IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity 27 (2014) 501–525 doi:10.1088/0951-7715/27/3/501

Towards a theory for diffusive coupling functions allowing persistent synchronization

Tiago Pereira[1], **Jaap Eldering**[1], **Martin Rasmussen**[1] and **Alexei Veneziani**[2]

[1] Department of Mathematics, Imperial College London, London SW7 2AZ, UK
[2] Centro de Matemática, Computação e Cognição, UFABC, Santo André, Brazil

E-mail: tiago.pereira@imperial.ac.uk, j.eldering@imperial.ac.uk and m.rasmussen@imperial.ac.uk

Received 8 August 2013, revised 15 December 2013
Accepted for publication 6 January 2014
Published 17 February 2014

Recommended by A E Motter

Abstract

We study synchronization properties of networks of coupled dynamical systems with interaction akin to diffusion. We assume that the isolated node dynamics possesses a forward invariant set on which it has a bounded Jacobian, then we characterize a class of coupling functions that allows for uniformly stable synchronization in connected complex networks—in the sense that there is an open neighbourhood of the initial conditions that is uniformly attracted towards synchronization. Moreover, this stable synchronization persists under perturbations to non-identical node dynamics. We illustrate the theory with numerical examples and conclude with a discussion on embedding these results in a more general framework of spectral dichotomies.

Keywords: synchronization, stability, persistence, networks
Mathematics Subject Classification: 05C82, 05C50, 34D06, 34D10, 34D20

(Some figures may appear in colour only in the online journal)

1. Introduction

Network synchronization is observed to occur in a broad range of applications in physics [35], neuroscience [6, 12, 20, 31] and ecology [8]. During the last 50 years, empirical studies of real complex systems have led to a deep understanding of the structure of networks [2, 21], and the interaction properties between oscillators, that is, the coupling function [18, 32, 36].

The stability of network synchronization is a balance between the isolated dynamics and the coupling function. Past research suggests that in networks of identical oscillators with interaction akin to diffusion, under mild conditions on the isolated dynamics, the coupling function dictates the synchronization properties of the network [19, 23–25, 36]. However, it still remains an open problem to describe the class of coupling functions that lead the network to persistent synchronization.

Our work contributes to the development of a general theory for coupling functions that allow for persistent synchronization for a connected complex network. The coupling functions under consideration appear in a variety of synchronization models on networks (such as the Kuramoto model [18] and its extensions [1, 5, 27]).

More precisely, we consider the dynamics of a network of n identical elements with interaction akin to diffusion, described by

$$\dot{x}_i = f(t, x_i) + \alpha \sum_{j=1}^{n} W_{ij} h(x_j - x_i), \tag{1}$$

where α is the overall coupling strength, and the matrix $W = (W_{ij})_{i,j\in\{1,\ldots,n\}}$ describes the interaction structure of the network, i.e. W_{ij} measures the strength of interaction between the nodes i and j. The function $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ describes the isolated node dynamics, and the coupling function $h : \mathbb{R}^m \to \mathbb{R}^m$ describes the diffusion-like interaction between nodes. We make the following two assumptions for these functions. Note that below and throughout the manuscript D_2 denotes the Jacobian with respect to the second argument.

Assumption A1. *The function f is continuous, and there exists an inflowing invariant open ball $U \subset \mathbb{R}^m$ such that f is continuously differentiable in U with*

$$\|D_2 f(t, x)\| \leqslant \varrho \quad \text{for all } t \in \mathbb{R} \text{ and } x \in U$$

for some $\varrho > 0$.

For instance, the Lorenz system has a bounded inflowing invariant ball, see section 3.2. In general, smooth nonlinear systems with compact attractors satisfy assumption A1. This assumption will be generalized in section 5 to also include noncompact sets U.

Assumption A2. *The coupling function h is continuously differentiable with $h(0) = 0$. We define $\Gamma := Dh(0)$ and denote the (complex) eigenvalues of Γ by β_i, $i \in \{1, \ldots, m\}$.*

The network structure plays a central role for the synchronization properties. We consider the intensity of the ith node $V_i = \sum_{j=1}^{n} W_{ij}$, and define the positive definite matrix $V := \mathrm{diag}(V_1, \ldots, V_n)$. Then the so-called *Laplacian* reads as

$$L = V - W.$$

Let λ_i, $i \in \{1, \ldots, n\}$, denote the eigenvalues of L. Note that $\lambda_1 = 0$ is an eigenvalue with eigenvector $\frac{1}{\sqrt{n}}(1, \ldots, 1)$. The multiplicity of this eigenvalue equals the number of connected components of the network.

The following assumption incorporates the coupling and structural network properties.

Assumption A3. *We suppose that*

$$\gamma := \min_{\substack{2\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}} \mathrm{Re}(\lambda_i \beta_j) > 0,$$

where $\mathrm{Re}(z)$ denotes the real part of a complex number z.

This assumption plays a central role in the results. In particular, γ controls the exponential decay to synchronization in theorem 1. See section 2.2 for a more detailed discussion.

The dynamics of such a diffusive model can be intricate. Indeed, even if the isolated dynamics possesses a globally stable fixed point, the diffusive coupling can lead to instability of the fixed point and the system can exhibit oscillatory behaviour [28].

Note that due to the diffusive nature of the coupling, if all oscillators start with the same initial condition, then the coupling term vanishes identically. This ensures that the globally synchronized state $x_1(t) = x_2(t) = \cdots = x_n(t) = s(t)$ is an invariant state for all coupling strengths α and all choices of coupling functions h. That is, the *diagonal manifold*

$$M := \left\{x_i \in \mathbb{R}^m \quad \text{for} \quad i \in \{1, \cdots, n\} : x_1 = \cdots = x_n\right\}$$

is invariant, and we call the subset

$$S := \left\{x_i \in U \subset \mathbb{R}^m \quad \text{for } i \in \{1, \cdots, n\} : x_1 = \cdots = x_n\right\} \subset M \tag{2}$$

the *synchronization manifold*. The main result of this paper is a proof that under the general conditions given above and α sufficiently large, the synchronization manifold S is uniformly exponentially stable.

Theorem 1 (Synchronization). *Consider the network of diffusively coupled equations (1) satisfying A1–A3. Then there exists a* $\rho = \rho(f, \Gamma)$ *such that for all coupling strengths*

$$\alpha > \frac{\rho}{\gamma},$$

the network is locally uniformly synchronized. This means that there exist a $\delta > 0$ *and a* $C = C(L, \Gamma) > 0$ *such that if* $x_i(t_0) \in U$ *and* $\|x_i(t_0) - x_j(t_0)\| \leqslant \delta$ *for any* $i, j \in \{1, \ldots, n\}$*, then*

$$\|x_i(t) - x_j(t)\| \leqslant C\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)}\|x_i(t_0) - x_j(t_0)\| \qquad \text{for all } t \geqslant t_0. \tag{3}$$

Hence, the synchronization manifold is locally uniformly exponentially attractive. The constant ρ depends on the bounds on the Jacobian $D_2 f$ as set out in assumption A1 and on the condition number $\kappa(\Gamma) = \|\Gamma\|\|\Gamma^{-1}\|$ of the matrix Γ (see (27) in the case Γ is diagonalizable). In the case that the Laplacian L and Γ are diagonalizable, C depends on the condition number of the similarity transformation that diagonalizes these matrices (see lemma 8 for details), so loosely speaking, it depends on how well the eigenvectors of L and Γ are orthogonal. If L and Γ are non-diagonalizable, then C is related to condition numbers as well, see the proof of lemma 9 for details. The size of δ can be estimated explicitly if more concrete details about the system are known, also see remark 15.

Our second main result shows that synchronization is persistent under perturbation of the isolated nodes. Consider a network of non-identical nodes described by

$$\dot{x}_i = f_i(t, x_i) + \alpha \sum_{j=1}^{n} W_{ij} h(x_j - x_i), \tag{4}$$

where $f_i(t, x_i) = f(t, x_i) + g_i(t, x_i)$. Note that, in this case, the synchronization manifold S is no longer invariant. We show in this paper that for small perturbation functions g_i, $i \in \{1, \ldots, n\}$, the synchronization manifold is stable in the sense that orbits starting near the synchronization manifold S remain in a neighbourhood of S.

Theorem 2 (Persistence). *Consider the perturbed network (4) of diffusively coupled equations fulfilling assumptions A1–A3, and suppose that*

$$\alpha > \frac{\rho}{\gamma}$$

as in theorem 1. Then there exist $\delta > 0$, $C > 0$ *and* $\varepsilon_g > 0$ *such that for all* ε_0*-perturbations satisfying*

$$\|g_i(t,x)\| \leqslant \varepsilon_0 \leqslant \varepsilon_g \qquad \text{for all} \quad t \in \mathbb{R}, \quad x \in U \text{ and } i \in \{1,\dots,n\} \tag{5}$$

and initial conditions satisfying $\|x_i(t_0) - x_j(t_0)\| \leqslant \delta$ *for any* $i, j \in \{1,\dots,n\}$*, the estimate*

$$\left\|x_i(t) - x_j(t)\right\| \leqslant C\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)}\left\|x_i(t_0) - x_j(t_0)\right\| + \frac{C\varepsilon_0}{\alpha\gamma - \rho} \qquad \text{for all } t \geqslant t_0 \tag{6}$$

holds.

Note that the additional term $C\varepsilon_0/(\alpha\gamma - \rho)$ can be made small either by controlling the perturbation size ε_0 or by increasing $\alpha\gamma$. This provides control of the network coherence in terms of the network properties and coupling strength.

If the Laplacian L is symmetric (i.e. the systems are mutually coupled), its spectrum is real and can be ordered as $0 = \lambda_1 < \lambda_2 \leqslant \lambda_3 \leqslant \dots \leqslant \lambda_n$. Moreover, consider $\beta := \min_{i\in\{1,\dots,m\}} \mathrm{Re}\beta_i$, and note that this implies

$$\gamma = \beta\lambda_2.$$

The following corollary to the above persistence result then shows that the enhancement of coherence in the network in terms of network connectivity depends on the spectral gap λ_2.

Corollary 3 (Synchronization error). *Consider the perturbed network (4) with symmetric Laplacian* L *and the average synchronization error*

$$e_s(t) = \frac{1}{n(n-1)}\sum_{i,j=1}^{n} \|x_i(t) - x_j(t)\| \qquad \text{for all } t \geqslant t_0,$$

where the initial conditions $x_i(t_0)$, $i \in \{1,\dots,n\}$*, are chosen as in theorem 2. Then whenever* $\alpha\gamma = \alpha\beta\lambda_2 > \rho$*, one has*

$$\limsup_{t\to\infty} e_s(t) \leqslant K\frac{\varepsilon_0}{\alpha\beta\lambda_2 - \rho},$$

where $K = K(\Gamma)$ *is independent of the network size.*

This corollary has excellent agreement with recent numerical simulations for the synchronization transition in complex networks of mutually coupled non-identical oscillators [26] and generalizes the case studied for nearly identical coupled systems [33].

The paper is organized as follows. In section 2, we discuss our assumptions, ideas of the proofs as well as how our results relate to previous contributions. In section 3, we illustrate our main synchronization result with a nonautonomous linear system and a coupled Lorenz system. Section 4 provides fundamental results on nonautonomous linear differential equations. In section 5, we provide auxiliary results to prove our main theorems in sections 6 and 7. Finally, in section 8, we discuss how to generalize this theory using the dichotomy spectrum and normal hyperbolicity.

Notation. We endow the vector space $\mathbb{R}^m$ with the Euclidean norm $\|x\| = \sqrt{\sum_{i=1}^m |x_i|^2}$ and the associated Euclidean inner product. In addition, we equip the vector space $(\mathbb{R}^m)^n = \mathbb{R}^{nm}$ with the norm

$$\|(x_1,\dots,x_n)\| := \max_{i=1,\dots,n} \|x_i\| \qquad \text{where } x_i \in \mathbb{R}^m\,. \tag{7}$$

Note that linear operators on the above spaces will be equipped with the induced operator norm. For a given invertible matrix $A \in \mathbb{R}^{d\times d}$, the *condition number* is defined by $\kappa(A) = \|A\|\|A^{-1}\|$. Note that the condition number depends on the underlying operator norm. Finally, the symbol I_d stands for the identity matrix in $\mathbb{R}^d$.

2. Discussion of the main results

This section is devoted to relating our results to the state of the art and to explaining the assumptions and the central ideas of the proofs.

2.1. State of the art

Recent research on synchronization has focused on the role of the coupling function for the stability of network synchronization. Notably, Pecora and collaborators have developed so-called *master stability functions*, which typically use Lyapunov exponents corresponding to the transversal directions of the synchronization manifold *as a stability criterion* [15, 24, 33]. In contrast to this approach, we estimate the contraction rate by dichotomy techniques. Our results show that the synchronization state is locally stable and persistent, and thus stable under small perturbations. This means that the phenomenon of bubbling [3, 30] and riddling [13] (which leads to synchronization loss) will not be observed under our conditions, in contrast to the maximum Lyapunov exponent.

Another aspect of our results is that the synchronization properties do not depend on diagonalization properties of the Laplacian. Recently, the master stability function has been extended to include non-diagonalizable Laplacians [22]. If one considers the master stability function with the maximum Lyapunov exponent as a stability criterion, one cannot guarantee that an open neighbourhood of the synchronization manifold will be attracted by the synchronization manifold, nor does it imply persistence of the synchronization. In our set-up, these properties follow naturally by means of persistence of exponential dichotomies, which is relevant in applications that are subject to noise and external influences. Note that the master stability function approach is applicable to a broader class of coupling functions than the ones we consider, but our approach is constructive and making use of further dichotomy techniques and normal hyperbolicity our results can be generalized further, as discussed later in section 8.

In addition, Pogromsky and Nijmeijer [29] use control techniques to show that if the coupling function is linear and given by a symmetric positive definite matrix, then the synchronization manifold is globally asymptotically stable for connected networks. Likewise, Belykh *et al* [4] develop a connection graph stability method to obtain global synchronization for the network, by assuming the existence of a quadratic Lyapunov function associated with the isolated system. In this paper, we tackle only local stability properties, but we consider a more general class of coupling functions. However, under additional conditions on the dynamics and coupling functions, it is possible to prove global stability with the techniques we have developed by applying the mean value theorem instead of using Taylor expansions of the vector field.

2.2. The assumptions

Assumption A1 concerns the existence of solutions and the boundedness of the Jacobian, this is the case if the isolated dynamics is dissipative. Assumption A2 makes it possible to characterize the stability of synchronization by the linearization of h. Assumption A3 guarantees that the eigenvalues of the tensor (or Kronecker) product $L \otimes \Gamma$ have real part bounded away from zero (except for the trivial eigenvalue).

These hypotheses basically imply that with a finite value of α, we are able to damp all the instabilities of the vector field and obtain a stable synchronization state. If, for example, assumption A3 is dropped, γ may become negative and synchronization may no longer be possible.

We illustrate the relevance of assumption A3 with the following example. Consider the isolated dynamics $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x) = -\varepsilon x$, for any $\varepsilon > 0$. Moreover, consider three coupled systems

$$\begin{aligned}\dot{x}_1 &= f(x_1) + 2\alpha\Gamma(x_2 - x_1) + \alpha\Gamma(x_3 - x_1),\\ \dot{x}_2 &= f(x_2) + 2\alpha\Gamma(x_3 - x_2)\\ \dot{x}_3 &= f(x_3) + \alpha\Gamma(x_1 - x_3)\end{aligned}$$

with

$$\Gamma = \begin{pmatrix} 2 & 1 \\ -17 & 0 \end{pmatrix} \qquad \text{and note that} \qquad L = \begin{pmatrix} 3 & -2 & -1 \\ 0 & 2 & -2 \\ -1 & 0 & 1 \end{pmatrix}.$$

The eigenvalues of L are $\lambda_1 = 0$, $\lambda_2 = 3 + i$ and $\lambda_3 = 3 - i$ and the eigenvalues of Γ are $\beta_1 = 1 + 4i$ and $\beta_2 = 1 - 4i$. Hence,

$$\gamma = -1,$$

and although the isolated dynamics has a stable trivial fixed point, for any $\alpha > \varepsilon$ the origin is unstable and there are trajectories of the coupled systems that escape any compact set. This shows that breaking condition A3 can have severe effects on the dynamics of the coupled systems.

Assumption A3 has not been considered in the literature to our best knowledge. In the following, we rephrase this condition in the following two special cases.

(i) *The spectrum of* Γ *is positive.* If Γ has a spectrum consisting of only real, positive eigenvalues, then A3 has a representation in terms of the Laplacian. In this case, this condition reads as

$$\mathrm{Re}(\lambda_i) > 0 \qquad \text{for all } i \neq 1,$$

since the Laplacian always has a zero eigenvalue. If the network is connected, this eigenvalue is simple, and by virtue of the disc theorem, a sufficient condition for all other eigenvalues to have positive real part is *positive interaction strength*, i.e. $W_{ij} > 0$ whenever i is connected to j, and zero otherwise.

(ii) *The Laplacian is symmetric.* This is the most studied case in the literature. Assume that the network is connected. Since the spectrum of the Laplacian is real, assumption A3 requires that the real part of the spectrum of Γ is positive and that the spectrum of the Laplacian is positive apart from the single zero eigenvalue (or alternatively, that the spectra of Γ and the Laplacian are both negative, but note that this is non-physical).

Note in general that the $\lambda_i\beta_j$ in assumption A3 describe the eigenvalues of $L\otimes\Gamma$ transverse to the synchronization manifold. Since the transverse dynamics is governed by $-L \otimes \Gamma$, we precisely obtain synchronization if these have positive real part. It is not sufficient that both the λ_i and the β_j have positive real part, as their product may still lie left of the imaginary axis.

2.3. Ideas of the proofs

The proofs of our main results rely on identifying the synchronization problem with a corresponding fixed point problem. We first concentrate on the case of diagonalizable Laplacians, where diagonal dominance (proposition 6) can be used to show that the synchronized state is uniformly asymptotically stable. To obtain the claim for general coupling functions, we make use of the persistence property associated with the equilibrium point (theorem 5). The main aspect here is to approximate the coupling function by a diagonalizable one while keeping control of the contraction rates. Finally, the proof for general Laplacians follows from the fact that the set of diagonalizable Laplacians is dense in the space of Laplacians. From these results and the persistence property the main claim follows.

3. Illustrations

Before proving the two main results of this paper, two examples are discussed.

3.1. Nonautonomous linear equations

Consider the nonautonomous linear equation

$$\dot{x} = A(t)x \tag{8}$$

where

$$A(t) = \begin{pmatrix} -1 - 9\cos^2(6t) + 12\sin(6t)\cos(6t) & 12\cos^2(6t) + 9\sin(6t)\cos(6t) \\ -12\sin^2(6t) + 9\sin(6t)\cos(6t) & -1 - 9\sin^2(6t) - 12\sin(6t)\cos(6t) \end{pmatrix}.$$

This is a prototypical example where the eigenvalues of the time-dependent matrices do not characterize the stability of a nonautonomous linear system. Indeed, the eigenvalues of $A(t)$ are -1 and -10, independent of $t \in \mathbb{R}$, and a direct computation shows that

$$x(t) = \begin{pmatrix} \mathrm{e}^{2t}(\cos(6t) + 2\sin(6t)) + 2\mathrm{e}^{-13t}(2\cos(6t) - \sin(6t)) \\ \mathrm{e}^{2t}(\cos(6t) - 2\sin(6t)) + 2\mathrm{e}^{-13t}(2\cos(6t) - \sin(6t)) \end{pmatrix}$$

is a solution of the system, which does not converge to 0 as $t \to \infty$.

Consider now two diffusively coupled systems

$$\dot{x}_1 = A(t)x_1 + \alpha\Gamma(x_2 - x_1),$$
$$\dot{x}_2 = A(t)x_2 + \alpha\Gamma(x_1 - x_2),$$

where Γ is a real 2×2 matrix. Theorem 1 yields that it is possible to synchronize these two systems for any coupling matrix with $\beta(\Gamma) > 0$. Consider the coupling matrix

$$\Gamma = \begin{pmatrix} \beta & 1 \\ 0 & \beta \end{pmatrix}.$$

Γ is in its Jordan form and non-diagonalizable. The transformation $y = x_1 - x_2$ leads to

$$\dot{y} = \big(A(t) - 2\alpha\Gamma\big)y. \tag{9}$$

Our main result shows that the trivial solution of (9) is stable if α is large enough.

We have integrated (9) using a sixth order Runge–Kutta method with step size 0.001. We have computed the critical coupling value α_c as a function of β. Using a bisection technique we estimate with precision 10^{-3} the critical value of α_c such that the trivial solution of equation (9) is stable. In figure 1 we plot the corresponding critical value $\rho_\mathrm{c} = \beta\alpha_\mathrm{c}$. Hence, we are able to analyse the dependence of ρ on Γ. The behaviour of ρ appears to be intricate. For large β, we obtain that ρ tends to a constant; however, as we decrease β, various changes in the behaviour can be observed. Although the problem is linear, the critical coupling strength depends nonlinearly on the parameter β. We analyse this dependence in more detail in section 6.1

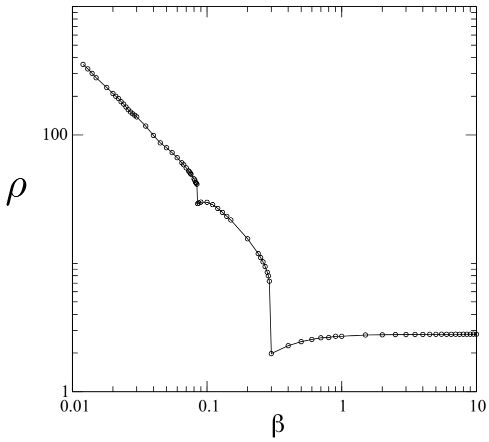
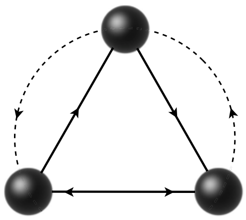


Figure 1. $\rho = \rho(f, \Gamma)$ as a function of β on a log–log scale for a fixed f given by equation (8). For small β the slope is -1 in good approximation.



$$W = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & a \\ a & 1 & 0 \end{pmatrix}$$

Figure 2. The network and its weight matrix. The matrix $L = V - W$ is non-diagonalizable for every $a \neq 1$; here we choose $a = \frac{1}{3}$.

3.2. The Lorenz system

Using the notation $x = (u, v, w)$, the Lorenz vector field is given by

$$f(x) = \begin{pmatrix} \sigma(v-u) \\ u(r-w)-v \\ -bw+uv \end{pmatrix},$$

where we choose the classical parameter values $\sigma = 10$, $r = 28$ and $b = \frac{8}{3}$. All trajectories of the Lorenz system eventually enter a compact set and therefore they exist for all positive times. Moreover, the trajectories accumulate in the neighbourhood of a chaotic attractor [34].

Consider the network of three coupled Lorenz systems

$$\dot{x}_i = f(x_i) + \alpha \sum_{j=1}^{3} W_{ij} H(x_j - x_i), \tag{10}$$

where the interaction matrix W is given as in figure 2.

We use two different nonlinear coupling functions; for the first, the associated matrix Γ is positive definite, whereas for the second, Γ is a Jordan block. The specific forms of the coupling functions can be seen in figure 3. We have integrated (10) using a sixth order Runge–Kutta method with step size 0.0001 and computed the critical coupling α_c as a function of β, and then plotted the value $\rho_c = \alpha_c\beta$ (see figure 3). The behaviour of ρ depends in an essential way on Γ. This behaviour is further discussed in section 6.1.

$$h(x) = \begin{pmatrix} \beta u + w^2 \\ uv + \beta \sin v \\ \beta w(1-u) \end{pmatrix} \quad \Gamma = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \beta \end{pmatrix} \qquad h(x) = \begin{pmatrix} \beta u + v \\ \beta \sin v + w \\ \beta w(1-u) \end{pmatrix} \quad \Gamma = \begin{pmatrix} \beta & 1 & 0 \\ 0 & \beta & 1 \\ 0 & 0 & \beta \end{pmatrix}$$

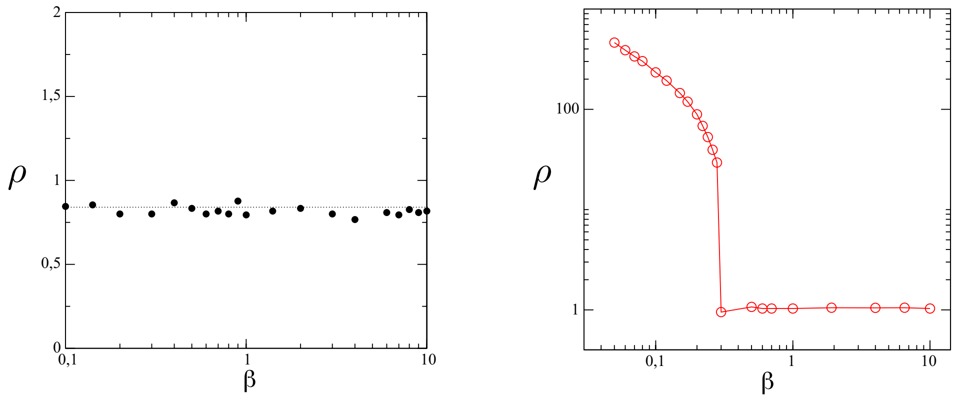


Figure 3. Simulation results for ρ for the two coupling functions. For the first case, see left side, $\Gamma = \beta I$ is positive definite for $\beta > 0$, and the behaviour of ρ does not depend significantly on β. For the second case, Γ is a Jordan block with eigenvalues equal to β. In this situation, for large values of β, the critical coupling ρ appears independent of β, as opposed to the small values of β. In that case, the critical coupling scales as $\rho \propto \beta^{-1}$.

4. Nonautonomous linear differential equations

Consider the m-dimensional linear differential equation

$$\dot{x} = A(t)x, \tag{11}$$

where $x \in \mathbb{R}^m$ and $A : \mathbb{R} \to \mathbb{R}^{m\times m}$ is a bounded and continuous matrix function. Recall that solutions of (11) can be written in terms of the evolution operator $\Phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m\times m}$; the solution for the initial condition $x(t_0) = x_0$ is given by

$$t \mapsto \Phi(t, t_0)x_0.$$

Definition 4 (Uniform exponential stability). *Consider the linear system (11) with evolution operator Φ. System (11) is said to be uniformly exponentially stable if there exists $K, \mu > 0$ such that*

$$\|\Phi(t, t_0)\| \leqslant K\mathrm{e}^{-\mu(t-t_0)} \qquad \text{for all } t \geqslant t_0. \tag{12}$$

The following persistence theorem guarantees that uniform exponential stability is persistent under perturbations. A proof can be found in [7, lecture 4, proposition 1].

Theorem 5 (Persistence of exponential stability). *Consider the linear system (11) and assume that for $K > 0$ and $\mu \in \mathbb{R}$, the evolution operator Φ satisfies the exponential estimate*

$$\|\Phi(t, t_0)\| \leqslant K\mathrm{e}^{-\mu(t-t_0)} \qquad \text{for all } t \geqslant t_0. \tag{13}$$

Consider a continuous matrix function $V : \mathbb{R} \to \mathbb{R}^{m\times m}$ such that

$$\delta := \sup_{t\in\mathbb{R}} \|V(t)\| < \infty.$$

Then the evolution operator $\hat{\Phi}$ of the perturbed equation

$$\dot{y} = \big(A(t) + V(t)\big)y$$

satisfies the exponential estimate

$$\|\hat{\Phi}(t,t_0)\| \leqslant K\mathrm{e}^{-\hat{\mu}(t-t_0)} \qquad \text{for all } t \geqslant t_0,$$

where $\hat{\mu} := \mu - \delta K$.

There are various criteria to obtain conditions for uniform exponential stability. We shall use the following criterion for diagonal dominant matrices, which can be found in [7, lecture 6, proposition 3].

Proposition 6 (Diagonal dominance criterion). *Consider the linear system (11) with complex time-dependent coefficient matrices* $A(t) = (A_{ij}(t))_{i,j=1,\dots,m}$, *and suppose that there exists a constant* $\mu > 0$ *such that*

$$\mathrm{Re}(A_{ii}(t)) + \sum_{\substack{j=1,\\ j\neq i}}^{m} |A_{ij}(t)| \leqslant -\mu < 0 \qquad \text{for all } t \in \mathbb{R} \text{ and } i \in \{1,\dots,m\}\,. \tag{14}$$

Then the evolution operator Φ *of (11) satisfies*

$$\|\Phi(t,t_0)\| \leqslant K\mathrm{e}^{-\mu(t-t_0)} \qquad \text{for all } t \geqslant t_0.$$

with $K = K(m) \geqslant 1$.

5. Auxiliary results

In this section, we obtain various exponential estimates for orbits near the synchronization manifold S of (1). First, we introduce a convenient splitting of coordinates along the synchronization manifold and complementary to it, and derive the equations with respect to these coordinates. Then we prove linear stability of the synchronization manifold. Here we distinguish between diagonalizable and non-diagonalizable Laplacians. The latter case will follow from approximation results on diagonalizable Laplacians and persistence of the exponential estimates. Finally, we introduce the concept of a tubular neighbourhood as a final ingredient to tackle the general proof of nonlinear stability.

In order to treat noncompact absorbing sets U in assumption A1, we reformulate this assumption as follows.

Assumption A1′. *The function* f *is continuous in the first argument and continuously differentiable in the second argument, and there exists an open simply connected set* $U \subset \mathbb{R}^m$ *with* C^1*-boundary that is* ε*-inflowing invariant for some* $\varepsilon > 0$*, i.e. for all* $x \in \partial U$ *with inward-pointing normal vector* q_x*, we have*

$$\langle q_x, f(t,x)\rangle \geqslant \varepsilon \qquad \text{for all } t \in \mathbb{R} \text{ and } x \in \partial U. \tag{15}$$

Moreover, there exists a $\Delta > 0$ *such that the Jacobian* $D_2 f$ *is uniformly continuous and bounded on* $B_\Delta(U) := \bigcup_{x\in U}\{y \in \mathbb{R}^m : \|x-y\| < \Delta\}$*, i.e. for some* $\varrho > 0$*, we have*

$$\|D_2 f(t,x)\| \leqslant \varrho \qquad \text{for all } t \in \mathbb{R} \text{ and } x \in B_\Delta(U).$$

Note that if the closure $\bar{U}$ is compact, then uniformity of the inflowing invariance condition as well as the uniform continuity of $D_2 f$ and existence of a bound ϱ follow automatically. In the noncompact case, we require uniform bounds on the Δ-enlarged neighbourhood $B_\Delta(U)$ for technical reasons.

We first obtain equations that govern the dynamics near the synchronization manifold. Using a tensor representation, we can write the nm-dimensional system (1) equations by means of a single equation. To this end, define

$$X := \mathrm{col}(x_1,\dots,x_n),$$

where col denotes the vector formed by stacking the column vectors x_i into a single column vector. Similarly, define

$$F(t, X) := \mathrm{col}(f(t, x_1), \ldots, f(t, x_n)).$$

We can analyse small perturbations away from the synchronization manifold in terms of the tensor representation

$$X = \mathbb{1} \otimes s + \xi, \tag{16}$$

where $\otimes$ is the tensor (or Kronecker) product and $\mathbb{1} = \mathrm{col}(1, \ldots, 1) \in \mathbb{R}^n$, which is the eigenvector of L corresponding to the eigenvalue zero. Note that $\mathbb{1} \otimes s$ defines the diagonal manifold with s the state of each node in the synchronous state of the network. We view ξ as a perturbation to the synchronized state.

The state space $\mathbb{R}^n \otimes \mathbb{R}^m$ can be canonically identified with $\mathbb{R}^{nm}$, which we will use for shorter notation. The coordinate splitting (16) is associated with a splitting of $\mathbb{R}^{nm}$ as the direct sum of subspaces

$$\mathbb{R}^{nm} = M \oplus N$$

with associated projections

$$\pi_M : \mathbb{R}^{nm} \to M, \qquad \pi_N : \mathbb{R}^{nm} \to N.$$

The subspaces $M, N \subset \mathbb{R}^{nm}$ are determined by embeddings from $\mathbb{R}^m$ and $\mathbb{R}^{(n-1)m}$, respectively, induced by the Laplacian L on $\mathbb{R}^n$.

Let us for the moment use the simplifying assumption that L is diagonalizable with eigenvectors $\mathbb{1}, v_2, \ldots, v_n$. Then the subspaces M, N have natural representations in terms of these eigenvectors as

$$M = \mathrm{span}(\mathbb{1}) \otimes \mathbb{R}^m, \qquad N = \mathrm{span}(v_2, \ldots, v_n) \otimes \mathbb{R}^m.$$

This means that we have 'natural' embeddings that induce coordinates on these subspaces:

$$\begin{aligned} \iota_M &: \mathbb{R}^m \to M, & s &\mapsto \mathbb{1} \otimes s = \mathrm{col}(s, \ldots, s), \\ \iota_N &: \mathbb{R}^{(n-1)m} \to N, & (y_2, \ldots, y_n) &\mapsto \sum_{j=2}^{n} v_j \otimes y_j. \end{aligned}$$

If we drop the assumption that L is diagonalizable, then we lose the natural choice of an embedding for N. Note, however, that N is still determined as the eigenspace of all non-zero eigenvalues.

Note that the norm on $\mathbb{R}^{nm}$ we chose is the maximum over the Euclidean norm on $\mathbb{R}^m$, see (7). The norm $\|\cdot\|$ on $\mathbb{R}^{nm}$ can be restricted to the subspaces M, N and induces norms on the 'coordinate' spaces $\mathbb{R}^m$ and $\mathbb{R}^{(n-1)m}$ by pullback under the embeddings. Then the induced norm on $s \in \mathbb{R}^m$ is given by

$$\|s\|_{\iota_M} = \|\iota_M(s)\| = \|\mathbb{1} \otimes s\|, \tag{17}$$

which is precisely the Euclidean norm. Similarly, ι_M induces an inner product on M. Henceforth, we shall identify $s \in \mathbb{R}^m$ with $\mathbb{1} \otimes s \in M$ under the isometry ι_M.

Using the representation (16) for $X \in \mathbb{R}^{nm}$, given an initial condition $X_0 = (s_0, \xi_0)$, the corresponding solution to (1) reads as $X(t) = (s(t), \xi(t))$. In the next result, we derive differential equations for these two components in a neighbourhood of the synchronization manifold.

Proposition 7. *The two components of the solution $X(t) = (s(t), \xi(t))$ satisfy the system of equations*

$$\mathbb{1} \otimes \dot{s} = \mathbb{1} \otimes f(t, s) + R_s(s, \xi), \tag{18}$$

$$\dot{\xi} = T(t, s)\xi + R_\xi(s, \xi), \tag{19}$$

where

$$T(t, s) = I_n \otimes D_2 f(t, s) - \alpha(L \otimes \Gamma) \tag{20}$$

and $R_* := R_s, R_\xi$ are the remainder functions such that for any $\varepsilon > 0$, there is a $\delta > 0$ such that for all $\|\xi\| \leqslant \delta$, one has $\|R_*(s, \xi)\| \leqslant \varepsilon\|\xi\|$.

Proof. By assumption A2, Taylor's theorem implies that given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$h(x) = \Gamma x + r(x) \qquad \text{with } \|r(x)\| \leqslant \varepsilon\|x\| \text{ whenever } \|x\| \leqslant \delta.$$

Now we define

$$R_h(X)_i = \sum_{j=1}^{n} W_{ij} r(x_i - x_j) = \sum_{j=1}^{n} W_{ij} r\big(p_i(\mathbb{1} \otimes s + \xi) - p_j(\mathbb{1} \otimes s + \xi)\big)$$
$$= \sum_{j=1}^{n} W_{ij} r\big(p_i(\xi) - p_j(\xi)\big),$$

where $p_i : \mathbb{R}^{nm} \to \mathbb{R}^m$ maps canonically to the ith component of the argument, $i \in \{1, \ldots, n\}$. The vectors $R_h(X)_i \in \mathbb{R}^m$, $i \in \{1, \ldots, n\}$ define a vector in $\mathbb{R}^{nm}$. Note that $R_h(X) = R_h(\xi)$ does not depend on $s \in M$ and satisfies the estimate

$$\|R_h(\xi)\| \leqslant \max_{i=1,\ldots,n} \left(\sum_{j=1}^{n} |W_{ij}| \right) \varepsilon\, 2\|\xi\| \qquad \text{whenever } \|\xi\| \leqslant \frac{\delta}{2}.$$

Recall that $L_{ij} = \delta_{ij} V_i - W_{ij}$, so the coupling term can then be rewritten as

$$\sum_{j=1}^{n} W_{ij} h(x_j - x_i) = -\sum_{j=1}^{n} L_{ij} \Gamma x_j + R_h(\xi)_i. \tag{21}$$

The Taylor expansion of $F(t, X)$ around $\mathbb{1} \otimes s$ reads as

$$F(t, \mathbb{1} \otimes s + \xi) = F(t, \mathbb{1} \otimes s) + D_2 F(t, \mathbb{1} \otimes s)\xi + R_F(t, s, \xi)$$
$$= \mathbb{1} \otimes f(t, s) + I_n \otimes D_2 f(t, s)\xi + R_F(t, s, \xi),$$

where $\|R_F(t, s, \xi)\| \leqslant \varepsilon\|\xi\|$ when $\|\xi\| \leqslant \delta$. An algebraic manipulation of (21) allows a representation in coordinates $(s, \xi) \in M \oplus N$ of the n equations forming (1):

$$\dot{X} = \mathbb{1} \otimes \dot{s} + \dot{\xi} = \mathbb{1} \otimes f(t, s) + I_n \otimes D_2 f(t, s)\xi - \alpha(L \otimes \Gamma)\xi + R_F(t, s, \xi) + \alpha R_h(\xi), \tag{22}$$

where we used $L\,\mathbb{1} = 0$. Hence, the term $(L \otimes \Gamma)(\mathbb{1} \otimes s)$ vanishes.

Next, we project the differential equation (22) onto the spaces M and N to obtain differential equations for s and ξ:

$$\mathbb{1} \otimes \dot{s} = \mathbb{1} \otimes f(t, s) + \pi_M(R_F(t, s, \xi) + \alpha R_h(\xi)),$$
$$\dot{\xi} = T(t, s)\xi + \pi_N(R_F(t, s, \xi) + \alpha R_h(\xi)),$$

where

$$T(t, s) = I_n \otimes D_2 f(t, s) - \alpha(L \otimes \Gamma).$$

Note that both $I_n \otimes D_2 f(t, s)$ and $L \otimes \Gamma$ preserve the subspaces M and N, since I_n and L preserve both $\mathrm{span}(\mathbb{1})$ and $\mathrm{span}(v_2, \ldots, v_n)$, so the projections can be dropped there. □

5.1. Diagonalizable Laplacians

We now prove stability of the linear flow (20) for $\xi \in N$, along any curve $s(t) \in S$, that is, we do not assume $s(t)$ to be a solution curve. We first treat the diagonalizable case and then the non-diagonalizable one. Then, in section 6, we use these results to prove stability of the fully nonlinear problem.

Lemma 8 (Diagonalizable case). *Consider the linearization of (19), given by*

$$\dot{\xi} = T(t, s(t))\xi, \qquad \xi \in N \tag{23}$$

with $s(t) \in U$, and the representations

$$L = P\Lambda P^{-1} \qquad \text{and} \qquad \Gamma = QBQ^{-1}$$

with $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$, such that $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $B = \mathrm{diag}(\beta_1, \ldots, \beta_m)$. Then there exists a $\rho > 0$ such that for all coupling strengths

$$\alpha > \frac{\rho}{\gamma},$$

the evolution operator Φ of (23) satisfies the estimate

$$\|\Phi(t, t_0)\| \leqslant K\kappa(P \otimes Q)\,\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)} \quad \text{for all } t \geqslant t_0,$$

with $K \geqslant 1$, and where $\kappa(P \otimes Q)$ denotes the condition number of $P \otimes Q$.

Note that (23) is well-defined on N since $T(t, s)$ preserves the splitting $M \oplus N$. Furthermore, for matrices $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$, using the properties of induced norm, we obtain[3]

$$\|P \otimes Q\| = \|P\|_\infty \|Q\|_2\,,$$

which implies that $\kappa(P \otimes Q) = \kappa_\infty(P)\kappa_2(Q)$.

Proof of lemma 8. Note that $O := P \otimes Q$ is an invertible matrix that diagonalizes $L \otimes \Gamma$, and the change of coordinates

$$\tilde{T}(t) = O^{-1}\, T(t, s(t))\, O = I_n \otimes Q^{-1}\, D_2 f(t, s(t))\, Q - \alpha\, \Lambda \otimes B \tag{24}$$

reduces $T(t)$ to m-block diagonal form. Thus, we have

$$\tilde{T}(t) = \bigoplus_{i=1}^{n} \tilde{T}_i(t) = \mathrm{diag}\big(\tilde{T}_1(t), \ldots, \tilde{T}_n(t)\big),$$

where

$$\tilde{T}_i(t) := \underbrace{Q^{-1}\, D_2 f(t, s(t))\, Q}_{\tilde{A}(t):=} - \alpha\, \lambda_i\, B \quad \text{for all } t \in \mathbb{R}.$$

Since for all $t \in \mathbb{R}$, the matrix $\tilde{T}(t)$ is block diagonal, the dynamics given by $\dot{Y} = \tilde{T}(t)Y$ preserves the splitting $\mathbb{R}^{nm} = \bigoplus_{i=1}^{n} \mathbb{R}^m$, and hence, its associated evolution operator $\tilde{\Phi}$ is also of the form

$$\tilde{\Phi}(t, t_0) = \bigoplus_{i=1}^{n} \tilde{\Phi}_i(t, t_0) \qquad \text{for all } t, t_0 \in \mathbb{R}, \tag{25}$$

[3] We have $\|(P \otimes Q)X\| = \max_i \|\sum_j P_{ij}(Qx_j)\|_2 \leqslant \|Q\|_2(\max_i \sum_j |P_{ij}|)(\max_j \|x_j\|_2) = \|Q\|_2\|P\|_\infty\|X\|$, from which '$\leqslant$' follows. Equality is reached for $X = \xi \otimes x$ when ξ and x realize the operator norms for P and Q, respectively.

where each $\tilde{\Phi}_i$ is the evolution operator of $\dot{y}_i = \tilde{T}_i(t)y_i$. Note that restricting T to N corresponds to restricting $\tilde{T}$ to the blocks $i \geqslant 2$. The dynamics in each block is determined by

$$\dot{y}_i = (\tilde{A}(t) - \alpha\,\lambda_i\,B)y_i. \tag{26}$$

Now define

$$\tilde{\rho} := \sup_{t\in\mathbb{R},\, s\in U} \left\|\tilde{A}(t)\right\|.$$

Note that the matrix $\tilde{A}(t)$ depends implicitly on $s(t) \in U$, so by assumption A1 we get the estimate

$$\tilde{\rho} \leqslant \kappa(Q)\varrho. \tag{27}$$

To apply proposition 6, we search for a condition on α such that

$$\mathrm{Re}\big(\tilde{A}_{kk} - \alpha\lambda_i\beta_k\big) + \sum_{\substack{1\leqslant j\leqslant m\\ j\neq k}} \left|\tilde{A}_{kj}(t)\right| < 0 \qquad \text{for all } k \in \{1,\ldots,m\}. \tag{28}$$

Since $\mathrm{Re}(\tilde{A}_{kk}) \leqslant |\tilde{A}_{kk}|$, it is therefore sufficient that

$$\alpha > \frac{\sum_{j=1}^{m} |\tilde{A}_{kj}|}{\mathrm{Re}(\lambda_i\beta_k)}$$

holds. Note that $\mathrm{Re}(\lambda_i\beta_k) \geqslant \gamma$, so if we define

$$\sum_{j=1}^{m} |\tilde{A}_{ij}| \leqslant c\tilde{\rho} =: \rho,$$

where $c > 0$ depends on the choice of the norm, then by the diagonal dominance criterion (proposition 6), the evolution operator $\tilde{\Phi}_i$ satisfies

$$\|\tilde{\Phi}_i(t,t_0)\| \leqslant K\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)} \qquad \text{for all } t \geqslant t_0. \tag{29}$$

Finally, using (25) and changing back to the original coordinates, we have

$$\begin{aligned} \|\Phi(t,t_0)\| &= \left\|O\big(\textstyle\bigoplus_{i\geqslant 2}\tilde{\Phi}_i(t,t_0)\big)O^{-1}\right\| \\ &\leqslant \kappa(O)\,\max_{i\geqslant 2}\left\|\tilde{\Phi}_i(t,t_0)\right\| \\ &\leqslant K\kappa(O)\,\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)} \qquad \text{for all } t \geqslant t_0. \end{aligned} \tag{30}$$

Note that O^{-1} maps M and N onto the first and last $n-1$ of the m-tuples in $\mathbb{R}^{nm}$, respectively, so the restriction to N reduces to a direct sum over $i \geqslant 2$ after conjugation with O, while we can simply estimate $\kappa(O|_{O^{-1}N}) \leqslant \kappa(O)$. □

5.2. Non-diagonalizable Laplacian

We now treat the case when the Laplacian is non-diagonalizable and Γ is diagonalizable. Note that if Γ is non-diagonalizable, the results follow from the density of diagonalizable matrices and the persistence property.

Lemma 9 (Non-diagonalizable Laplacian). *Consider the situation of lemma 8 without the condition that the Laplacian is diagonalizable. Then there exists a $\bar{\rho} > 0$ such that for all coupling strengths*

$$\alpha > \frac{\bar{\rho}}{\gamma},$$

the evolution operator Φ of (23) satisfies the estimate

$$\|\Phi(t,t_0)\| \leqslant \bar{C}\mathrm{e}^{-(\alpha\gamma-\bar{\rho})(t-t_0)} \qquad \text{for all } t \geqslant t_0,$$

where $\bar{C} = \bar{C}(\Gamma, L) \geqslant 1$.

The proof of this lemma makes use of persistence of exponential dichotomies and the density of diagonalizable Laplacians. We first establish the following auxiliary result.

Proposition 10. *Let $\varepsilon > 0$ and J be a complex Jordan block of dimension m. Consider*

$$\tilde{J} = J + E,$$

where $E = \mathrm{diag}(0, \varepsilon, 2\varepsilon, \ldots, (m-1)\varepsilon)$. Then there exists an $R \in \mathbb{R}^{m\times m}$ such that $R^{-1}\tilde{J}R$ is diagonal and

$$\|R^{-1}ER\| = b\varepsilon$$

with the constant $b = b(m)$.

Proof. Note that $\tilde{J}$ is diagonalizable, since all the eigenvalues are distinct. All corresponding transformations R are matrices of eigenvectors, upper triangular and can be computed explicitly. We normalize the eigenvectors such that for $\ell, j \in \{1, \ldots, m\}$

$$R_{\ell j} := \begin{cases} \dfrac{(j-1)!}{(j-\ell)!}\varepsilon^{\ell-1} & \text{for all } \ell \leqslant j, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that the elements R^{-1}_{ik} with $i, k \in \{1, \ldots, m\}$ of the inverse of R read as

$$R^{-1}_{ik} = \begin{cases} \dfrac{(-1)^{i+k}}{(i-1)!(k-i)!}\varepsilon^{-(k-1)} & \text{for all } i \leqslant k, \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$(R^{-1}ER)_{ij} = \sum_{k,\ell} R^{-1}_{ik}E_{k\ell}R_{\ell j} = \varepsilon\frac{(-1)^i(j-1)!}{(i-1)!}\sum_{k=i}^{j}\frac{(-1)^k(k-1)}{(j-k)!(k-i)!}.$$

Note that $(R^{-1}ER)_{ii} = (i-1)\varepsilon$. Likewise, we have $(R^{-1}ER)_{i,i+1} = -i\varepsilon$. Moreover, if $j > i+1$ then $(R^{-1}ER)_{ij} = 0$, since

$$\sum_{k=i}^{j}\frac{(-1)^k(k-1)}{(j-k)!(k-i)!} = (-1)^i\sum_{l=1}^{j-i}\frac{(-1)^l}{(l-1)!(j-i-l)!} = 0.$$

Therefore, $\max_{1\leqslant i\neq m}\sum_{j=1}^{m}|(R^{-1}ER)_{ij}| = \max\{(2m-3), m-1\}\varepsilon$, and the result follows. □

Now we are ready to prove our approximation result.

Proposition 11. *Let L be a Laplacian with simple eigenvalue zero and $\mathbb{1}$ its associated eigenvector. Then for any $\varepsilon > 0$, there exists a matrix $\tilde{L}$ with simple eigenvalue zero and $\mathbb{1}$ its associated eigenvector such that*

(i) $\tilde{L} = P\tilde{\Lambda}P^{-1}$ *with a diagonal matrix* $\tilde{\Lambda} \in \mathbb{R}^{n\times n}$ *and*
(ii) $\left\|P^{-1}(\tilde{L} - L)P\right\| \leqslant \varepsilon$.

Proof. We only need to prove the statement if L is non-diagonalizable. We decompose L in its complex Jordan canonical form

$$L = OJO^{-1},$$

where J is a block diagonal matrix. The first block corresponds to the simple eigenvalue zero, so the first row contains only zeros, that is, $J = \mathrm{diag}(0, J_1, \ldots, J_k)$, where J_i are Jordan blocks corresponding to non-zero eigenvalues. Without loss of generality, we consider $k = 1$.

Define $v := O^{-1}\mathbb{1}$. By hypothesis, we have $L\mathbb{1} = 0$, so

$$Jv = 0. \tag{31}$$

As each Jordan block has its own invariant subspace, (31) implies $v = (1, 0, \ldots, 0)$. Define $E := \mathrm{diag}(0, \varepsilon, 2\varepsilon, \ldots, (n-1)\varepsilon)$ and note that

$$Ev = 0. \tag{32}$$

Consider the matrix

$$\tilde{L} = O(J + E)O^{-1},$$

which is diagonalizable. Moreover, by (31) and (32), we obtain that $\tilde{L}$ has zero as a simple eigenvalue with associated eigenvector $\mathbb{1}$. By proposition 10, we obtain

$$J + E = R\tilde{\Lambda}R^{-1},$$

and hence the matrix $P = OR$ diagonalizes $\tilde{L}$. For this reason,

$$P^{-1}(\tilde{L} - L)P = P^{-1}(OEO^{-1})P = R^{-1}ER,$$

and the result follows by proposition 10. □

Proof of lemma 9. As in the diagonalizable case, we consider the linearized equation (23) for $\xi \in N$ along any curve $s(t) \in U$. By proposition 11, there is a diagonalizable matrix $\tilde{L}$ in an arbitrary neighbourhood of the Laplacian L. We rewrite (23) as

$$\dot{\xi} = \left[I_n \otimes D_2 f(t, s(t)) - \alpha\tilde{L} \otimes \Gamma\right]\xi + \alpha\left[(\tilde{L} - L) \otimes \Gamma\right]\xi. \tag{33}$$

Note that this is a small perturbation of the same equation with diagonalizable Laplacian $\tilde{L}$, so we can apply the results from section 5.1. Recall that $\Gamma = QBQ^{-1}$ and $\tilde{L} = P\tilde{\Lambda}P^{-1}$ (see proposition 11). Moreover, consider the change of variables $\zeta = (P^{-1} \otimes Q^{-1})\xi$. We obtain

$$\dot{\zeta} = \left[I_n \otimes Q^{-1}D_2 f(t, s(t))Q - \alpha\tilde{\Lambda} \otimes B\right]\zeta + \alpha\left[P^{-1}(\tilde{L} - L)P \otimes B\right]\zeta. \tag{34}$$

We treat $\alpha\left(P^{-1}(\tilde{L} - L)P \otimes B\right)\zeta$ as a perturbation of the equation

$$\dot{\zeta} = \left(I_n \otimes Q^{-1}D_2 f(t, s(t))Q - \alpha(\tilde{\Lambda} \otimes B)\right)\zeta. \tag{35}$$

It follows from the proof of lemma 8 (see (29) for details) that the evolution operator $\tilde{\Phi}$ of (35) satisfies

$$\|\tilde{\Phi}(t, t_0)\| \leqslant K\mathrm{e}^{-(\alpha\gamma - \rho)(t - t_0)},$$

where K does not depend on n as (35) is block diagonal. Theorem 5 (the persistence theorem) implies that the condition

$$\alpha\|P^{-1}(\tilde{L} - L)P \otimes B\| < \frac{\alpha\gamma - \rho}{K} \tag{36}$$

leads to an exponential stability estimate for the perturbed equation (33). By proposition 11 (ii), we can choose $\tilde{L}$ such that $\|P^{-1}(\tilde{L} - L)P\| \leqslant \varepsilon/\|B\|$, so (36) is satisfied if taking $\varepsilon < (\alpha\gamma - \rho)/(\alpha K)$. Hence, setting $\bar{\rho} := \rho + \alpha K\varepsilon$, then for all $\alpha > \bar{\rho}/\gamma$ the linear flow $\Phi(t, t_0)$ for (33) satisfies

$$\|\Phi(t, t_0)\| \leqslant K\kappa(P \otimes Q)\mathrm{e}^{-(\alpha\gamma - \bar{\rho})(t - t_0)} \qquad \text{for all } t \geqslant t_0,$$

where the condition number is due to transforming back to the original variables ξ. □

To analyse the solution curves $(s(t), \xi(t))$ of the nonlinear system (18), (19) we introduce the concept of a tubular neighbourhood.

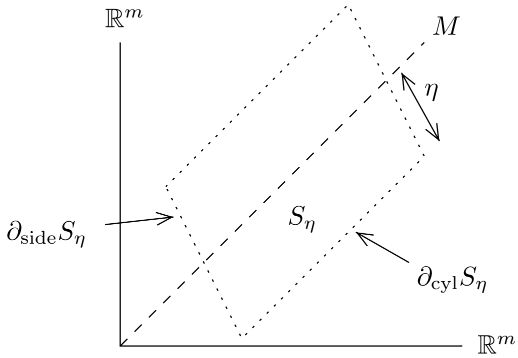


Figure 4. Tubular neighbourhood S_η for $n = 2$.

Definition 12 (η-tubular neighbourhood). *Let $S = \mathbb{1} \otimes U \subset M$ be a subset of the diagonal manifold. Then the set*

$$S_\eta = \left\{ \mathbb{1} \otimes s + \xi : s \in U \text{ and } \xi \in N, \text{ where } \|\xi\| < \eta \right\} \tag{37}$$

for a given $\eta > 0$ is called the η-tubular neighbourhood of S.

See figure 4 for a schematic illustration of this definition. Note that the directions along N in which the tubular neighbourhood stretches out do not need to be orthogonal to M.

Assumption A1$'$ says that the single-node system has a uniformly inflowing invariant set $U \subset \mathbb{R}^m$. A similar result holds in a neighbourhood of the synchronization manifold S in the coupled network, since the following lemma implies that if the solution curve $(s(t), \xi(t))$ leaves S_η, then it must do so by $\|\xi(t)\|$ growing larger than η.

Lemma 13. *Consider assumption A1$'$ with the ε-inflowing invariant set $U \subset \mathbb{R}^m$. Let $\dot{X} = F(t, X)$ describe the dynamics of n uncoupled copies of this system and let $G : \mathbb{R} \times \mathbb{R}^{nm} \to \mathbb{R}^{nm}$ be a perturbation such that for some $r > 0$ and $\delta > 0$, one has*

$$\sup_{t \in \mathbb{R}, X \in S_r} \|G(t, X)\| \leqslant \delta < \frac{\varepsilon}{\|\pi_M\|}.$$

Then there exists an $\eta \in (0, r]$ such that solution curves $(s(t), \xi(t))$ of $\dot{X} = F(t, X) + G(t, X)$ can only leave the tubular neighbourhood S_η through

$$\partial_{\text{cyl}} S_\eta := \{\mathbb{1} \otimes s + \xi : s \in U \text{ and } \|\xi\| = \eta\}.$$

Proof. Choose η such that $0 < \eta \leqslant r$. The boundary of S_η consists of two parts:

$$\partial S_\eta = \partial_{\text{cyl}} S_\eta \cup \partial_{\text{side}} S_\eta,$$

where $\partial_{\text{side}} S_\eta := \{\mathbb{1} \otimes s + \xi : \|\xi\| \leqslant \eta \text{ and } s \in \partial U\}$.

We consider the dynamics on $\partial_{\text{side}} S_\eta$. Let q be the inward-pointing normal vector at $s \in \partial U$. Locally we have $\partial_{\text{side}} S_\eta = \partial S \oplus N$, so $F + G$ points inwards at $\mathbb{1} \otimes q + \xi$ precisely if its projection onto M along N has positive inner product with q. Note that we use the isometry ι_M from (17) to endow M with the inner product $\langle \cdot, \cdot \rangle_M$ induced from $\langle \cdot, \cdot \rangle_{\mathbb{R}^m}$, but no inner product on $\mathbb{R}^{nm}$ is used (nor defined).

If η is chosen sufficiently small, then S_η is contained within the product space $B_\Delta(U)^n$ where we have uniform bounds $\|D_2F\| \leqslant \varrho$ and $\|G\| \leqslant \delta$. It follows that

$$\begin{aligned}&\langle \iota_M(q), \pi_M[F(t, \mathbb{1}\otimes s+\xi)+G(t, \mathbb{1}\otimes s+\xi)]\rangle_M\\ &\qquad = \langle q, f(t,s)\rangle_{\mathbb{R}^m} + \langle q, \iota_M^{-1}\circ\pi_M[D_2F(t, \mathbb{1}\otimes s+\tau\,\xi)\xi + G(t, \mathbb{1}\otimes s+\xi)]\rangle_{\mathbb{R}^m}\\ &\qquad \geqslant \varepsilon - \|\pi_M\|(\|D_2F\|\|\xi\| + \|G\|)\\ &\qquad \geqslant \varepsilon - \|\pi_M\|(\varrho\eta+\delta),\end{aligned}$$

where we applied the mean value theorem with $\tau \in (0,1)$ as interpolation variable. Since $\|\pi_M\|\delta < \varepsilon$, there exists an $\eta > 0$ such that $F+G$ points inwards everywhere at $\partial_{\text{side}} S_\eta$. □

Finally, we shall make use of the following lemma, which is a variant on Gronwall's lemma.

Lemma 14. *Let $x(t) \in \mathbb{R}$ satisfy the integral inequality*

$$x(t) \leqslant C\mathrm{e}^{-\mu(t-t_0)}x_0 + \int_{t_0}^{t} C\mathrm{e}^{-\mu(t-\tau)}\big(\alpha x(\tau)+\beta\big)\,\mathrm{d}\tau, \tag{38}$$

with $C, \mu > 0$ and $x_0, \alpha, \beta \geqslant 0$, whenever $x \leqslant \delta$.

If $\tilde{\mu} := \mu - C\alpha > 0$ and $x_0 < \frac{1}{C}\big(\delta - \frac{\beta}{\tilde{\mu}}\big)$, then $x(t)$ is bounded by

$$x(t) \leqslant C\mathrm{e}^{-\tilde{\mu}(t-t_0)}\Big(x_0 - \frac{\beta}{\tilde{\mu}}\Big) + \frac{C\beta}{\tilde{\mu}} \qquad \text{for all } t \geqslant t_0, \tag{39}$$

and in particular $x(t) < \delta$ holds for all $t \geqslant t_0$.

Proof. The integral inequality is equivalent to the differential inequality

$$\dot{x}(t) \leqslant -\mu x(t) + C\big(\alpha x(t)+\beta\big), \qquad x(t_0) = Cx_0,$$

so by a standard application of Gronwall's lemma we obtain (39), as long as the solution satisfies $x(t) \leqslant \delta$. Now assume by contradiction that this assumption is violated. Then there exists a $t_1 \geqslant t_0$ such that $x(t) = \delta$ for the first time at $t = t_1$. However, the assumption $x(t) \leqslant \delta$ is true up to time t_1, so by the previous estimates and the assumption that $x_0 < \frac{1}{C}\big(\delta - \frac{\beta}{\tilde{\mu}}\big)$ it follows that $x(t_1) < \delta$. This contradiction completes the proof. □

6. Synchronization

In the previous section we have established all auxiliary results to prove our main theorem on synchronization (theorem 1), which will be restated for convenience.

Theorem (Synchronization). *Consider the network of diffusively coupled equations (1) satisfying A1–A3. Then there exists a $\rho = \rho(f, \Gamma)$ such that for all coupling strengths*

$$\alpha > \frac{\rho}{\gamma},$$

the network is locally uniformly synchronized. This means that there exist a $\delta > 0$ and a $C = C(L, \Gamma) > 0$ such that if $x_i(t_0) \in U$ and $\|x_i(t_0) - x_j(t_0)\| \leqslant \delta$ for any $i, j \in \{1, \ldots, n\}$, then

$$\|x_i(t) - x_j(t)\| \leqslant C\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)}\|x_i(t_0) - x_j(t_0)\| \qquad \text{for all } t \geqslant t_0.$$

Proof. Set

$$X(t_0) = \mathbb{1}\otimes s(t_0) + \xi(t_0) := \mathrm{col}\big(x_1(t_0), \cdots, x_n(t_0)\big)$$

where $x_i(t_0) \in U$ and $U \subset \mathbb{R}^m$ is ε-inflowing invariant. Due to the uniformity assumptions in A1′, there exists a slightly enlarged neighbourhood $B_{\Delta/2}(U) \supset U$ that is still $\varepsilon/2$-inflowing invariant. We set $S = \mathbb{1} \otimes B_{\Delta/2}(U)$. If we choose the distance bound $\|x_i(t_0) - x_j(t_0)\| \leqslant \delta$ sufficiently small (depending on the angle between M and N), then $s(t_0) \in S$ holds, while we also have $\|\xi(t_0)\| \leqslant \|\pi_N\|\delta$.

By lemma 13 there exists a tubular neighbourhood S_η of positive size $\eta > 0$ over S that is inflowing invariant on the 'side' and contained within $B_\Delta(U)^n \subset \mathbb{R}^{nm}$, so the uniform assumptions of A1′ hold.

Now lemmas 8 and 9 together imply that there exists a $\rho > 0$ such that for $\alpha > \frac{\rho}{\gamma}$, the evolution operator $\Phi(t, t_0)$ for ξ satisfies an exponential estimate with decay rate $-(\alpha\gamma - \rho)$. The nonlinear remainder of the flow of ξ can be bounded by an arbitrarily small linear term when $\|\xi\|$ is small, as controlled by η. By variation of constants, equation (19) for ξ is equivalent to

$$\xi(t) = \Phi(t, t_0)\xi(t_0) + \int_{t_0}^{t} \Phi(t, \tau) R_\xi(s(\tau), \xi(\tau))\, \mathrm{d}\tau. \tag{40}$$

Now we assume that $\|\xi(t)\| \leqslant \eta$ for all $t \geqslant t_0$ and estimate

$$\|\xi(t)\| \leqslant C\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)}\|\pi_N\|\delta + \int_{t_0}^{t} C\mathrm{e}^{-(\alpha\gamma-\rho)(t-\tau)}\varepsilon(\eta)\, \mathrm{d}\tau.$$

Hence, when we choose $\delta < \frac{\eta}{C\|\pi_N\|}$ and $\varepsilon(\eta)$ sufficiently small, then we can apply lemma 14 with $\beta = 0$ and conclude that

$$\|\xi(t)\| \leqslant C\mathrm{e}^{-\tilde{\mu}(t-t_0)}\|\pi_N\|\delta \qquad \text{for all } t \geqslant t_0,$$

with $\tilde{\mu} = \alpha\gamma - \rho - C\varepsilon(\eta)$. Thus, if we choose $\tilde{\rho} = \rho + C\varepsilon(\eta)$, then for all $\alpha > \frac{\tilde{\rho}}{\gamma}$ the complete solution curve $(s(t), \xi(t))$ for the nonlinear system is contained in S_η for all $t \geqslant t_0$ and converges to the synchronization manifold S with decay rate $-(\alpha\gamma - \tilde{\rho})$. The explicit estimate for $\|x_i(t) - x_j(t)\|$ can be recovered from

$$\left\|x_i(t) - x_j(t)\right\| \leqslant 2\|x_i(t) - s(t)\| \leqslant 2\|\xi(t)\|$$

and the fact that δ can be chosen smaller to satisfy $\|x_i(t) - x_j(t)\| < \eta$ for all $t \geqslant t_0$. □

Remark 15. Explicit estimates for the size of δ in theorem 1 can be found when more details of the system are known. For example, if the second derivative of f is bounded, i.e.

$$\left\|D_2^2 f(t, x)\right\| \leqslant \sigma \qquad \text{for all } t \in \mathbb{R} \text{ and } x \in U,$$

and the coupling function is linear, i.e. $h(x) = \Gamma x$, then δ can be estimated as

$$\delta = \frac{\alpha\gamma - \rho}{4\sigma C\|\pi_N\|}. \tag{41}$$

Note that for convenience, we ignore effects on the size of δ introduced by estimates at the boundary of the synchronization manifold. Under these assumptions the remainder R_ξ in (40) consists of R_F, the nonlinearities of f, and can be estimated as $\|R_F(t, s, \xi)\| \leqslant \sigma\|\xi\|^2$ using mean value theorem arguments. To conclude the argument, fix $\delta = \eta/(2C\|\pi_N\|)$ and follow the proof of theorem 1.

6.1. Behaviour of ρ as a function of Γ

Our approach is constructive and allows one to estimate the bounds for $\rho = \rho(f, \Gamma)$ whenever specific information on the function h is provided. By lemma 9, it is clear that the diagonalization properties of the Laplacian have no effect on the bounds for ρ. In the following, we only discuss symmetric Laplacians L. As an illustration, we look at two cases for Γ.

(i) Γ *is symmetric*. There exists an orthogonal matrix Q such that $\Gamma = QBQ^{-1}$. Note that $\kappa(Q) = 1$ (i.e. the condition number with respect to the Euclidean norm). From (27), it follows that

$$\rho \leqslant \tilde{c}_Q$$

for some $\tilde{c} > 0$. The bound for ρ is independent of Γ for this reason. Note that this can be observed in the left panel of figure 3.

(ii) Γ *is non-diagonalizable.* To treat the non-diagonalizable case, we employ the above perturbation techniques we developed for the Laplacian, i.e. we approximate Γ by a diagonalizable matrix $\tilde{\Gamma}$. Note that Γ can be represented in its Jordan form $\Gamma = QJQ^{-1}$, and without loss of generality we may assume that J is a single Jordan block. We can write $\tilde{J} = J + E$, where E is an ε-perturbation diagonal matrix as in proposition 10. The approximation $\tilde{\Gamma}$ reads as $\tilde{\Gamma} = Q(J + E)Q^{-1}$, and as in proposition 10, if P denotes the matrix that diagonalizes $J + E$ (i.e. $\tilde{B} = P^{-1}(J + E)P$ is diagonal), then $\tilde{\Gamma} = QP\tilde{B}P^{-1}Q^{-1}$. Hence,

$$\rho \leqslant c_Q\kappa(QP) \leqslant c_Q\kappa(Q)\kappa(P).$$

By proposition 10, it is easy to check that

$$\kappa(P) = \|P\|\|P^{-1}\| \leqslant \frac{d}{\varepsilon^{m-1}},$$

where $d > 0$ does not depend on ε. The aim is to minimize ρ, which means minimizing $\kappa(P)$. The perturbation size ε should be of the same order as β, since the real parts of the eigenvalues of $J + E$ must be positive. This can be obtained, for instance, by choosing $\varepsilon = r\beta$ for some fixed $r \in (0, 1)$. This yields the following bound

$$\rho \leqslant \frac{k}{\beta^{m-1}},$$

where k is a constant.

Note the different behaviour for the bound as a function of β between the case when Γ is symmetric and when Γ is non-diagonalizable. This helps to explain the nonlinear behaviour observed in figure 1 and in the right panel of figure 3.

7. Persistence

As in the previous section, we make use of the auxiliary results from section 5 in order to prove our main theorem on persistence (theorem 2), which will be restated for convenience.

Theorem (Persistence). *Consider the perturbed network (4) of diffusively coupled equations fulfilling assumptions A1–A3, and suppose that*

$$\alpha > \frac{\rho}{\gamma}$$

as in theorem 1. Then there exist $\delta > 0$, $C > 0$ *and* $\varepsilon_g > 0$ *such that for all* ε_0*-perturbations satisfying*

$$\|g_i(t,x)\| \leqslant \varepsilon_0 \leqslant \varepsilon_g \qquad \text{for all } t \in \mathbb{R},\ x \in U \text{ and } i \in \{1,\ldots,n\}$$

and initial conditions satisfying $\|x_i(t_0) - x_j(t_0)\| \leqslant \delta$ *for any* $i, j \in \{1,\ldots,n\}$*, the estimate*

$$\left\|x_i(t) - x_j(t)\right\| \leqslant C\mathrm{e}^{-(\alpha\gamma-\rho)(t-t_0)}\left\|x_i(t_0) - x_j(t_0)\right\| + \frac{C\varepsilon_0}{\alpha\gamma-\rho} \qquad \text{for all } t \geqslant t_0$$

holds.

Note that the proof of this theorem does not specifically depend on the fact that the perturbations g_i of the nodes are decoupled; the function G below can depend arbitrarily on the total state X (or can be subject to random perturbations).

Proof of theorem 2. Denote by

$$G(t,X) = \mathrm{col}(g_1(t,x_1),\ldots,g_n(t,x_n))$$

the perturbation for the network and note that $\|G\| \leqslant \varepsilon_0$. As in the proof of theorem 1, lemma 13 guarantees that there exists an η-tubular neighbourhood S_η such that solutions of the complete system for (s,ξ) cannot escape along s, when ε_g, η are sufficiently small.

The perturbed network equation for $X = (s,\xi)$ in S_η now reads as

$$\dot{X} = F(t,X) - \alpha L \otimes \Gamma\xi + R_h(\xi) + G(t,X),$$

where R_h is the Taylor remainder associated with the coupling function h. Projecting this equation onto the synchronization manifold yields an equation for the component s of X. On the other hand, the differential equation for ξ is given by

$$\dot{\xi} = T(t,s(t))\xi + R(t,s(t),\xi) + \pi_N(G(t, \mathbb{1}\otimes s + \xi)), \tag{42}$$

see proposition 7. Let $\varepsilon(\eta)$ denote a Lipschitz constant within S_η of R with respect to ξ, which does not depend on t.

In the same way as in the proof of theorem 2, we obtain a variation of constants formula for solutions of (42),

$$\xi(t) = \Phi(t,t_0)\xi(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\big[R(\tau,s(\tau),\xi(\tau)) + \pi_N(G(\tau, \mathbb{1}\otimes s(\tau) + \xi(\tau)))\big]\,\mathrm{d}\tau.$$

With initial conditions $\|\xi(t_0)\| \leqslant \|\pi_N\|\delta$, lemmas 8 and 9, and the assumption that

$$\|\xi(t)\| \leqslant \delta_1 < \eta \qquad \text{for all } t \geqslant t_0,$$

this leads to the estimate

$$\|\xi(t)\| \leqslant C\mathrm{e}^{-\mu t}\|\pi_N\|\delta + \int_{t_0}^{t} C\mathrm{e}^{-\mu(t-\tau)}(\varepsilon(\delta_1)\|\xi(\tau)\| + \|\pi_N\|\varepsilon_0)\,\mathrm{d}\tau\,,$$

where $\mu = \alpha\gamma - \rho$. We choose $\delta < \frac{\eta}{C\|\pi_N\|}$ and δ_1, ε_g sufficiently small and apply lemma 14 with $\alpha = \varepsilon(\delta_1)$, $\beta = \|\pi_N\|\varepsilon_0$ to find that

$$\|\xi(t)\| \leqslant C\mathrm{e}^{\tilde{\mu}(t-t_0)}\|\pi_N\|\left(\delta - \frac{\varepsilon_0}{C\tilde{\mu}}\right) + \frac{C\|\pi_N\|\varepsilon_0}{\tilde{\mu}} \qquad \text{for all } t \geqslant t_0, \tag{43}$$

where $\tilde{\mu} = \alpha\gamma - \rho - C\varepsilon(\delta_1)$. As in the proof of theorem 1, we choose $\tilde{\rho} = \rho + C\varepsilon(\delta_1)$ instead of ρ and the estimate for $\|x_i(t) - x_j(t)\|$ follows from (43) by adapting δ. □

In particular, note that asymptotically, the bound in (43) converges to $\frac{C\|\pi_N\|\varepsilon_0}{\alpha\gamma-\tilde{\rho}}$. Furthermore, it follows from the details of lemma 8 that the constant C depends on the Laplacian L only through its condition number $\kappa(P)$.

Finally, we can prove corollary 3 from the introduction.

Proof of corollary 3. This corollary is a direct consequence of our persistence result. For simplicity, we now endow the space $\mathbb{R}^{nm}$ with the Euclidean norm

$$\|X\|_2 = \left(\sum_{i=1}^{n} \|x_i\|_2^2\right)^{1/2} \qquad \text{for all } X = \mathrm{col}(x_1, \ldots, x_n) \in \mathbb{R}^{nm}.$$

Note that in view of (43), for large times, we obtain

$$\|\xi\|_2 = \left(\sum_{i=1}^{n} \|s - x_i\|_2^2\right)^{1/2} \leqslant \frac{2K\kappa_2(P \otimes Q)\|G\|_2}{\mu} \tag{44}$$

where the contraction rate μ is given by $\mu = \alpha\gamma - \rho$. For simplicity, we omit the arguments of the functions s, x, G and ξ.

Moreover, $\kappa_2(P \otimes Q) \leqslant \kappa_2(P)\kappa_2(Q)$, and since the Laplacian is symmetric, it can be diagonalized by an orthogonal similarity transformation, which implies that $\kappa_2(P) = 1$ together with $\|\pi_N\|_2 = 1$. Moreover, by the equivalence of norms we obtain

$$\|G\|_2 \leqslant \sqrt{n}\|G\| \leqslant \sqrt{n}\varepsilon_0,$$

Replacing this estimate in (44) we obtain

$$\left(\sum_{i=1}^{n} \|s - x_i\|_2^2\right)^{1/2} \leqslant \frac{\tilde{K}\sqrt{n}\varepsilon_0}{\mu}, \tag{45}$$

where $\tilde{K} = 2K\kappa_2(Q)$. We scale equation (45) to obtain

$$\left(\frac{1}{n}\sum_{i=1}^{n} \|s - x_i\|_2^2\right)^{1/2} \leqslant \frac{\tilde{K}\varepsilon_0}{\mu}, \tag{46}$$

and applying the sum of squares inequality

$$\frac{1}{n}\sum_{i=1} a_i \leqslant \sqrt{\frac{1}{n}\sum_{i=1}^{n} a_i^2}$$

leads to

$$\frac{1}{n}\sum_{i=1}^{n} \|s - x_i\|_2 \leqslant \frac{\tilde{K}\varepsilon_0}{\mu}. \tag{47}$$

Together with the triangle inequality we obtain

$$\begin{aligned}
\frac{1}{n(n-1)}\sum_{i,j=1}^{n} \left\|x_j - x_i\right\|_2 &\leqslant \frac{1}{n(n-1)}\sum_{i,j=1, i\neq j}^{n} \left(\left\|s - x_j\right\|_2 + \|s - x_i\|_2\right) \\
&\leqslant \frac{1}{n(n-1)}\left(\sum_{j=1}^{n}\sum_{i\neq j} \left\|s - x_j\right\|_2 + \sum_{i=1}^{n}\sum_{j\neq i} \|s - x_i\|_2\right) \\
&= \frac{2}{n}\sum_{i=1}^{n} \|s - x_i\|_2 \leqslant \frac{2\tilde{K}\varepsilon_0}{\mu}.
\end{aligned}$$

This finishes the proof of this corollary. □

8. Generalizations

Although our set-up is very general and includes nonautonomous systems and non-diagonalizable Laplacians, the assumptions we make are only sufficient for synchronization, but not necessary. For instance, let $(u, v) = x \in \mathbb{R}^2$ and consider as isolated dynamics $\dot{x} = f(x)$ with $f(x) = (u, u - v)$, and

$$\begin{aligned} \dot{x}_1 &= f(x_1) + \alpha\Gamma(x_2 - x_1) \\ \dot{x}_2 &= f(x_2) + \alpha\Gamma(x_1 - x_2) \end{aligned} \qquad \text{with} \qquad \Gamma = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Note that in this situation Γ has an eigenvalue zero, so assumption A3 is violated. However, this coupled system synchronizes for $\alpha > 1/2$. This happens as all instabilities occur due to the first variable, and the coupling Γ acts solely on this variable. For a numerical example of a chaotic system displaying synchronization with only one variable coupled, see [24].

The boundedness of the Jacobian $D_2 f$ in assumption A1$'$ and assumption A3 are used in lemma 8 to obtain uniform exponential stability of the linear system (23). For this purpose, we use the diagonal dominance criterion, see (28) in the proof of lemma 8. It is clear that one could get uniform exponential stability without the two above-mentioned assumptions. Note that under reasonable assumptions, a necessary and sufficient condition for uniform exponential stability (and thus persistent synchronization) is that the dichotomy spectrum of (23) is contained in the negative half line [17] (see [11] for a comparative study of numerical methods to approximate the dichotomy spectrum).

For persistent synchronization, we thus only require a dichotomy spectrum in the directions transverse to the synchronization manifold. Instead we can impose the stricter condition of normal hyperbolicity (see [10, 14] and e.g. [16] in the context of synchronization of networks). That is, we also require that any exponential contraction tangent to the synchronization manifold is weaker than in the transverse directions. In other words, the spectra in the normal and tangential directions must be disjoint and the normal spectrum must be strictly below the tangential one. In our explicit set-up, this so-called spectral gap condition translates to

$$\rho - \alpha\gamma < -r\rho \qquad \text{with } r \geqslant 1.$$

Under these assumptions we find a stronger form of persistence. Under arbitrary C^1-small perturbations, solutions not only converge into a neighbourhood of the synchronization manifold but an invariant manifold[4]

$$\tilde{S} = \{x_i = h_i(s), s \in U \subset \mathbb{R}^m, 1 \leqslant i \leqslant n\}$$

close to S persists to which these solutions converge. Moreover, a stronger 'shadowing' or 'isochrony' property holds that any solution curve $X(t)$ that converges to $\tilde{S}$, actually converges at exponential rate $\tilde{\mu}$ to a unique solution curve $X_{\tilde{S}}(t)$ on $\tilde{S}$ in the sense that there exists a C such that for all $t \geqslant 0$

$$\left\| X(t) - X_{\tilde{S}}(t) \right\| \leqslant C\mathrm{e}^{-\mu t},$$

with μ close to $\alpha\gamma - \rho$.

Acknowledgments

TP was supported by a Marie Curie IIF Fellowship (Project 303180), JE was supported by the ERC Advanced Grant 267382, and MR and JE were supported by an EPSRC Career

[4] Both smoothness and uniqueness of this manifold are subtle issues. In general the invariant manifold cannot be expected to be smoother than C^r. If the synchronization manifold has a boundary (where it is only forward invariant), then non-uniqueness follows from local modifications that have to be made to apply the persistence theorem, see [16]. Note that both results hold also when the synchronization manifold S is noncompact, see [9, theorem 3.1 and chapter 4].

Acceleration Fellowship (2010–2015). The authors also thank CNPq and the Marie Curie IRSES staff exchange project DynEurBraz. They thank the referees and editors for valuable comments.

References

[1] Acebron J A, Bonilla L L, Vicente C J P, Ritort F and Spigler R 2005 The Kuramoto model: a simple paradigm for synchronization phenomena *Rev. Mod. Phys.* **77** 137–85

[2] Albert R and Barabasi A L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97

[3] Ashwin P, Buescu J and Stewart I 1994 Bubbling of attractors and synchronisation of chaotic oscillators *Phys. Lett.* A **193** 126–39

[4] Belykh V N, Belykh I V and Hasler M 2004 Connection graph stability method for synchronized coupled chaotic systems *Physica* D **195** 159–87

[5] Blasius B and Tönjes R 2005 Quasiregular concentric waves in heterogeneous lattices of coupled oscillators *Phys. Rev. Lett.* **95** 084101

[6] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nature Rev. Neurosci.* **10** 186

[7] Coppel W A 1978 *Dichotomies in Stability Theory* (*Springer Lecture Notes in Mathematics* vol 629) (Berlin: Springer)

[8] Earn D J D, Levin S A and Rohani P 2000 Coherence and conservation *Science* **290** 1360–4

[9] Eldering J 2013 *Normally Hyperbolic Invariant Manifolds—the Noncompact Case* (*Atlantis Series in Dynamical Systems* vol 2) (Berlin: Springer)

[10] Fenichel N 1971/1972 Persistence and smoothness of invariant manifolds for flows *Indiana Univ. Math. J.* **21** 193–226

[11] Froyland G, Hüls T, Morriss G P and Watson T M 2013 Computing covariant Lyapunov vectors, Oseledets vectors, and dichotomy projectors: a comparative numerical study *Physica* D **247** 18–39

[12] Gregoriou G G, Gotts S J, Zhou H and Desimone R 2009 High-frequency, long-range coupling between prefrontal and visual cortex during attention *Science* **324** 1207–10

[13] Heagy J F, Carroll T L and Pecora L M 1994 Experimental and numerical evidence for riddled basins in coupled chaotic systems *Phys. Rev. Lett.* **73** 3528–31

[14] Hirsch M W, Pugh C C and Shub M 1977 *Invariant Manifolds* (*Springer Lecture Notes in Mathematics* vol 583) (Berlin: Springer)

[15] Huang L, Chen Q, Lai Y-C and Pecora L M 2009 Generic behavior of master-stability functions in coupled nonlinear dynamical systems *Phys. Rev.* E **80** 036204

[16] Josić K 2000 Synchronization of chaotic systems and invariant manifolds *Nonlinearity* **13** 1321–36

[17] Kloeden P E and Rasmussen M 2011 *Nonautonomous Dynamical Systems* (*Mathematical Surveys and Monographs* vol 176) (Providence, RI: American Mathematical Society)

[18] Kuramoto Y 1984 *Chemical Oscillations, Waves, and Turbulence* (Berlin: Springer)

[19] Li Z and Chen G R 2004 Design of coupling functions for global synchronization of uncertain chaotic dynamical networks *Phys. Lett.* A **326** 333–9

[20] Milton J and Jung P (ed) 2003 *Epilepsy as a Dynamic Disease* (Berlin: Springer)

[21] Newman M 2010 *Networks: An Introduction* (Oxford: Oxford University Press)

[22] Nishikawa T and Motter A E 2006 Synchronization is optimal in nondiagonalizable networks *Phys. Rev.* E **73** 065106(R)

[23] Orosz G, Moehlis J and Ashwin P 2009 Designing the dynamics of globally coupled oscillators *Prog. Theor. Phys.* **122** 611–30

[24] Pecora L M and Carroll T L 1998 Master stability functions for synchronized coupled systems *Phys. Rev. Lett.* **80** 2109–12

[25] Pereira T 2010 Hub synchronization in scale-free networks *Phys. Rev.* E **82** 036201

[26] Pereira T, Eroglu D, Bagci G B, Tirnakli U and Jensen H J 2013 Connectivity-driven coherence in complex networks *Phys. Rev. Lett.* **110** 234103

[27] Petkoski S and Stefanovska A 2012 Kuramoto model with time-varying parameters *Phys. Rev.* E **86** 046212

[28] Pogromsky A, Glad T and Nijmeijer H 1999 On diffusion driven oscillations in coupled dynamical systems *Int. J. Bifur. Chaos* **9** 629–44

[29] Pogromsky A and Nijmeijer H 2001 Cooperative oscillatory behavior of mutually coupled dynamical systems *IEEE Trans. Circuits Systems I* **48** 152–62

[30] Restrepo J G, Ott E and Hunt B R 2004 Spatial patterns of desynchronization bursts in networks *Phys. Rev.* E **69** 066215

[31] Singer W 1999 Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24** 49–65

[32] Stankovski T, Duggento A, McClintock P V E and Stefanovska A 2012 Inference of time-evolving coupled dynamical systems in the presence of noise *Phys. Rev. Lett.* **109** 024101

[33] Sun J, Bollt E M and Nishikawa T 2009 Master stability functions for coupled nearly identical dynamical systems *Europhys. Lett.* **85** 60011

[34] Viana M 2000 What's new on Lorenz strange attractors? *Math. Int.* **22** 6–19

[35] Wiesenfeld K, Colet P and Strogatz S H 1998 Frequency locking in Josephson arrays: connection with the Kuramoto model *Phys. Rev.* E **57** 1563–9

[36] Wu C W 2007 *Synchronization in Complex Networks of Nonlinear Dynamical Systems* (Singapore: World Scientific)