

# Distributed Synchronization of Heterogeneous Oscillators on Networks with Arbitrary Topology

Enrique Mallada, Randy A. Freeman, and Ao Tang

## Abstract

Many network applications rely on the synchronization of coupled oscillators. For example, such synchronization can provide networked devices with a common temporal reference necessary for coordinating actions or decoding transmitted messages. In this paper, we study the problem of using distributed control to achieve both phase and frequency synchronization of a network of coupled heterogeneous nonlinear oscillators. Not only do our controllers guarantee zero phase error in steady state under arbitrary frequency heterogeneity, but they also require little knowledge of the oscillator nonlinearities and network topology. Furthermore, we provide a global convergence analysis, in the absence of noise and propagation delay, for the resulting nonlinear system whose phase vector evolves on the  $n$ -torus.

## Index Terms

Synchronization, coupled oscillators, control of networks, distributed control, nonlinear control.

## I. Introduction

Achieving temporal coordination among different networked devices is a fundamental requirement for the successful operation of many engineering systems. For example, it is necessary in communication systems for recovering transmitted messages [1], in sensor networks for coordinating wake up cycles [2] or achieving temporal measurement coherence [3], and in computer networks for preserving the causality of distributed events [4]. Almost ubiquitously, such coordination is accomplished by providing each node of the network with its own local oscillator and then compensating its phase and frequency (using information received from other devices on the network) to achieve a common temporal reference.

E. Mallada is with Caltech, CMS Dept. 1200 E California Blvd, Pasadena, CA 91125, USA, mallada@caltech.edu.

R. Freeman is with Northwestern U., EECS Dept., 2145 Sheridan Rd., Evanston, IL 60208-3118, USA, free-man@eecs.northwestern.edu. His work was supported in part by a grant from the Office of Naval Research.

A. Tang is with Cornell U., School of ECE, Ithaca, NY 14853, USA, atang@ece.cornell.edu. His work was supported in part by a grant from the Office of Naval Research.

Legacy applications such as public switched telephone networks and cellular networks use a centralized hierarchical synchronization scheme with high-precision oscillators having relative frequency errors ranging from 0.01 to 4.6 parts per million (ppm) [5, 6]. For several reasons, however, these traditional synchronization architectures have become increasingly unsuitable for newer applications such as wireless sensor networks. For example, traditional methods can break down with the failure of only a few nodes. In addition, many newer applications use inexpensive oscillators having errors as high as 100 ppm [7]. Thus, a synchronization protocol designed for these newer applications should satisfy two essential requirements: it should be distributed and independent of the network topology (each node should use only its neighbors' oscillator information to adjust its own oscillator), and it should be robust to wide variations and uncertainty in the specifications of the oscillators used throughout the network.

A variety of synchronization algorithms have been proposed along these lines, jointly inspired by collective synchronization in physics and biology [8–11] and cooperative control in engineering networks [12, 13]. One possible solution is to use monotonically increasing time sources (e.g. clocks) and update their times based on offset information [14–22] to achieve a common absolute time reference (clock synchronization). This is suitable for applications in computer networks where a reference to an absolute time is needed (e.g. distributed databases). Another solution is to use periodic time sources (e.g. oscillators) interconnected with phase comparators [23–25] or pulse-coupling [26–28], where the objective is to achieve common relative time reference (phase synchronization) that allows temporal coordination within the network (e.g. waking up simultaneously).<sup>1</sup>

While the theoretical study of clock synchronization is fairly mature, with solutions that can provide zero offset error synchronization on networks with arbitrary heterogeneous frequencies [4] and asynchronous updates [22], little is known about the phase synchronization counterpart. For example, most phase synchronization solutions present nonzero steady state phase differences in the presence of frequency heterogeneity [23, 24, 26–28] with convergence guarantees limited to idealized scenarios such as homogeneous frequencies [29]. The only exception is [25] which can guarantee phase synchronization for complete graph topologies. Thus whether or not such systems can synchronize for arbitrary networks and arbitrary frequency heterogeneity has remained as an open question [24].

In this paper, we provide a positive answer to this question under very general conditions. We propose two distributed controllers that can achieve phase synchronization for a network of arbitrarily interconnected oscillators, under mild assumptions on the oscillator and phase comparator characteristics. For example, we allow the instantaneous frequency of each oscillator to be a highly uncertain nonlinear function of the local control input, a model consistent with most analog oscillators (such as voltage-controlled oscillators or

<sup>1</sup>Although one can use clock synchronization to achieve phase synchronization by simply mapping the linear times onto the circle using a modulo operator, this approach can lead to undesirable transients if the phases keep wrapping around the circle as the linear times synchronize. For this reason we consider these as separate synchronization problems, each suitable for different application areas.

CMOS oscillators). Also, unlike existing work, we allow the set of oscillator frequencies to be bounded, so that each oscillator may operate within a prescribed frequency range, even during the transient part of the response. Finally, we allow flexibility in the choice of the phase comparator responses, rather than assuming as in [25] that they are sinusoidal. We only require that the measured phase difference is noiseless and can be obtained without propagation delay.

The main contribution of the paper is a novel nonlinear convergence analysis that leverages recent results on the stability of equilibria of homogeneous-frequency coupled oscillators [29]. In particular, our controllers are based on a Hamiltonian dynamic system defined on the graph in which each local minimum of the energy function represents a synchronized trajectory. Each controller employs a different mechanism to dissipate energy and thereby converge to a synchronized solution. Furthermore, we show that any trajectories that are synchronized in frequency but not in phase must be unstable.

## II. Notation and terminology

We let  $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$  denote the unit circle, regarded as the Lie group of angle addition. We equip  $\mathbb{T}$  with the usual Riemannian metric which defines the distance between two points to be the length of the shorter of the two arcs connecting them (so that  $\text{diam}(\mathbb{T}) = \pi$ ). For  $p \in \mathbb{N}$ , we let  $\mathbb{T}^p$  denote the Cartesian product of  $p$  circles. For  $i \in \{1, \dots, p\}$ , we let  $\partial_i$  denote the unit vector field pointing in the counterclockwise direction on the  $i^{\text{th}}$  factor of  $\mathbb{T}^p$  (which we write simply as  $\partial$  when  $p = 1$ ). Because these unit vector fields form an ordered basis for the tangent space of  $\mathbb{T}^p$  at each point, we can represent tangent vectors for  $\mathbb{T}^p$  as elements of  $\mathbb{R}^p$ , that is, as coordinate vectors with respect to this basis. Moreover, all Jacobian matrices of mappings defined on  $\mathbb{T}^p$  will be representations of the differential with respect to this basis.

All graphs in this paper will be simple, undirected, connected graphs having  $n$  vertices (with  $2 \leq n < \infty$ ) and  $m$  edges (with  $m \geq n - 1$ ). We represent such a graph  $\mathbb{G}$  as a pair  $\mathbb{G} = (\mathcal{V}, \mathcal{E})$  for a vertex set  $\mathcal{V}$  and edge set  $\mathcal{E}$ . We label and order the vertices and edges, writing  $\mathcal{V} = \{1, \dots, n\}$  and  $\mathcal{E} = \{1, \dots, m\}$ , where each edge  $k \in \mathcal{E}$  is an unordered pair of distinct vertices  $k = \{i, j\} \subset \mathcal{V}$ . For each vertex  $i \in \mathcal{V}$ , we let  $\mathcal{N}_i$  denote the following indexed set of neighbors of  $i$ :

$$\mathcal{N}_i = \{(j, k) \in \mathcal{V} \times \mathcal{E} : k = \{i, j\}\}. \quad (1)$$

Thus  $(j, k) \in \mathcal{N}_i$  if and only if  $(i, k) \in \mathcal{N}_j$ , that is, if and only if edge  $k$  connects vertices  $i$  and  $j$ .

## III. Problem statement and results

We consider a network of controlled oscillators in which each oscillator shares the current value of its phase with its immediate neighbors. The purpose of the controller design is to guarantee both frequency and phase synchronization of the interconnected system. We adopt the classical phase-locked loop (PLL) structure for each controlled oscillator [23, 30]. This structure consists of three components connected in feedback, as illustrated in Fig. 1: a base oscillator, a phase comparator, and a loop filter. The base oscillator is a physical device (such as a voltage-controlled oscillator) whose frequency is determined dynamically by

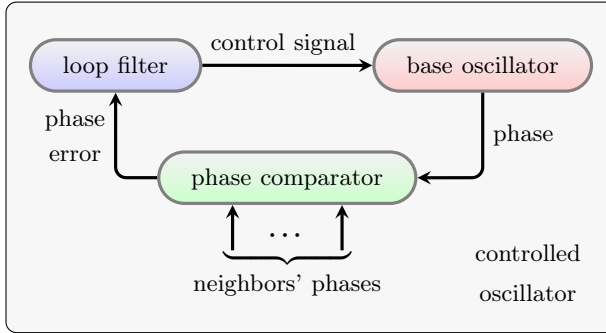


Fig. 1: PLL components of a controlled oscillator.

means of a control signal; the phase of each controlled oscillator system is simply the phase of its base oscillator. The phase comparator produces a phase error signal by comparing its own phase with the phases of its neighbors. Finally, the loop filter produces the control signal from the phase error.

We represent the network of oscillators by a graph  $\mathbb{G} = (\mathcal{V}, \mathcal{E})$  in which each vertex is a controlled oscillator and each edge indicates an exchange of phases between neighboring vertices. We represent the phase of oscillator  $i$  at time  $t \in \mathbb{R}$  by  $\varphi_i(t) \in \mathbb{T}$ , and we write the phase vector signal for the entire network as the column vector

$$\varphi = [\varphi_1 \dots \varphi_n]^T \in \mathbb{T}^n. \quad (2)$$

We define a synchronization measure  $\delta\varphi$  to be the diameter of the finite subset of  $\mathbb{T}$  consisting of the  $n$  phases  $\varphi_i$ :

$$\delta\varphi = \text{diam}\{\varphi_1, \dots, \varphi_n\} \in [0, \pi]. \quad (3)$$

Note that if  $\delta\varphi < \frac{2\pi}{3}$ , then the  $n$  phases  $\varphi_i$  all lie within an arc of length  $\delta\varphi$ . Therefore a small value of  $\delta\varphi$  represents a tight clustering of the oscillator phases, and these phases are all identical when  $\delta\varphi = 0$ . This diameter (3) is a worst-case measure of synchronization, rather than an average measure, as a single outlier can make this diameter large. The following definition describes our design goal:

**Definition 1.** Let  $\mathcal{I} \subset \mathbb{R}$  be an interval. The network of oscillators achieves almost-global synchronization within  $\mathcal{I}$  when for almost all initial states (including both the initial oscillator phases and any initial loop filter states), the trajectories of the system satisfy:

asymptotic frequency synchronization: there exists a constant  $\omega^* \in \mathcal{I}$  such that  $\dot{\varphi}_i(t) \rightarrow \omega^*$  as  $t \rightarrow \infty$  for each  $i \in \mathcal{V}$ ;

asymptotic phase synchronization:  $\delta\varphi(t) \rightarrow 0$  as  $t \rightarrow \infty$ ;

constrained frequencies:  $\dot{\varphi}_i(t) \in \mathcal{I}$  for all  $t \geq 0$  and each  $i \in \mathcal{V}$ ;

internal boundedness: all loop filter states (if any) are bounded in forward time.

The constraint interval  $\mathcal{I}$  used in this definition characterizes the desired range of frequencies for the oscillators. Choosing an appropriate interval  $\mathcal{I}$  in the design is thus useful for preventing the oscillators

from moving too fast, too slow, or reversing direction.

#### A. Base oscillator and phase comparator

In common models of controlled analog oscillators, the instantaneous frequency is some uncertain monotonic nonlinear function of the control input. For example, the graph of this function is the “tuning curve” seen on data sheets of many voltage-controlled oscillators, where it is understood that this curve is typical rather than exact. Motivated by such models, we assign an uncertain frequency function  $\chi_i$  to the base oscillator in each vertex  $i$ , where  $\chi_i : \mathcal{U} \rightarrow \mathbb{R}$  is a strictly increasing  $C^1$  function on a known open interval domain  $\mathcal{U} \subset \mathbb{R}$ . The angular velocity of the phase of the oscillator at time  $t$  (that is, its instantaneous frequency) is then given by  $\chi_i(u_i(t))$ , where  $u_i(\cdot)$  is a  $\mathcal{U}$ -valued control signal. Hence our model for the base oscillator is the differential equation

$$\dot{\varphi}_i = \chi_i(u_i). \quad (4)$$

In this oscillator model, the oscillator stops whenever  $\chi_i(u_i) = 0$  and reverses direction whenever  $\chi_i(u_i)$  changes sign. However, an oscillator need not admit such behavior, as the image  $\chi_i(\mathcal{U})$  need not contain zero.

As we see from the model (4), the base oscillators are nonlinear and heterogeneous. In addition, we do not need precise knowledge of the frequency functions  $\chi_i$  to complete our design and guarantee almost-global synchronization. In fact, as we will see when we list our assumptions in Section III-E below, all we need to know about the functions  $\chi_i$  in the unconstrained case  $\mathcal{I} = \mathbb{R}$  is that they are  $C^1$  with positive derivatives on  $\mathcal{U}$ , and that the intersection of their images  $\chi_i(\mathcal{U})$  is nonempty.

The phase comparator in vertex  $i$  calculates a linear combination of functions of phase differences to produce a phase error  $e_i$ :

$$e_i(\varphi) = \sum_{(j,k) \in \mathcal{N}_i} a_k f(\varphi_j - \varphi_i), \quad (5)$$

where the constants  $a_k$  are positive edge weights and  $f : \mathbb{T} \rightarrow \mathbb{R}$  is the phase coupling function. For example, the coupling function used in many classical PLL designs is the sine function  $f = \sin$ .

#### B. Example: the Kuramoto model

Suppose that the frequency function  $\chi_i$  for each base oscillator  $i$  has the simple affine form  $\chi_i(u_i) = \omega_i + u_i$ , where the constant  $\omega_i \in \mathbb{R}$  represents the nominal oscillator frequency. Suppose further that the phase coupling function is the sine function  $f = \sin$  and that all loop filters are constant unity gains so that  $u_i \equiv e_i$ . Then the controlled phase equation becomes

$$\dot{\varphi}_i = \omega_i + \sum_{(j,k) \in \mathcal{N}_i} a_k \sin(\varphi_j - \varphi_i). \quad (6)$$

This model of a network of coupled oscillators has been studied extensively, and we refer the reader to the survey paper [31]. In particular, if the graph  $\mathbb{G}$  is complete, and if all edge weights  $a_k$  are the same, then this is the famous Kuramoto model of coupled oscillators [32].

The oscillator network characterized by the dynamics in (6) fails to meet the design goal of Definition 1. Indeed, the existence of a synchronized trajectory for the model (6) in which  $\delta\varphi \equiv 0$  implies that the nominal frequencies  $\omega_i$  are all identical. If these frequencies were known precisely, then we could simply cancel them out via control by setting  $u_i = e_i + \omega^* - \omega_i$  to obtain a model of the form (6) in which  $\omega_i = \omega^*$  for all  $i \in \mathcal{V}$ . However, even in this case of identical frequencies, almost-global synchronization has been proved only for special classes of connected graphs, such as complete graphs or trees [31]. Instead, we are interested in oscillators having unknown heterogeneous frequencies on arbitrary connected graphs. In this case one can choose sufficiently large edge weights  $a_k$  to guarantee “practical” phase synchronization in which  $\delta\varphi$  becomes small, provided it does not start off too large [31]. However, choosing large edge weights makes it less likely that the frequencies  $\dot{\varphi}_i$  will be constrained to a desired interval  $\mathcal{I}$  during the transient. In any case, we see that we must depart from this standard model (6) to meet the design goal of Definition 1, and we do so by choosing a non-sinusoidal coupling function  $f$  in (5) and a nonlinear dynamic loop filter.

### C. Toward synchronization: a Hamiltonian system

The loop filter in vertex  $i$  produces the control signal  $u_i$  from the phase error  $e_i$ . It is well known that introducing integral action into the loop filter can compensate offset mismatches for networks of heterogeneous oscillators [19, 21–23, 25, 30]. Thus as a first attempt at achieving almost-global synchronization of the oscillator network, we simply make the loop filter a scaled integrator:

$$\dot{\gamma}_i = c_i e_i(\varphi) \quad (7)$$

$$u_i = \zeta(\gamma_i), \quad (8)$$

where  $\gamma_i(\cdot)$  is the real-valued internal filter state,  $c_i$  is a positive parameter, and  $\zeta : \mathbb{R} \rightarrow \mathcal{U}$  is a scaling function which squeezes the value of  $\gamma_i$  into the domain  $\mathcal{U}$  of the frequency function  $\chi_i$ . We do not assume that the oscillators have access to the global time variable  $t$ . As a result, the differential  $dt$  used in the construction of  $d\gamma_i/dt$  in (7) is unknown, and we account for this by assuming that the positive constant  $c_i$  is unknown.

The system resulting from (4) and (7)–(8) is

$$\dot{\varphi}_i = \chi_i(\zeta(\gamma_i)) \quad (9)$$

$$\dot{\gamma}_i = c_i e_i(\varphi), \quad (10)$$

where  $e_i$  is the phase error in (5). The first thing we notice about this system (9)–(10) is that it admits a synchronized solution at any frequency  $\omega^* \in \mathbb{R}$  which belongs to the image of each function  $\chi_i \circ \zeta$ . Indeed, choose any initial states such that  $\delta\varphi(0) = 0$  and  $\chi_i(\zeta(\gamma_i(0))) = \omega^*$  for each  $i \in \mathcal{V}$ ; then  $\dot{\varphi}_i \equiv \omega^*$  and  $\dot{\gamma}_i \equiv 0$  for every  $i \in \mathcal{V}$ , which implies  $\delta\varphi \equiv 0$ . However, solutions starting from other initial conditions will not converge to such synchronized trajectories, which means this system still fails to meet the design goal of Definition 1.

The second thing we notice about the system (9)–(10) is that it is a Hamiltonian system with position variables  $\varphi_i$  and momentum variables  $\gamma_i/c_i$ . Indeed, assuming the phase coupling function  $f$  is odd, the  $n$ -vector of phase errors  $e_i$  is the negative gradient of a potential function of the phase vector  $\varphi$  (as we will show in Section IV-A). Furthermore, each function  $(\chi_i \circ \zeta)/c_i$ , being a scalar function of a scalar variable, is trivially the gradient of a potential function of its argument  $\gamma_i$ . The sum of these potential functions is a Hamiltonian associated with the dynamics (9)–(10). Moreover, as we will see in Section IV-A, if we let the position variables be the phases  $\varphi_i$  measured relative to an appropriate rotating frame, then the system is still Hamiltonian, but now the Hamiltonian is proper and nonnegative and is thus a Lyapunov function candidate. If we can guarantee that synchronized trajectories represent the only local minima of this Lyapunov function, then we can perturb these Hamiltonian dynamics with dissipation terms in the controller to achieve almost-global synchronization. This is the control design strategy we will pursue in this paper.

#### D. The PI loop filter: a perturbed Hamiltonian system

We propose two different perturbed Hamiltonian systems in this paper, the first now and the second later on in section III-F. For our first perturbation, we add a proportional term to the integral control (8), resulting in a proportional-integral (PI) loop filter of the form

$$\dot{\gamma}_i = c_i e_i(\varphi) \quad (11)$$

$$u_i = \zeta(e_i(\varphi) + \gamma_i). \quad (12)$$

Such a loop filter (without our nonlinear scaling function  $\zeta$ ) can be found in classical PLL designs [30] as well as in various coupled oscillator systems [23, 25]. More generally, the error term  $e_i$  in (12) can be replaced with a scaled error term  $\kappa_i e_i$ , where the “proportional gain”  $\kappa_i$  is a positive constant (or even a positive function of  $\gamma_i$ ), with virtually no change in the convergence proof. We have left this gain  $\kappa_i$  out of the analysis to simplify notation, but the flexibility it adds will be important for tuning the performance of the system. We might also introduce a corresponding “integral gain,” but for the analysis this can be absorbed into the constant  $c_i$ .

To summarize, each controlled oscillator in the network has second-order dynamics of the form

$$\dot{\varphi}_i = \chi_i(\zeta(e_i(\varphi) + \gamma_i)) \quad (13)$$

$$\dot{\gamma}_i = c_i e_i(\varphi) \quad (14)$$

with

$$e_i(\varphi) = \sum_{(j,k) \in \mathcal{N}_i} a_k f(\varphi_j - \varphi_i). \quad (15)$$

The state space of each such oscillator is the cylinder  $\mathbb{T} \times \mathbb{R}$ . We next present conditions under which this oscillator system exhibits almost-global synchronization within a given interval  $\mathcal{I}$ .

### E. Assumptions and main result

We assume that each vertex knows its neighbors in the graph  $\mathbb{G}$  (as they need to exchange phase information), but otherwise the graph is unknown. However, we do assume the following:

(A1) the graph  $\mathbb{G}$  is connected, and there exists a known upper bound on the number  $n$  of vertices in  $\mathbb{G}$ .

The frequency functions  $\chi_i$  of the base oscillators are uncertain; we merely assume that they satisfy the following:

(A2) the frequency functions  $\chi_i : \mathcal{U} \rightarrow \mathbb{R}$  are all  $C^1$  with positive derivatives  $\chi'_i : \mathcal{U} \rightarrow (0, \infty)$ , and are such that

$$\bigcap_{i \in \mathcal{V}} \chi_i(\mathcal{U}) \neq \emptyset. \quad (16)$$

It is clear from the oscillator model (4) that (16) is necessary for the existence of phase trajectories having synchronized frequencies. This condition (16) implies the existence of some common interval of possible base oscillator frequencies, and thus places an inherent limit on the extent to which these oscillators can differ from each other. Indeed, if one oscillator can only produce frequencies in the kilohertz range and another only in the megahertz range, then there is no possibility of synchronization.

We next assume that our loop filter scaling function  $\zeta$  satisfies:

(A3) the scaling function  $\zeta : \mathbb{R} \rightarrow \mathcal{U}$  is  $C^1$  with positive derivative  $\zeta' : \mathbb{R} \rightarrow (0, \infty)$ , and is such that

$$\bigcap_{i \in \mathcal{V}} \chi_i(\zeta(\mathbb{R})) \neq \emptyset \quad (17)$$

and

$$\bigcup_{i \in \mathcal{V}} \chi_i(\zeta(\mathbb{R})) \subset \mathcal{I}, \quad (18)$$

where  $\mathcal{I}$  is the constraint interval from Definition 1.

It is clear from (4) that (18) constrains the frequencies  $\dot{\varphi}_i$  to the interval  $\mathcal{I}$  as required by the design goal in Definition 1. If  $\mathcal{I}$  is large enough to contain the union of the images  $\chi_i(\mathcal{U})$ , then we can always satisfy assumption (A3) by choosing the scaling function  $\zeta$  to be a diffeomorphism onto  $\mathcal{U}$  (so that (16) and (17) are the same). However, if  $\mathcal{I}$  is not that large, then assumption (A3) states that we have found a solution to the problem of designing  $\zeta$  to satisfy both (17) and (18) based on some a priori knowledge about the set of possible frequency functions  $\chi_i$ . Such a design problem could very well have no solution if  $\mathcal{I}$  is too small.

We let  $f' : \mathbb{T} \rightarrow \mathbb{R}$  denote the derivative of the phase coupling function  $f$  in the direction of the unit vector field  $\partial$ . We make two assumptions on this function  $f$ :

(A4) the phase coupling function  $f : \mathbb{T} \rightarrow \mathbb{R}$  is  $C^1$  and odd, that is,  $f(-\theta) = -f(\theta)$  for all  $\theta \in \mathbb{T}$ ;

(A5) there is a constant  $b \in (0, \frac{\pi}{n-1}]$  such that  $f'(\theta) > 0$  whenever  $\cos(\theta) > \cos(b)$  and  $f'(\theta) < 0$  whenever  $\cos(\theta) < \cos(b)$ , for any  $\theta \in \mathbb{T}$ .



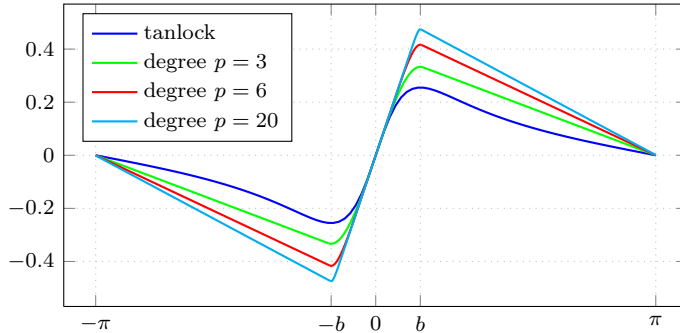


Fig. 2: Examples of phase-coupling functions  $f$  when  $b = 0.5$ .

Assumptions (A4)–(A5) also appeared in [33, 34] and play an important role in the convergence analysis: (A4) allows the interpretation of (4) and (7)–(8) as a Hamiltonian system, while (A5) is needed to guarantee convergence to the desired solution.

Note that to choose such a parameter  $b$ , we must use our assumed knowledge in (A1) of a known upper bound on  $n$ . Examples of functions  $f$  which satisfy assumptions (A4)–(A5) are shown in Fig. 2 for  $b = 0.5$ . The first example is given by the  $C^\omega$  formula

$$f(\theta) = [1 - \cos(b)] \frac{\sin(\theta)}{1 - \cos(b) \cos(\theta)}. \quad (19)$$

This function is related to the characteristic of certain “tanlock” phase comparators [30], and it generates the sinusoidal coupling  $f = \sin$  when  $b = \frac{\pi}{2}$ . The other examples are  $C^1$  and piecewise polynomial of various degrees  $p \geq 1$ , each having a derivative given by  $f'(\theta) = 1 - (|\theta|/b)^{p-1}$  on a certain arc containing  $[-b, b]$  and a constant derivative on its complement. All of these examples are normalized to have unit derivative at zero, which means they should result in similar performance for small deviations around a stable synchronized trajectory. When  $b$  is small (which we require when  $n$  is large), the magnitude of the derivative of the tanlock function is small on the arc  $[b, \pi]$  when compared to the magnitude of the derivatives of the piecewise-polynomial functions. As a result, the piecewise-polynomial functions might provide faster convergence to a synchronized state when some initial phase differences are greater than  $b$  (due to their larger gains for large phase differences).

The final assumption is on the choice of the edge weights  $a_k$ :

(A6) each edge weight  $a_k$  is chosen at random from a continuous probability distribution on the interval  $(0, \infty)$ .

This assumption allows us to state that, with probability one, we avoid an unknown zero-measure set of bad edge weight vectors  $a = [a_1 \dots a_m]^T \in \mathbb{R}^m$  for which our stability analysis does not guarantee convergence.

**Theorem 2.** Assume (A1)–(A6). Then with probability one in the selection of edge weights in (A6), the network of oscillators with vertex dynamics (13)–(15) achieves almost-global synchronization within  $\mathcal{I}$ .

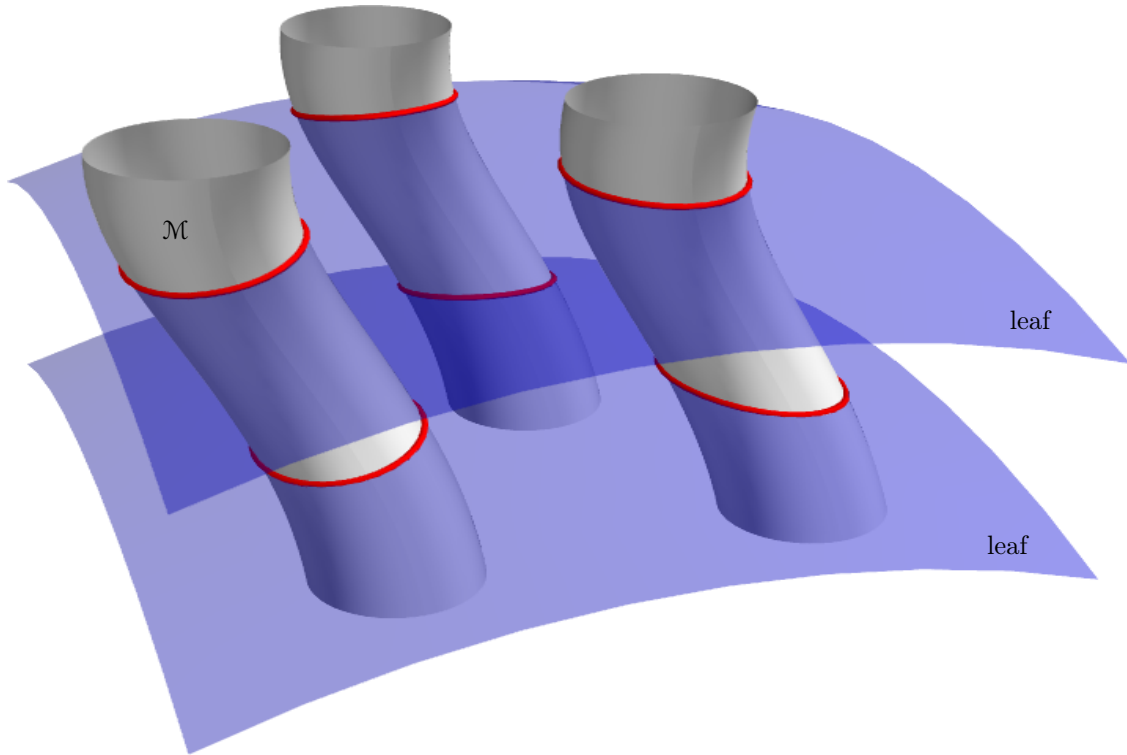


Fig. 3: The invariant set  $\mathcal{M}$  (shown in gray) is a collection of isolated cylinders. The intersection of each invariant leaf of the foliation (shown in blue) with  $\mathcal{M}$  is a collection of isolated circles (shown in red).

Note that when  $n \geq 4$ , assumption (A5) rules out the sinusoidal phase-coupling function  $f = \sin$ . In fact, the requirement  $b \leq \frac{\pi}{n-1}$  is within a factor of two of being necessary. Indeed, one can show that if  $b > \frac{2\pi}{n}$ , then for some connected graphs having  $n$  vertices there exists a nonempty open set of bad edge weight vectors such that trajectories of (13)–(15) starting from a set of initial states having positive measure achieve asymptotic frequency synchronization but not asymptotic phase synchronization.

We will prove Theorem 2 in Section IV. In this proof, we first construct a Lyapunov function for the system, and then apply the Krasovskii-LaSalle invariance theorem to show that all trajectories converge to a certain invariant subset  $\mathcal{M}$  of the state space  $\mathbb{T}^n \times \mathbb{R}^n$ . We then determine the structure of this set  $\mathcal{M}$ , showing first that it is the disjoint union of isolated embeddings of the cylinder  $\mathbb{T} \times \mathbb{R}$ . We next show that the state space admits a foliation into invariant  $(2n-1)$ -dimensional submanifolds, and that the intersection of each leaf of this foliation with the set  $\mathcal{M}$  is the disjoint union of isolated embeddings of the circle  $\mathbb{T}$  (as conceptualized in Fig. 3). Each such circle represents a periodic trajectory with a constant synchronized frequency  $\omega^* \in \mathcal{I}$  (so that  $\dot{\varphi}_i \equiv \omega_i^*$ ). One of the cylinders of the set  $\mathcal{M}$  is good, in that all of the periodic trajectories on this cylinder are also synchronized in phase (so that  $\delta\varphi \equiv 0$ ). The remaining cylinders of  $\mathcal{M}$  (if any) are bad, in that all of the periodic trajectories on these cylinders are out of phase. Finally, using a local linearization analysis, we show that the phase-synchronized trajectories are exponentially stable (relative to each leaf of the foliation), whereas the out-of-phase trajectories are exponentially unstable.

Because we show that these periodic trajectories are isolated on each leaf (a key step in the proof), we conclude that almost all trajectories achieve asymptotic synchronization in both frequency and phase.

#### F. The dual controller: a different perturbed Hamiltonian system

The approach we take in the controller design in (13)–(15) is to perturb the Hamiltonian system (9)–(10) by adding a dissipation term to the dynamics of each position variable. An alternative, dual approach is to add a dissipation term to the dynamics of each momentum variable. We will do this by including a second comparator in addition to the phase comparator, that is, by introducing a second set of error signals  $z_i$  in addition to the phase errors  $e_i$  in (5). The resulting controller is more complex, but further analysis might reveal that it has performance advantages over the controller (13)–(15).

For this dual controller, each pair of neighboring vertices agrees on an orientation of the edge connecting them. Thus each vertex  $i$  can partition its neighbor set as  $\mathcal{N}_i = \mathcal{N}_i^+ \cup \mathcal{N}_i^-$ , where  $(j, k) \in \mathcal{N}_i^+$  when  $i$  is the head of the edge  $k = \{i, j\}$  and  $(j, k) \in \mathcal{N}_i^-$  when  $i$  is the tail of the edge  $k$ . Furthermore, we make an additional assumption on the frequency functions  $\chi_i$  of the base oscillators:

(A7) the interval  $\mathcal{U}$  and the intervals  $\chi_i(\mathcal{U})$  for each  $i \in \mathcal{V}$  contain only positive real numbers.

In other words, we assume that the graph of each  $\chi_i(\cdot)$  lies completely in the first quadrant, which is true for many analog oscillators. This assumption guarantees that the functions  $\zeta$  and  $\chi_i$  take on only positive values. Consequently, for each ordered pair of vertices  $ij$  we can define the ratio

$$\rho_{ij}(\gamma_i, \gamma_j) = \frac{\zeta(\gamma_j)\chi_i(\zeta(\gamma_i))}{\zeta(\gamma_i)\chi_j(\zeta(\gamma_j))} > 0. \quad (20)$$

Using (9), that is, using the base oscillator model (4) together with the integrator loop filter (7)–(8), we can calculate (20) as

$$\rho_{ij}(\gamma_i, \gamma_j) = \frac{\zeta(\gamma_j)}{\zeta(\gamma_i)} \left[ \frac{d\varphi_j}{d\varphi_i} \right]^{-1}. \quad (21)$$

We then define a new error signal  $z_i$  for vertex  $i$  as

$$\begin{aligned} z_i(\gamma) = & \sum_{(j,k) \in \mathcal{N}_i^+} d_k [\zeta(\gamma_j) - \rho_{ij}(\gamma_i, \gamma_j)\zeta(\gamma_i)] \\ & + \sum_{(j,k) \in \mathcal{N}_i^-} d_k [\rho_{ij}^{-1}(\gamma_i, \gamma_j)\zeta(\gamma_j) - \zeta(\gamma_i)], \end{aligned} \quad (22)$$

where  $\gamma = [\gamma_1 \dots \gamma_n]^T \in \mathbb{R}^n$  is the vector of loop filter states and the constants  $d_k$  are an additional set of positive edge weights. For vertex  $i$  to compute this error variable  $z_i$ , it must receive the signal  $\zeta(\gamma_j)$  from each of its neighbors. Also, we see from (21) that it must calculate the derivative of each neighbor's phase  $\varphi_j$  with respect to its own phase  $\varphi_i$ , a calculation which does not require knowledge of the global time variable  $t$ .

Using these new error signals  $z_i$ , we define the dual perturbed Hamiltonian system as

$$\dot{\varphi}_i = \chi_i(\zeta(\gamma_i)) \quad (23)$$

$$\dot{\gamma}_i = c_i e_i(\varphi) + c_i z_i(\gamma). \quad (24)$$

Thus we have retained the simple integrator loop filter in (7)–(8), but have changed its input from  $e_i$  to the sum  $e_i + z_i$ . This amounts to modifying Fig. 1 by putting a “ $\gamma$ -comparator” in parallel with the phase comparator.

Although the at first sight the physical interpretation of  $z_i$  is obscured by the implementation details, substituting (20) and (23) into (22) gives

$$\begin{aligned} z_i(\gamma) &= \sum_{(j,k) \in \mathcal{N}_i} \eta_k(\gamma) [\chi_j(\zeta(\gamma_j)) - \chi_i(\zeta(\gamma_i))] \\ &= \sum_{(j,k) \in \mathcal{N}_i} \eta_k(\gamma) [\dot{\varphi}_j - \dot{\varphi}_i] \end{aligned} \quad (25)$$

with

$$\eta_k(\gamma) = d_k \frac{\zeta(\gamma_j)}{\chi_j(\zeta(\gamma_j))}, \text{ where } j \text{ is the tail of edge } k, \quad (26)$$

for each  $k \in \mathcal{E}$ . Therefore (22) can be interpreted as an indirect estimate of the frequency mismatch (25) between oscillator  $i$  and its neighbors. As we will see in the proof of the following theorem in Section V below, this frequency error terms  $z_i$  are the right perturbation that need to be added to the dynamics of the momentum variables to guarantee the dissipation of the Hamiltonian energy.

**Theorem 3.** Assume (A1)–(A7). Then with probability one in the selection of edge weights in (A6), the network of oscillators with vertex dynamics (22)–(24) achieves almost-global synchronization within  $\mathcal{I}$ .

The proof of this theorem is similar to that of Theorem 2, and we will highlight the main differences in Section V.

#### IV. Proof of Theorem 2

For any  $p \in \mathbb{N}$ , we let  $\mathbf{0}_p$  and  $\mathbf{1}_p$  denote the column vectors of  $p$  zeros and  $p$  ones, respectively, and we let  $I_p$  denote the  $p \times p$  identity matrix. Any  $\ell \times p$  matrix with integer elements will represent either an  $\mathbb{R}$ -linear map from  $\mathbb{R}^p$  to  $\mathbb{R}^\ell$  or a  $\mathbb{Z}$ -linear map from  $\mathbb{T}^p$  to  $\mathbb{T}^\ell$ , depending on context. In particular, if  $M$  is an  $\ell \times p$  matrix with integer elements and  $t \mapsto x(t)$  is a curve in  $\mathbb{T}^p$ , then  $t \mapsto y(t) = Mx(t)$  is a curve in  $\mathbb{T}^\ell$  (with  $M$  representing a  $\mathbb{Z}$ -linear map), and furthermore  $\dot{y}(t) = M\dot{x}(t)$  for all  $t$  (with  $M$  now representing an  $\mathbb{R}$ -linear map).

We first introduce some notation for writing the overall coupled dynamics (13)–(15) in a compact form.

We define state vectors

$$\varphi = [\varphi_1 \dots \varphi_n]^T \in \mathbb{T}^n \quad (27)$$

$$\gamma = [\gamma_1 \dots \gamma_n]^T \in \mathbb{R}^n \quad (28)$$

along with the following diagonal matrices:

$$A = \text{diag}\{a_1, \dots, a_m\} \in \mathbb{R}^{m \times m} \quad (29)$$

$$C = \text{diag}\{c_1, \dots, c_n\} \in \mathbb{R}^{n \times n}. \quad (30)$$

For convenience we define  $\sigma_i = \chi_i \circ \zeta$  for each  $i$ , and we note from assumptions (A2)–(A3) that each  $\sigma_i$  is  $C^1$  with a positive derivative. We next define mappings  $F : \mathbb{T}^m \rightarrow \mathbb{R}^m$  and  $\Sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$  by

$$F(\theta) = [f(\theta_1) \dots f(\theta_m)]^T \in \mathbb{R}^m \quad (31)$$

$$\Sigma(y) = [\sigma_1(y_1) \dots \sigma_n(y_n)]^T \in \mathbb{R}^n, \quad (32)$$

where  $\theta = [\theta_1 \dots \theta_m]^T \in \mathbb{T}^m$  and  $y = [y_1 \dots y_n]^T \in \mathbb{R}^n$ . Note that  $\Sigma$  is a  $C^1$  diffeomorphism onto its image  $\Sigma(\mathbb{R}^n)$ . Finally, we let  $B \in \{-1, 0, 1\}^{n \times m}$  be an oriented incidence matrix for the graph  $\mathbb{G}$ . Using the fact from (A4) that  $f$  is odd, we can write the dynamics (13)–(15) as

$$\dot{\varphi} = \Sigma(-BAF(B^T\varphi) + \gamma) \quad (33)$$

$$\dot{\gamma} = -CBAF(B^T\varphi). \quad (34)$$

For convenience we define  $e : \mathbb{T}^n \rightarrow \mathbb{R}^n$  by

$$e(\varphi) = [e_1(\varphi) \dots e_n(\varphi)]^T = -BAF(B^T\varphi), \quad (35)$$

which enables us to write (33)–(34) as

$$\dot{\varphi} = \Sigma(e(\varphi) + \gamma) \quad (36)$$

$$\dot{\gamma} = Ce(\varphi). \quad (37)$$

Note that  $\mathbf{1}_n^T e \equiv 0$  because  $\mathbf{1}_n^T B = 0$  (a property of any oriented incidence matrix). This property also implies  $e(\varphi) = e(\varphi + \mathbf{1}_n \theta)$  for any  $\varphi \in \mathbb{T}^n$  and any  $\theta \in \mathbb{T}$ .

#### A. Global Lyapunov function

In this section we construct a Lyapunov function for the system (36)–(37), which is the Hamiltonian we described in Section III-C for the system (9)–(10).

Because  $f$  is odd from (A4), the integral of the 1-form  $f \cdot \langle \partial, \cdot \rangle$  around any smooth closed curve in  $\mathbb{T}$  is zero. Thus this 1-form is the differential of a smooth function  $\Psi : \mathbb{T} \rightarrow \mathbb{R}$ , which is unique up to an additive constant (which we choose so that the minimum value of  $\Psi$  on  $\mathbb{T}$  is zero). Therefore  $\frac{d}{dt}(\Psi \circ x) \equiv f(x)\dot{x}$  for

any curve  $x : \mathbb{R} \rightarrow \mathbb{T}$ . We then define  $V : \mathbb{T}^m \rightarrow [0, \infty)$  as the sum

$$V(\theta) = \sum_{k \in \mathcal{E}} a_k \Psi(\theta_k), \quad (38)$$

where  $\theta = [\theta_1 \dots \theta_m]^T \in \mathbb{T}^m$ . It follows that

$$\frac{d}{dt} V(B^T \varphi) = F^T(B^T \varphi) A B^T \dot{\varphi} = -e^T(\varphi) \dot{\varphi}. \quad (39)$$

From (17) there exists a frequency  $\varpi \in \bigcap_i \sigma_i(\mathbb{R})$ . Therefore the function  $W : \mathbb{R}^n \rightarrow [0, \infty)$  defined as

$$W(\gamma) = \sum_{i \in \mathcal{V}} \frac{1}{c_i} \int_{\sigma_i^{-1}(\varpi)}^{\gamma_i} [\sigma_i(s) - \varpi] ds \quad (40)$$

is proper (because  $\sigma_i$  is increasing) and has a derivative given by

$$\frac{d}{dt} W(\gamma) = [\Sigma^T(\gamma) - \varpi \mathbf{1}_n^T] C^{-1} \dot{\gamma}. \quad (41)$$

We now obtain a Lyapunov function by adding (38) and (40): we define  $U : \mathbb{T}^n \times \mathbb{R}^n \rightarrow [0, \infty)$  by

$$U(\varphi, \gamma) = V(B^T \varphi) + W(\gamma), \quad (42)$$

which is a proper function on  $\mathbb{T}^n \times \mathbb{R}^n$ . Because  $\mathbf{1}_n^T e \equiv 0$ , the derivative of (42) along trajectories of (36)–(37) is

$$\begin{aligned} \dot{U} &= -e^T(\varphi) \Sigma(e(\varphi) + \gamma) + \Sigma^T(\gamma) e(\varphi) \\ &= -e^T(\varphi) \int_0^1 \Sigma'(e(\varphi)s + \gamma) ds \cdot e(\varphi), \end{aligned} \quad (43)$$

where  $\Sigma'(\cdot)$  denotes the diagonal Jacobian matrix

$$\Sigma'(\gamma) = \text{diag}\{\sigma'_1(\gamma_1), \dots, \sigma'_n(\gamma_n)\}. \quad (44)$$

Because  $\Sigma'(\cdot)$  is positive definite, we have  $\dot{U} \leq 0$  and  $\dot{U} = 0$  if and only if  $e(\varphi) = 0$ . It follows from the Krasovskii-LaSalle invariance theorem that all trajectories of the system (36)–(37) converge to the largest invariant set  $\mathcal{M}$  contained within the set  $\Phi \times \mathbb{R}^n$ , where  $\Phi \subset \mathbb{T}^n$  denotes the set

$$\Phi = e^{-1}(\{0\}) = \{\varphi \in \mathbb{T}^n : B A F(B^T \varphi) = 0\}. \quad (45)$$

Note that because  $B^T \mathbf{1}_n = 0$ , this set  $\Phi$  has the symmetry property  $\Phi = \Phi + \mathbf{1}_n \mathbb{T}$ . Also, it follows from (A4) that  $F(0) = 0$ , which means  $\Phi$  contains all points of the form  $\mathbf{1}_n \theta$  for  $\theta \in \mathbb{T}$ . The next step in the proof is to investigate the structure of this set  $\mathcal{M}$ .

## B. Structure of the set $\mathcal{M}$

In this section we show that the largest invariant set  $\mathcal{M}$  contained within the set  $\Phi \times \mathbb{R}^n$  is  $\mathcal{M} = \Phi \times \Gamma$ , where  $\Phi$  in (45) is the zero set of  $e$  and  $\Gamma \subset \mathbb{R}^n$  is the set

$$\Gamma = \{\gamma \in \mathbb{R}^n : B^T \Sigma(\gamma) = 0\}. \quad (46)$$

We begin by exploring the structure of these sets  $\Phi$  and  $\Gamma$ . First, we show that  $\Phi$  is the disjoint union of isolated embeddings of the circle  $\mathbb{T}$ . Next, we show that  $\Gamma = \alpha(\mathbb{R})$ , where  $\alpha : \mathbb{R} \rightarrow \mathbb{R}^n$  is a  $C^1$  curve in  $\mathbb{R}^n$ .

Moreover, if we let  $q \in \mathbb{R}^n$  denote the unit vector in the direction of  $C^{-1}\mathbf{1}_n$ , then this curve  $\alpha$  is such that  $q^T\alpha(\cdot)$  is the identity map on  $\mathbb{R}$ . Finally, we show that  $\mathcal{M} = \Phi \times \Gamma$ , that is, that  $\mathcal{M}$  is the disjoint union of isolated embeddings of the cylinder, as illustrated in Fig. 3.

We partition  $\mathbb{T}^n$  using the following matrices:

$$R = \begin{bmatrix} \mathbf{0}_{n-1}^T \\ I_{n-1} \end{bmatrix} \in \{0, 1\}^{n \times (n-1)} \quad (47)$$

$$S = \begin{bmatrix} -\mathbf{1}_{n-1}^T \\ I_{n-1} \end{bmatrix} \in \{-1, 0, 1\}^{n \times (n-1)}. \quad (48)$$

Clearly the state  $\varphi$  satisfies the identity

$$\varphi = \mathbf{1}_n \varphi_1 + RS^T \varphi, \quad (49)$$

which defines the direct sum  $\mathbb{T}^n = \mathbf{1}_n \mathbb{T} \oplus R\mathbb{T}^{n-1}$ . Here the first summand represents the first component angle and the second one represents the remaining angles measured relative to the first. Note that  $S^T R = I_{n-1}$ , that  $S^T \mathbf{1}_n = 0$ , and that  $B^T = B^T R S^T$ . Also, because  $\mathbb{G}$  is connected from (A1), we have  $\text{rank}(B) = n - 1$  and thus the columns of  $B^T R$  are independent.

Because  $\varphi \in \Phi$  if and only if  $RS^T \varphi \in \Phi$ , it follows from (49) that  $\Phi = RS^T \Phi + \mathbf{1}_n \mathbb{T}$ . It also follows that for any  $\mu \in \mathbb{T}^{n-1}$ , we have  $\mu \in S^T \Phi$  if and only if  $R\mu \in \Phi$ , that is, if and only if  $e(R\mu) = 0$ . We next show that the points in the set  $S^T \Phi$  are isolated, which implies that  $\Phi$  is the disjoint union of isolated embeddings of the circle  $\mathbb{T}$ .

Using the above partition of  $\mathbb{T}^n$ , we define two symmetric matrix functions  $L : \mathbb{T}^{n-1} \rightarrow \mathbb{R}^{n \times n}$  and  $L^\flat : \mathbb{T}^{n-1} \rightarrow \mathbb{R}^{(n-1) \times (n-1)}$  by

$$L(\mu) = B A F'(B^T R \mu) B^T \quad \text{and} \quad L^\flat(\mu) = R^T L(\mu) R \quad (50)$$

for  $\mu \in \mathbb{T}^{n-1}$ , where  $F'(\cdot)$  denotes the diagonal Jacobian matrix

$$F'(\theta) = \text{diag}\{f'(\theta_1), \dots, f'(\theta_m)\} \quad (51)$$

with  $\theta = [\theta_1 \dots \theta_m]^T \in \mathbb{T}^m$ . Here  $L(\mu)$  represents a weighted Laplacian matrix for the graph  $\mathbb{G}$  in which the weights can have positive, negative, or zero values. Also, because  $B = S R^T B$  we have

$$L(\mu) R = S L^\flat(\mu) \quad \text{and} \quad L(\mu) = S L^\flat(\mu) S^T \quad (52)$$

for all  $\mu \in \mathbb{T}^{n-1}$ . Note that  $L(\mu)$  is congruent to the block diagonal matrix  $\text{diag}\{0, L^\flat(\mu)\}$ . The proof of the following theorem is in Appendix A:

**Theorem 4.** There is a set  $\mathcal{Z} \subset \mathbb{R}^m$  having zero Lebesgue measure such that if  $a = [a_1 \dots a_m]^T \notin \mathcal{Z}$ , then the matrix  $L^\flat(\mu)$  in (50) is invertible for all  $\mu \in S^T \Phi$ .

**Corollary 5.** If  $a \notin \mathcal{Z}$  then the points in  $S^T \Phi$  are isolated.

Proof: Define the mapping  $P : \mathbb{T}^{n-1} \rightarrow \mathbb{R}^{n-1}$  by setting  $P(\mu) = R^T B A F(B^T R \mu)$  so that  $P^{-1}(\{0\}) = S^T \Phi$ . The Jacobian matrix for  $P$  is just  $L^b(\mu)$ , which by Theorem 4 is invertible for all  $\mu \in S^T \Phi$ . The result follows from the inverse function theorem.  $\blacksquare$

We do not provide a method for computing the set  $\mathcal{Z}$  of bad edge weight vectors; instead, we rely on the random edge weight selection in (A6) to avoid this zero-measure set. It is clear from (45) and (50) that this set  $\mathcal{Z}$  depends only on the graph  $\mathbb{G}$  (through  $B$ ) and the phase coupling function  $f$  (through  $F$ ). In principle we could remove the dependence on the unknown graph by taking the countable union of all such sets  $\mathcal{Z}$  over all connected graphs to be our zero-measure set of bad edge weights, but this larger set would still be difficult to compute. Thus from now on we will assume  $a \notin \mathcal{Z}$ , which happens with probability one according to (A6)

To explore the structure of the set  $\Gamma$  in (46), we partition  $\mathbb{R}^n$  by defining  $q \in \mathbb{R}^n$  and  $Q \in \mathbb{R}^{n \times (n-1)}$  as

$$q = \frac{C^{-1} \mathbf{1}_n}{\|C^{-1} \mathbf{1}_n\|} \quad \text{and} \quad Q = CS(S^T C^2 S)^{-\frac{1}{2}}. \quad (53)$$

Using the fact that  $\mathbf{1}_n^T S = 0$ , it is straightforward to show that the  $n \times n$  matrix  $[q \ Q]$  is orthogonal; thus the state  $\gamma$  satisfies the identity

$$\gamma = qq^T \gamma + QQ^T \gamma, \quad (54)$$

which defines the direct sum  $\mathbb{R}^n = q\mathbb{R} \oplus Q\mathbb{R}^{n-1}$  via orthogonal projections. We define the open interval  $J \subset \mathbb{R}$  as

$$J = \bigcap_{i \in \mathcal{V}} \sigma_i(\mathbb{R}), \quad (55)$$

which is nonempty from (17). Because  $s\mathbf{1}_n \in \Sigma(\mathbb{R}^n)$  for all  $s \in J$ , we can define the function  $\beta : J \rightarrow \mathbb{R}$  as

$$\beta(s) = \sum_{i \in \mathcal{V}} q_i \sigma_i^{-1}(s) = q^T \Sigma^{-1}(s\mathbf{1}_n), \quad (56)$$

where the constants  $q_i$  are the components of the vector  $q$ . Note that each term in the sum in (56) is strictly increasing in  $s$ . Suppose  $\{s_\ell\}$  is a sequence in  $J$  such that  $s_\ell \rightarrow \sup J$  as  $\ell \rightarrow \infty$ . Then there exists  $i \in \mathcal{V}$  such that  $s_\ell \rightarrow \sup \sigma_i(\mathbb{R})$  as  $\ell \rightarrow \infty$ , which means  $\sigma_i^{-1}(s_\ell) \rightarrow \infty$  as  $\ell \rightarrow \infty$ . Thus  $\beta(s_\ell) \rightarrow \infty$  as  $\ell \rightarrow \infty$ , which means  $\sup \beta(J) = \infty$ . Similar reasoning yields  $\inf \beta(J) = -\infty$ , and we conclude that  $\beta$  is invertible. Therefore we can define a  $C^1$  curve  $\alpha : \mathbb{R} \rightarrow \mathbb{R}^n$  as follows:

$$\alpha(r) = \Sigma^{-1}(\beta^{-1}(r)\mathbf{1}_n), \quad (57)$$

which satisfies  $q^T \alpha(r) = \beta(\beta^{-1}(r)) = r$  for all  $r \in \mathbb{R}$ . Because  $\text{rank}(B) = n - 1$ , it follows from (46) that  $\gamma \in \Gamma$  if and only if  $\Sigma(\gamma) = x\mathbf{1}_n$  for some  $x \in J$ , that is, if and only if

$$\gamma = \Sigma^{-1}(x\mathbf{1}_n) = \Sigma^{-1}(\beta^{-1}(\beta(x))\mathbf{1}_n) = \alpha(\beta(x)) \quad (58)$$

for some  $x \in J$ . Therefore  $\Gamma = \alpha(\beta(J)) = \alpha(\mathbb{R})$ .



Our remaining task in this section is to show that  $\mathcal{M} = \Phi \times \Gamma$ . First we calculate the derivative of  $e(\varphi)$  in (35) using (36):

$$\begin{aligned} \dot{e} &= -BAF'(B^T\varphi)B^T\Sigma(e(\varphi) + \gamma) \\ &= -L(S^T\varphi)\Sigma(e(\varphi) + \gamma). \end{aligned} \quad (59)$$

We see from (36)–(37) that  $\dot{\varphi} = \Sigma(\gamma)$  and  $\dot{\gamma} = 0$  on the set  $\Phi \times \mathbb{R}^n$ , which means the second derivative  $\ddot{\varphi}$  is zero:

$$\begin{aligned} 0 = \ddot{\varphi} &= -\Sigma'(\gamma)L(S^T\varphi)\Sigma(\gamma) \\ &= -\Sigma'(\gamma)SL^{\flat}(S^T\varphi)S^T\Sigma(\gamma) \end{aligned} \quad (60)$$

on  $\Phi \times \mathbb{R}^n$ . Now  $\Sigma'(\cdot)$  in (44) is positive definite,  $L^{\flat}(S^T\varphi)$  is invertible from Theorem 4, and  $S$  in (48) has independent columns; therefore (60) implies  $S^T\Sigma(\gamma) = 0$ . Because  $B^T = B^TR S^T$ , this in turn implies  $B^T\Sigma(\gamma) = 0$ , and we conclude that  $\mathcal{M} \subset \Phi \times \Gamma$ . We observe that  $\Phi \times \Gamma$  is itself invariant under the dynamics (36)–(37), and it follows that  $\mathcal{M} = \Phi \times \Gamma$ .

To summarize the results of this section, we have found that  $(\varphi, \gamma) \in \mathcal{M}$  if and only if  $e(\varphi) = 0$  and  $\gamma = \alpha(r)$  for some  $r \in \mathbb{R}$ , in which case  $r = q^T\gamma$ . Furthermore, the zero set of  $e$  in (45) is  $RS^T\Phi + \mathbf{1}_n\mathbb{T}$ , where the points in  $S^T\Phi$  are isolated.

### C. Global analysis: frequency synchronization

To study the synchronization properties of our system (36)–(37), we first observe that  $q^T\dot{\gamma} \equiv 0$ , which means the state space admits a foliation whose leaves are the invariant manifolds  $\mathbb{T}^n \times \Xi_r$ , where

$$\Xi_r = qr + Q\mathbb{R}^{n-1} = \{\gamma \in \mathbb{R}^n : q^T\gamma = r\} \quad (61)$$

for a constant parameter  $r \in \mathbb{R}$ . Note that because  $\Gamma = \alpha(\mathbb{R})$  and  $q^T\alpha(\cdot)$  is the identity map, the intersection  $\Gamma \cap \Xi_r$  is the singleton  $\{\alpha(r)\}$ . Thus the intersection of each invariant leaf  $\mathbb{T}^n \times \Xi_r$  with  $\mathcal{M} = \Phi \times \Gamma$  is just  $\Phi \times \{\alpha(r)\}$ , which we have shown to be the disjoint union of isolated embeddings of the circle  $\mathbb{T}$  (depicted as red circles in Fig. 3).

We now fix  $r \in \mathbb{R}$  and examine the dynamics on  $\mathbb{T}^n \times \Xi_r$ , noting that  $\gamma \equiv qr + QQ^T\gamma$  on this invariant manifold. Now that  $r$  is fixed, we will write  $\alpha_r = \alpha(r)$ . Because  $\alpha_r = qr + QQ^T\alpha_r$ , we have

$$\gamma - \alpha_r \equiv QQ^T(\gamma - \alpha_r) \quad (62)$$

on  $\mathbb{T}^n \times \Xi_r$ . We next define the projected state variables

$$w_1 = S^T\varphi \in \mathbb{T}^{n-1} \quad \text{and} \quad w_2 = Q^T(\gamma - \alpha_r) \in \mathbb{R}^{n-1}. \quad (63)$$

Taking time derivatives of these variables, and using (62) and the fact that  $e(\varphi) = e(Rw_1)$ , we obtain

$$\dot{w}_1 = S^T\Sigma(e(Rw_1) + Qw_2 + \alpha_r) \quad (64)$$

$$\dot{w}_2 = Q^TCe(Rw_1). \quad (65)$$

This is an autonomous system in the projected states  $(w_1, w_2)$ , and its equilibria are precisely all points of the form  $(\mu^*, 0)$  for vectors  $\mu^* \in S^T\Phi$ . Each equilibrium represents a frequency-synchronized solution of (36)–(37) with  $\dot{\varphi} \equiv \beta^{-1}(r)\mathbf{1}_n$  and  $\gamma \equiv \alpha_r$ . The equilibrium with  $\mu^* = 0$  represents a phase-synchronized trajectory with  $\delta\varphi \equiv 0$ , and all other equilibria represent out-of-phase trajectories.

Furthermore, each trajectory of the system (64)–(87) converges to an equilibrium, which means each trajectory of the system (36)–(37) achieves asymptotic frequency synchronization. Indeed, we have shown that all trajectories of the system (36)–(37) converge to the set  $\mathcal{M} = \Phi \times \Gamma$  in forward time. Therefore  $\gamma$  converges to the singleton set  $\Gamma \cap \Xi_r = \{\alpha_r\}$ , which means  $w_2$  converges to zero. In addition,  $\varphi$  converges to the set  $\Phi$ , which means  $w_1$  converges to the set  $S^T\Phi$ . Corollary 5 states that points in  $S^T\Phi$  are isolated, and we conclude that  $w_1$  converges to one of these points.

In the next section, we will perform a local linearization analysis at each equilibrium of the system (64)–(87) to determine its stability.

#### D. Local analysis: phase synchronization

We compute the linear approximation of the dynamics (64)–(87) at an equilibrium  $(\mu^*, 0)$  as follows:

$$\dot{w}_1 \approx -S^T\Sigma'(\alpha_r)L(\mu^*)R(w_1 - \mu^*) + S^T\Sigma'(\alpha_r)Qw_2 \quad (66)$$

$$\dot{w}_2 \approx -Q^TCL(\mu^*)R(w_1 - \mu^*), \quad (67)$$

or using (52),

$$\dot{w}_1 \approx -S^T\Sigma'(\alpha_r)SL^b(\mu^*)(w_1 - \mu^*) + S^T\Sigma'(\alpha_r)Qw_2 \quad (68)$$

$$\dot{w}_2 \approx -Q^TCSL^b(\mu^*)(w_1 - \mu^*). \quad (69)$$

If we define the  $(n-1) \times (n-1)$  matrices

$$X = S^T\Sigma'(\alpha_r)S > 0 \quad (70)$$

$$Y = Q^TCS = (S^TC^2S)^{\frac{1}{2}} > 0 \quad (71)$$

$$Z = S^T\Sigma'(\alpha_r)Q, \quad (72)$$

then we can write this approximation more compactly as

$$\dot{w}_1 \approx -XL^b(\mu^*)(w_1 - \mu^*) + Zw_2 \quad (73)$$

$$\dot{w}_2 \approx -YL^b(\mu^*)(w_1 - \mu^*). \quad (74)$$

Note from (53) and (71) that  $Q = CSY^{-1}$  so that  $S = C^{-1}QY$ , which means  $Z$  in (72) satisfies

$$Y^{-1}Z = Q^TC^{-1}\Sigma'(\alpha_r)Q > 0. \quad (75)$$

Thus the linearization (73)–(74) satisfies the assumptions in the following theorem, whose proof is in Appendix B:

Theorem 6. Let  $\Lambda \in \mathbb{R}^{2p \times 2p}$  be the block matrix

$$\Lambda = \begin{bmatrix} -XL & Z \\ -YL & 0 \end{bmatrix}, \quad (76)$$

where  $L, X, Y, Z \in \mathbb{R}^{p \times p}$  satisfy:

- 1)  $L$  is symmetric,
- 2)  $X + X^T > 0$ ,
- 3)  $Y$  is symmetric and invertible, and
- 4)  $Y^{-1}Z$  is symmetric with  $Y^{-1}Z \geq 0$ .

If  $L$  has a negative eigenvalue, then  $\Lambda$  has an eigenvalue with a positive real part. If instead  $Z$  is invertible and  $L > 0$ , then  $\Lambda$  is Hurwitz.

We can complete the local linearization analysis by applying Theorem 6 together with the following theorem:

Theorem 7. The Laplacian  $L(\mu)$  has  $n - 1$  positive eigenvalues for  $\mu = 0$ , and it has at least one negative eigenvalue for any nonzero  $\mu \in S^T\Phi$ .

The proof of Theorem 7, which can be found in [29], relies on the properties of the phase coupling function  $f$  defined in Assumption (A5). The basic intuition is that for  $\mu^* \neq 0$ , choosing  $b$  small enough guarantees a negative eigenvalue in  $L(\mu^*)$  independently of the choice of  $\mu^* \in S^T\Phi$ .

Let us now consider an equilibrium  $(\mu^*, 0)$  of the nonlinear system (64)–(87). Recall that  $L(\mu^*)$  is congruent to  $\text{diag}\{0, L^b(\mu^*)\}$ ; thus Theorem 7 together with Sylvester's law of inertia imply that  $L^b(\mu^*) > 0$  when  $\mu^* = 0$  and that  $L^b(\mu^*)$  has a negative eigenvalue for all nonzero  $\mu^* \in S^T\Phi$ . Therefore if  $\mu^* = 0$ , which represents an in-phase steady-state solution, then it follows from Theorem 6 that this equilibrium is exponentially stable. Likewise, if  $\mu^* \neq 0$ , which represents an out-of-phase steady-state solution, then this equilibrium is exponentially unstable. Because all out-of-phase equilibria are both isolated and exponentially unstable, we conclude (say from [35, Proposition 1], for example) that the set  $\mathcal{S}_r \subset \mathbb{T}^n \times \Xi_r$  of initial states from which trajectories converge to out-of-phase steady-state solutions has zero measure in  $\mathbb{T}^n \times \Xi_r$  (regarded here as a  $(2n - 1)$ -dimensional manifold). It then follows from Tonelli's theorem that the set  $\mathcal{S} = \bigcup_{r \in \mathbb{R}} \mathcal{S}_r$  has zero measure in  $\mathbb{T}^n \times \mathbb{R}^n$ . In other words, the system achieves asymptotic phase synchronization from almost every initial state.

## V. Proof of Theorem 3

We define the diagonal matrix function

$$H(\gamma) = \text{diag}\{\eta_1(\gamma), \dots, \eta_m(\gamma)\} \in \mathbb{R}^{m \times m}, \quad (77)$$

so that we can write the dynamics (23)–(24) as

$$\dot{\varphi} = \Sigma(\gamma) \quad (78)$$

$$\dot{\gamma} = Ce(\varphi) - CBH(\gamma)B^T\Sigma(\gamma). \quad (79)$$

Because  $\mathbf{1}_n^T e \equiv 0 = \mathbf{1}_n^T B$ , the derivative of (42) along trajectories of (78)–(79) is

$$\dot{U} = -\Sigma^T(\gamma)BH(\gamma)B^T\Sigma(\gamma). \quad (80)$$

Now  $H(\cdot)$  is positive definite; hence  $\dot{U} \leq 0$  and  $\dot{U} = 0$  if and only if  $B^T\Sigma(\gamma) = 0$ . It follows from the Krasovskii-LaSalle invariance theorem that all trajectories of the system (78)–(79) converge to the largest invariant set  $\mathcal{M}$  contained within the set  $\mathbb{T}^n \times \Gamma$ , where  $\Gamma \subset \mathbb{R}^n$  is from (46). We next show that again  $\mathcal{M} = \Phi \times \Gamma$ , just as in the proof of Theorem 2.

It follows from (79) that this systems also satisfies  $q^T\dot{\gamma} \equiv 0$  and thus admits a foliation whose leaves are the invariant manifolds  $\mathbb{T}^n \times \Xi_r$ . Because  $B^T\Sigma(\gamma)$  is the constant zero on the set  $\Gamma$ , its derivative is zero on  $\mathcal{M}$ :

$$\begin{aligned} 0 &\equiv B^T\Sigma'(\gamma)\dot{\gamma} \equiv B^T\Sigma'(\gamma)QQ^T\dot{\gamma} \\ &\equiv B^TR S^T\Sigma'(\gamma)CS(S^TC^2S)^{-\frac{1}{2}}Q^T\dot{\gamma}. \end{aligned} \quad (81)$$

Because the columns of  $B^TR$  and  $S$  are independent and because the diagonal matrix  $\Sigma'(\gamma)C$  is positive definite, it follows that  $Q^T\dot{\gamma} \equiv 0$  and thus  $\dot{\gamma} \equiv 0$  on  $\mathcal{M}$ . It then follows from (79) that  $e(\varphi) \equiv 0$  on  $\mathcal{M}$ , and we conclude that  $\mathcal{M} \subset \Phi \times \Gamma$ . We observe that  $\Phi \times \Gamma$  is itself invariant under the dynamics (78)–(79), and it follows that  $\mathcal{M} = \Phi \times \Gamma$ .

We now fix  $r \in \mathbb{R}$  and examine the dynamics on the invariant manifold  $\mathbb{T}^n \times \Xi_r$ . Using (62), we see that the derivatives of the projected state variables (63) along trajectories of (78)–(79) are

$$\dot{w}_1 = S^T\Sigma(Qw_2 + \alpha_r) \quad (82)$$

$$\begin{aligned} \dot{w}_2 &= Q^TCe(Rw_1) \\ &\quad - Q^TCBH(Qw_2 + \alpha_r)B^T\Sigma(Qw_2 + \alpha_r). \end{aligned} \quad (83)$$

As in the proof of Theorem 2, the equilibria of this system are precisely all points of the form  $(\mu^*, 0)$  for vectors  $\mu^* \in S^T\Phi$ . The equilibrium with  $\mu^* = 0$  represents a fully synchronized trajectory with  $\delta\varphi \equiv 0$ , and all other equilibria represent out-of-phase, frequency-synchronized trajectories. Because  $\mathcal{M} = \Phi \times \Gamma$ , because  $\Gamma \cap \Xi_r$  is a singleton, and because the points of  $S^T\Phi$  are isolated, we see that each trajectory of the system (82)–(83) converges to an equilibrium.

We now perform the local linearization analysis at each equilibrium  $(\mu^*, 0)$  of the system (82)–(83). Similarly, to (73)–(74) we can approximate (82)–(83) around  $(\mu^*, 0)$  using

$$\dot{w}_1 \approx Zw_2 \quad (84)$$

$$\dot{w}_2 \approx -YL^b(\mu^*)(w_1 - \mu^*) - YL_\eta^b Zw_2 \quad (85)$$

where

$$L_\eta^b(\alpha_r) = R^T L_\eta(\alpha_r) R \quad \text{and} \quad L_\eta(\alpha_r) = B H(\alpha_r) B^T, \quad (86)$$

with the matrices  $L_\eta(\alpha_r)$  and  $L_\eta^b(\alpha_r)$  also satisfying a condition analogous to (52).

We can now use a dual version of Theorem 6 based on [36, Theorem 5, Corollary 2] which we state below.

Lemma 8. Suppose  $\Lambda \in \mathbb{C}^{q \times q}$  has no eigenvalues on the imaginary axis, suppose  $H \in \mathbb{C}^{q \times q}$  is an invertible Hermitian matrix, and suppose  $\Lambda H + H \Lambda^* \geq 0$ , where  $\Lambda^*$  denotes the conjugate transpose of  $\Lambda$ . Then the number of eigenvalues of  $\Lambda$  having positive real part is the same as the number of positive eigenvalues of  $H$ .

The dual version of Theorem 6 is then:

Theorem 9. Let  $\Lambda \in \mathbb{R}^{2p \times 2p}$  be the block matrix

$$\Lambda = \begin{bmatrix} 0 & Z \\ -Y L_1 & -Y L_2 Z \end{bmatrix}, \quad (87)$$

where  $L_1, L_2, Y, Z \in \mathbb{R}^{p \times p}$  satisfy:

- 1)  $L_1$  is symmetric and invertible,
- 2)  $L_2$  is symmetric with  $L_2 > 0$ ,
- 3)  $Y$  is symmetric and invertible, and
- 4)  $Y^{-1}Z$  is symmetric with  $Y^{-1}Z > 0$ .

If  $L_1$  has a strictly negative eigenvalue, then  $\Lambda$  has an eigenvalue with a strictly positive real part. If instead  $L_1 > 0$ , then  $\Lambda$  is Hurwitz.

The proof is in Appendix C.

Thus, we can use Theorem 7 together with Theorem 9 to show that, for the system (84)–(85), the equilibrium with  $\mu^* = 0$  is exponentially stable and any other equilibria with  $\mu^* \neq 0$  is exponentially unstable. The rest of the proof follows that of Theorem 2.

## VI. Concluding remarks

We have presented two distributed controllers for the phase and frequency synchronization of heterogeneous nonlinear oscillators, and we have shown that each guarantees almost-global convergence on arbitrary connected graphs. Our solutions can be readily implemented using analog oscillators and phase comparators, and our analysis holds under very general assumptions on the system. In particular, unlike most existing work, we neither require the set of admissible frequencies to be unbounded, nor assume any special network topology.

## Appendix

## A. Proof of Theorem 4

We let  $\mathcal{T}$  denote the finite collection of all  $m \times m$  diagonal matrices  $\Delta = \text{diag}\{\delta_1, \dots, \delta_m\}$  such that for all  $k \in \mathcal{E}$ , either  $\delta_k = f'(0)$  or  $\delta_k = f'(\pi)$ . For each such matrix  $\Delta \in \mathcal{T}$ , we define the closed set

$$\mathcal{P}_\Delta = \{a \in \mathbb{R}^m : \det(R^T B \text{diag}(a) \Delta B^T R) = 0\}, \quad (88)$$

where  $\text{diag}(a) = \text{diag}\{a_1, \dots, a_m\}$  denotes the diagonal matrix whose diagonal entries are the  $m$  elements of  $a$ . Now  $\Delta$  is invertible by assumption (A5), and furthermore the columns of  $B^T R$  are independent; it follows that  $\mathcal{P}_\Delta \neq \mathbb{R}^m$  (take  $\text{diag}(a) = \Delta^{-1}$ ), which means  $\mathcal{P}_\Delta$  is a closed algebraic set having zero measure. Thus

$$\mathcal{P} = \bigcup_{\Delta \in \mathcal{T}} \mathcal{P}_\Delta \quad (89)$$

is also a closed algebraic set having zero measure. Therefore the set  $\mathcal{O} = \mathbb{R}^m \setminus \mathcal{P}$  is a nonempty open semialgebraic set. Next we define the mapping  $\mathcal{H} : \mathbb{T}^{n-1} \times \mathcal{O} \rightarrow \mathbb{R}^{n-1}$  by

$$\mathcal{H}(\mu, a) = R^T B \text{diag}(a) F(B^T R \mu). \quad (90)$$

The Jacobian matrix of  $\mathcal{H}$  is

$$D\mathcal{H}(\mu, a) = \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial \mu}(\mu, a) & \frac{\partial \mathcal{H}}{\partial a}(\mu, a) \end{bmatrix}, \quad (91)$$

where

$$\frac{\partial \mathcal{H}}{\partial \mu}(\mu, a) = R^T B \text{diag}(a) F'(B^T R \mu) B^T R \quad (92)$$

$$\frac{\partial \mathcal{H}}{\partial a}(\mu, a) = R^T B \text{diag}(F(B^T R \mu)). \quad (93)$$

If we define the matrix

$$\mathcal{J}(\mu, a) = \begin{bmatrix} I \\ \text{diag}(a) \text{diag}(F(B^T R \mu))^+ \\ \cdot [F'(0) - F'(B^T R \mu)] B^T R \end{bmatrix}, \quad (94)$$

where  $(\cdot)^+$  denotes the Moore–Penrose pseudoinverse, then

$$D\mathcal{H}(\mu, a) \cdot \mathcal{J}(\mu, a) = R^T B \text{diag}(a) \Delta(\mu) B^T R, \quad (95)$$

where  $\Delta(\mu)$  is the diagonal matrix

$$\begin{aligned} \Delta(\mu) &= F'(B^T R \mu) + \text{diag}(F(B^T R \mu)) \text{diag}(F(B^T R \mu))^+ \\ &\quad \cdot [F'(0) - F'(B^T R \mu)]. \end{aligned} \quad (96)$$

It follows from assumptions (A4)–(A5) that  $f(\theta) = 0$  if and only if  $\theta \in \{0, \pi\}$ , so for any  $\mu \in \mathbb{T}^{n-1}$ , the matrix  $\Delta(\mu)$  in (96) belongs to  $\mathcal{T}$ . It follows from the definition of  $\mathcal{O}$  that the matrix in (95) is invertible, and

we conclude that  $D\mathcal{H}(\mu, a)$  has rank  $n - 1$  for all  $(\mu, a) \in \mathbb{T}^{n-1} \times \mathcal{O}$ . Thus  $\mathcal{H} \pitchfork \{0\}$ ,<sup>2</sup> and it follows from the parametric transversality theorem<sup>3</sup> [37, Theorem 6.35] that there exists a set  $\mathcal{Y} \subset \mathcal{O}$  having zero measure such that if  $a \in \mathcal{O} \setminus \mathcal{Y}$  then  $\mathcal{H}_a \pitchfork \{0\}$ , where  $\mathcal{H}_a$  denotes the mapping  $\mu \mapsto \mathcal{H}(\mu, a)$ . Choose  $\mathcal{Z} = \mathcal{P} \cup \mathcal{Y}$ ; we have thus shown that for all  $a \in \mathbb{R}^m \setminus \mathcal{Z}$ , if  $\mu$  is such that  $\mathcal{H}(\mu, a) = 0$ , then the matrix in (92) is invertible.

Suppose  $a \in \mathbb{R}^m$  is the edge weight vector so that  $A = \text{diag}(a)$ , suppose  $a \in \mathbb{R}^m \setminus \mathcal{Z}$ , and suppose  $\mu \in S^T \Phi$ . Then  $R\mu \in \Phi$  which means  $\mathcal{H}(\mu, a) = 0$ , and it follows that  $L^b(\mu)$  in (50), which is the same as the matrix in (92), is invertible.

## B. Proof of Theorem 6

Because both  $Y$  and  $Y^{-1}Z$  are symmetric matrices, we see that the following holds:

$$Y^{-1}ZY = (Y^{-1}Z)^T Y = Z^T Y^{-1} Y = Z^T. \quad (97)$$

We will investigate the stability of the zero solution of the system  $\dot{x} = \Lambda x$ . Partition the state as  $x = [x_1^T \ x_2^T]^T$  with  $x_1, x_2 \in \mathbb{R}^p$ , and consider the quadratic function

$$\Upsilon(x) = \Upsilon(x_1, x_2) = x_1^T L x_1 + x_2^T Y^{-1} Z x_2 \quad (98)$$

whose derivative along trajectories of  $\dot{x} = \Lambda x$  is

$$\begin{aligned} \dot{\Upsilon}(x) &= -2x_1^T L X L x_1 + 2x_1^T L Z x_2 - 2x_2^T Y^{-1} Z Y L x_1 \\ &= -x_1^T L (X + X^T) L x_1 \leq 0. \end{aligned} \quad (99)$$

First suppose  $L$  has an eigenvalue  $\lambda < 0$  with unit eigenvector  $v$ . The standard Chetaev instability conditions [38, Theorem 4.3] do not guarantee exponential instability, so we will instead exploit the linear-quadratic structure to strengthen the Chetaev result. The cone

$$\mathcal{C} = \{x \in \mathbb{R}^{2p} : \Upsilon(x) < 0\}, \quad (100)$$

is nonempty because

$$\Upsilon(v, 0) = v^T L v = \lambda v^T v = \lambda < 0. \quad (101)$$

For any  $x \in \mathcal{C}$  we have

$$0 \leq x_2^T Y^{-1} Z x_2 < -x_1^T L x_1 = |x_1^T L x_1|, \quad (102)$$

which means

$$|\Upsilon(x)| \leq |x_1^T L x_1| + x_2^T Y^{-1} Z x_2 < 2|x_1^T L x_1|. \quad (103)$$

<sup>2</sup>If  $M$  and  $N$  are smooth manifolds, if  $f : N \rightarrow M$  is smooth, and if  $S$  is an embedded submanifold of  $M$ , then  $f$  is transverse to  $S$ , written  $f \pitchfork S$ , when for every  $p \in f^{-1}(S)$  we have  $T_{f(p)}M = T_{f(p)}S + df_p(T_pN)$ , where  $df_p$  denotes the differential of  $f$  at  $p$ .

<sup>3</sup>This transversality theorem is based on Sard's theorem, which holds only for sufficiently smooth functions. One can verify that in our case, this application of Sard's theorem is valid when the phase coupling function  $f$  is merely  $C^1$ .

Let  $L^+$  denote the Moore–Penrose pseudoinverse of  $L$ , which is nonzero because  $L$  has a nonzero eigenvalue  $\lambda$ . Then we have

$$|x_1^T L x_1| = |x_1^T L L^+ L x_1| \leq \|L^+\| \cdot \|L x_1\|^2. \quad (104)$$

Let  $\sigma > 0$  denote the minimum eigenvalue of  $X + X^T$ , so that

$$\dot{\Upsilon}(x) = -x_1^T L (X + X^T) L x_1 \leq -\sigma \|L x_1\|^2. \quad (105)$$

It follows that for any  $x \in \mathcal{C}$  we have

$$\dot{\Upsilon}(x) < \frac{\sigma}{2\|L^+\|} \Upsilon(x) < 0. \quad (106)$$

Therefore  $\mathcal{C}$  is positively invariant and  $\Upsilon(x(t))$  exhibits exponential growth towards  $-\infty$  as  $t \rightarrow \infty$  from any initial state  $x(0) \in \mathcal{C}$ . Hence  $\|x(t)\| = \|\exp(\Lambda t)x(0)\|$  exhibits exponential growth as  $t \rightarrow \infty$  when  $x(0) \in \mathcal{C}$ , and we conclude that  $\Lambda$  must have an eigenvalue with a positive real part.

Next suppose that  $Z$  is invertible and  $L > 0$ ; then the largest invariant set contained in the set  $\{\dot{\Upsilon}(x) = 0\}$  is the origin  $x = 0$ . In this case  $\Upsilon$  is positive definite, hence asymptotic stability follows from the Krasovskii–LaSalle invariance theorem.

### C. Proof of Theorem 9

For the instability part of Theorem 6, we do not require the matrices  $L_1$  or  $Z$  to be invertible, even though they happen to be so in our application. However, the above Chetaev-style proof of Theorem 6 does not carry over to a proof of the instability part of Theorem 9; instead, we will apply Lemma 8. To apply this lemma to matrix  $\Lambda$  in (87), we must show that  $\Lambda$  has no eigenvalues on the imaginary axis. Suppose  $\Lambda v = j\lambda v$  for some  $\lambda \in \mathbb{R}$  and  $v \in \mathbb{C}^{2p}$ , and partition  $v$  as  $v = [v_1^T \ v_2^T]^T$  with  $v_1, v_2 \in \mathbb{C}^p$ . Then

$$Z v_2 = j\lambda v_1 \quad (107)$$

$$-Y L_1 v_1 - Y L_2 Z v_2 = j\lambda v_2. \quad (108)$$

If  $\lambda = 0$ , then  $v_1 = v_2 = 0$  because  $L_1$ ,  $Y$ , and  $Z$  are all invertible. Otherwise we solve for  $v_2$  in (107) and substitute into (108) to obtain

$$Y[\lambda^2 Y^{-1} Z^{-1} - L_1 - j\lambda L_2] v_1 = 0. \quad (109)$$

If we write  $v_1 = x + jy$  for  $x, y \in \mathbb{R}^p$ , then because  $Y$  is invertible we obtain

$$\begin{aligned} 0 &= [\lambda^2 Y^{-1} Z^{-1} - L_1 - j\lambda L_2](x + jy) \\ &= [Xx + \lambda L_2 y] + j[-\lambda L_2 x + Xy], \end{aligned} \quad (110)$$

where we have defined  $X = \lambda^2 Y^{-1} Z^{-1} - L_1$ . Setting the imaginary part to zero in (110) and solving for  $x$  yields

$$x = \frac{1}{\lambda} L_2^{-1} X y. \quad (111)$$



Setting the real part to zero in (110) and substituting (111) for  $x$  yields

$$0 = \left[ \frac{1}{\lambda} X L_2^{-1} X + \lambda L_2 \right] y. \quad (112)$$

Now (??) implies  $ZY = YZ^T$ , which means

$$\begin{aligned} (Y^{-1}Z^{-1})^T &= [(ZY)^{-1}]^T = [(ZY)^T]^{-1} \\ &= [YZ^T]^{-1} = [ZY]^{-1} \\ &= Y^{-1}Z^{-1}, \end{aligned} \quad (113)$$

and we conclude that  $X$  is symmetric. Thus the matrix  $X L_2^{-1} X + \lambda^2 L_2$  is symmetric and positive definite, and therefore (112) implies  $y = 0$ . Hence  $x = 0$  from (111), which means  $v_1 = 0$  and thus also  $v_2 = 0$  from (107). Therefore, if  $\Lambda v = j\lambda v$  for some  $\lambda \in \mathbb{R}$  and  $v \in \mathbb{C}^{2p}$  then  $v = 0$ , and we conclude that  $\Lambda$  has no eigenvalues on the imaginary axis.

Let  $H \in \mathbb{R}^{2p \times 2p}$  be the block diagonal matrix

$$H = \begin{bmatrix} -L_1^{-1} & 0 \\ 0 & -Z^{-1}Y \end{bmatrix}, \quad (114)$$

which is symmetric and invertible by assumption. Then a straightforward calculation yields

$$\Lambda H + H \Lambda^T = \begin{bmatrix} 0 & 0 \\ 0 & 2YL_2Y \end{bmatrix} \geq 0, \quad (115)$$

and the result follows from Lemma 8.

#### References

- [1] S. Bregni, "A historical perspective on telecommunications network synchronization," *Communications Magazine*, IEEE, vol. 36, no. 6, pp. 158–166, 1998.
- [2] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *Networking*, IEEE/ACM Transactions on, vol. 12, no. 3, pp. 493–506, 2004.
- [3] B. Sundararaman, U. Buy, and A. D. Kshemkalyani, "Clock synchronization for wireless sensor networks: a survey," *Ad Hoc Networks*, vol. 3, no. 3, pp. 281–323, 2005.
- [4] E. Mallada, X. Meng, M. Hack, L. Zhang, and A. Tang, "Skewless network clock synchronization," in *Network Protocols (ICNP)*, 2013 21st IEEE International Conference on, 2013.
- [5] I. Rec, "G. 812 timing requirements of slave clocks suitable for use as node clocks in synchronization networks," Geneva, June, 1998.
- [6] I. T. Union, "Itu-t recommendation g.811: timing characteristics of primary reference clocks," ITU, Tech. Rep., 97.
- [7] H. Karl and A. Willig, *Protocols and architectures for wireless sensor networks*. Wiley. com, 2007.
- [8] A. T. Winfree, "Biological rhythms and the behavior of populations of coupled oscillators," *Journal of theoretical biology*, vol. 16, no. 1, pp. 15–42, 1967.

- [9] Y. Kuramoto, “Cooperative dynamics of oscillator community a study based on lattice of rings,” *Progress of Theoretical Physics Supplement*, vol. 79, pp. 223–240, 1984.
- [10] R. E. Mirollo and S. H. Strogatz, “Synchronization of pulse-coupled biological oscillators,” *SIAM Journal on Applied Mathematics*, vol. 50, no. 6, pp. 1645–1662, 1990.
- [11] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*. Courier Dover Publications, 2003.
- [12] A. Jadbabaie, J. Lin, and A. S. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” *Automatic Control, IEEE Transactions on*, vol. 48, no. 6, pp. 988–1001, 2003.
- [13] W. Ren, R. W. Beard, and E. M. Atkins, “Information consensus in multivehicle cooperative control,” *Control Systems, IEEE*, vol. 27, no. 2, pp. 71–82, 2007.
- [14] F. Tong and Y. Akaiwa, “Theoretical analysis of interbase-station synchronization systems,” *Communications, IEEE Transactions on*, vol. 46, no. 5, pp. 590–594, 1998.
- [15] O. Simeone and G. Scutari, “Pulse-coupled distributed PLLs in heterogeneous wireless networks,” in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, 2007, pp. 1770–1775.
- [16] O. Simeone, U. Spagnolini, G. Scutari, and Y. Bar-Ness, “Physical-layer distributed synchronization in wireless networks and applications,” *Physical Communication*, vol. 1, no. 1, pp. 67–83, 2008.
- [17] C. H. Rentel and T. Kunz, “A mutual network synchronization method for wireless ad hoc and sensor networks,” *Mobile Computing, IEEE Transactions on*, vol. 7, no. 5, pp. 633–646, 2008.
- [18] L. Schenato and F. Fiorentin, “Average TimeSynch: a consensus-based protocol for clock synchronization in wireless sensor networks,” *Automatica*, vol. 47, no. 9, pp. 1878–1886, 2011.
- [19] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, “A PI consensus controller for networked clocks synchronization,” in *Proceedings of the 2008 IFAC World Congress*, (Seoul, South Korea), 2008, pp. 10 289–10 294.
- [20] R. Carli and S. Zampieri, “Networked clock synchronization based on second order linear consensus algorithms,” in *Proceedings of the 49th IEEE Conference on Decision and Control*, (Atlanta, Georgia), Dec. 2010, pp. 7259–7264.
- [21] R. Carli, E. D’Elia, and S. Zampieri, “A PI controller based on asymmetric gossip communications for clocks synchronization in wireless sensors networks,” in *Proceedings of the Joint 50th IEEE Conference on Decision and Control and European Control Conference*, (Orlando, Florida), Dec. 2011, pp. 7512–7517.
- [22] R Carli and S Zampieri, “Network Clock Synchronization Based on the Second-Order Linear Consensus Algorithm,” *Automatic Control, IEEE Transactions on*, vol. 59, no. 2, pp. 409–422, 2014.
- [23] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. H. Strogatz, “Distributed synchronization in wireless networks,” *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 81–97, Sep. 2008.

- [24] R. Carareto, F. M. Orsatti, and J. R. C. Piqueira, “Architectures, stability and optimization for clock distribution networks,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4672–4682, Dec. 2012.
- [25] J. Lunze, “Complete synchronization of Kuramoto oscillators,” *J. Phys. A: Math. Theor.*, vol. 44, no. 42, p. 425 102, 2011.
- [26] Y.-W. Hong and A. Scaglione, “Time synchronization and reach-back communications with pulse-coupled oscillators for uwb wireless ad hoc networks,” in *Ultra Wideband Systems and Technologies, 2003 IEEE Conference on*, IEEE, 2003, pp. 190–194.
- [27] D. Lucarelli, I.-J. Wang, et al., “Decentralized synchronization protocols with nearest neighbor communication,” in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, ACM, 2004, pp. 62–68.
- [28] X. Y. Wang, R. K. Dokania, and A. Apsel, “Pco-based synchronization for cognitive duty-cycled impulse radio sensor networks,” *Sensors Journal*, IEEE, vol. 11, no. 3, pp. 555–564, 2011.
- [29] E. Mallada and A. Tang, “Synchronization of weakly coupled oscillators: coupling, delay and topology,” *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 50, p. 505 101, Dec. 2013.
- [30] F. M. Gardner, *Phaselock Techniques*, 2nd ed. New York: Wiley, 1979.
- [31] F. Dörfler and F. Bullo, “Synchronization in complex networks of phase oscillators: A survey,” *Automatica*, May 2014.
- [32] Y. Kuramoto, “Self-entrainment of a population of coupled non-linear oscillators,” in *International Symposium on Mathematical Problems in Theoretical Physics*, H. Araki, Ed., vol. 39, ser. *Lecture Notes in Physics*, Springer, 1975, pp. 420–422.
- [33] E. Mallada and A. Tang, “Synchronization of phase-coupled oscillators with arbitrary topology,” in *Proceedings of the 2010 American Control Conference*, (Baltimore, Maryland), Jun. 2010, pp. 1777–1782.
- [34] —, “Distributed clock synchronization: Joint frequency and phase consensus,” in *Proceedings of the Joint 50th IEEE Conference on Decision and Control and European Control Conference*, (Orlando, Florida), Dec. 2011, pp. 6742–6747.
- [35] R. A. Freeman, “A global attractor consisting of exponentially unstable equilibria,” in *Proceedings of the 2013 American Control Conference*, (Washington, D.C.), Jun. 2013, pp. 4862–4867.
- [36] D. Carlson and H. Schneider, “Inertia theorems for matrices: the semi-definite case,” *Bull. Amer. Math. Soc.*, vol. 68, no. 5, pp. 479–483, 1962.
- [37] J. M. Lee, *Introduction to Smooth Manifolds*, ser. *Graduate Texts in Mathematics*. New York: Springer, 2012, vol. 218.
- [38] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, New Jersey: Prentice Hall, 2002.