates. For simplicity, we restrict ourselves here to the description of two clusters [77]. A common measure of coherence within a population of phase oscilla-
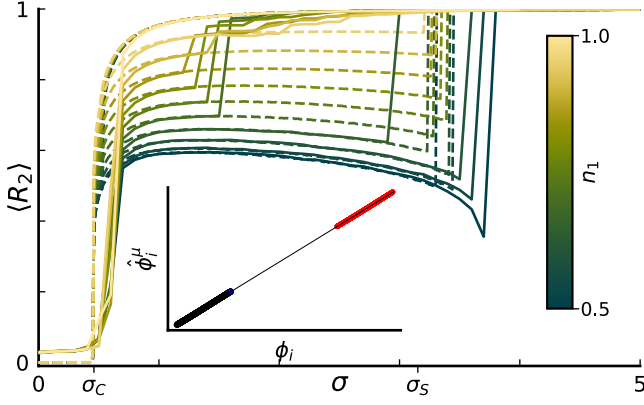
FIG. 3. Bifurcation diagram for the adaptive Kuramoto model (6)–(7) (full system) with $N = 1000$ oscillators (solid lines) and the reduced system (13) (dashed lines), showing the time average (over $T = 1000$) of the second Kuramoto-Daido order parameter R_2 as a function of the coupling strength σ. For the full system a two-cluster state is initially prepared with a fraction of n_1 oscillators placed in the first cluster, and $n_2 = 1 - n_1$ oscillators in the second cluster at $\sigma = 1$, for the reduced system at $\sigma = 0.5$ with values of n_1 from 0.5 to 0.95 in steps of $\Delta n_1 = 0.05$. Additionally the full system is simulated for $n_1 = 1$ to visualize hysteresis. The analytical estimates of upper and lower bounds of two-cluster existence at $\sigma_S = 3.152$ and $\sigma_C = 0.460$ are marked on the axis. Inset: oscillator phases ϕ_i for the full system vs. the approximation $\hat{\phi}_i^\mu = \vartheta_\mu \omega_i$ of the reduced system for $n_1 = 0.5$ and $\sigma = 2$; colors denote the two clusters. Other parameters: $\hat{\omega} = 0.25$, $\beta = -0.53\pi$ and $\epsilon = 0.01$.

tors is provided by the mth Kuramoto-Daido order parameter $R_m(t) = \frac{1}{N} \left| \sum_{j=1}^{N} e^{mi\phi_j(t)} \right|$ [8, 66]. Here, we need the first moment restricted to a single cluster μ: $r_\mu \equiv R_{m=1,\mu}(t)$. In the continuum limit $N \to \infty$ the resulting mesoscopic dynamics of system (6)–(7) is governed by

$$\dot{\vartheta}_\mu = 1 + \frac{\sigma}{v_\mu \vartheta_\mu} \left(\cos\left(\frac{\vartheta_\mu n_\mu}{4} \right) - r_\mu \right)$$
$$\times \left[\varkappa_{\mu\mu} n_\mu r_\mu + \varkappa_{\mu\nu} n_\nu r_\nu \cos f \right], \quad (5a)$$

$$\dot{f} = \Omega_{12} - \sigma r_1 r_2 \left(n_1 \kappa_{21} + n_2 \kappa_{12} \right) \sin f, \quad (5b)$$

$$\dot{\varkappa}_{\mu\nu} = -\epsilon \left[\varkappa_{\mu\nu} + r_\mu r_\nu \sin\left(f_\mu - f_\nu + \beta \right) \right]. \quad (5c)$$

Here, $\mu, \nu \in \{1, 2\}$ are the indices for the clusters, $f = f_1 - f_2$ is the phase difference of the two clusters, $\Omega_{12} = \Omega_1 - \Omega_2$, $v_\mu = \frac{1}{N_\mu} \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu)^2$ is the variance of the natural frequencies of a cluster. The first Kuramoto order parameter for a cluster is approximated by $r_\mu = \sin z / z$ with $z = \vartheta_\mu n_\mu / 4$. See [56] for details.

Figure 3 shows a comparison of the high-dimensional adaptive Kuramoto system (6)–(7) and the reduced system (13). We use $N = 1000$ oscillators with a fixed realization of natural frequencies. Two-cluster configurations are generated by varying the relative number of oscillators in the first cluster from $n_1 = 0.5$ to $n_1 = 0.95$

for which we prepare special initial conditions that result in the desired state. For the reduced system, we proceed analogously.

Figure 3 shows that the dynamics of the two-cluster state is captured very well by the collective coordinate framework. The solid lines (full system) overlap with the dashed lines (reduced system). The single-step first-order transition, also seen in Fig. 1(c), is well explained by the merging of two clusters in the reduced system. In both systems the multicluster structure ceases to exist beyond a certain coupling strength σ_c. The critical values for the onset of cluster synchronization σ_c and full synchrony σ_s are well approximated by a perturbative approach for $\epsilon \ll 1$ to the reduced system (13). We obtain $\sigma_c \approx 0.460$ and $\sigma_s \approx 3.152$, see [56]. In particular, the analytic result shows that multicluster states exist only for an intermediate range of σ, which agrees with the observations in Fig. 1.

In the inset of Fig. 3 the excellent agreement between the oscillator phases of the full system and the collective coordinate ansatz (8) is shown for $n_1 = 0.5$.

Fig. 3 shows that two-cluster states in the reduced system are stable for a much larger interval in σ than in the full system. This discrepancy may be linked to the observation that the stability of multicluster states is mainly determined by intracluster links [67]. Such intracluster effects are not captured by our mean-field ansatz. However, the reduced system provides important insights into the existence of partially synchronized clusters from which the stability can be studied numerically employing the full system. This is relevant to the two transition scenarios observed in Fig. 1: the reduced two-cluster system withstands full synchronization for larger values of the coupling strength the more equal the respective cluster sizes are with a maximum at $n_1 = n_2 = 0.5$. Hence a homogeneous collection of clusters of similar size will remain stable for a wide range of coupling strengths whereas heterogeneous nucleation with a dominant initial cluster will entrain further oscillators upon increasing the coupling strength.

In summary, we have shown two qualitatively different transitions to synchronization induced by the interplay of an adaptive network structure and finite size inhomogeneities in the natural frequency distributions: a single-step transition and a multi-step transition. In the multi-step transition, a single large cluster (nucleus) forms around an inhomogeneity in the frequency distribution and successively grows until full synchronization is reached. In contrast, in the single-step synchronization transition, multiple equally sized clusters (nuclei) form around multiple inhomogeneities, grow and coexist stably. Each cluster moves with its mean frequency, which results in a higher difference of the average phase velocity between the clusters than between two freely moving oscillators. This higher difference inhibits the synchronization of the clusters for a significant range in the coupling

strength during the transition, before full synchronization is reached abruptly in a first-order transition.

The described nucleation phenomena are very similar to heterogeneous nucleation induced by local impurities known, e.g., from cloud formation [68], crystal growth [69] or Ostwald ripening in equilibrium and nonequilibrium systems [49]. Due to this relation, our results provide an intriguing bridge between synchronization transitions in finite-size dynamical complex networks and thermodynamic phase transition. Our numerical investigation has been complemented by a mean-field theory capable of describing multicluster states in the presence of an arbitrary frequency distribution and adaptive coupling weights. By this, we contribute to the research on mean-field models of coupled phase oscillators [62] where only recently first steps have been undertaken to include adaptive coupling [70]. Remarkably, our reduced mean-field model provides an excellent approximation of the macroscopic as well as the microscopic multicluster dynamics. The multi-step transition with a continually changing size of the main cluster (nucleus) and the importance of the stability of each cluster is only partially captured by the mean-field approach introduced in this work. This limitation, however, could be overcome by generalizing methods on partial synchronization in pure phase oscillator systems [63] and systems with inertia [22, 71, 72].

Appendix

A: Numerical schemes

Integration schemes

For the numerical integration of the full Kuramoto system with $N = 1000$ oscillators, we use a procedure that combines a classical Runge-Kutta scheme for the oscillator phases with Euler's method for the coupling weights. We use a fixed $\Delta t = 0.01$ as the step size. Let u_i be the state vector of our system at time $t = i\Delta t$. In order to calculate $u_{i+1} = u_i + k$, with k being the vector of change of our state, we split the calculation into an oscillator k_ϕ and a coupling k_κ part. Since $\epsilon \ll 1$, the dynamics of the coupling is slower than the dynamics of the phases. For the slow dynamics of the coupling we employ a simple Euler scheme with

$$k_\kappa = \dot{k}_\kappa \left(u_i \right) \Delta t.$$

For the fast oscillator part k_ϕ we use a more accurate fourth-order Runge-Kutta scheme. Here $\dot{k}_\phi \left(u \right)$ corresponds to the equation of motion (2) of the main text. Using an Euler step for the coupling significantly speeds up the computation and allows for the extensive sweeping protocols over many realisations. This is justified since the dynamics is slow and the coupling can be assumed to be constant during the fine-scale Runge-Kutta steps for the fast oscillator dynamics.

For the numerical integration of the full Kuramoto system with $N = 50$ oscillators, the integration is performed using a fifth order Runge-Kutta method and the time step is chosen at every step by estimating the error via comparison with a Runge-Kutta of fourth order and enforcing its smallness.

Parameter sweeping protocols

For the parameter sweeping in Fig. 1 of the main text, we use different protocols depending on the type of analysis done. For most parts of our study, we consider systems of $N = 50$ adaptively coupled oscillators oscillators. Large systems with $N = 1000$ are used to show the accuracy of our mean-field approach for two-cluster states, however, the sweeping approach has been adjusted as described in the subsequent section.

Up-sweep protocol: For systems with $N = 50$ oscillators, we initiate each simulation with random initial conditions and natural frequencies at $\sigma = 0$. Then we simulate the system for a total of 10^4 time units. The last 7000 time units of those are used to calculate the synchronization index S, see definition in the main text. For the calculation of S, we evaluate if the mean phase velocities $\left\langle \dot{\phi}_i \right\rangle$ of the oscillators are the same allowing for a small difference of $\left\langle \dot{\phi}_i \right\rangle - \left\langle \dot{\phi}_j \right\rangle < 10^{-3}$. The total number of pairings for which this is true divided by the total number of oscillator pairs N^2 determines the synchronization index S. The final state ϕ_i reached after these 10^4 time units is used as the initial condition for the simulation at a value of the increased coupling strength $\sigma = \Delta\sigma = 0.01$. This process repeats during the up-sweep until $\sigma = 6$.

Down-sweep protocol: The down-sweep is performed in a similar manner, starting from the fully synchronized state at $\sigma = 6$, decreasing σ in steps of $\Delta\sigma = -0.01$ after every 10^4 time units and using the previously calculated state as the new initial condition. We analyse the down-sweep until $\sigma = 0$.

Parameter sweeping for two-cluster states

The scans through σ shown in Fig. 3 of the main text for systems with $N = 1000$ oscillators are achieved as follows. To make sure the analysis is performed for the desired two-cluster state with a particular n_1, the system is initiated with oscillators belonging to a cluster being grouped together and their intracluster coupling maximized to $\varkappa_{\mu\mu} = -\sin\beta$, whereas the intercluster coupling was minimized to $\varkappa_{\mu\nu} = 0$. The reduced system uses the

same initial values for the couplings as well as $\vartheta_\mu = 0.1$ for each cluster, with $f_\mu = 0$ or 0.3π respectively.

With the initial state prepared in the desired two-cluster state the simulation is started at an estimated value of σ that is close to the border of desynchronization. We choose $\sigma_0 = 0.5$ for the reduced system and $\sigma_0 = 1$ for the full system. With this, the protocols described above for sweeping through the values of σ are applied with a step size of $\Delta\sigma = 0.01$ for the reduced system and $\Delta\sigma = 0.1$ for the full system. Due to its reduced complexity and the ten times higher resolution in σ, the reduced system for the collective coordinates only requires a shorter transient time of $T_0 = 500$ time units before the order parameter is calculated from the simulation of further $T = 1500$ time units. The full phase oscillator system, Eqs.(1)–(2) of the main text, for $N = 1000$ oscillators requires more care in choosing the integration times. Let σ_J be the current value of σ of the simulation after J steps. First the system is run for a short time of $T_1 = 100$ time units and the order parameter is calculated over $T = 1000$ time units. After this we calculate the change in the order parameter compared to the previous value at σ_{J-1}, $\Delta R_2(J) = \langle R_2 \rangle (\sigma_J) - \langle R_2 \rangle (\sigma_{J-1})$. This change is compared with the change in the previous step $\Delta R_2(J-1)$. If $\Delta R_2(J) < 2\Delta R_2(J-1)$ we assume the state to be stable, proceed as outlined above and simulate the system at a higher value of $\sigma_{J+1} = \sigma_J + \Delta\sigma$. If on the other hand $\Delta R_2(J) \geq 2\Delta R_2(J-1)$, the system is assumed to undergo a change in its state and is simulated for a further $T_2 = 10^4$ time units to allow for relaxation towards its new equilibrium state before the order parameter is estimated. To ensure that the transition to the new state has not been missed due to the smaller transient times the previous state for σ_{J-1} is simulated again for T_2 time units. Should this repeated simulation reveal the state to be unstable already at σ_{J-1}, we use the newly simulated state as an initial condition to continue the sweep with σ_J.

B: Equiprobably drawn frequencies

To test if the occurrence of the two distinct scenarios is indeed caused by finite size inhomogeneities in the realization of the natural frequency distribution, we simulate the system with equiprobably drawn frequencies. In particular, we simulate (1)-(2) of the main text using an equidistant draw of the natural frequencies $\omega_i = -0.25 + 0.5(i-1)/(N-1)$ for $N = 50$ oscillators. Figure 4 shows the results of 30 simulations for different initial conditions. The system exhibits a high degree of multistability for different initial conditions leading to different cluster structures as σ is increased. Most of the 30 initial conditions follow the path of the single-step transition to synchrony, with two or three rather evenly sized clusters. For some initial conditions, however, we
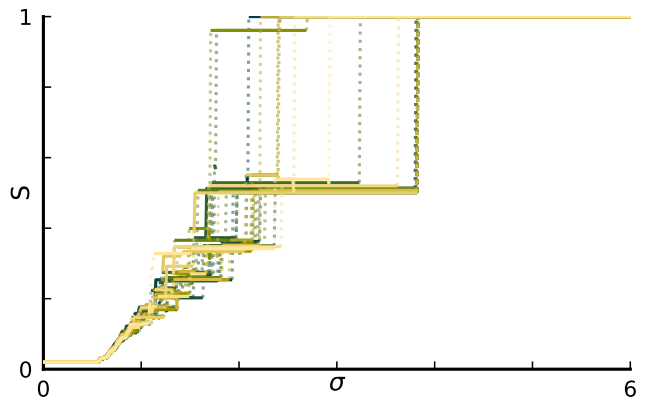

FIG. 4. The synchronization index S for increasing values of the coupling strength σ. The simulations use an equidistant distribution of natural frequencies $\omega_i = -0.25 + 0.5(i-1)/(N-1)$. Shown are 30 simulations from different random initial conditions for $N = 50$ oscillators each, where σ is varied from 0 to 6 in steps of size $\Delta\sigma = 0.01$. The protocol is the same as in Fig. 1(a) of the main text. Other parameters: $\beta = -0.53\pi$, $\epsilon = 0.01$.

observe the formation of a single dominant cluster with mean-frequency close to zero leading to the multi-step transition. In the absence of finite size fluctuations in the realization of the natural frequencies, fluctuations in the initial conditions lead to the differentiation in the synchronization process into the different paths. Hence, when approximating the continuum limit without nucleation seeds stemming from inhomogeneities in the realization of the natural frequency distribution, typically homogeneous nucleation takes place with a single-step transition.

C: Multimodal frequency distributions

In order to support the findings in the main text, we simulate the system (1)–(2) of the main text for $N = 50$ oscillators with different frequency distributions modelling the distributions displayed in the inset of Fig. 1(a) of the main text. To model different nucleation seeds we consider multimodal frequency distributions with each peak representing one nucleation site. In particular we consider a multimodal distribution consisting of M separated Gaussians of the form

$$\rho(\omega) = \sum_{m=1}^{M} \frac{1}{M} \mathcal{N}(\mu_m, s).$$

Here $\mathcal{N}(\mu_m, s)$ is a normalized Gaussian distribution with mean μ_m and standard deviation s.

The associated paths to synchronization are shown in Fig. 5. We examine distributions with up to $M = 3$ and choose the parameters for the Gaussian distributions in such a way that the peaks are clearly separated and decay
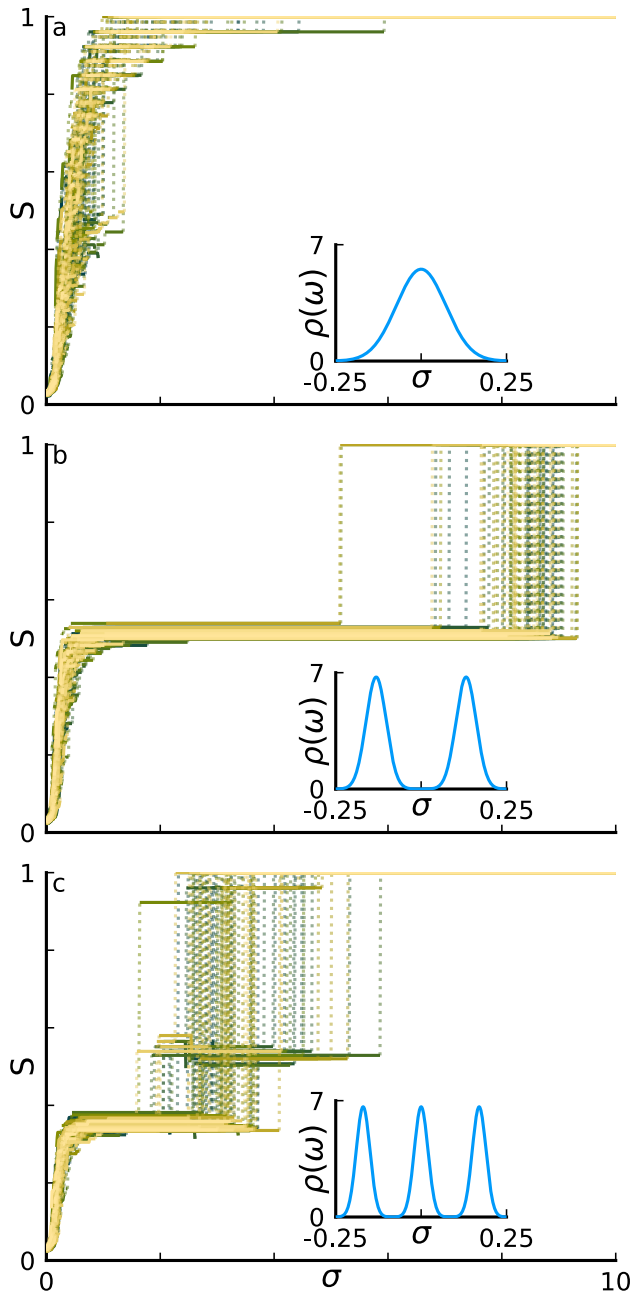
FIG. 5. Paths to synchrony for different multimodal Gaussian distributions. The procedure for these plots is the same as in Fig. 1 of the main text (up-sweep). (a) Unimodal Gaussian distribution with a standard deviation of $s = 0.0722$ and mean $\mu = 0$, corresponding to the multi-step transition with a single synchronization seed. (b) Bimodal Gaussian distribution of equal standard deviation $s = 0.0296$ with mean values of $\mu = \pm 0.132$. This distribution corresponds to the case of the single-step transition with two synchronization seeds. (c) Trimodal Gaussian distribution of equal standard deviation $s = 0.020$. One peak is centered at 0 whereas the other two have a mean of $\mu = \pm 0.170$, respectively. The insets show the respective underlying probability density function of the natural frequencies. Each panel consists of 100 runs with different realizations of the natural frequencies and initial conditions.

sufficiently fast within the interval $\omega \in [-0.25, 0.25]$. The distributions are shown as insets in each panel. Fig. 5 (a) shows the resulting paths for a single peak with a standard deviation of $s = 0.0722$. All realizations of the natural frequencies follow the path of the multi-step transition where a big cluster emerges around the mean of the frequency distribution that progressively grows by absorbing more of the outer oscillators. The case of two peaks with a standard deviation of $s = 0.0296$ is shown in Fig. 5 (b). One can see clearly that all realisations follow the path of the single-step transition. As explained in the main text this is due to the emergence of frequency clusters of roughly equal size with mean cluster phase velocities symmetrical around 0. These clusters merge abruptly at sufficiently large coupling strength σ forming a single cluster. This corresponds to the observation that synchronization nuclei form around inhomogeneities in the realization of the natural frequency distribution and determine the transition path as σ is increased. A single seed leads to the multi-step transition, whereas two seeds lead to the single-step transition. For a triple peak distribution, shown in fig. 5 (c), the synchronization process is more complex. It starts as expected with the creation of three frequency-synchronized clusters, but as the coupling strength is increased not all simulations show a sudden merging of all three clusters. In fact some of the simulations again show either the creation of a single dominant cluster leading to higher synchronization index S or the emergence of two big clusters that subsequently merge at even higher coupling strengths σ.

D: Hysteresis

In Figure 6 we show the same data as in Fig. 1 of the main text (up-sweep) as well as the behavior of the time averaged order parameter $\langle R_2 \rangle$ as σ is decreased via the down-sweep protocol. Additionally we show the analytical one-cluster solution derived from Eq. (27) as a dashed line. For a large range in σ, the synchronous states is shown to be co-stable with various multicluster states. The fully synchronized state is stable until small values of σ, where both down-sweep simulations show a sudden and sharp decline in their order parameter below a critical value of σ. As soon as the fully frequency synchronized state disappears most of the other coexisting multicluster states disappear as well, compare with Fig. 3 of the main text. The system abruptly approaches an asynchronous state. The hysteretic behavior presented in 6 is a common feature of first-order phase transitions. The analytical solution predicts the order parameter and the border to desynchronization very well, with a small error towards higher order parameters for lower values of σ. We attribute this error to the assumption of infinitely many oscillators used in the derivation of the analytic solution.
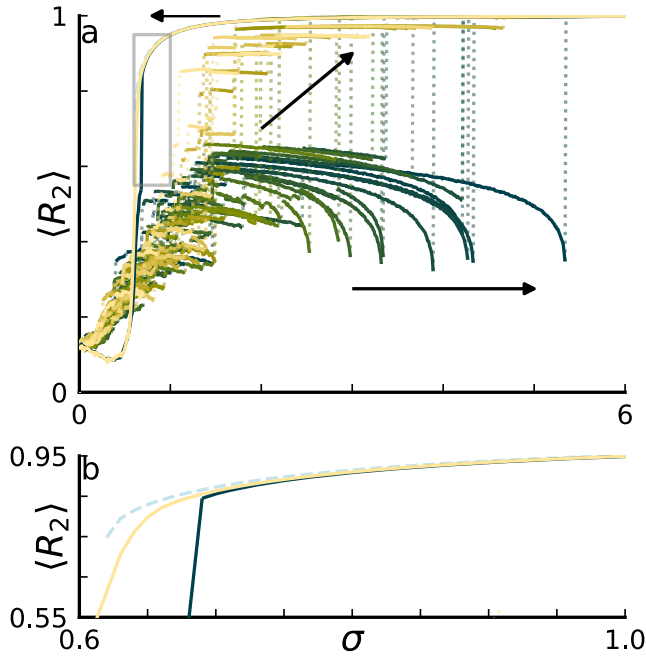
FIG. 6. Hysteric transition in system (6)–(7) with $N = 50$ oscillators. (a) Up-sweep: Shown are the time averaged second moment of the order parameter $\langle R_2 \rangle$ as a function of the coupling strength σ for 100 simulations with $N = 50$ oscillators. Each run is initiated with random initial conditions and σ is increased in steps of $\Delta\sigma = 0.01$. For each run natural frequencies are drawn independently from a uniform distribution $\omega_i \in [-0.25, 0.25]$. Down-sweep: Additionally shown are the synchronization indices for two simulations with decreasing σ starting from the final state of the up-sweep protocol. The natural frequencies of these down-sweeps are chosen such that one corresponds to the single-step transition path and one to the multi-step transition path. The light-blue dashed line corresponds to the stable branch of the analytical solution of Eq. (27). (b) We provide a blow-up for the region depicted as a grey rectangle in panel (a). Other parameters: $\beta = -0.53\pi$, $\epsilon = 0.01$.

E: Antipodal symmetry

In order to motivate why it is sufficient to plot the absolute values $|\kappa_{ij}|$ in Fig. 2 of the main text, we show that the system

$$\frac{d\phi_i}{dt} = \omega_i - \frac{\sigma}{N} \sum_{j=1}^{N} \kappa_{ij} \sin(\phi_i - \phi_j), \qquad (6)$$

$$\frac{d\kappa_{ij}}{dt} = -\epsilon \left(\kappa_{ij} + \sin(\phi_i - \phi_j + \beta) \right) \qquad (7)$$

is invariant under the transformation

$$\phi_i \rightarrow \tilde{\phi}_i = \phi_i + \pi,$$
$$\kappa_{lm} \rightarrow \tilde{\kappa}_{lm} = -\kappa_{lm} \text{ where either } l = i \text{ or } m = i,$$

for a fixed $i = 1, \dots, N$.

For equation (6) it is sufficient to examine the term in the sum which yields

$$\tilde{\kappa}_{ij} \sin\left(\tilde{\phi}_i - \phi_j \right) = -\kappa_{ij} \sin\left(\phi_i + \pi - \phi_j \right)$$
$$= \kappa_{ij} \sin\left(\phi_i - \phi_j \right).$$

Similarly, equation (7) remains invariant under the transformation as can be readily seen by comparing with the transformed equation

$$-\frac{d\tilde{\kappa}_{ij}}{dt} = -\epsilon \left(-\tilde{\kappa}_{ij} - \sin(\tilde{\phi}_i - \phi_j + \beta) \right).$$

This transformation can be applied to any of the N oscillators, changing the sign of a particular κ_{lm}.

F: Derivation of multicluster mean-field equations

As outlined in the main text, we use an ansatz that describes each of the M clusters with three collective variables and reduces the oscillator-oscillator coupling to a cluster-cluster coupling. This is done by first assigning each oscillator i to a cluster μ with $\mathcal{C}_\mu$ being the index set of all oscillators in cluster μ. Second, we assume a set of collective coordinates $\vartheta_\mu(t), f_\mu(t)$ and $\varkappa_{\mu\nu}(t)$ that describe the temporal dependency of the phases and the coupling as follows

$$\phi_i \approx \tilde{\phi}_i^\mu = \vartheta_\mu(\omega_i - \Omega_\mu) + f_\mu, \qquad (8)$$
$$\kappa_{ij} \approx \tilde{\kappa}_{ij}^{\mu\nu} = \varkappa_{\mu\nu}. \qquad (9)$$

The dynamical description of each oscillator is split up into two parts. f_μ serves as a phase variable for the cluster of the oscillator, with the relative position within the cluster being described by ϑ_μ and the relative frequency difference $(\omega_i - \Omega_\mu)$. The coupling variables κ_{ij} are reduced from an oscillator-oscillator coupling to a cluster-cluster coupling $\varkappa_{\mu\nu}$. Inserting the ansatz Eq. (8)–(9) into the original system Eq. (6)–(7) yields a non-zero $N + N^2$ dimensional error vector $\mathbf{E} = \left(E_{\phi_1^1} \dots E_{\phi_N^M}, E_{\kappa_{1,1}^{1,1}} \dots E_{\kappa_{N,N}^{M,M}} \right)$. Introducing the difference in natural frequencies $\Delta\omega_{ij} := \omega_i - \omega_j$ and the phase difference $\Delta\phi_{ij}^{\mu\nu} := \tilde{\phi}_i^\mu - \tilde{\phi}_j^\nu = \vartheta_\mu(\omega_i - \Omega_\mu) + f_\mu - \vartheta_\nu(\omega_j - \Omega_\nu) - f_\nu$ we write the error corresponding to the dynamics of the phase and of the coupling as

$$E_{\phi_i^\mu} = \dot{\vartheta}_\mu (\omega_i - \Omega_\mu) + \dot{f}_\mu - \omega_i$$
$$+ \frac{\sigma}{N} \sum_{j \in \mathcal{C}_\mu} \varkappa_{\mu\mu} \sin(\vartheta_\mu \Delta\omega_{ij})$$
$$+ \frac{\sigma}{N} \sum_{\nu \neq \mu} \sum_{j \in \mathcal{C}_\nu} \varkappa_{\mu\nu} \sin\left(\Delta\phi_{ij}^{\mu\nu} \right),$$
$$E_{\kappa_{ij}^{\mu\nu}} = \dot{\varkappa}_{\mu\nu} + \epsilon \left[\varkappa_{\mu\nu} + \sin\left(\Delta\phi_{ij}^{\mu\nu} + \beta \right) \right].$$

The error is minimized by requiring that it is orthogonal to the manifold of the ansatz (8)–(9) parametrized by

the collective coordinates. The tangent space is spanned by the partial derivatives of the ansatz functions with respect to the collective coordinates. In particular, if we denote the ansatz (8)–(9) as $u = \left(\tilde{\phi}, \tilde{\kappa}\right)$ and the collective coordinates $\mathbf{c} = (\vartheta_\mu, f_\mu, \kappa_{\mu\nu})$, the minimization of the error amounts to $\left(\frac{\partial}{\partial c_i} u\right)^T \cdot \mathbf{E} = 0$ for all $i = 1, 2, 3$. This yields the desired temporal evolution equations for the collective coordinates. Projecting onto $\partial u / \partial \vartheta_\mu$ yields

$$
\begin{aligned}
\dot{\vartheta}_\mu \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu)^2 &= \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu) \omega_i \\
&- \frac{\sigma}{N} \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu) \sum_{j \in \mathcal{C}_\mu} \varkappa_{\mu\mu} \sin (\vartheta_\mu \Delta\omega_{ij}) \\
&- \frac{\sigma}{N} \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu) \sum_{\nu \neq \mu} \sum_{j \in \mathcal{C}_\nu} \varkappa_{\mu\nu} \sin \left(\Delta\phi_{ij}^{\mu\nu}\right). \quad (10)
\end{aligned}
$$

Here, we used that, by definition, $\sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu) = 0$ to eliminate $\dot{f}_\mu$ from this equation. Similarly, the projections onto $\partial u / \partial f_\mu$ and onto $\partial u / \partial \kappa_{\mu\nu}$ yield

$$
\begin{aligned}
\dot{f}_\mu = \; & \Omega_\mu - \frac{\sigma}{N_\mu N} \sum_{i,j \in \mathcal{C}_\mu} \varkappa_{\mu\mu} \sin (\vartheta_\mu \Delta\omega_{ij}) \\
& - \frac{\sigma}{N_\mu N} \sum_{i \in \mathcal{C}_\mu} \sum_{\nu \neq \mu} \sum_{j \in \mathcal{C}_\nu} \varkappa_{\mu\nu} \sin \left(\Delta\phi_{ij}^{\mu\nu}\right),
\end{aligned} \quad (11)
$$

$$
\dot{\varkappa}_{\mu\nu} = -\epsilon \left[\varkappa_{\mu\nu} + \frac{1}{N_\mu N_\nu} \sum_{i \in \mathcal{C}_\mu} \sum_{j \in \mathcal{C}_\nu} \sin \left(\Delta\phi_{ij}^{\mu\nu} + \beta\right) \right]. \quad (12)
$$

Equations (10)-(12) approximate the dynamics for a finite number of oscillators and can be analyzed for any given set of natural frequencies. To further simplify these equations, we consider the continuum limit $N \to \infty$, where sums of the form $\frac{1}{N} \sum_{i=1}^N g(\omega_i)$ are evaluated as $\int g(\omega)\rho(\omega) \, d\omega$ for an arbitrary function g. We mainly consider a uniform frequency distribution with

$$
\rho(\omega) = \begin{cases} 2, & \text{for } \omega \in [-0.25, 0.25], \\ 0, & \text{otherwise.} \end{cases}
$$

For the analysis of two cluster states we consider a distribution with disjoint support. We consider

$$
\rho_1(\omega) = \begin{cases} \frac{2}{n_1} & \text{for } \omega \in [-0.25, -0.25 + 0.5n_1], \\ 0 & \text{otherwise,} \end{cases}
$$

$$
\rho_2(\omega) = \begin{cases} \frac{2}{n_2} & \text{for } \omega \in [0.25 - 0.5n_2, 0.25]. \\ 0 & \text{otherwise,} \end{cases}
$$

with $n_2 = 1 - n_1$. The $N_1 = n_1 N$ oscillators with frequencies drawn from $\rho_1(\omega)$ are assigned to cluster 1, and the remaining $N_2 = n_2 N$ oscillators with frequencies drawn from $\rho_2(\omega)$ are assigned to cluster 2. The probability density of the full system can be recovered as a mixture of the two cluster probability densities $\rho(\omega) = n_1\rho_1(\omega) + n_2\rho_2(\omega)$ In this case the order parameter for a single cluster is $r_\mu = \frac{4}{n_\mu \vartheta_\mu} \sin\left(\frac{n_\mu \vartheta_\mu}{4}\right)$. Introducing $f = f_1 - f_2$, we arrive at equations (5a)-(5c) in the main text for the collective coordinates which we recall here for completeness

$$
\begin{aligned}
\dot{\vartheta}_\mu = 1 &+ \frac{\sigma}{\mathrm{v}_\mu \vartheta_\mu} \left(\cos\left(\frac{\vartheta_\mu n_\mu}{4}\right) - r_\mu \right) \\
&\times [\varkappa_{\mu\mu} n_\mu r_\mu + \varkappa_{\mu\nu} n_\nu r_\nu \cos f],
\end{aligned} \quad (13a)
$$

$$
\dot{f} = \Omega_{12} - \sigma r_1 r_2 (n_1 \kappa_{21} + n_2 \kappa_{12}) \sin f, \quad (13b)
$$

$$
\dot{\varkappa}_{\mu\nu} = -\epsilon [\varkappa_{\mu\nu} + r_\mu r_\nu \sin (f_\mu - f_\nu + \beta)], \quad (13c)
$$

with $\Omega_{12} = \Omega_1 - \Omega_2$ and the variance of the natural frequencies of a cluster $\mathrm{v}_\mu = \frac{1}{N_\mu} \sum_{i \in \mathcal{C}_\mu} (\omega_i - \Omega_\mu)^2$.

G: Perturbative approximation for the two-cluster state

In this section, we derive an explicit expression for the two-cluster states emerging in (13). To do this, we use a perturbative approach with respect to the small parameter ϵ. Numerical simulations show that intercluster interactions, mediated via the coupling $\varkappa_{\mu\nu}$, can be considered as small. We assume that each cluster in first approximation can be treated as if it were isolated up to a small perturbation due to intercluster interaction. Further, we assume that the phase difference between the two clusters f grows linearly in time corresponding to a relative rotational motion with constant phase velocity Ω'. This suggests the following expansion of the collective coordinates in orders of ϵ

$$
\vartheta_\mu(t) = \vartheta_\mu^{(0)} + \epsilon \vartheta_\mu^{(1)}(t) + \mathcal{O}\left(\epsilon^2\right), \quad (14)
$$

$$
\varkappa_{\mu\nu}(t) = \varkappa_{\mu\nu}^{(0)}(t) + \epsilon \varkappa_{\mu\nu}^{(1)}(t) + \mathcal{O}\left(\epsilon^2\right), \quad (15)
$$

$$
f(t) = \Omega't + \epsilon f^{(1)}(t) + \mathcal{O}\left(\epsilon^2\right). \quad (16)
$$

Using a Taylor expansion, the order parameter is expanded as

$$
\begin{aligned}
r_\mu(t) = \; & \frac{4}{\vartheta_\mu^{(0)} n_\mu} \sin\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) \\
& + \epsilon \frac{\vartheta_\mu^{(1)}}{\vartheta_\mu^{(0)}} \left[\cos\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) - \frac{4}{\vartheta_\mu^{(0)} n_\mu} \sin\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) \right] \\
& + \mathcal{O}\left(\epsilon^2\right) \\
= \; & r_\mu^{(0)} + \epsilon r_\mu^{(1)}(t) + \mathcal{O}\left(\epsilon^2\right).
\end{aligned}
$$

In order to derive equations for the expansion coefficients, we insert the expansion (14) into Eqs.(13) and equate

powers of ϵ. For the coupling variables this yields

$$\dot{\varkappa}_{\mu\nu}^{(0)} = 0,$$
$$\dot{\varkappa}_{\mu\nu}^{(1)} = -\left(\varkappa_{\mu\nu}^{(0)} + r_\mu^{(0)} r_\nu^{(0)} \sin\left(\pm\Omega' t + \beta\right)\right),$$

where Ω' and $-\Omega'$ correspond to $\varkappa_{12}$ and $\varkappa_{21}$, respectively. Note that for $\nu = \mu$ we set $\Omega' = 0$. To ensure constant intracluster coupling we require $\varkappa_{\mu\mu}^{(0)} = -\left(r_\mu^{(0)}\right)^2 \sin(\beta)$. Setting $\kappa_{\mu\nu}^{(0)} = 0$ ensures that the intercluster coupling is small compared to the intracluster coupling. For $\mu \neq \nu$ the intercluster coupling in the non-synchronized case is readily integrated yielding

$$\varkappa_{\mu\nu}^{(1)}(t) = \frac{\pm 1}{\Omega'} r_\mu^{(0)} r_\nu^{(0)} \cos\left(\pm\Omega' t + \beta\right). \qquad (17)$$

Similarly, we find for ϑ_μ

$$0 = 1 + \frac{\sigma}{v_\mu \vartheta_\mu^{(0)}} \left(\cos\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) - r_\mu^{(0)}\right) \\ \cdot \left[\varkappa_{\mu\mu}^{(0)} n_\mu r_\mu^{(0)} + \varkappa_{\mu\nu}^{(0)} n_\nu r_\nu^{(0)} \cos\left(\Omega' t\right)\right], \qquad (18)$$

$$\dot{\vartheta}_\mu^{(1)} = \frac{\vartheta_\mu^{(1)} \sigma}{v_\mu \left(\vartheta_\mu^{(0)}\right)^2} \left(\frac{8\sin\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right)}{\vartheta_\mu^{(0)} n_\mu} - 2\cos\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) \right.\\ \left. - \frac{\vartheta_\mu^{(0)} n_\mu \sin\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right)}{4}\right) \\ \cdot \left[n_\mu \kappa_{\mu\mu}^{(0)} r_\mu^{(0)} + \varkappa_{\mu\nu}^{(0)} n_\nu r_\nu^{(0)} \cos\left(\Omega' t\right)\right] \\ + \frac{\sigma}{v_\mu \vartheta_\mu^{(0)}} \left(\cos\left(\frac{\vartheta_\mu^{(0)} n_\mu}{4}\right) - r_\mu^{(0)}\right) \\ \cdot \left[\varkappa_{\mu\mu}^{(1)} n_\mu r_\mu^{(0)} + \varkappa_{\mu\mu}^{(0)} n_\mu r_\mu^{(1)} + \varkappa_{\mu\nu}^{(1)} n_\nu r_\nu^{(1)} \cos\left(\Omega' t\right)\right]. \qquad (19)$$

A time-independent $\vartheta_\mu^{(0)}$ can only be achieved by setting $\varkappa_{\mu\nu}^{(0)} = 0$ as above. Using these solutions to calculate the intercluster dynamics given by f yields

$$\epsilon \dot{f}^{(1)} = \Omega_1 - \Omega_2 - \Omega' \\ - \frac{\epsilon\sigma}{\Omega'} \left(r_1^{(0)} r_2^{(0)}\right)^2 \left[\frac{n_2}{2}\left(\sin\left(2\Omega' t + \beta\right) - \sin\beta\right)\right. \qquad (20) \\ \left. + \frac{n_1}{2}\left(\sin\left(\beta - 2\Omega' t\right) - \sin\beta\right)\right].$$

For self-consistency with the ansatz (14), we choose Ω' in such a way that only sinusoidal terms are left in the previous equation. All dynamics which is linear in time is absorbed providing us with an explicit formula to derive Ω', which reads

$$0 = \left(\Omega'\right)^2 + \Omega'\left(\Omega_2 - \Omega_1\right) - \frac{\epsilon\sigma}{2}\left(r_1^{(0)} r_2^{(0)}\right)^2 \sin\beta. \qquad (21)$$

H: Estimating upper and lower bounds for cluster synchronization

Eq. (21) can be used to estimate the upper bound for the existence of the two-cluster solution. Solving the quadratic equation for real valued Ω' implies that

$$(\Omega_1 - \Omega_2)^2 \geq -2\epsilon\sigma \left(r_1^{(0)} r_2^{(0)}\right)^2 \sin\beta. \qquad (22)$$

For the chosen parameter values of $\epsilon = 0.01$, $\beta = -0.53\pi$ and $\Omega_1 - \Omega_2 = -0.25$, we calculate the highest value of σ for which the condition is no longer satisfied. A rough estimate using $r_1^{(0)} = r_2^{(0)} = 1$ yields $\sigma_S = 3.14$. Note that the approximate values for $r_\mu^{(0)}$ depend on n_1 and σ explicitly, and implicitly through $\vartheta_\mu^{(0)}$, see Eq. (18). For increasing values of σ starting at our rough estimate, we find the value for σ for which condition (22) is no longer satisfied. This turns out to be about $\sigma_S \approx 3.152$ for both $n_1 = 0.5$ and $n_1 = 0.95$ with an accuracy of $\Delta\sigma = 0.001$.

To estimate a lower bound for the cluster synchronization we use Eq. (18) and analyse if it is solvable. Again checking for decreasing values of σ using a step size of $\Delta\sigma = 0.001$, we arrive at $\sigma_C \approx 0.460$ as a lower bound for the two-cluster states. Here again the same value is attained for both $n_1 = 0.5$ and $n_1 = 0.95$.

I: Single Cluster synchronization with collective coordinates

In the case when all oscillators are frequency synchronized, we can use a simpler collective coordinate ansatz with just a single variable. The ansatz is as follows

$$\phi_i \approx \hat{\phi}_i = \vartheta\omega_i \qquad (23)$$
$$\kappa_{ij} \approx -\sin\left(\phi_i - \phi_j + \beta\right). \qquad (24)$$

Here ϑ describes the distance of a given oscillator with natural frequency ω_i from the mean phase of the frequency synchronized cluster and κ_{ij} are approximated by their steady-state values. Proceeding analogously as in Section we determine the error introduced by the ansatz of the phases and couplings, and obtain

$$E_{\phi_i} = \dot{\vartheta}\omega_i - \omega_i - \frac{\sigma}{N}\sum_{j=1}^N \sin\left(\vartheta\Delta\omega_{ij} + \beta\right)\sin\left(\vartheta\Delta\omega_{ij} + \alpha\right)$$

$$E_{\kappa_{ij}} = \dot{\vartheta}\Delta\omega_{ij} \cos\left(\vartheta\Delta\omega_{ij} + \beta\right).$$

Calculating the projection of the error on the tangent space of the ansatz space and forcing it to be zero yields

an equation of motion for our collective coordinate

$$\dot{\vartheta}\left(\sum_{i=1}^{N}\omega_i^2 - \sum_{i,j=1}^{N}\Delta\omega_{ij}^2\cos^2\left(\vartheta\Delta\omega_{ij}+\beta\right)\right) =$$

$$\sum_{i=1}^{N}\omega_i^2 + \frac{\sigma}{N}\sum_{i=1}^{N}\omega_i\sum_{j=1}^{N}\sin(\vartheta\Delta\omega_{ij}+\beta)\sin(\vartheta\Delta\omega_{ij}). \quad (25)$$

Setting $\dot{\vartheta} = 0$ yields a condition for the existence of a fully synchronized state. We obtain

$$0 = 1 - \frac{\sigma}{v^2 N^2}\sum_i \omega_i \sum_j \frac{1}{2}\cos\left(2\vartheta\Delta\omega_{ij}+\beta\right). \quad (26)$$

Here $v = \frac{1}{N}\sum_{i=1}^{N}\omega_i$ is the standard deviation of the natural frequencies. Again choosing a uniform distribution of frequencies between -0.25 and 0.25 provides a simplified equation to determine ϑ

$$1 = 48\sigma\sin(\beta)\frac{\sin\left(\frac{\vartheta}{2}\right)}{\vartheta}\frac{\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right) - \sin\left(\frac{\vartheta}{2}\right)}{\vartheta^2}. \quad (27)$$

This equation can be used to analyze the transition from a fully synchronized system to a desynchronized one. In Fig. 6, we plot the stable branch of the analytical solution for the single cluster state along with the simulation result for the full system. Note that (27) for the reduced system gives rise to one stable and one unstable solution that vanish in a fold bifurcation, see Ref. [58] from the main text. We observe a very good agreement of the numerical and the analytical results. Together with the findings presented in the two previous sections, the emergence of various transition pathways to synchrony and the hysteretic behavior can be accurately explained.

J: Frequency clustering in the synchronization transition

To examine the clustering during the transition to synchrony in more detail, we present the distribution of the mean phase velocities $\rho\left(\left\langle\dot{\phi}_i\right\rangle\right)$ for 100 different simulations encoded in color versus the overall coupling strength σ in Fig. 7. The results are displayed in logarithmic scale. Each picture uses the data from 100 simulation runs and for each run the mean phase velocity is shifted by the average frequency of that particular run. We use the upsweep protocol described in Sec. with a transient time of $T_0 = 10^5$ time units and an averaging time of $T = 10^4$ time units. Figure 7(a) shows the statistics from 100 simulations with different natural frequencies from a uniform distribution $\omega_i \in [-0.25, 0.25]$ for each simulation. Figure 7(b) displays the data from 100 simulations using a fixed realization of natural frequencies that leads to a multi-step first-order transition, similar to the ones
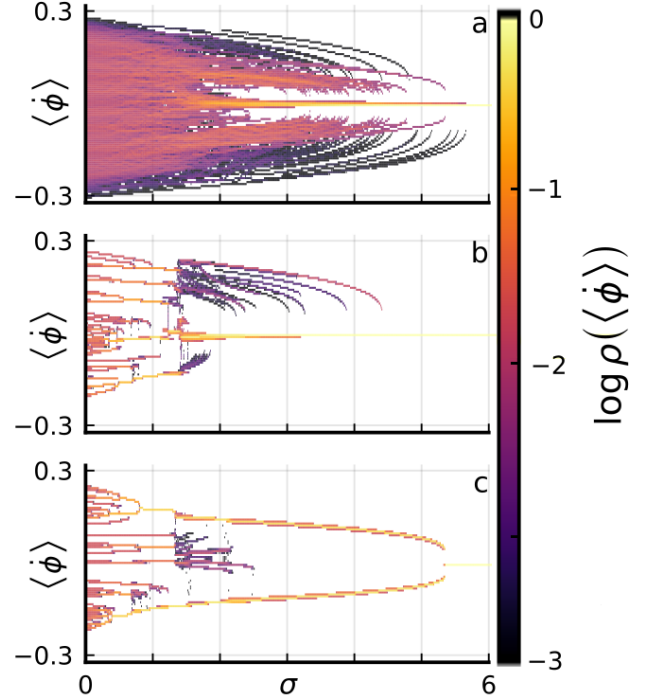

FIG. 7. Evolution of the mean phase velocity $\left\langle\dot{\phi}_i\right\rangle$ as a function of increasing σ. For each value of σ the color code denotes the natural logarithm of the relative number of oscillators found at a specific mean phase velocity. Each picture shows results for 100 simulations with 50 oscillators each. (a) 100 different realizations of natural frequencies, (b) using fixed frequencies with 100 different initial conditions leading to a multi-step transition, (c) frequencies leading to the single-step transition. In the plot the frequencies are shifted vertically by the mean value of the realization of the natural frequencies such that each simulation is centered around 0. The simulation procedure is the same as in Fig. 1 of the main text. Other parameters: $\beta = -0.53\pi$, $\epsilon = 0.01$.

shown in Fig. 1(b) of the main paper. The fixed realization leads to predetermined paths and allows for an analysis of the synchronization process in more detail. Note the emergence of a big group of oscillators around $\left\langle\dot{\phi}_i\right\rangle = 0$ for small values of σ. This corresponds to the emergence of a large cluster for all simulations, exactly as shown in Fig. 2 of the main text. Figure 7(c) shows the data for 100 simulations of a fixed realization of natural frequencies that lead to a single-step first-order transition, similar to the ones shown in Fig. 1(c) of the main paper. Note the two branches of oscillators located symmetrically around 0. These represent the two clusters of almost equal size. Until these two clusters merge there are no oscillators with mean phase velocities around 0. These two clusters are also observable in Fig. 2(g) of the main text.

* jan.fialkowski@campus.tu-berlin.de

† rico.berner@physik.hu-berlin.de

[1] H. E. Stanley: *Phase transitions and critical phenomena* (Clarendon Press, Oxford, 1971).

[2] M. E. J. Newman: *The structure and function of complex networks*, SIAM Review **45**, 167 (2003).

[3] S. Boccaletti, A. N. Pisarchik, C. I. del Genio, and A. Amann: *Synchronization: From Coupled Systems to Complex Networks* (Cambridge University Press, Cambridge, 2018).

[4] H. Haken: *Synergetics, An Introduction* (Springer, Berlin, 1983), 3rd ed.

[5] E. Schöll: *Nonequilibrium Phase Transitions in Semiconductors* (Springer, Berlin, 1987).

[6] F. A. Rodrigues, T. K. D. M. Peron, P. Ji, and J. Kurths: *The Kuramoto model in complex networks*, Phys. Rep. **610**, 1 (2016).

[7] D. Ghosh, M. Frasca, A. Rizzo, S. Majhi, S. Rakshit, K. Alfaro-Bittner, and S. Boccaletti: *The synchronized dynamics of time-varying networks*, Phys. Rep. **949**, 1 (2022).

[8] Y. Kuramoto: *Chemical Oscillations, Waves and Turbulence* (Springer-Verlag, Berlin, 1984).

[9] A. Pikovsky, M. Rosenblum, and J. Kurths: *Synchronization: a universal concept in nonlinear sciences* (Cambridge University Press, Cambridge, 2001), 1st ed.

[10] D. Pazó: *Thermodynamic limit of the first-order phase transition in the kuramoto model*, Phys. Rev. E **72**, 046211 (2005).

[11] J. Gómez-Gardeñes, S. Gómez, A. Arenas, and Y. Moreno: *Explosive synchronization transitions in scale-free networks*, Phys. Rev. Lett. **106**, 128701 (2011).

[12] S. Boccaletti, J. A. Almendral, S. Guan, I. Leyva, Z. Liu, I. Sendiña-Nadal, Z. Wang, and Y. Zou: *Explosive transitions in complex networks' structure and dynamics: Percolation and synchronization*, Phys. Rep. **660** (2016).

[13] R. M. D'Souza, J. Gómez-Gardeñes, J. Nagler, and A. Arenas: *Explosive phenomena in complex networks*, Adv. Phys. **68**, 123 (2019).

[14] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas: *Paths to synchronization on complex networks*, Phys. Rev. Lett. **98**, 034101 (2007).

[15] X. Zhang, X. Hu, J. Kurths, and Z. Liu: *Explosive synchronization in a general complex network*, Phys. Rev. E **88**, 010802 (2013).

[16] M. K. S. Yeung and S. H. Strogatz: *Time delay in the Kuramoto model of coupled oscillators*, Phys. Rev. Lett. **82**, 648 (1999).

[17] S. H. Strogatz: *From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators*, Physica D **143**, 1 (2000).

[18] W. S. Lee, E. Ott, and T. M. Antonsen: *Large coupled oscillator systems with heterogeneous interaction delays*, Phys. Rev. Lett. **103**, 044101 (2009).

[19] M. Breakspear, S. Heitmann, and A. Daffertshofer: *Generative models of cortical oscillations: neurobiological implications of the Kuramoto model*, Front. Hum. Neurosci. **4**, 190 (2010).

[20] A. Nabi and J. Moehlis: *Single input optimal control for globally coupled neuron networks*, J. Neural Eng. **8**, 065008 (2011).

[21] C. Bick, M. Goodfellow, C. R. Laing, and E. A. Martens: *Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions: a review*, J. Math. Neurosci. **10**, 9 (2020).

[22] S. Olmi, A. Navas, S. Boccaletti, and A. Torcini: *Hysteretic transitions in the Kuramoto model with inertia*, Phys. Rev. E **90**, 042905 (2014).

[23] L. Tumash, S. Olmi, and E. Schöll: *Effect of disorder and noise in shaping the dynamics of power grids*, Europhys. Lett. **123**, 20001 (2018).

[24] L. Tumash, S. Olmi, and E. Schöll: *Stability and control of power grids with diluted network topology*, Chaos **29**, 123105 (2019).

[25] H. Taher, S. Olmi, and E. Schöll: *Enhancing power grid synchronization and stability through time delayed feedback control*, Phys. Rev. E **100**, 062306 (2019).

[26] F. Hellmann, P. Schultz, P. Jaros, R. Levchenko, T. Kapitaniak, J. Kurths, and Y. Maistrenko: *Network-induced multistability through lossy coupling and exotic solitary states*, Nat. Commun. **11**, 592 (2020).

[27] S. Jain and S. Krishna: *A model for the emergence of cooperation, interdependence, and structure in evolving networks*, Proc. Natl. Acad. Sci. **98**, 543 (2001).

[28] C. Kuehn: *Multiscale dynamics of an adaptive catalytic network*, Math. Model. Nat. Phenom. **14**, 402 (2019).

[29] T. Gross, C. J. D. D'Lima, and B. Blasius: *Epidemic dynamics on an adaptive network*, Phys. Rev. Lett. **96**, 208701 (2006).

[30] S. R. Proulx, D. E. L. Promislow, and P. C. Phillips: *Network thinking in ecology and evolution*, Trends Ecol. Evol. **20**, 345 (2005).

[31] W. Gerstner, R. Kempter, J. L. von Hemmen, and H. Wagner: *A neuronal learning rule for sub-millisecond temporal coding*, Nature **383**, 76 (1996).

[32] C. Meisel and T. Gross: *Adaptive self-organization in a realistic neural network model*, Phys. Rev. E **80**, 061917 (2009).

[33] K. Mikkelsen, A. Imparato, and A. Torcini: *Emergence of slow collective oscillations in neural networks with spike-timing dependent plasticity*, Phys. Rev. Lett. **110**, 208101 (2013).

[34] E. A. Martens and K. Klemm: *Transitions from trees to cycles in adaptive flow networks*, Front. Phys. **5**, 62 (2017).

[35] T. Gross and B. Blasius: *Adaptive coevolutionary networks: a review*, J. R. Soc. Interface **5**, 259 (2008).

[36] L. Horstmeyer and C. Kuehn: *Adaptive voter model on simplicial complexes*, Phys. Rev. E **101**, 022305 (2020).

[37] R. Gutiérrez, A. Amann, S. Assenza, J. Gómez-Gardeñes, V. Latora, and S. Boccaletti: *Emerging meso- and macroscales from synchronization of adaptive networks*, Phys. Rev. Lett. **107**, 234103 (2011).

[38] D. V. Kasatkin, S. Yanchuk, E. Schöll, and V. I. Nekorkin: *Self-organized emergence of multi-layer structure and chimera states in dynamical networks with adaptive couplings*, Phys. Rev. E **96**, 062211 (2017).

[39] R. Berner, E. Schöll, and S. Yanchuk: *Multiclusters in networks of adaptively coupled phase oscillators*, SIAM J. Appl. Dyn. Syst. **18**, 2227 (2019).

[40] R. Berner, J. Sawicki, and E. Schöll: *Birth and stabilization of phase clusters by multiplexing of adaptive networks*, Phys. Rev. Lett. **124**, 088301 (2020).

[41] P. Feketa, A. Schaum, and T. Meurer: *Synchronization and multi-cluster capabilities of oscillatory networks with*

adaptive coupling, IEEE Trans. Autom. Control **66**, 3084 (2020).

[42] R. Berner, S. Vock, E. Schöll, and S. Yanchuk: *Desynchronization transitions in adaptive networks*, Phys. Rev. Lett. **126**, 028301 (2021).

[43] L. Lücken, O. V. Popovych, P. A. Tass, and S. Yanchuk: *Noise-enhanced coupling between two oscillators with long-term plasticity*, Phys. Rev. E **93**, 032210 (2016).

[44] V. Röhr, R. Berner, E. L. Lameu, O. V. Popovych, and S. Yanchuk: *Frequency cluster formation and slow oscillations in neural populations with plasticity*, PLoS ONE **14**, e0225094 (2019).

[45] J. Sawicki, R. Berner, T. Löser, and E. Schöll: *Modelling tumor disease and sepsis by networks of adaptively coupled phase oscillators*, Front. Netw. Physiol. **1**, 730385 (2022).

[46] R. Berner, J. Sawicki, M. Thiele, T. Löser, and E. Schöll: *Critical parameters in dynamic network modeling of sepsis*, Front. Netw. Physiol. **2**, 904480 (2022).

[47] R. Berner, S. Yanchuk, and E. Schöll: *What adaptive neuronal networks teach us about power grids*, Phys. Rev. E **103**, 042315 (2021).

[48] N. T. K. Thanh, N. Maclean, and S. Mahiddine: *Mechanisms of nucleation and growth of nanoparticles in solution*, Chem. Rev. **114**, 7610 (2014).

[49] L. Schimansky-Geier, C. Zülicke, and E. Schöll: *Domain formation due to ostwald ripening in bistable systems far from equilibrium*, Z. Phys. B **84**, 433 (1991).

[50] Note that the choice of $\hat{\omega}$ is arbitrary as it can be compensated by the rescaling of σ, ϵ, and time t.

[51] T. Aoki and T. Aoyagi: *Co-evolution of phases and connection strengths in a network of phase oscillators*, Phys. Rev. Lett. **102**, 034101 (2009).

[52] R. Berner, J. Fialkowski, D. V. Kasatkin, V. I. Nekorkin, S. Yanchuk, and E. Schöll: *Hierarchical frequency clusters in adaptive networks of phase oscillators*, Chaos **29**, 103134 (2019).

[53] See the supplemental material at ... for details on the numerical implementation, the synchronization transition with equiprobable natural frequencies, the synchronization transition with multimodal natural frequencies, the hysteretic behavior, the antipodal symmetry of the adaptive network model, the derivation and analysis of the mean-field equations and for details on the clustering behavior during the transition to synchrony.

[54] For the finite time evaluation in the numerical analysis, we use the threshold $\langle \dot{\phi}_i \rangle - \langle \dot{\phi}_j \rangle \leq 10^{-3}$.

[55] The integration is performed using a fifth order Runge-Kutta method and the time step is chosen at every step by estimating the error via comparison with a Runge-Kutta of fourth order and enforcing its smallness.

[56] See Supplemental Material at http... for

[57] Y. Maistrenko, B. Lysyansky, C. Hauptmann, O. Burylko, and P. A. Tass: *Multistability in the kuramoto model with synaptic plasticity*, Phys. Rev. E **75**, 066207 (2007).

[58] G. A. Gottwald: *Model reduction for networks of coupled oscillators*, Chaos **25**, 053111 (2015).

[59] L. D. Smith and G. A. Gottwald: *Chaos in networks of coupled oscillators with multimodal natural frequency distributions*, Chaos **29**, 093127 (2020).

[60] E. J. Hancock and G. A. Gottwald: *Model reduction for kuramoto models with complex topologies*, Phys. Rev. E **98**, 012307 (2018).

[61] S. Watanabe and S. H. Strogatz: *Integrability of a globally coupled oscillator array*, Phys. Rev. Lett. **70**, 2391 (1993).

[62] E. Ott and T. M. Antonsen: *Low dimensional behavior of large systems of globally coupled oscillators*, Chaos **18**, 037113 (2008).

[63] L. D. Smith and G. A. Gottwald: *Model reduction for the collective dynamics of globally coupled oscillators: From finite networks to the thermodynamic limit*, Chaos **30**, 093107 (2020).

[64] Note that the ansatz (8) is only valid for fully connected networks [60] but the numerical results show that this simplification is justified.

[65] Note that the case of an arbitrary number of clusters can be achieved within the framework of collective coordinates [59, 60].

[66] H. Daido: *Generic scaling at the onset of macroscopic mutual entrainment in limit-cycle oscillators with uniform all-to-all coupling*, Phys. Rev. Lett. **73**, 760 (1994).

[67] P. Feketa, A. Schaum, and T. Meurer: *Stability of cluster formations in adaptive Kuramoto networks*, IFAC-PapersOnLine **54**, 14 (2021).

[68] H. R. Pruppacher and J. D. Klett: *Microphysics of Clouds and Precipitation*, Atmospheric and Oceanographic Sciences Library (Springer, Dordrecht, 2010).

[69] J. W. Mullin: *Crystallization* (Elsevier, 2001).

[70] M. A. Gkogkas, C. Kuehn, and C. Xu: *Mean field limits of co-evolutionary heterogeneous networks*, arXiv:2202.01742 (2022).

[71] H. Tanaka, A. J. Lichtenberg, and S. Oishi: *First order phase transition resulting from finite inertia in coupled oscillator systems*, Phys. Rev. Lett. **78**, 2104 (1997).

[72] P. Ji, T. K. D. M. Peron, P. J. Menck, F. A. Rodrigues, and J. Kurths: *Cluster explosive synchronization in complex networks*, Phys. Rev. Lett. **110**, 218701 (2013).

[73] Note that the choice of $\hat{\omega}$ is arbitrary as it can be compensated by the rescaling of σ, ϵ, and time t.

[74] For the finite time evaluation in the numerical analysis, we use the threshold $\langle \dot{\phi}_i \rangle - \langle \dot{\phi}_j \rangle \leq 10^{-3}$.

[75] The integration is performed using a fifth order Runge-Kutta method and the time step is chosen at every step by estimating the error via comparison with a Runge-Kutta of fourth order and enforcing its smallness.

[76] Note that the ansatz (8) is only valid for fully connected networks [60] but the numerical results show that this simplification is justified.

[77] Note that the case of an arbitrary number of clusters can be achieved within the framework of collective coordinates [59, 60]