Ott-Antonsen attractiveness for parameter-dependent oscillatory systems

Bastian Pietras[1, 2, a)] and Andreas Daffertshofer[1, b)]

[1)] *MOVE Research Institute Amsterdam & Institute for Brain and Behavior Amsterdam, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, van der Boechorststraat 9, Amsterdam 1081 BT, The Netherlands*

[2)] *Department of Physics, Lancaster University, Lancaster LA1 4YB United Kingdom*

(Dated: 8 September 2016)

The Ott-Antonsen (OA) ansatz [Chaos 18, 037113 (2008), Chaos 19, 023117 (2009)] has been widely used to describe large systems of coupled phase oscillators. If the coupling is sinusoidal and if the phase dynamics does not depend on the specific oscillator, then the macroscopic behavior of the systems can be fully described by a low-dimensional dynamics. Does the corresponding manifold remain attractive when introducing an intrinsic dependence between an oscillator's phase and its dynamics by additional, oscillator specific parameters? To answer this we extended the OA ansatz and proved that parameter-dependent oscillatory systems converge to the OA manifold given certain conditions. Our proof confirms recent numerical findings that already hinted at this convergence. Furthermore we offer a thorough mathematical underpinning for networks of so-called theta neurons, where the OA ansatz has just been applied. In a final step we extend our proof by allowing for time-dependent and multi-dimensional parameters as well as for network topologies other than global coupling. This renders the OA ansatz an excellent starting point for the analysis of a broad class of realistic settings.

PACS numbers: 05.45.Xt, 05.45.-a, 89.75.-k, 84.35.+i

[a)]Electronic mail: b.pietras@vu.nl

[b)]Electronic mail: a.daffertshofer@vu.nl

Coupled phase oscillators are being widely used to describe synchronization phenomena. The study of their collective dynamics has experienced a major breakthrough by the results by Ott and Antonsen[1–3]. The asymptotic behavior of the mean field of infinitely many coupled oscillators can be cast into a reduced, low-dimensional system of ordinary differential equations. The evolution is hence captured by the so-called Ott-Antonsen (OA) manifold.

Very recently, the OA ansatz has been applied to networks of theta neurons, see, e.g., Refs. 4–10. A particular property of coupled, inhomogeneous theta neurons is that both the phase of a single neuron as well as its dynamics depend on a parameter, which establishes an intrinsic relation between them. While numerical results suggest the attractiveness of the OA manifold in the presence of such a parameter dependence, it has as to yet not been proven whether the dynamics really converges to it. For a certain class of parameter dependencies we here extend the existing theory of the OA ansatz and show that the OA manifold continues to asymptotically attract the mean field dynamics.

Parameter-dependent systems and their description through the OA ansatz have been considered by, e.g., Strogatz and co-workers[11], Wagemaker and co-workers[12], and So and Barreto[13]. There, parameters seemingly did not yield a correlation between an oscillator's phase and its dynamics but a rigorous proof for this is still missing. We explicitly address this last point. In particular, we prove a conjecture later formulated by Montbrió and co-workers[7] on the attractiveness of the OA manifold for parameter-dependent systems. The case of parameters serving as mere auxiliary variables readily follows from our result – we will refer to this as "weak" parameter-dependence[14]. By showing that a network of theta neurons can be treated as a parameter-dependent oscillatory system, our result establishes an immediate link to networks of quadratic integrate-and-fire (QIF) neurons: That is, the so-called Lorentzian ansatz as an equivalent approach to the OA ansatz is analytically substantiated. By this we may exert an important impact in mathematical neuroscience.

Finally, we extend the parameter-dependence for more general classes of networks. First, we address non-autonomous systems and show that our proof can be applied to time-varying parameters. An important example here is a biologi-

cally realistic approach to oscillatory systems proposed by Winfree[15]. Second, we include multiple distributed parameters illustrated by coupled limit-cycle oscillators with shear. Third, we apply our proof to networks with different coupling topologies including non-local coupling by using an heterogeneous mean field approach.

I.   INTRODUCTION

The Kuramoto model can be considered the most seminal description of globally coupled networks of phase oscillators. It has been investigated in great detail but its various extensions still make it *the* model-to-work-with when it comes to the study of network dynamics[16,17]. We adopt the notion of Montbrió, Pazó, and Roxin[7] and write the Kuramoto-like model as

$$\dot{\theta}_j = \omega_j + \mathrm{Im}\left[H\mathrm{e}^{-i\theta_j}\right] \ , \tag{1}$$

where the phase dynamics of the j-th oscillator ($j = 1, \ldots, N$) depends on its natural frequency ω_j and a driving complex-valued field H. The latter can depend on time t, on the mean field $z(t) = \sum_{j=1}^{N} \mathrm{e}^{i\theta_j(t)}$, and on other auxiliary variables, but not on the (index of) oscillator, i.e. it remains identical for all oscillators $j = 1, \ldots, N$. Given the right-hand side of (1), the oscillators are *sinusoidally coupled*.

In the thermodynamic limit ($N \to \infty$) the OA ansatz yields solutions for the dynamical evolution of the corresponding distribution function (of all the oscillators), which are attracted towards a reduced manifold of states[1,2]. Central to this is the description of the system via its distribution density $\rho(\theta, \omega, t)$. The quantity $\rho(\theta, \omega, t)\, d\theta\, d\omega$ is the fraction of oscillators whose phases are in the range $[\theta, \theta+d\theta]$ and have natural frequencies in $[\omega, \omega+d\omega]$ at time t. The distribution function ρ obeys the continuity equation

$$\partial_t \rho + \partial_\theta \left(\rho\nu\right) = 0 \tag{2}$$

with velocity field

$$\nu(\theta, t) = \omega + \mathrm{Im}\left[H(t)\mathrm{e}^{-i\theta}\right] \ . \tag{3a}$$

The latter can equivalently be written as[11,12]

$$\nu(\theta, t) = f e^{i\theta} + h + f^* e^{-i\theta} \ . \tag{3b}$$

In agreement with the assumptions on H we require that the functions f and h may explicitly depend on time t, on the (now continuum form of the) mean field $z(t) = \int_{-\infty}^{\infty} \int_{0}^{2\pi} \rho e^{i\theta} d\theta d\omega$, and on other auxiliary variables, but not on the the phase θ itself.

Asymptotic attractiveness of the OA manifold, given by distribution functions of the form

$$\rho(\theta, \omega, t) = \frac{g(\omega)}{2\pi} \left\{ 1 + \left[\sum_{n=1}^{\infty} \alpha(\omega, t)^n e^{in\theta} + \text{c.c.} \right] \right\} \tag{4}$$

that satisfy the normalization condition

$$\int_{-\infty}^{\infty} \int_{0}^{2\pi} \rho(\theta, \omega, t) \ d\theta \ d\omega = 1 \ , \tag{5}$$

has been proven for continuous frequency distribution functions $g(\omega)$ of non-zero width and for H being independent of θ; c.c. stands for complex conjugate. Other requirements include $|\alpha(\omega, t)| \leq 1$, and some analytic continuity conditions.[1,2]

In what follows we extend this approach by rigorously proving the asymptotic attractiveness of the OA manifold in the case of H and ω depending on an additional parameter η that may also influence θ. Equivalently, we include a time- and η-dependence of f and h in (3b). By this, we allow for an intrinsic relation between θ, H, and ω, or θ, f, and h, respectively. As of today, the attractiveness of the OA manifold in the (time- and) parameter-dependent case has only been hypothesized[11,18] but not proven.

II.   PARAMETER-DEPENDENT SYSTEMS

When including additional parameters at the oscillator level, the dynamics (1) becomes

$$\dot{\theta}_j = \Omega(\omega_j, \eta_j) + \text{Im}\left[H(\eta_j, t) \ e^{-i\theta_j} \right] \ . \tag{6}$$

The natural frequency Ω of oscillator j may therefore deviate from ω_j, which promotes further heterogeneity among oscillators. Moreover the driving field H may depend on η_j. The right-hand side of (6) expresses a certain dependence on the (index of the) j-th oscillator. Hence, such a dependence is no longer exclusive to the sinusoidal coupling, but also affects the natural frequency $\Omega(\omega_j, \eta_j)$ and the driving field $H(\eta_j, t)$.

4

When considering η a random variable, we may regard η_j to be drawn from a distribution function $g(\eta)$. Likewise ω_j may be drawn from a (different) distribution function. The oscillator-specific parameter η_j may change this distribution function in the oscillator's favor. Therefore, we here incorporate a joint distribution $g(\omega, \eta)$ in the normalization condition (5). In general, ω and η are not independent and the joint distribution consists of two nested distributions. We hence replace $\Omega(\omega_j, \eta_j)$ by $\omega(\eta_j)$. Then, in the continuum limit (6) reads:

$$\partial_t \theta(\eta, t) = \omega(\eta, t) + \mathrm{Im}\left[H(\eta, t)\, \mathrm{e}^{-i\theta}\right] . \tag{7}$$

The relation through η becomes now even more evident as the temporal derivative of θ has become partial.

Again, one can introduce a distribution function $\rho(\theta, \omega, \eta, t)$, which now additionally depends on η. And again, this distribution function satisfies the continuity equation (2) with velocity field (7). In line with the parameter-independent case, in which the distribution function $g(\omega)$ of the natural frequencies ω had non-zero width[1,2], we assume that the distribution function $g(\eta)$ of the parameter η also has non-zero width. The frequency ω, thus, cannot be constant but depends on η. Likewise, the driving field H depends on η. Importantly, these two terms exhibit so an implicit dependence on θ, such that the proof for the attractiveness of the OA manifold as has been derived in Ref. 2 may no longer hold. However, there is strong numerical incentive that the OA manifold fully covers the long-term behavior of the dynamics of the population of parameter-dependent phase oscillators; see, e.g., Refs. 4–10, 18–23.

In the following we demonstrate the proof of this conjecture for a particular class of parameter-dependent systems. We consider η to follow a Lorentzian distribution and assume that ω depends linearly on η, i.e. $\omega(\eta, t) = a \cdot \eta + c$, where, without loss of generality, we set $a = 1$ and consider $c = c(t) \in L_{1,\mathrm{loc}}(\mathbb{R})$ a locally integrable, and in particular piecewise smooth, function. Our line of argument follows closely that of Ott and Antonsen[2] but we extend their results whenever necessary. We would like to note that our findings remain valid for a larger class of distribution functions as has been depicted in detail in Ref. 3. We will comment on this and consider more general η-dependencies of ω in Sections IV and V.

Let $g(\eta)$ be a Lorentzian centered around $\eta = \eta_0$ with width Δ, i.e. $g(\eta) \sim L(\eta_0, \Delta)$. For the aforementioned linear dependency $\omega(\eta, t) = a \cdot \eta + c$, we have $\tilde{g}(\omega) = \hat{g}(\eta) \sim L(\eta_0 + c, \Delta)$ with frequency $\omega = \omega(\eta)$ that, in general, will depend on η. In this case ω is fully described

by (the distribution of) η and the distribution density reduces to $\rho(\theta, \omega, \eta, t) = \rho(\theta, \eta, t)$.[24] This can be expanded as a Fourier series in θ similar to Eqs.(5 & 6) in Ref. 2, where it is further decomposed into $\rho(\theta, \eta, t) = \hat{g}(\eta)/(2\pi) \cdot [1 + \rho_+(\theta, \eta, t) + \rho_-(\theta, \eta, t)]$. Next to the assumption that the analytic continuation of ρ_+ (ρ_-) into $\mathrm{Im}(\theta) > 0$ ($\mathrm{Im}(\theta) < 0$) has no singularities and decays to zero as $\mathrm{Im}(\theta) \to +\infty$ ($\mathrm{Im}(\theta) \to -\infty$), we exploit the symmetry of the Fourier expansion and focus on ρ_+. In particular, we expect ρ_+ to fulfill these conditions initially, i.e. $\rho_+(\theta, \eta, 0)$ can be continued into the complex η-plane, is analytic in $\mathrm{Im}(\eta) < 0$ and decays to zero for $\mathrm{Im}(\eta) \to -\infty$. These conditions are satisfied for all $t > 0$.[1]

We can further decompose ρ_+ into two parts, $\rho_+ = \hat{\rho}_+ + \hat{\rho}'_+$, where $\hat{\rho}'_+$ lies on the OA manifold and follows the dynamics given by Eq.(9) in Ref. 2. For the sake of completeness, this dynamics prescribes the evolution of the Fourier coefficients $\hat{\rho}'_+$ to the form $\hat{\rho}'_n(\eta, t) = [\alpha(\eta, t)]^n$, and reads

$$\partial_t \alpha + i\eta\alpha + \frac{1}{2}\left(H\alpha^2 - H^*\right) = 0 . \tag{8}$$

The quantity $\hat{\rho}_+$, on the other hand, is a solution of

$$\partial_t \hat{\rho}_+ + \partial_\theta \left\{ \left[\omega + \frac{1}{2i}\left(H e^{-i\theta} - H^* e^{i\theta}\right) \right] \hat{\rho}_+ \right\} = 0 . \tag{9}$$

Both the frequency ω and the field H may depend explicitly on η. To guarantee that the dynamics (6), whose state at time t can be represented by the afore-defined order parameter $z(t)$ in its continuous form,

$$z(t) = \int_{-\infty}^{\infty} \int_0^{2\pi} \rho(\theta, \eta, t) e^{i\theta} d\theta d\eta , \tag{10}$$

is asymptotically attracted by the OA manifold, it suffices to show that

$$\lim_{t \to +\infty} \int_{-\infty}^{+\infty} \hat{\rho}_+(\theta, \eta, t) \hat{g}(\eta) d\eta = 0 \tag{11}$$

holds. Before showing this, however, we would first like to remark that, without loss of generality, the center of the Lorentzian frequency distribution $\hat{g}(\eta) \sim L(\eta_0 + c, \Delta)$ can be considered zero since we may introduce a change of variables, $\tilde{\theta} = \theta - (\eta_0 t + C(t))$, where $C(t)$ is an antiderivative of $c(t)$. Furthermore, we can adjust (11) by substituting $\hat{g}$ by g.

If $\hat{\rho}_+$ is analytic in the lower half η-plane and decays to zero as $\mathrm{Im}(\eta) \to -\infty$, one can multiply (9) by $g(\eta) d\eta$ and integrate the result by employing the residue theorem. Hence, the integrals can be evaluated at the residue of the enclosed pole of $g(\eta)$ at $\eta = -i\Delta$. We

find

$$\partial_t \hat{\rho}_+(\theta, -i\Delta, t) + \partial_\theta \left\{ -i\Delta \cdot \hat{\rho}_+(\theta, -i\Delta, t) \right.$$

$$\left. + \frac{1}{2i} \left[\int_{-\infty}^{+\infty} H(\eta, t)\hat{\rho}_+(\theta, \eta, t)g(\eta)d\eta \, \mathrm{e}^{-i\theta} - \int_{-\infty}^{+\infty} H^*(\eta, t)\hat{\rho}_+(\theta, \eta, t)g(\eta)d\eta \, \mathrm{e}^{i\theta} \right] \right\} = 0 \ .$$

The two remaining integrals can be determined provided that H and H^* have no singularities in the lower half η-plane and do not increase "too" fast for $\mathrm{Im}(\eta) \to -\infty$. Since g is a Schwartz function, we only need H to diverge at most sub-exponentially. For common choices of H, as listed in Ref. 2, these requirements are met indeed, which yields

$$\partial_t f_+(\theta, t) + \partial_\theta \left[v(\theta, t) f_+(\theta, t) \right] = 0 \ , \tag{12}$$

$$v(\theta, t) = -i \left[\Delta + \frac{1}{2} \left(\mathrm{e}^{-i\theta} H(t) - \mathrm{e}^{i\theta} H^*(t) \right) \right] \ . \tag{13}$$

Here we substituted $f_+(\theta, t) = \hat{\rho}_+(\theta, -i\Delta, t)$ and $H(t) = H(-i\Delta, t)$. These equations agree exactly with Eqs.(17 & 18) in Ref. 2. Hence, following the same reasoning around Eqs.(19-31) in Ref. 2 one can conclude that (11) is fulfilled. To underscore the line of argument, we would like to give a short sketch of the proof. First, by introducing a conformal transformation of the upper half complex θ-plane into the unit disc via $w = \mathrm{e}^{i\theta}$, one can rewrite (12 & 13) as

$$\frac{d}{dt} \tilde{f}_+(w, t) + \tilde{f}_+(w, t)\partial_w \tilde{v}(w, t) = 0 \ , \tag{14}$$

where $\tilde{f}_+$ and $\tilde{v}$ are the transformed functions from (12 & 13), and $d/dt = \partial/\partial t + \tilde{v}\partial/\partial w$. (14) can be integrated using the method of characteristics for linear and homogeneous partial differential equations[25]. Here we require $\tilde{f}_+ \in C^2(\mathbb{R})$ but $\tilde{v}$ does not need to be continuous. This yields

$$\tilde{f}_+(w, t) = \tilde{f}_+(W(w, 0), 0) \exp\left[-\mu(w, t) \right] \ , \tag{15}$$

as solution with

$$\mu(w, t) = \int_0^t \partial_{w'} \tilde{v}(w', t')|_{w'=W(w, t')} \, dt' \ , \tag{16}$$

and the characteristics are given by

$$\partial_{t'} W(w, t') = \tilde{v}(W(w, t'), t') \ , \tag{17}$$

with final condition $W(w, t) = w$. Finally, in order to show that $\tilde{f}_+(w, t) \to 0$ for $t \to \infty$, which, by (15), we prove that

$$\lim_{t\to\infty} \mathrm{Re} \left[\mu(w, t) \right] = +\infty \ . \tag{18}$$

7

The details for the rather lengthy computation can be found in Ref. 2. We here we would only like to mention that the integral in (16) is split into three distinct parts, each of which is evaluated and while two of them remain bounded, the third diverges at the rate Δt, presuming $\Delta > 0$. This eventually completes the proof and underlines the importance that the distribution function $g(\eta)$ must have non-zero width Δ. We would also like to note that in the final step of the proof the continuity of v is required, i.e. H in (13) must be continuous. If one includes, e.g., square functions in the time-dependent parts of the frequency term and/or driving field, one is confronted with jump discontinuities, which become present in the right-hand side of (13) either directly or indirectly via the order parameter $z(t)$. A closer look at Ref. 2, however, confirms that for small jumps the reasoning can be guaranteed and for proper choices of a time constant T their Eq.(31) holds. Thus, we can argue that OA attractiveness will be maintained even in the case of discontinuities, which also confirms our rather long assumption for $c(t)$ to be in $L_{1,\text{loc}}(\mathbb{R})$ in the linear dependence of $\omega(\eta) = a\eta + c$.

So far we only considered a Lorentzian distribution and some linear dependence of ω on η. However, our result can be extended to a much broader class of distribution functions $g(\eta)$, non-linear dependencies $\omega(\eta)$, or even joint distributions $g(\omega, \eta)$ in the case of $\Omega(\omega, \eta)$; see Section IV below. Hence, it is proper to say that the asymptotic attractiveness of the OA manifold for parameter-dependent systems of coupled phase oscillators is generic. Note that the proof remains identical if $\theta = \theta(t)$ does not depend on the parameter η, that is, when there is no correlation between specific oscillators and their dynamics. We call this case "weak" parameter-dependence, which has been coined in several earlier studies, e.g., Refs. 11–13, and 18, where parameters were introduced as auxiliary variables. Our result therefore confirms the attractiveness of the OA manifold also in this case, as has simplifyingly been taken for granted in the afore-cited studies.

III.   NETWORKS OF QIF AND THETA NEURONS

As mentioned above, there is a variety of recent papers that showed numerically how the dynamics of networks of theta neurons is time asymptotically attracted by the OA manifold[4–6]. Recently, Montbrió and co-workers studied how the macroscopic dynamics of a network of quadratic integrate-and-fire (QIF) neurons is described by a low-dimensional system by using a so-called Lorentzian ansatz[7]. By transforming the QIF neurons into

a network of theta neurons, their Lorentzian ansatz does resemble the OA ansatz with parameter-dependent frequency and driving field, as considered in Section II.

To be more precise, the dynamics of the membrane potential V_j of a QIF neuron may be described by

$$\dot{V}_j = V_j^2 + I_j \ , \quad \text{if } V_j \geq V_p \ , \text{ then } V_j \leftarrow V_r \ , \tag{19}$$

for $j = 1, \ldots, N$. Here, I_j denotes an input current, V_p a peak value, and V_r a reset value. Once the membrane potential V_j reaches V_p, the neuron emits a spike, and V_j will be reset to V_r. Commonly, the limit $V_p = -V_r \to \infty$ is considered. The input current I_j consists of a neuron-specific quenched component η_j, a common time-dependent input $I(t)$ and a coupling term $Js(t)$, combining the synaptic weight J and a smooth mean synaptic activation $s(t)$, resulting in

$$I_j = \eta_j + Js(t) + I(t) \ . \tag{20}$$

The latter two time-dependent components are identical for all neurons in the network. In order to describe the macroscopic behavior of the network, Montbrió and co-workers used the Lorentzian ansatz

$$\rho(V|\eta, t) = \frac{1}{\pi} \frac{x(\eta, t)}{[V - y(\eta, t)]^2 - x(\eta, t)^2} \ , \tag{21}$$

with center $y(\eta, t)$ and time-dependent half-width $x(\eta, t)$, which turns out to exhibit the long-term solution for the distribution of the membrane potentials. The properties $x(\eta, t)$ and $y(\eta, t)$ that define the distribution function (21) are also closely linked to the firing rate of the neuronal population and to the mean membrane potential, respectively. While the Lorentzian ansatz applies to the (membrane voltage) dynamics of QIF neurons, we are here primarily interested in the phase dynamics. Using $V_j = \tan(\theta_j/2)$ one can transform (19 & 20) into theta neurons,[26]

$$\dot{\theta}_j = (1 - \cos \theta_j) + (1 + \cos \theta_j) \left[\eta_j + J \cdot s(t) + I(t) \right] \ . \tag{22}$$

In (22) the time-independent injected current η_j is drawn from a distribution function $g(\eta)$. For the sake of legibility we abbreviate the non-autonomous part of (22) as

$$J \cdot s(t) + I(t) = c(t) - 1 \ .$$

Rearranging terms and considering the thermodynamic limit, one can rewrite (22) as

$$\partial_t \theta(\eta, t) = \nu(\theta, \eta, t) = \Omega(\eta, t) + \text{Im} \left[H(\eta, t) \mathrm{e}^{-i\theta} \right] \tag{23}$$

9

with $H(\eta, t) = i(-1 + \eta + Js + I) = i(\eta + c - 2)$ and $\Omega(\eta, t) = \eta + c$; cf. Ref. 7.

To apply our result from above, one has to show that H does not diverge exponentially when $\text{Im}(\eta) \to -\infty$, and that $c(t)$ possesses an antiderivative. On the one hand, for the components of $c(t)$ with $s(t)$ being smooth and $I(t)$ piecewise smooth and (locally) integrable, there will always exist an antiderivative of $c(t)$. On the other hand, we have $H(\eta) = i\eta + const$, such that H grows only linearly for $\text{Im}(\eta) \to -\infty$. Hence, we find that the OA manifold does asymptotically attract the macroscopic behavior of a network of coupled theta neurons. Due to the existence of a conformal mapping between the quantity $w(\eta, t) = x(\eta, t) + iy(\eta, t)$ and the function $\alpha(\eta, t)$ defining the OA manifold $(4)^{27}$, see also Eq.(15) in Ref. 7, we have also proven the attractiveness of the Lorentzian ansatz (21) for a network of QIF neurons.

IV.   GENERAL PARAMETER DISTRIBUTIONS

As already mentioned in Section II, the assumptions of a linear relation between ω and η and of η being drawn from a Lorentzian can be loosened in many respects. We first consider $g(\eta)$ to still be a Lorentzian centered around $\eta = \eta_0$ with width Δ, i.e. $g(\eta) \sim L(\eta_0, \Delta)$. The linear dependency $\omega(\eta, t) = a \cdot \eta + c$ may be generalized by considering both $a = a(t)$ and $c = c(t)$ time-dependent. Then, by the common transformation properties for Lorentzian (Cauchy) distributions, ω follows a Lorentzian of the form $g(\omega) \sim L(a\eta_0 + c, \Delta|a|)$. Let $a \neq 0$ be constant. Then a similar change of variables, $\tilde{\theta} = \theta - (a\eta_0 t + C(t))$, with $C(t)$ being the antiderivative of $c(t)$, keeps the distribution function centered around 0. Without loss of generality we set $a = 1$; even if $a = a(t)$ and $a(t) > 0$ or $a(t) < 0$ for all $t > 0$, the rescaling of θ retrieves that we can stick to our assumption $a = 1$. If, however, a changes sign at, e.g., $t = t_0$, then the scale parameter $\Delta|a|$ tends to zero for $t \to t_0$. Due to (4) also $\rho(\theta, \omega, t)$ will exhibit a δ-peak at $t = t_0$. In this case our results are not readily applicable[3]. However, if $\dot{a}(t_0) \neq 0$, then we can shift the initial time to zero, $t_0 \mapsto 0$. Whenever $\rho_+(\theta, \omega, t_0)$ satisfies the necessary initial conditions, the OA manifold will remain attracting for all $t > t_0$, given that $t_0 = \max\{t \in \mathbb{R} \mid a(t) = 0\}$.

We proceed with more general cases of frequency and parameter distributions. In Ref. 3, the authors elegantly extend the original proof, which considers only Lorentzian frequency distributions: Instead of demanding analytic continuity of both the frequency distribution

10

$g(\omega)$ and the initial condition into the whole lower ω-half plane, it suffices that g and the initial condition have analytic continuations into a strip S defined by $0 \geq \mathrm{Im}(\omega) > -\sigma$ and $-\infty \leq \mathrm{Im}(\omega) \leq +\infty$ with $\sigma > 0$, where neither of them has singularities and both approach zero as $|\omega| \to \infty$. Thereby the class of applicable distribution functions includes Gaussians, sech-distributions, and many more, and even multimodal distributions can be incorporated as long as these functions have finite non-zero widths; see references in Ref. 3. This approach can be adopted and used in our η-parameter-dependent case. For this let us assume again individual oscillators given by (6). As mentioned in Section II, we might be confronted with a nesting of the distributions $\tilde{g}(\omega)$ and $g(\eta)$ for ω and η. In particular, the latter may determine the first in an oscillator-specific way. That is the reason why the resulting distribution function $\hat{g}(\eta)$ can become arbitrarily complicated. However, as long as the analytic continuations of $\tilde{g}$ and g into the strip S (for some $\sigma > 0$ as defined above) do not have singularities, and neither $\tilde{g}$ nor g features a δ-peak in their time-evolutions, also $\hat{g}$ will behave as required. An additional requirement is that the product $H(\eta, t)\hat{g}(\eta)$ satisfies these conditions, too. This means that we have to find a strip $S' \subset S$, defined by $0 < \sigma' \leq \sigma$, in which $H\hat{g}$ has an analytic continuation, does not have singularities, its time evolution does not feature δ-peaks (if necessary we have to reset the initial time point after such a peak), and that we require $|H(\eta_r + i\eta_i, t)\hat{g}(\eta_r + i\eta_i)| \to 0$ for $|\eta_r| \to \infty$ and $0 > \eta_i > -\sigma'$. In particular, H must not grow faster than $\hat{g}$ decays, such that the OA manifold continues to capture the long-term dynamics of the system.

Revisiting the example from Section III, where $H(\eta) = i(\eta + c - 2)$ and $\hat{g}(\eta) \sim \mathrm{L}(\eta_0 + c, \Delta)$, we find that $\hat{g}$ decays exponentially for $|\eta_r| \to \infty$ such that H must not increase at an exponential rate. In fact, H does not have any singularities in the whole complex η-plane (except for $|\eta| \to \infty$), and $H(\eta_r + i\eta_i) = -\eta_i + i\eta_r + const = \mathcal{O}(\eta_r)$ for $|\eta_r| \to \infty$. Consequently, for large $|\eta_r|$, the product $H\hat{g}$ will be dominated by $\hat{g}$ such that all assumptions are fulfilled. Hence, we can confirm again the attractiveness of the OA manifold.[28]

We would like to remark that initial conditions on the oscillator distribution function, $\rho(\theta, \eta, 0)$, play an important role. If they fail to be satisfied, this may hinder the OA manifold to attract the dynamics. For an example we would like to refer to Appendix C of Ref. 29, in which the specific time point has to be determined appropriately in order to set up promising initial conditions.

V. APPLICATIONS – REALISTIC SETTINGS

So far, we only considered non-independent frequency and parameter distributions, $\tilde{g}(\omega)$ and $g(\eta)$, respectively. In general, however, one cannot take this "simple" dependence for granted. The additional parameter might be multi-dimensional, i.e. $\eta \in \mathbb{R}^n$ with $n > 1$. When considering the thermodynamic limit of infinitely many coupled oscillators, the dynamics (6) may obey

$$\partial_t \theta(\eta, t) = \Omega(\omega, \eta, t) + \text{Im}\left[H(\eta, t)\, e^{-i\theta}\right] . \tag{24}$$

Employing the OA ansatz for this system one has to encounter distribution functions given by

$$\rho(\theta, \omega, \eta, t) = \frac{g(\omega, \eta)}{2\pi}\left\{1 + \left[\sum_{k=1}^{\infty} \alpha(\omega, \eta, t)^k e^{ik\theta} + \text{c.c.}\right]\right\}$$
$$\int_{\mathbb{R}^n} \int_{-\infty}^{\infty} \int_0^{2\pi} \rho(\theta, \omega, t)\, d\theta\, d\omega\, d\eta = 1 ; \tag{25}$$

the joint distribution $g(\omega, \eta)$ is a major modification to the setting considered before. Does the OA manifold remain attracting? (24) suggests the phase $\theta = \theta(\eta, t)$ to depend on the parameter η in line with our notion of parameter-dependent systems. But it is unclear whether the OA manifold is attracting even without this particular correlation between phase, natural frequency, and driving field. If, however, the OA attractiveness can be proven for systems with generalized natural frequency Ω and driving field H as in (24), this will allow for a further and even broader extension of the existing theory. In the following we first list a few examples for which numerical simulations have been reported and that give strong incentive that the OA ansatz may indeed be valid. We will show how our proof can be adopted, thereby confirm the OA attractiveness, and set the numerical results on solid ground. Last, we provide some general properties of Ω and H for which the OA ansatz holds.

We start with the Winfree model[15] which is an early mathematical description of synchronization phenomena in large populations of biological oscillators. Rewritten in terms of (24) this model takes the form

$$\partial_t \theta = \Omega(\omega, \eta, t) + \text{Im}\left[H(\eta, t)e^{-i\theta}\right]$$
$$\Omega(\omega, \eta, t) = \omega + \sigma\eta(t), \quad H(\eta, t) = e^{-i\beta}\eta(t), \quad \text{and} \quad \eta(t) = \varepsilon h(t) , \tag{26}$$

where $h(t)$ is a smooth function depending only on the mean field $z(t)$ but not on the phase itself[30]. In particular, this model contains time-dependent parameters, see also Ref. 31.

Next, we consider reaction-diffusion systems with heterogeneous, self-oscillating elements. In particular, we study the mean-field version of the complex Ginzburg-Landau equation, whose equation describes a population of globally coupled limit-cycle oscillators. Hence, we can rewrite the dynamics in form of (24). By introducing a shear (or nonisochronicity) parameter η as an additional random variable and transforming the system through a phase reduction, the governing equations in the continuum limit read[19–21]:

$$\partial_t \theta = \Omega(\omega, \eta, t) + \mathrm{Im}\left[H(\eta, t)\mathrm{e}^{-i\theta}\right]$$
$$\Omega(\omega, \eta, t) = \omega + K\eta \quad \text{and} \quad H(\eta, t) = Kz(1 - i\eta) \,, \tag{27}$$

where K denotes the coupling strength and $z = z(t)$ is the order parameter. The frequency ω and the shear η are drawn from a joint distribution $g(\omega, \eta)$. In contrast to Section II, we explicitly allow the additional parameter η to be drawn from another frequency distribution. For the joint distribution one has to address two scenarios. Either, the random variables are independent, such that the joint distribution can be split into $g(\omega, \eta) = g_1(\omega)g_2(\eta)$, or they are not. Iatsenko and co-workers, who independently investigated the Kuramoto model with both distributed natural frequencies ω and distributed coupling strengths η, coined the term *uncorrelated* joint distributions when the two random variables ω and η are independent, as opposed to *correlated* joint distributions, see Ref. 22, 23, and 32. Furthermore, frequency-weighted coupling[33,34], i.e. the driving field additionally depends on ω, $H = H(\omega, \eta, t)$, can be approached with the formalism introduced above.

As a third point, we will deal with systems that are not all-to-all coupled but exhibit some particular (and sparse) network topology. Therefore, these networks can barely be studied analytically. Although it was conjectured and numerically illustrated by Barlev, Antonsen, and Ott[35] in 2011 that the OA ansatz can be extended for uniform in-degree, Erdös-Rényi, and scale-free networks, a thorough proof has as to yet not been delivered. However, the upcoming branch of heterogeneous mean fields[36] presents a promising loophole to overcome this obstacle of intricate network topologies. We will prove that heterogeneous mean field models indeed fall in a category whose mean field dynamics can be described along the OA ansatz. Given a network with a particular degree distribution, it is possible to introduce so-called degree-block variables, whose dynamics govern the evolution of all nodes which

have the same degree k. This approach reveals the same equations as the annealed networks approximation[17,37], which can hence be considered equivalent. Recent studies considered the heterogeneous mean fields of the Kuramoto model, e.g., on scale-free[38–40] and random Erdös-Rényi networks[38]. The starting point is a specifically coupled Kuramoto network with coupling strength K and adjacency matrix $A = (a_{ij})$ with $i, j = 1, \ldots, N$,

$$\dot{\theta}_j = \omega_j + K \sum_{k=1}^{N} a_{jk} \sin(\theta_k - \theta_j) \, . \tag{28}$$

We can cluster various node dynamics by replacing the adjacency term with an expectation value for their node degree η_j. Ideally, the underlying topology exhibits some well-defined degree distribution $P(\eta)$. In the continuum limit $N \to \infty$, these node degrees are substituted in the phase dynamics as weighted, distributed coupling strengths, so that the governing dynamics read

$$\partial_t \theta(\eta, t) = \Omega(\omega, \eta, t) + \text{Im}\left[H(\eta, t)e^{-i\theta}\right]$$
$$\Omega(\omega, \eta, t) = \omega \quad \text{and} \quad H(\eta, t) = K\eta z(t) \, , \tag{29}$$

where ω and η are drawn from a joint distribution $g(\omega, \eta) = P(\eta)g_1(\omega)$. This setup is amenable to, e.g., random fields, as has been presented in Ref. 40 where oscillators are enforced through local fields, which find their way into the specific forms for Ω and H.

In all these different classes of parameter-dependent networks, we will show how the OA attractiveness can be regained.

A. Winfree model

As said, the Winfree model describes macroscopic synchronization phenomena of large oscillator systems whose individual nodes are naturally pulse-coupled with one another. The introduction of phase response curves (PRC) allows for quantifying how the phase of an oscillator responds to the pulse-like perturbations from the other oscillators. The general form of model reads at the single node level

$$\dot{\theta}_j = \omega_j + Q(\theta_j)\frac{\varepsilon}{N} \sum_{k=1}^{N} P(\theta_k) \, , \tag{30}$$

where ε denotes the coupling strength, Q is the PRC and P is a pulse-like signal. Following the notation of Pazó and Montrbrió in Ref. 30, we consider PRCs with sinusoidal shape,

$$Q(\theta) = \sigma - \sin(\theta + \beta) \, , \tag{31}$$

with an offset parameter σ, and a phase-lag β. Moreover, we assume the pulse-like signal to be smooth,

$$P(\theta) = P_n(\theta) = a_n(1 + \cos\theta)^n \ , \tag{32}$$

with $n \in \mathbb{N}_{\geq 1}$ controlling the width of the pulses, and a_n is a normalizing constant. In the thermodynamic limit, we regain (26) as

$$\partial_t\theta = \omega + \varepsilon\sigma h(t) + \mathrm{Im}\left[\varepsilon e^{-i\beta}h(t)e^{-i\theta}\right] \ , \tag{33}$$

where the coupling function incorporates the smooth mean field

$$h(t) = h_n(t) = \int_0^{2\pi} P_n(\theta)d\theta = 1 + 2(n!)^2 \sum_{k=1}^{n} \frac{\mathrm{Re}(z^k)}{(n+k)!(n-k)!} \tag{34}$$

with z the common (Kuramoto) order parameter (10). The frequency $\Omega(\omega, t) = \omega + c(t)$ with $c(t) = \varepsilon\sigma h(t)$ has a form identical to Section II, where ω follows a Lorentzian frequency distribution $g(\omega)$. Since the order parameter $z(t)$ is bounded with $|\mathrm{Re}(z)| \leq 1$, we have $h(t) \geq 0$ for all $t \geq 0$. Furthermore, the driving field does not depend on additional parameters, so that our proof can be directly applied, confirming that the OA ansatz holds and the OA manifold indeed captures the long-term dynamics of the Winfree model.

An alternative proof for the case of time-dependent frequency and driving field can be found in Ref. 31. However, as we have depicted in Section IV, our proof generalizes their findings and extends them to a broader class of frequency distribution functions $g(\omega)$. Of particular interest in the non-autonomous extension is also the matter of discontinuities. Recall that in Section III we introduced a time-dependent input current $I(t)$, see (20), which can, e.g., take the form of a square function with jump-discontinuities. Our proof applies to this specific feature and confirms existing numerical results[7].

B. Limit-cycle oscillations with shear

Investigating collective synchronization usually addresses networks of coupled elementary oscillatory units. The dynamics of these units may be described as normal form

$$\dot\varrho = \varrho(1 - \varrho^2) \ , \quad \dot\theta = \omega + \eta(1 - \varrho^2) \ , \tag{35}$$

where ϱ denotes the radius and ω determines the frequency of rotation on the stable limit cycle with $\varrho(t) \equiv 1$. The parameter η quantifies the shear, or non-isochronicity, of the flow,

i.e. how strongly perturbations away from the limit cycle modify the phase dynamics. When we consider an all-to-all coupled population of $N \gg 1$ of these oscillatory units, we arrive at the mean-field version of the complex Ginzburg-Landau equation with dissipative coupling

$$\dot{z}_j = z_j \left[1 + i \left(\omega_j + \eta_j \right) - (1 + i\eta_j) |z_j|^2 \right] + \frac{K}{N} \sum_{k=1}^{N} (z_k - z_j) \; ; \qquad (36)$$

$z_j = \varrho_j e^{i\phi_j}$. Heterogeneity among the population is promoted by having the frequency ω_j and shear parameters η_j drawn from a distribution function $g(\omega, \eta)$. In the weakly coupled case, i.e. the coupling strength $|K|$ is small, a phase reduction allows us to describe the dynamics of the system by their phases only. In the continuum limit $N \to \infty$, we can introduce the phase distribution function $\rho(\theta, \omega, \eta, t)$. Note that ω and η are independent, so that neither of them is redundant. Accordingly, the order parameter z takes now the form

$$z(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{0}^{2\pi} \rho(\theta, \omega, \eta, t) e^{i\theta} \, d\theta d\omega d\eta \; . \qquad (37)$$

Thus, the phase dynamics reads

$$\partial_t \theta = \omega + K\eta + \text{Im} \left[K z(t)(1 - i\eta) e^{-i\theta} \right] \; , \qquad (38)$$

and the phase distribution function satisfies the continuity equation

$$\partial_t \rho + \partial_\theta \left(v\rho \right) = 0 \; , \qquad (39)$$

with v the right-hand side of (38), see also Refs. 19 and 21. Using the notion of (24), the frequency and the driving field are both time-varying and depend on the additional shear parameter η:

$$\Omega(\omega, \eta, t) = \omega + K\eta \; , \quad H(\eta, t) = K z(t)(1 - i\eta) \; . \qquad (40)$$

To assure that the OA manifold indeed exhibits the mean field dynamics of this system with shear, we have to adapt our proof from Section IV for the joint distribution $g(\omega, \eta)$.

The general idea is again to decompose the distribution function ρ in Fourier space into

$$\rho(\theta, \omega, \eta, t) = \frac{g(\omega, \eta)}{2\pi} \left[1 + \rho_+(\theta, \omega, \eta, t) + \rho_-(\theta, \omega, \eta, t) \right] \qquad (41)$$

and use symmetry assumptions to focus on ρ_+, which again will be decomposed into $\rho_+ = \hat{\rho}_+ + \hat{\rho}'_+$. While $\hat{\rho}'_+$ lies on the OA manifold and has Fourier coefficients $\hat{\rho}'_{+,n} = [\alpha(\omega, \eta, t)]^n$, $\hat{\rho}_+$ solves

$$\partial_t \hat{\rho}_+ + \partial_\theta \left\{ \left[\Omega(\omega, \eta, t) + \frac{1}{2i} \left(H(\eta, t) e^{-i\theta} - H(\eta, t)^* e^{i\theta} \right) \right] \hat{\rho}_+ \right\} = 0 \; . \qquad (42)$$

The assumptions on the analytic continuation properties of Section IV hold – in particular we need analytic continuations with respect to both ω and η into strips S_ω and S_η. Hence we have to show that

$$\lim_{t\to\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{\rho}_+(\theta,\omega,\eta,t) \; g(\omega,\eta) \; d\omega d\eta = 0 \; . \tag{43}$$

Discussing general solutions of (43) given an arbitrary joint distribution function are beyond the scope of this paper. However, for particular $g(\omega,\eta)$ we can affirm the attractiveness of the OA manifold for these parameter-dependent systems. To begin with, we use the assumption of Montbrió and Pazó that the joint distribution can be written as the product of two Lorentzians[19],

$$g(\omega,\eta) = g_1(\omega)g_2(\eta) = \frac{\delta/\pi}{(\omega-\omega_0)^2 + \delta^2} \frac{\gamma/\pi}{(\eta-\eta_0)^2 + \gamma^2} \; . \tag{44}$$

Multiplying (42) with $g(\omega,\eta)$ and integrating over (ω,η), we can use Fubini's theorem (on the assumption of integrability of $\Omega g \hat{\rho}_+$ and $H g \hat{\rho}_+$) and compute the double integral by changing the order of integration. First, we can evaluate the integral over ω by applying the residue theorem as in Section II and then move on to the second integral, which reads

$$\partial_t \hat{\rho}_+(\theta,\omega_0 - i\delta, -i\gamma, t) =$$
$$-\int_{-\infty}^{\infty} \partial_\theta \left\{ \left[\Omega(\omega_0 - i\delta, \eta, t) + \frac{1}{2i} \left(H(\eta,t)e^{-i\theta} - H(\eta,t)^* e^{i\theta} \right) \right] g_2(\eta) \hat{\rho}_+(\theta, \omega_0 - i\delta, \eta, t) \right\} d\eta \; .$$

While the term $\int \Omega g_2 \hat{\rho}_+$ can be evaluated at the pole $\eta = \eta_0 \pm i\gamma$ ($\pm$ depending on the contour of integration, which again depends on the coupling K, see also Ref. 19) , we have to assure that the product $H(\eta,t)g_2(\eta)$ vanishes for $\text{Im}(\eta) \to \pm\infty$. Indeed, the linear growth of H in η, see (40), will be dominated by the exponential decay of g_2, such that the residue theorem can be applied here, too, which results finally in (12)&(13), from which the claim follows as presented in Section II. As has been shown in Section IV, the restrictions to unimodal Lorentzians can be dropped and the OA attractiveness is sustained. Here we can even handle δ-functions as long as one of the partial distribution functions has finite width: due to the special form of $\Omega(\omega,\eta,t)$, the OA ansatz holds for homogeneous frequencies $\omega_j = \omega$ while the shear is heterogeneous and the coupling $K > 0$ does not vanish.

The case in which the joint distribution $g(\omega,\eta)$ is no longer uncorrelated, i.e. if the first equality in (44) fails, demands a more careful investigation in order to estimate the long-time evolution of $\hat{\rho}_+$. Although the ultimate goal is to categorize adequate joint distributions that

allow for the OA ansatz, there might appear a variety of uncertainties for a general proof. For instance, to the best of our knowledge it is an open problem whether and how singularities can appear in joint distributions given smooth marginal distributions. This issue becomes even more intricate in the case for multi-dimensional parameters $\eta \in \mathbb{R}^n$, $n \in \mathbb{N}$. However, there are certain approaches using the OA ansatz for parameter-dependent systems with correlated joint distributions, which we would like to briefly revise.

The introduction of shear into the oscillator system shows how an additional parameter can be treated as a random variable and thereby changing the natural frequency and driving field of the original Kuramoto model. A more fundamental approach has been presented by Petkoski and co-workers in Refs. 22, 23, 31, and 32: Given the Kuramoto model with heterogeneous natural frequencies, they assume the coupling strengths to be drawn from a distribution function. That is, their model reads

$$\dot{\theta}_j = \omega_j + \frac{K_j}{N} \sum_{k=1}^{N} \sin(\theta_k - \theta_j) \tag{45}$$

with (ω, K) following a joint distribution $g(\omega, K)$. Given the strong resemblance between their numerical simulations and the predictions via the OA ansatz, the authors realized that the latter "formulas were derived on the assumption of at least asymptotic validity of the OA ansatz."[22] They also investigated necessary initial conditions with respect to their analytic continuation and applicability to the OA ansatz. Unfortunately, they did not prove this their system dynamics (45) does not belong the classes of systems considered in the proofs by Ott and Antonsen[1–3]. Recall, a general characterization of correlated joint distribution $g(\omega, K) \neq g_1(\omega)g_2(K)$ that are applicable for the extended OA ansatz is hardly feasible. However, for three examples used in literature we can prove that the OA manifold defines the asymptotic evolution of the whole system.

First, let $g(\omega, K) \sim \delta(K - k) \left[\omega^2 + e^{-\omega^2} \right]^{-1}$, see Fig.1 in Ref. 23. The specific form with the δ-function in K reduces system (45) to the common Kuramoto model with heterogeneous frequencies $\omega \propto g_1(\omega) = \left[\omega^2 + e^{-\omega^2} \right]^{-1}$, which can be dealt with along the proof of the original OA ansatz.

The other two examples are more elaborate in that the joint distribution functions are given by[22]

$$g(\omega, K) = (1 - p)\delta(K - K_1)L(\omega; \omega_0, \gamma_1) + p\delta(K - K_2)L(\omega; -\omega_0, \gamma_2) , \tag{46}$$

with $p \in (0, 1]$, and

$$g(\omega, K) = \Gamma(K) \sum_{n=1}^{N_q} q_n L(\omega; \omega_n, \gamma_n) \quad \text{with} \quad \sum_{n=1}^{N_q} q_n(K) = 1 \ . \tag{47}$$

Here, $L(\omega; \omega_n, \gamma_n)$ denotes a Lorentzian with width $\gamma_n > 0$ and centered around $\omega = \omega_n$, and $\Gamma(K)$ is a multimodal-δ-function. For properly chosen $q_{1,2}$ the distributions (47) can be regarded a generalization of (46) so that it suffices to consider the former. For simplicity, let $N_q = 2$, i.e. $g(\omega, K)$ be a bimodal joint distribution. Inserting $g(\omega, K)$ in the definition of the order parameter (37), we can decompose the latter into $z(t) = q_1 z_1(t) + q_2 z_2(t)$ with $q_1 + q_2 = 1$. Put differently, we can view our system as two all-to-all coupled populations with population-specific coupling strengths $K_{1,2}$. Given that the frequency distributions are Lorentzians of finite width $\gamma_{1,2}$, the results for two-population/bimodal Kuramoto models as in Refs. 41–43 can be readily applied, which confirms the attractiveness of the OA manifold for this kind of joint distributions. Note that we do not require $q_n \in [0, 1]$ but may choose, e.g., $q_1 = \delta/(\delta - \xi) > 1$ and $q_2 = -\xi/(\delta - \xi) < 0$. Then, the bimodal distribution results from one Lorentzian being subtracted from the other one, which, in principle allows the central minimum between the two peaks to converge to zero[42]. The case of multiple Kuramoto populations with specific coupling strengths can be approached by transforming the system into one global system whose oscillators' frequencies follow a multimodal distribution consisting of weighted inhomogeneous unimodal distributions, which can mirror the underlying coupling topology across populations[44].

Admittedly, the aforementioned examples are not exhaustive let alone complete. They represent a concise set of a broad variety of joint distribution functions. Nevertheless, we believe that our results may be a major breakthrough for the applicability of the OA ansatz for systems with more intricate distribution functions. First extensions concentrated on a multiple-population-approach and have been presented in Refs. 13, 45, and 46. Skardal and Restrepo[45] focused on hierarchical synchrony effects in modular networks and investigated how local and global synchrony evolve differently by allowing for different subpopulation sizes, heterogeneous intra- and inter-population coupling strengths as well as population-specific frequency distributions. Reformulating their approach results in (47). This establishes rigorously the agreement of the predictions by the OA ansatz and their numerical results. So and co-workers, on the other hand, aimed for synchronization criteria in a network of two coupled populations with static and time-varying coupling topologies[46]. Their

governing equations can be cast into (46) when additionally considering $K = K(t)$ to be time-dependent. Also, they numerically determined macroscopic chaos by assuming a single network with bimodally distributed natural frequencies[13]. Combining our results from this section together with the preceding part where we incorporated non-autonomicity, we again corroborate the numerous numerical findings by providing the ingredients to prove the implicit assumption that the OA ansatz holds for these kinds of parameter-dependent and non-autonomous systems.

C. Heterogeneous mean field models

While the general case of uncorrelated joint distributions has already been covered in the preceding Section V B, we would like to concentrate on the specific derivation of the heterogeneous mean field model. Recall the standard Kuramoto model on a given network,

$$\dot{\theta}_j = \omega_j + K \sum_{k=1}^{N} a_{jk} \sin(\theta_k - \theta_j) \; , \tag{48}$$

where K is the coupling strength and the adjacency matrix is given by $A = (a_{ij})_{i,j=1,\dots,N}$. We substitute the adjacency values $a_{jk} \in \{0,1\}$ by their expectation values $\langle a_{jk} \rangle \in [0,1]$, which are given by

$$\langle a_{jk} \rangle = \frac{\eta_j \eta_k}{N \langle \eta \rangle} \; . \tag{49}$$

Introducing the complex order parameter as

$$z = \frac{1}{N \langle \eta \rangle} \sum_{k=1}^{N} \eta_k \mathrm{e}^{i\theta_k} \; ,$$

the dynamics for all nodes with the same degree η_k read

$$\dot{\theta}_k = \omega_k + K\eta_k \mathrm{Im}(z\mathrm{e}^{-\theta_k}) \; .$$

In this special form, in which the single nodes are replaced by block-degree variables, we returned to the all-to-all coupling. For a given degree distribution $P(\eta)$ property (49) also holds in the continuum limit $N \to \infty$ where the governing dynamics read

$$\partial_t \theta(\eta, t) = \omega + \mathrm{Im} \left[K\eta z(t) \mathrm{e}^{-i\theta} \right] \; , \tag{50}$$

with ω and η being drawn from a joint distribution $g(\omega, \eta) = P(\eta)g_1(\omega)$. As above we can introduce a phase distribution function $\rho(\theta, \omega, \eta, t)$, which fulfills the continuity equation

$\partial_t \rho + \partial_\theta(v\rho) = 0$ with v the right-hand side of (50). Note, however, that depending on the underlying network topology and its degree distribution $P(\eta)$, one has to choose the domain of η properly. In the case of a scale-free network, the degree distribution follows $P(\eta) \propto \eta^{-\gamma}$ with $\gamma > 1$. Hence the normalization conditions for the distribution function ρ obey

$$\int_1^\infty \int_0^{2\pi} \rho(\theta, \omega, \eta, t) \, d\theta d\eta = g_1(\omega) \quad \text{and} \quad \int_{-\infty}^\infty \int_0^{2\pi} \rho(\theta, \omega, \eta, t) \, d\theta d\omega = P(\eta) \; .$$

We can apply the OA ansatz as above. By the same reasoning as in Section V B, we can so prove the OA attractiveness for heterogeneous mean field models, rendering also non-globally coupled oscillator networks applicable to have their mean field dynamics evolved on a low-dimensional manifold.

Before elaborating more on coupling schemes other than global coupling, we briefly discuss further topological network effects such as nodal correlations between in and out degrees, correlations between nodal frequencies and degrees, and degree as well as so-called frequency assortativity in the formation of links. Recent numerical findings by Restrepo, Ott, and Skardal[47,48] exploited assortative networks and gave strong incentive to believe that their dimensionality reduction techniques along the OA ansatz do capture the dynamics of the full network. An assortativity function $a_{\mu' \to \mu}$ represents the probability that a link exists from an oscillator with target property μ' to one with property μ. Using this one can indicate an exact instruction on how to construct a network model of the form (48). The nodal properties μ are chosen in such a way that the network displays, e.g., a particular degree[47] ($\mu = k$), frequency[48] ($\mu = \omega_0$), or even a combined ($\mu = \{k, \omega_0\}$) assortativity. Key ingredient for relating this to the OA ansatz is the reformulation of the order parameter. We first define

$$z(\mu, t) = \sum_{\mu'} P_{\mu'} a_{\mu' \to \mu} \iint \rho_{\mu'}(\theta, \omega, t) e^{i\theta} d\theta d\omega \; , \tag{51}$$

where $\rho_{\mu'}(\theta, \omega, t) = \rho(\theta, \omega, \mu', t)$ is the common phase distribution function with target property μ', see (25), $P_{\mu'} = P_p(\mu')$ is a normalized target property distribution, and $a_{\mu' \to \mu} = a(\mu' \to \mu)$ the assortativity function; for details see Refs. 47 and 48. Then, we can integrate over all possible properties μ – note that we write the sum over the target properties μ' also in integral form – and we arrive at the order parameter

$$z(t) = \frac{1}{\langle \eta \rangle} \iint P_p(\mu) P_p(\mu') a(\mu' \to \mu) \iint \rho(\theta, \omega, \mu', t) e^{i\theta} d\theta d\omega \, d\mu' d\mu, \tag{52}$$

with $\langle \eta \rangle$ the average degree. Given a specific degree distribution $P(\eta)$ one finally ends up with (50), from which we can follow the lines of argument as presented above to complete the proof. The addition of assortativity in the network topology enriches the existing theory further. It discloses many new qualitative effects on the dynamics such as transitions between steady state, periodic, quasiperiodic attractors, and even macroscopic chaos may emerge without external driving or time-varying parameters.

D.   Non-local coupling

Two months before Ott and Antonsen published their ansatz, Ko and Ermentrout investigated the creation of partially locked states in a network of identical all-to-all coupled oscillators due to inhomogeneous coupling[49]. Instead of heterogeneity of the oscillators' frequencies, it was the coupling heterogeneity that led to partial synchronization. Carlo Laing analytically investigated this network of globally coupled oscillators with coupling strengths drawn from a power-law distribution[50] along the line of the OA ansatz – recall the resemblance to the heterogeneous mean field approach for scale-free networks. Assuming "nearly" identical oscillators, i.e. the frequencies ω were drawn from a Lorentzian with width $0 < \Delta \ll 1$, he could verify the earlier results that were derived via a self-consistency argument[49], and extend them by including a thorough bifurcation analysis. Our findings in Section V B put these results on a solid mathematical ground.

Of particular interest is Laing's work on a ring of oscillators[50,51]. For a given ring topology, the typical coupling scheme is neither local neighbor-to-neighbor, nor global coupling. Instead, the oscillators are non-locally coupled via a coupling kernel G. We assume that each oscillator $j = 1, \ldots, N$ has some fixed spatial position $x_j \in [-\pi, \pi]$, a natural frequency ω_j drawn from a continuous distribution function $g(\omega)$ with non-zero width, and interacts with the others depending on the distance between their sites modulo periodic boundary conditions. The governing dynamics read

$$\dot{\theta}_j = \omega_j + \frac{2\pi}{N} \sum_{j=1}^{N} G(x_k - x_j) \sin(\theta_k - \theta_j + \alpha) , \tag{53}$$

where α is a phase-lag parameter and $G \colon \mathbb{R} \to \mathbb{R}$ a continuous even and 2π-periodic coupling function[52]. We retrieve global coupling, if $G \neq 0$ is constant. Commonly used coupling functions G are of exponential form $G(x) \sim \mathrm{e}^{-\kappa|x|}$ with $\kappa > 0$, or of trigonometric form

$G(x) = 1/2\pi(1 + A\cos x + B\sin x)$ with $A > 0, B \geq 0$. The reflection symmetry of G is lost for $B \neq 0$. In the continuum limit, the velocity field (7) becomes

$$\partial_t \theta = \omega + \mathrm{Im}\left[H(x,t)\mathrm{e}^{-i\theta}\right] \; ,$$

$$H(x,t)\mathrm{e}^{i\alpha} = \int_{-\pi}^{\pi} G(x-y) \int_{-\infty}^{\infty} \int_{0}^{2\pi} \rho(\theta,y,\omega,t)\mathrm{e}^{i\theta} \; d\theta d\omega dy \; . \tag{54}$$

While the inner two integrals have the form of a local complex order parameter $z(y,t)$, measuring the synchronization degree of oscillators around y, we can interpret the last integral as a convolution of the local order parameter with the (spatial coupling) kernel G. In particular, we can regard the dynamics $\partial_t \theta(x,t)$ of an oscillator at position x as being controlled by the local mean field $H(x,t)$. Unlike the case of global coupling, the order parameter has become space-dependent and thus the driving field. However, a similar "physical picture" as for global coupling is valid: practically we deal with an assembly of independent oscillators under the control of a common forcing field[52,53]. We now go a step further and interpret the space variable x as a subpopulation index[54]. Equivalent to the block-degree variables in the heterogeneous mean field approach, we consider the subpopulation index as a parameter that follows a particular, in this case a uniform, distribution function. Hence, (54) represents the governing dynamics of a parameter-dependent system, for which we proved the OA attractiveness in the preceding sections.

E. External forcing and time delay

Already in their original work, Ott and Antonsen proposed that their ansatz extends to external forcing and the incorporation of time delays. However, recent results that leaned against the OA ansatz for tackling more intricate issues of external forcing and/or time delays went beyond the reach of the original proof. Therefore, we first revisit the existing theory and revise the proof appropriately with the concepts introduced above.

Ott and Antonsen approached the forced Kuramoto model,

$$\dot{\theta}_j = \omega_j + \frac{K}{N}\sum_{k=1}^{N}\sin(\theta_k - \theta_j) + \eta\sin(\varpi t - \theta_j) \; , \tag{55}$$

in Ref. 1. Rearranging terms, moving in a rotating frame, $\theta \to \theta + \varpi t$, and considering the

23

thermodynamic limit, the velocity field reads

$$\partial_t \theta = \Omega(\omega, \varpi) + \text{Im}\left[H(\omega, K, \eta, t)e^{-i\theta}\right] \ ,$$

$$\Omega(\omega, \varpi) = \omega - \varpi \ , \quad H(\omega, K, \eta, t) = Kz(t) + \eta \ ,$$

(56)

with $z(t)$ the common Kuramoto order parameter. While Ott and Antonsen provided a proof for systems with constants K, ϖ, and η, conjoining thereby the numerical findings and the extensive analysis by Childs and Strogatz[55], we generalized their proof extensively in Section II. By this, the additional parameters ϖ and η that characterize the forcing can be both random and time-dependent variables. This adaptation renders a more detailed analysis of, for instance, the circadian rhythm problem possible. One extension has been published very recently addressing the east-west asymmetry of jet-lag[56], where a discontinous phase quantity p is added to model the travel across time-zones. The adapted model reads

$$\dot{\theta}_j = \omega_j + \frac{K}{N}\sum_{k=1}^{N}\sin(\theta_k - \theta_j) + \eta\sin(\varpi t - \theta_j + p(t)) \ ,$$

(57)

where p jumps from one constant value to another depending on the corresponding time-zone. In particular, $p(t)$ is locally integrable, which allows for a thorough analytic description of how the human organism may adapt after several cross-time-zone travels. This extends the existing work where the authors solely focused on the recovery dynamics of circadian rhythms after a single travel "shock".

To address the presence of time delays, let us first concentrate on time-delayed coupling, i.e. the response of oscillator j at time t depends on the state of another oscillator k at time $t - \tau_{kj}$. Here, τ_{kj} is some specific delay time for the interaction. In general, the single oscillator dynamics may be given by

$$\dot{\theta}_j(t) = \omega_j + \frac{K}{N}\sum_{k=1}^{N}\sin(\theta_k(t - \tau_{kj}) - \theta_j(t)).$$

(58)

There already exists a plethora of studies considering the case in which τ_{kj} follows some given distribution function $h(\tau)$, e.g., Refs. 50, 57, and 58. That the OA ansatz also holds in this case, has been proven by Ott and Antonsen in Ref. 2, generalizing their original idea of identical time delays, $\tau_{kj} = \tau'$ for all $j, k = 1, \ldots, N$, i.e. $h(\tau) = \delta(\tau - \tau')$, see Ref. 1. The driving field H of the original velocity field (3a) is replaced by

$$H = K\int_{\Omega} h(\tau)z(t - \tau)d\tau \ ,$$

24

where $\Omega \subset \mathbb{R}$ is the domain of the time delay distribution h and $z(t)$ the common Kuramoto order parameter. Slightly more elaborate and not covered by Ott and Antonsen's original proof is the extension to so-called coupling adaptation[59]. The coupling strength is no longer constant but slowly adapts depending on the current coupling strength and the delayed order parameter. As long as the function that models the adaptation process is locally integrable, our extended proof guarantees the OA attractiveness for such time-varying parameter-dependent systems. For this reason we believe that the mainly numerical work by Skardal and co-workers[59] can also be analytically substantiated. This will not only contribute to exploring the underlying phenomena of explosive synchronization[60,61], but also enhance the modeling of information processing and memory effects, for which network adaptation is crucial[62–64].

VI.   RELAXATION DYNAMICS

As discussed, we allow time-varying parameters to affect the oscillator dynamics. The change of parameters comes with its time scale(s). The change can be periodic. This periodicity may also influence the evolution of the mean field and thereby the OA manifold. Therefore, the relation between this periodicity and the characteristic time of the system to approach the manifold needs to be investigated. If the relaxation dynamics onto the manifold is way slower than the characteristic time scale of the time-varying manifold itself, then our findings will remain true for the limit $t \to \infty$. They are, however, of minor interest for describing the transient behavior of the mean field. Several numerical results[7,22,23,31] suggest that the relaxation to the OA manifold is reasonably fast, in some cases even instantaneous. To address this analytically, we briefly recall the proof for the attractiveness from Section II. After having Fourier expanded the phase distribution function $\rho(\theta, \eta, t)$, and then decomposed the positive Fourier modes into a part that already lies on the manifold, $\hat{\rho}'_+$, and a residual part $\hat{\rho}_+$, we showed how the latter converged to zero in a weak sense, cf. (11). We can extract the relaxation time to the OA manifold from out of the proof: From (12)&(13) we obtain a solution $f_+(\theta, t) = \hat{\rho}'_+(\theta, -i\sigma, t)$, with $\sigma' > \sigma > 0$ where $\hat{\rho}'_+(\theta, \eta, t)$ admits an analytic continuation into the strip $S = \{\eta \in \mathbb{C} \mid -\infty \leq \mathrm{Re}(\eta) \leq \infty , \ 0 \geq \mathrm{Im}(\eta) \geq -\sigma'\}$;

the solution (15) obeys

$$\tilde{f}_+(w,t) = \tilde{f}_+(W(w,0),0) \exp\left[-\mu(w,t)\right] \ ,$$

hence the relaxation time τ is by definition

$$const \cdot \exp(-t/\tau) = \exp\left[-\mu(w,t)\right] \quad \Rightarrow \quad \tau = \frac{t}{\text{Re}\left[\mu(w,t)\right]} \ . \tag{59}$$

Put differently, $\text{Re}\left[\mu(w,t)\right]$ scales with σt, such that $\tau = 1/\sigma$. The wider the frequency distribution becomes, the larger σ can be chosen. Thus, one may argue that the characteristic time scale decreases with increasing heterogeneity among the single oscillators. This relation has already been noted for a particular example of a Lorentzian frequency distribution by Ott and Antonsen in Ref. 1. It has been investigated in more detail by Petkoski and Stefanovska for the non-autonomous Kuramoto model[31]. Interestingly, there is an intrinsic relation between the frequency inhomogeneity and the coupling strength. Therefore, at critical coupling strengths, which distinguish different dynamical regimes, the relaxation times tend to infinity, which has been reported independently by Petkoski et al.[31] and Yoon et al.[39] for the full Kuramoto network, its non-autonomous version and the heterogeneous mean field model.

For the non-autonomous case we would like to mention that the proof presented in Section II entirely holds for continuous time-varying parameters. Introducing discontinuities in either the frequency Ω and/or the driving field H, however, will eventually lead to a non-continuous right-hand side of (13) – due to H itself, or via the order parameter z, which absorbs the time-varying part of Ω and influences H directly or indirectly. While employing the method of characteristics still can be performed, estimating the integral in (16) cannot exploit the continuity assumption and a proper evaluation has to be circumvented. In spite of this sinister outlook, numerical results remain promising; for instance, the simulations in Ref. 7 with a square input function (Fig.2a,c,e,g). A possible way to overcome this obstacle might be to approximate the jumps by smooth sigmoid functions, which might be valid as long as the height of the jumps is lower than their length. Another more rigorous approach might be to find weak solutions for (12 & 13) and estimate their long-time behavior. There, a starting point could be the very recent results by Dietert, Fernandez and co-workers, who investigated stability properties of different dynamical regimes of the Kuramoto model in a mathematically rigorous way, confirming the exponential decay to the manifold[65–67]. More details are way beyond the scope of our paper.

Interestingly, the approach by Dietert and others is based on the idea of "Landau damping" in plasma physics. Strogatz, Mirollo and Matthews were the first who incorporated this concept in order to understand relaxation dynamics of the Kuramoto model[68,69]. They showed that for frequency distributions $g(\omega)$ supported on the whole real axis, the decay towards the incoherent state is exponentially fast for coupling strengths below the critical threshold, $K < K_c$. If $g(\omega)$ has compact support, i.e. g is non-zero only on a compact interval $[-\gamma, \gamma] \subset \mathbb{R}$, $0 < \gamma < \infty$, the rate may be considerably slower, even polynomial. In the example they used to illustrate their result, the authors assumed the frequencies ω to be distributed uniformly on $\mathcal{I} = [-\gamma, \gamma]$, i.e. $g(\omega) = 1/2\gamma$ if $\omega \in \mathcal{I}$, and 0 otherwise. The jump discontinuities of g on $\partial\mathcal{I}$, however, prohibited an analytic continuation of g into a strip S in the lower complex ω-plane, contradicting the required conditions for applying the OA ansatz[3]. That is why the proofs above cannot be applied here, and our argumentation about the relaxation times remains unaffected.

Last but not least, we would like to add that decay times typically depend on initial conditions. Pikovsky and Rosenblum pointed out that for identical macroscopic, i.e. mean field, initial conditions the microscopic initial states can lead to very different transient dynamics towards the OA manifold, see Section 3.2 in Ref. 18. A more thorough investigation about this specific topic has not been undergone, yet, but might shed light on the underlying dynamics of the microscopic variables of large oscillatory systems in contrast to its mean field behavior.

VII. DISCUSSION AND CONCLUSION

The OA ansatz has proven rather fruitful for investigating the macroscopic behavior of systems of coupled phase oscillators in terms of a low-dimensional system. Although parameter dependence has already been mentioned in Ott and Antonsen's original work, parameters were merely considered auxiliary variables and the velocity field was required to incorporate the phase only through a sinusoidal coupling term.

Our main result was to prove that the η-dependence sustains the time-asymptotic attractiveness of the OA manifold for systems of coupled oscillators. For this we required that the driving field H does not have singularities in the complex η-plane and that it diverges at most sub-exponentially for $\text{Im}(\eta) \to -\infty$, next to the conditions in the original Ott &

Antonsen formulation[1,2]. Furthermore, we assumed the frequency $\omega(\eta, t)$ to be linear in η. We were able to depict the proof step by step. Subsequently we loosened the restrictive assumptions and showed that our results remain valid for a much broader class of distribution functions $g(\eta)$ as well as more complex dependencies of the driving field $H(\eta)$ and the natural frequencies $\omega(\eta)$ on the parameter η.

Although the main idea of introducing a common parameter η was to correlate the driving field and the natural frequency with their specific oscillator, our proof is identical for the case when η does only influence the mean field dynamics. By this, we have proved the claim in Ref. 3 that the OA manifold remains attractive in the "weak" parameter-dependent case when H depends on "other non-phase-oscillator variables obeying auxiliary dynamical systems."

Common choices of H and ω usually fulfill the aforementioned assumptions as stated in Section II. That is, our result can be immediately applied in a variety of circumstances. Here, we highlighted an application in mathematical neuroscience. By this, our findings strengthen the theory of coupled theta neurons: The many recent numerical findings of Ref. 7 and the references therein are finally set in a solid mathematical framework. Moreover, the link between QIF neurons and theta neurons has been underscored by proving the attractiveness of the Lorentzian ansatz.

We generalized and extended existing proofs for non-autonomous systems. In particular we addressed the Winfree model, which is biologically more realistic than the Kuramoto model and therefore closer to applications. We also addressed coupled oscillatory systems with an additional shear parameter, another important tool to render the Kuramoto model more realistic. The major novelty was our rigorous proof of the OA attractiveness for systems with uncorrelated joint distribution functions when more parameters than only the natural frequencies are treated as a random variable. This finding opened the way for networks with specific underlying coupling topologies other than the restrictive global coupling. Using the heterogeneous mean field approach, we showed how these networks can be treated along the OA ansatz. First steps were also taken in the direction of correlated joint distributions.

All in all, we consider the explicit dependence on an additional parameter η of both the oscillator's phase and the (non-sinusoidal) components an important extension introducing an intrinsic relation between phase, frequency, and driving field of an oscillator. The latter two are correlated with the phase so that the η-dependence does not allow for applying the

original theory.

Still, there are several open problems concerning the mean field dynamics of an oscillatory system and its description by a low-dimensional system. A first urgent one is the case of δ-peaked frequency distributions. Numerical simulations[70] and heuristic arguments hint at convergence of the OA manifold, where a proper mathematical derivation is omitted under the pretence of "nearly identical oscillators"[51,71,72]. A thorough proof would render the OA ansatz rigorously applicable to "chimera states", a topic that is particularly en vogue; see, e.g., the recent review paper by Panaggio and Abramscite[73]. Importantly, such a proof has to circumvent the main argument of Ott and Antonsen's original proof, where the width $\Delta > 0$ of the distribution $g(\omega)$ allowed for a consequent evaluation of the mean field dynamics. On the other hand, Pikovsky and Rosenblum[74] already showed that more complicated dynamics can emerge from the OA manifold when describing the system along the Watanabe-Strogatz (WS) ansatz[75]. Deviations from the OA ansatz appear only if the WS constants of motion are not uniformly distributed over the whole domain, but only over a compact subset. Given (a) the direct correspondence between the constants of motion and the initial conditions of phases in the OA ansatz[18,75], and (b) the necessary requirements on (analytic continuation properties of) the initial conditions, it may be worth investigating the influence of nonuniform distributions of the constants of motion and whether this may hinder the initial conditions of phases to satisfy the requirements of the OA ansatz.

Another intriguing open problem is whether the mean field dynamics is attracted by a low-dimensional manifold when the parameter-dependence of the frequency and driving field is extended by an explicit dependence on the phase. A recent example is given by Laing[76], who considered the driving field H to follow a dynamics that explicitly depends on the phase θ. This system exhibits partial synchronization patterns, which are also covered by the OA ansatz, but any attempt to apply the OA ansatz has been avoided "due to the dynamics of the extra variables."[76]

When the coupling term incoporates higher harmonics, see, e.g., Refs. 77 and 78, no low-dimensional analytic solution for the mean field evolution has been found. This is another open question whether further generalizations of the work of Ott and Antonsen [1] can be rigorously manifested. We believe that our current proof for parameter-dependent networks is a good starting point for tackling these important issues.

ACKNOWLEDGMENTS

REFERENCES

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

[2] E. Ott and T. M. Antonsen, "Long time evolution of phase oscillator systems," Chaos **19**, 023117 (2009).

[3] E. Ott, B. R. Hunt, and T. M. Antonsen, "Comment on 'Long time evolution of phase oscillator systems' [Chaos 19, 023117 (2009)]," Chaos **21**, 025112 (2011).

[4] T. B. Luke, E. Barreto, and P. So, "Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons," Neural Comput. **25**, 3207–3234 (2013).

[5] P. So, T. B. Luke, and E. Barreto, "Networks of theta neurons with time-varying excitability: Macroscopic chaos, multistability, and final-state uncertainty," Physica **267D**, 16–26 (2014).

[6] C. R. Laing, "Derivation of a neural field model from a network of theta neurons," Phys. Rev. E **90**, 010901 (2014).

[7] E. Montbrió, D. Pazó, and A. Roxin, "Macroscopic description for networks of spiking neurons," Phys. Rev. X **5**, 021028 (2015).

[8] C. R. Laing, "Phase oscillator network models of brain dynamics," Computational Models of Brain and Behavior. Wiley-Blackwell (2016).

[9] Á. Byrne, M. J. Brookes, and S. Coombes, "A mean field model for movement induced changes in the beta rhythm," submitted.

[10] S. Coombes and Á. Byrne, "Next generation neural mass models," submitted to: Lecture Notes in Nonlinear Dynamics in Computational Neuroscience.

[11] S. A. Marvel, R. E. Mirollo, and S. H. Strogatz, "Identical phase oscillators with global sinusoidal coupling evolve by möbius group action," Chaos **19**, 043104 (2009).

[12]A. Wagemakers, E. Barreto, M. A. F. Sanjuán, and P. So, "Control of collective network chaos," Chaos **24**, 023127 (2014).

[13]P. So and E. Barreto, "Generating macroscopic chaos in a network of globally coupled phase oscillators," Chaos **21**, 033127 (2011).

[14]Parameter-dependent systems comprise a wide class of systems, from which we here only choose a single family. This family represents a rather weak parameter-dependent system. However, we refrain from this notion since weak parameter-dependence would imply that parameter changes have little to no considerable effect. Here, the original proof by Ott and Antonsen has to be changed, such that the parameter effect can be strong. We use the attribute "weak" to highlight that specific oscillator does not depend on the additional parameter but its mean field dynamics only.

[15]A. T. Winfree, "Biological rhythms and the behavior of populations of coupled oscillators," Journal of theoretical biology **16**, 15–42 (1967).

[16]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**, 137–185 (2005).

[17]F. A. Rodrigues, T. K. D. Peron, P. Ji, and J. Kurths, "The Kuramoto model in complex networks," Physics Reports **610**, 1 – 98 (2016).

[18]A. Pikovsky and M. Rosenblum, "Dynamics of heterogeneous oscillator ensembles in terms of collective variables," Physica **240D**, 872–881 (2011).

[19]E. Montbrió and D. Pazó, "Shear diversity prevents collective synchronization," Phys. Rev. Lett. **106**, 254101 (2011).

[20]D. Pazó and E. Montbrió, "The Kuramoto model with distributed shear," EPL (Europhysics Letters) **95**, 60007 (2011).

[21]E. Montbrió and D. Pazó, "Collective synchronization in the presence of reactive coupling and shear diversity," Phys. Rev. E **84**, 046206 (2011).

[22]D. Iatsenko, S. Petkoski, P. V. E. McClintock, and A. Stefanovska, "Stationary and traveling wave states of the Kuramoto model with an arbitrary distribution of frequencies and coupling strengths," Phys. Rev. Lett. **110**, 064101 (2013).

[23]D. Iatsenko, P. V. E. McClintock, and A. Stefanovska, "Glassy states and super-relaxation in populations of coupled phase oscillators," Nature Comm. **5**, 4118 (2014).

[24]Alternatively, the dependence $\omega(\eta)$ may be constituted by considering ρ as a conditional

probability density $\rho(\theta, \omega, t | \eta) = \rho(\theta, \omega | \eta, t)$ in line with Ref. 7.

[25] L. C. Evans, *Partial differential equations* (American Mathematical Society, Providence (R.I.), 1998).

[26] G. B. Ermentrout and N. Kopell, "Parabolic bursting in an excitable system coupled with a slow oscillation," SIAM Journal on Applied Mathematics **46**, 233–253 (1986).

[27] We substituted ω by η in line with our arguments in Section II. However, we do assume an implicit dependence $\omega = \omega(\eta)$.

[28] To give a brief idea of the proof, it is important to note that, next to the assumption that in S' the product $H\hat{g}$ decays to zero for $|w| \to \infty$, the crucial point for proving the attractiveness of the OA manifold is that $\sigma' > 0$. To be more precise, given the integral expression (equivalent to) (11), the idea is to shift the path of integration from the real η-axis to the line $\eta_r + i\eta_i$ with $0 > \eta_i > -\sigma'$, $-\infty \le \eta_r \le \infty$, for details see Ref. 3. This leads directly to (12 & 13) from where one can complete the proof along the known formalism outlined in Section II.

[29] E. Ott, J. H. Platig, T. M. Antonsen, and M. Girvan, "Echo phenomena in large systems of coupled oscillators," Chaos **18**, 037115 (2008).

[30] D. Pazó and E. Montbrió, "Low-dimensional dynamics of populations of pulse-coupled oscillators," Phys. Rev. X **4**, 011009 (2014).

[31] S. Petkoski and A. Stefanovska, "Kuramoto model with time-varying parameters," Phys. Rev. E **86**, 046212 (2012).

[32] S. Petkoski, D. Iatsenko, L. Basnarkov, and A. Stefanovska, "Mean-field and mean-ensemble frequencies of a system of coupled oscillators," Phys. Rev. E **87**, 032908 (2013).

[33] H. Wang and X. Li, "Synchronization and chimera states of frequency-weighted Kuramoto-oscillator networks," Phys. Rev. E **83**, 066214 (2011).

[34] C. Xu, Y. Sun, J. Gao, T. Qiu, Z. Zheng, and S. Guan, "Synchronization of phase oscillators with frequency-weighted coupling," Scientific reports **6** (2016).

[35] G. Barlev, T. M. Antonsen, and E. Ott, "The dynamics of network coupled phase oscillators: An ensemble approach," Chaos **21**, 025103 (2011).

[36] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," Nature Physics **8**, 32–39 (2012).

[37] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks," Rev. Mod. Phys. **80**, 1275–1335 (2008).

[38] B. Coutinho, A. Goltsev, S. Dorogovtsev, and J. Mendes, "Kuramoto model with frequency-degree correlations on complex networks," Phys. Rev. E **87**, 032106 (2013).

[39] S. Yoon, M. S. Sindaci, A. Goltsev, and J. Mendes, "Critical behavior of the relaxation rate, the susceptibility, and a pair correlation function in the Kuramoto model on scale-free networks," Phys. Rev. E **91**, 032814 (2015).

[40] M. Lopes, E. Lopes, S. Yoon, J. Mendes, and A. Goltsev, "Synchronization in the random field Kuramoto model on complex networks," arXiv preprint arXiv:1605.04733 (2016).

[41] E. A. Martens, E. Barreto, S. Strogatz, E. Ott, P. So, and T. Antonsen, "Exact results for the Kuramoto model with a bimodal frequency distribution," Phys. Rev. E **79**, 026204 (2009).

[42] D. Pazó and E. Montbrió, "Existence of hysteresis in the Kuramoto model with bimodal frequency distributions," Phys. Rev. E **80**, 046215 (2009).

[43] B. Pietras, N. Deschle, and A. Daffertshofer, "Coupled networks and networks with bimodal frequency distributions are equivalent," arXiv preprint arXiv:1602.08368 (2016).

[44] B. Pietras, N. Deschle, and A. Daffertshofer, "On the equivalence of coupled and multi-modal networks," to-be-submitted.

[45] P. S. Skardal and J. G. Restrepo, "Hierarchical synchrony of phase oscillators in modular networks," Phys. Rev. E **85**, 016208 (2012).

[46] P. So, B. C. Cotton, and E. Barreto, "Synchronization in interacting populations of heterogeneous oscillators with time-varying coupling," Chaos **18**, 037114 (2008); E. Barreto, B. Hunt, E. Ott, and P. So, "Synchronization in networks of networks: The onset of coherent collective behavior in systems of interacting populations of heterogeneous oscillators," Phys. Rev. E **77**, 036107 (2008).

[47] J. G. Restrepo and E. Ott, "Mean-field theory of assortative networks of phase oscillators," EPL **107**, 60006 (2014).

[48] P. S. Skardal, J. G. Restrepo, and E. Ott, "Frequency assortativity can induce chaos in oscillator networks," Phys. Rev. E **91**, 060902 (2015).

[49] T.-W. Ko and G. B. Ermentrout, "Partially locked states in coupled oscillators due to inhomogeneous coupling," Phys. Rev. E **78**, 016203 (2008).

[50] C. R. Laing, "The dynamics of chimera states in heterogeneous Kuramoto networks," Physica **238D**, 1569–1588 (2009).

[51] C. R. Laing, "Chimera states in heterogeneous networks," Chaos **19**, 013113 (2009).

[52]O. E. Omel'chenko, "Coherence–incoherence patterns in a ring of non-locally coupled phase oscillators," Nonlinearity **26**, 2469 (2013).

[53]Y. Kuramoto and D. Battogtokh, "Coexistence of coherence and incoherence in nonlocally coupled phase oscillators," Nonlinear Phenom. Complex Syst. **5**, 380–5 (2002).

[54]M. Wolfrum, O. E. Omel'chenko, S. Yanchuk, and Y. L. Maistrenko, "Spectral properties of chimera states," Chaos **21**, 013112 (2011).

[55]L. M. Childs and S. H. Strogatz, "Stability diagram for the forced Kuramoto model," Chaos **18**, 043128 (2008).

[56]Z. Lu, K. Klein-Cardea, S. Lee, T. M. Antonsen, M. Girvan, and E. Ott, "Resynchronization of circadian oscillators and the east-west asymmetry of jet-lag," Chaos **26**, 094811 (2016).

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

[58]S. Petkoski, A. Spiegler, T. Proix, P. Aram, J.-J. Temprado, and V. K. Jirsa, "Heterogeneity of time delays determines synchronization of coupled oscillators," Phys. Rev. E **94**, 012209 (2016).

[59]P. S. Skardal, D. Taylor, and J. G. Restrepo, "Complex macroscopic behavior in systems of phase oscillators with adaptive coupling," Physica **267D**, 27 – 35 (2014).

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

[61]M. M. Danziger, O. I. Moskalenko, S. A. Kurkin, X. Zhang, S. Havlin, and S. Boccaletti, "Explosive synchronization coexists with classical synchronization in the Kuramoto model," Chaos **26**, 065307 (2016).

[62]P. D. Lellis, M. D. Bernardo, F. Sorrentino, and A. Tierno, "Adaptive synchronization of complex networks," Int. J. Comput. Math. **85**, 1189–1218 (2008).

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

[64]Y.-H. Eom, S. Boccaletti, and G. Caldarelli, "Concurrent enhancement of percolation and synchronization in adaptive networks," Scientific Reports **6**, 27111 (2016).

[65]H. Dietert, "Stability and bifurcation for the Kuramoto model," Journal de Mathématiques Pures et Appliquées **105**, 451–489 (2016).

[66]B. Fernandez, D. Gérard-Varet, and G. Giacomin, "Landau damping in the Kuramoto

model," in *Annales Henri Poincaré* (Springer, 2015) pp. 1–31.

[67] H. Dietert, B. Fernandez, and D. Gérard-Varet, "Landau damping to partially locked states in the Kuramoto model," arXiv preprint arXiv:1606.04470 (2016).

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

[69] S. H. Strogatz, "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators," Physica **143D**, 1–20 (2000).

[70] O. E. Omel'chenko, M. Wolfrum, and C. R. Laing, "Partially coherent twisted states in arrays of coupled phase oscillators," Chaos **24**, 023102 (2014).

[71] E. A. Martens, C. Bick, and M. J. Panaggio, "Chimera states in two populations with heterogeneous phase-lag," Chaos (in press); arXiv preprint arXiv:1606.01871 (2016).

[72] C. R. Laing, "Disorder-induced dynamics in a pair of coupled heterogeneous phase oscillator networks," Chaos **22**, 043104 (2012).

[73] M. J. Panaggio and D. M. Abrams, "Chimera states: coexistence of coherence and incoherence in networks of coupled oscillators," Nonlinearity **28**, R67 (2015).

[74] A. Pikovsky and M. Rosenblum, "Partially integrable dynamics of hierarchical populations of coupled oscillators," Phys. Rev. Lett. **101**, 264103 (2008).

[75] S. Watanabe and S. H. Strogatz, "Constants of motion for superconducting Josephson arrays," Physica **74D**, 197–253 (1994).

[76] C. R. Laing, "Chimeras in networks with purely local coupling," Phys. Rev. E **92**, 050904 (2015).

[77] P. S. Skardal, E. Ott, and J. G. Restrepo, "Cluster synchrony in systems of coupled phase oscillators with higher-order coupling," Phys. Rev. E **84**, 036208 (2011).

[78] Y. Terada and T. Aoyagi, "Dynamics of two populations of phase oscillators with different frequency distributions," Phys. Rev. E **94**, 012213 (2016).