



Kempton, L., Herrmann, G., & Di Bernardo, M. (2017). Distributed optimisation and control of graph Laplacian eigenvalues for robust consensus via an adaptive multi-layer strategy. *International Journal of Robust and Nonlinear Control*. DOI: 10.1002/rnc.3808

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1002/rnc.3808](https://doi.org/10.1002/rnc.3808)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Wiley at <http://onlinelibrary.wiley.com/doi/10.1002/rnc.3808/abstract>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

Distributed optimisation and control of graph Laplacian eigenvalues for robust consensus via an adaptive multilayer strategy

L. C. Kempton^{1,*},, G. Herrmann² and M. di Bernardo^{1,3}

¹*Department of Engineering Mathematics, University of Bristol, Bristol, UK*

²*Department of Mechanical Engineering, University of Bristol, Bristol, UK*

³*Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy*

SUMMARY

Functions of eigenvalues of the graph Laplacian matrix \mathbf{L} , especially the extremal non-trivial eigenvalues, the algebraic connectivity λ_2 and the spectral radius λ_n , have been shown to be important in determining the performance in a host of consensus and synchronisation applications. In this paper, we focus on formulating an entirely distributed control law for the control of edge weights in an undirected graph to solve a constrained optimisation problem involving these extremal eigenvalues.

As an objective for the distributed control law, edge weights must be found that minimise the spectral radius of the graph Laplacian, thereby maximising the robustness of the network to time delays under a simple linear consensus protocol. To constrain the problem, we use both local weight constraints that weights must be non-negative, and a global connectivity constraint, maintaining a designated minimum algebraic connectivity. This ensures that the network remains sufficiently well connected.

The distributed control law is formulated as a multilayer strategy, using three layers of successive distributed estimation. Adequate timescale separation between the layers is of paramount importance for the proper functioning of the system, and we derive conditions under which the distributed system converges as we would expect for the centralised control or optimisation system to converge. © 2017 The Authors. *International Journal of Robust and Nonlinear Control* published by John Wiley & Sons Ltd.

Received 15 April 2016; Revised 24 February 2017; Accepted 24 February 2017

KEY WORDS: distributed control and optimisation; robust consensus; multilayer networks; singular perturbation theory; graph Laplacian eigenvalues

1. INTRODUCTION

In the control of multi-agent systems, both the structure and strength of connections between individuals are instrumental in determining the performance of the system as a whole. The totality of connections and their strengths uniquely defines a graph Laplacian matrix \mathbf{L} , and the spectrum of \mathbf{L} , $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, concisely reveals a number of useful properties of the network system. Importantly, it has been shown that functions of these eigenvalues often define or can be used to bound bulk properties of the network, including the performance of tasks ranging from consensus [1–4] to synchronisation [5–7]. These tasks are particularly important in the applications of clock synchronisation in wireless networks [8, 9], energy management in smart grids [10, 11] and formation control in multi-agent systems [12, 13].

*Correspondence to: Louis Charles Kempton, Bristol Centre for Complexity Sciences, Bristol University, 1-9 Old Park Hill, Bristol BS2 8BB, UK.

†E-mail: lk8915@bristol.ac.uk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

To illustrate this fundamental concept, we can consider the classic example of a group of agents approaching average consensus in an undirected network under the simple linear weighted consensus protocol [1]:

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} w_{\{i,j\}}(x_j - x_i) \quad (1)$$

Each agent i in the system updates its state x_i according to the weighted sum of the differences between its own state and those of its neighbours x_j , defining the neighbourhood of i , \mathcal{N}_i , as the set of agents with which i can communicate. The strength of the connection between agents i and j is modelled through the weight $w_{\{i,j\}}$. This control law is seen to be distributed as each agent only receives information from its neighbours. The collection of all agents' dynamics can be concisely represented by the weighted graph Laplacian matrix \mathbf{L} , as defined in Section 2:

$$\dot{\mathbf{x}} = -\mathbf{L}(\mathbf{w})\mathbf{x} \quad (2)$$

It is well known that in this system, consensus is achieved if and only if the network is connected, which is equivalent to the second smallest eigenvalue of \mathbf{L} , λ_2 the algebraic connectivity, being strictly greater than zero [14]. Furthermore, the exponential rate at which the consensus value is approached is governed exactly by λ_2 , which represents the convergence rate of the slowest mode. Defining the disagreement vector $\mathbf{y}(t) = \mathbf{x}(t) - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbf{x}(0)$ as the disagreement between the state at time t and the arithmetic mean of $\mathbf{x}(0)$ yields

$$\|\mathbf{y}(t)\|_2 \leq \|\mathbf{y}(0)\|_2 e^{-\lambda_2 t} \quad (3)$$

In this simple system, both the sufficient condition for consensus to be achieved and the measure of the performance are governed by the algebraic connectivity λ_2 of the graph Laplacian. This quantity is a function of all weights in the network and is therefore a global property of the network.

When the simple consensus protocol is affected by a homogeneous time delay of τ ,

$$\dot{\mathbf{x}}(t) = -\mathbf{L}(\mathbf{w})\mathbf{x}(t - \tau) \quad (4)$$

the largest eigenvalue of the graph Laplacian, λ_n , becomes critically important to the stability of the system. In a result from [1], it is shown that for the consensus solution to be stable, an upper bound must hold on the spectral radius λ_n of the graph Laplacian and consequently, that a network with lower spectral radius will be stable under longer time delays. Specifically, it is found that for consensus to be robust to the time delay, the spectral radius of the graph Laplacian must satisfy the following inequality:

$$\lambda_n < \frac{\pi}{2\tau} \quad (5)$$

Because of their importance in network systems, much research has been undertaken on methods to design networks with desirable graph Laplacian eigenvalues. These methods can broadly be categorised by which eigenvalues, or functions of eigenvalues, are controlled and by the technique utilised to modify edge weights. Semi-definite programming (SDP) has proven to be a powerful technique for optimising a number of functions of graph Laplacian eigenvalues, including the algebraic connectivity, and the total effective graph resistance (see [15] and the references therein). Furthermore, when the graph Laplacian is state dependent, SDPs still prove to be an effective tool, and the problem of maximising the algebraic connectivity of a mobile robot network is achieved using a centralised algorithm in [16]. A distributed solution to this problem using repeated solution of local approximate SDPs is presented in [17]. The research area of connectivity maintenance in mobile robot networks [12, 13, 18–22] has developed a number of decentralised methods to control the algebraic connectivity of a state-dependent graph Laplacian. In particular, an adaptive method is presented in [21] for estimating the algebraic connectivity in a completely decentralised manner, which we utilise in this paper. This method has been used for maintenance of connectivity in a mobile network [21], distributed optimisation of the algebraic connectivity [23] and is extended for switching topologies and adapting control gains in [24].

In general, most methods in the literature for controlling the spectral radius λ_n of the graph Laplacian or the eigenratio λ_n/λ_2 follow centralised approaches, with the notable exception of the algorithm presented in [25], which uses decentralised power iteration methods to estimate the eigenvectors associated with λ_2 and λ_n , in order to then adapt edge weights. However, this method breaks down when $\lambda_2 \approx \lambda_3$, or $\lambda_n \approx \lambda_{n-1}$, and so cannot further optimise the network once this point is reached. In [26] and [27], the eigenratio is minimised through simulated annealing. And in [28], mixed integer nonlinear optimisation is utilised to solve eigenvalue optimisation problems including maximising the algebraic connectivity subject to a constraint on the spectral radius. Edge rewiring based on the eigenvectors of the graph Laplacian is used in [29] to minimise the eigenratio λ_n/λ_2 .

In this paper, we take a different approach and formulate a decentralised method for estimating the spectral radius of the graph Laplacian [23], on the basis of the application and extension of the algebraic connectivity estimator presented in [21]. We use these estimators in parallel to evaluate the extremal eigenvalues of the graph Laplacian and inform an adaptive strategy, driving the edge weights of a network to a state that improves the robustness to time delay of the simple consensus protocol. This is achieved by adapting the edge weights so as to minimise the spectral radius λ_n of the graph Laplacian, whilst ensuring that the network remains connected through the enforcement of a lower bound on the algebraic connectivity, λ_2 . We find that this approach is more robust when estimating eigenvalues close to non-uniqueness compared with [25], allowing the graph Laplacian eigenvalues to be optimised past the first point when $\lambda_2 \approx \lambda_3$ or $\lambda_n \approx \lambda_{n-1}$, which can lead to significant improvements in performance.

The distributed optimisation and control strategy that we propose consists of multiple layers of distributed estimation. The edge weight dynamics rely on estimates of the graph Laplacian eigenvalues, and these estimators themselves require further estimated variables to function correctly. An outline of this three-layer structure is illustrated in Figure 1. In Sections 3 to 5 of the paper, we will first present each layer independently, and then in Section 6, we specify how the three layers are interconnected, and derive sufficient conditions for convergence of the distributed strategy, using the framework of multilevel singular perturbation theory [30]. In Section 7, we will illustrate the theoretical results on a set of representative numerical examples, confirming the effectiveness of the approach.

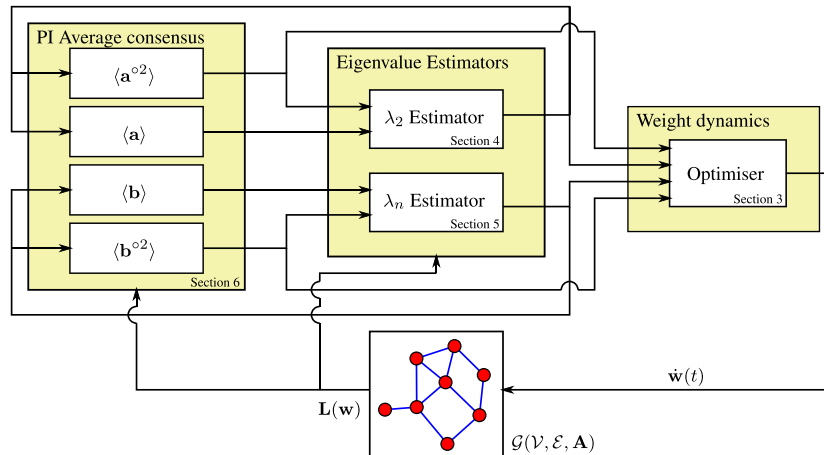


Figure 1. General overview of the distributed estimator-optimiser system. Each subsystem is contained within a white box, and groups of subsystems that operate in the same timescale share a yellow box. In this diagram, the rightmost subsystem is the slow system, then, reading to the left, we have the fast systems and finally, the ultra-fast subsystems. Dependencies between subsystems are illustrated via arrows. The vectors \mathbf{a} and \mathbf{b} are estimates of the eigenvectors associated with λ_2 and λ_n , respectively, and the proportional-integral (PI) average consensus subsystems are used to estimate the mean of the components $\langle \mathbf{a} \rangle = \frac{1}{n} \sum_i a_i$, $\langle \mathbf{b} \rangle = \frac{1}{n} \sum_i b_i$ and the mean of the components squared: $\langle \mathbf{a}^{\circ 2} \rangle = \frac{1}{n} \sum_i a_i^2$, $\langle \mathbf{b}^{\circ 2} \rangle = \frac{1}{n} \sum_i b_i^2$. These estimated means are required by the eigenvalue estimators, as described in Section 4. [Colour figure can be viewed at wileyonlinelibrary.com]

2. PROBLEM FORMULATION

We define the weighted, undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{A})$ on n nodes, such that the vertex set $\mathcal{V} = \{1, \dots, n\}$, and with $m = |\mathcal{E}|$ edges, where the edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ contains no self loops $\{i, i\} \notin \mathcal{E}$. An edge $\{i, j\} \in \mathcal{E}$ if nodes i and j can communicate (pass local variables), which in turn defines the neighbourhood of the i th node, $\mathcal{N}_i : j \in \mathcal{N}_i \iff \{i, j\} \in \mathcal{E}$. To each edge, we assign a non-negative real weight $w_{\{i,j\}} \geq 0$ that defines the weighted adjacency matrix $\mathbf{A} = [a_{i,j}]_{n \times n}$

$$a_{i,j} = \begin{cases} w_{\{i,j\}} & \text{if } \{i, j\} \in \mathcal{E} \\ 0 & \text{if } \{i, j\} \notin \mathcal{E} \\ 0 & \text{if } i = j \end{cases} \quad (6)$$

and the weighted graph Laplacian $\mathbf{L} = [l_{i,j}]_{n \times n}$

$$l_{i,j} = \begin{cases} -w_{\{i,j\}} & \text{if } \{i, j\} \in \mathcal{E} \\ 0 & \text{if } \{i, j\} \notin \mathcal{E} \\ \sum_{k \in \mathcal{N}_i} w_{\{i,k\}} & \text{if } i = j \end{cases} \quad (7)$$

with non-negative diagonal elements $l_{i,i}$ referred to as the weighted degree of node i . The graph Laplacian is symmetric and positive semidefinite, which can clearly be seen using Gersgorin discs [31]: \mathbf{L} is diagonally dominant with non-negative elements on the diagonal, and so we can lower bound the spectrum of \mathbf{L} by 0 and upper bound the spectrum by twice the maximum weighted degree, $\Delta = \max_i \{l_{i,i}\}$ [2]. The row sums of \mathbf{L} all equal 0; therefore, the consensus mode $\mathbf{1}$ is an eigenvector of \mathbf{L} with associated eigenvalue $\lambda_1 = 0$. Henceforth, we can order the eigenvalues of \mathbf{L} , $0 = \lambda_1 \leq \lambda_2 \leq \dots \lambda_n \leq 2\Delta$.

It is well known that the second smallest eigenvalue λ_2 , known as the algebraic connectivity, is non-zero if and only if the network is connected. That is, for each pair of nodes in the network, there is a path between them that traverses only edges with non-zero weight. Moreover, the algebraic connectivity serves as a lower bound on the edge connectivity and vertex connectivity [14, 32].

The graph Laplacian may also be defined another way, by constructing a weight vector \mathbf{w} and oriented incidence matrix \mathbf{P} [33], which will be a useful notation:

Definition 1

Let $q : \{1, \dots, m\} \rightarrow \mathcal{E}$ be a bijective function from the set of integers from 1 to m , to the edge set, so that we may define an ordering on the vector of edge weights $\mathbf{w} = [w_i]_{m \times 1}$, where $w_i = w_{q(i)}$. We may then define the unoriented incidence matrix $\mathbf{M} = [m_{i,j}]_{m \times n}$

$$m_{i,j} = \begin{cases} 1 & \text{if } j \in q(i) \\ 0 & \text{if } j \notin q(i) \end{cases} \quad (8)$$

Similarly, we may define the oriented incidence matrix, $\mathbf{P} = [p_{i,j}]_{m \times n}$

$$p_{i,j} = \begin{cases} -1 & \text{if } j = \min\{q(i)\} \\ +1 & \text{if } j = \max\{q(i)\} \\ 0 & \text{if } j \notin q(i) \end{cases} \quad (9)$$

so that $\mathbf{L}(\mathbf{w}) = \mathbf{P}^\top \text{diag}\{\mathbf{w}\}\mathbf{P}$.

In this paper, we consider, as a specific example, the problem of minimising the largest eigenvalue of the graph Laplacian λ_n so that a consensus process taking place over the network will be more robust to time delays as noted in [1, 34], whilst bounding the algebraic connectivity away from zero, $\lambda_2 \geq \kappa > 0$, so that the network remains sufficiently well connected. Moreover, greater algebraic connectivity is associated with an increased convergence rate in consensus problems with time delay (see, e.g. theorem 3 and figure 5 in [34]).

This optimisation problem can be formulated as

$$\underset{\mathbf{w}}{\text{minimise}} \quad \lambda_n(\mathbf{L}(\mathbf{w})) \quad (10)$$

$$\begin{aligned} \text{subject to} \quad & \lambda_2(\mathbf{L}(\mathbf{w})) - \kappa \geq 0 \\ & w_{\{i,j\}} \geq 0 \quad \forall \{i,j\} \in \mathcal{E} \end{aligned} \quad (11)$$

which is a convex optimisation problem: the objective function, the spectral radius of the Laplacian matrix, is a convex function of edge weights; and the constraint that the algebraic connectivity be greater than a desired value is also convex (the algebraic connectivity is a concave function of edge weights). These facts can be observed by noting that, respectively, the maximum and the negative of the minimum are symmetric convex functions of the spectrum of the graph Laplacian matrix, and hence, the spectral radius and the negative of the algebraic connectivity are convex spectral functions [15, 35].

Moreover, we wish to solve the problem (10, 11) in an entirely distributed manner. By which, we mean that with each edge, $\{i, j\}$ is associated a set of variables $\chi_{\{i,j\}}$, which includes the edge weight $w_{\{i,j\}} \in \chi_{\{i,j\}}$, and to each node, i is associated a set of variables ξ_i . The dynamics of these sets of variables may be dependent only on locally available variables, and all edges and all nodes behave identically so that no edge or node is a leader:

$$\dot{\chi}_{\{i,j\}} = \mathbf{f}(\chi_{\{i,j\}}, \xi_i, \xi_j) \quad \forall \{i, j\} \in \mathcal{E} \quad (12)$$

$$\dot{\xi}_i = \mathbf{g}(\xi_i, \xi_j, \chi_{\{i,j\}} \quad \forall j \in \mathcal{N}_i) \quad \forall i \in \mathcal{V} \quad (13)$$

Furthermore, we restrict the size of these sets reflecting a limit on the memory and communication capacity of individual agents, so that the number of variables at each edge and node remains small even as the network becomes arbitrarily large. This means that nodes cannot simply accumulate information about the entire network and solve the problem with perfect knowledge of the network. Thus, the distributed estimator system with states $\chi_{\{i,j\}}$ and ξ_i uses one-hop neighbour information only.

In this paper, this distributed optimisation problem is solved using a multilayer strategy, shown in Figure 1, with each layer performing successive stages of distributed estimation. Starting with the slowest subsystem, we formulate weight dynamics, which solve the optimisation problem in Section 3. This solution is initially centralised in the respect that it is dependent on global variables, specifically the eigenvalues of the graph Laplacian and their partial derivatives. To render this strategy completely distributed, we implement two eigenvalue estimators that converge at a faster rate: one for the algebraic connectivity [21], described in Section 4.1, and one for the spectral radius, expounded in Section 4.2. Each of these estimators requires access to the mean of the estimator vector and to the estimator vector's magnitude. To estimate these functions in a distributed manner, four proportional–integral (PI) average consensus subsystems [36], as given in Section 5, are utilised. After discussing each of the subsystems individually, we prove in Section 6, using a singular perturbation approach, that given sufficient timescale separation between layers, and under certain assumptions, the entire distributed system is stable and converges onto the solution to the optimisation problem. This result is demonstrated numerically in Section 7, and we also show the performance of the system when the assumptions we make for proving convergence are relaxed.

3. WEIGHT ADAPTATION

To begin, we propose a system governing the weight dynamics, where each node and edge perform the following updates

$$\dot{w}_{\{i,j\}} = -\frac{\partial \lambda_n(\mathbf{w})}{\partial w_{\{i,j\}}} + \frac{1}{2} \left(\mu_{i,j}^2 + \frac{\partial \lambda_2(\mathbf{w})}{\partial w_{\{i,j\}}} v_{\{i,j\}}^2 \right) \quad (14)$$

$$\dot{\mu}_{\{i,j\}} = -w_{\{i,j\}} \mu_{\{i,j\}}, \quad \mu_{\{i,j\}}(0) > 0 \quad (15)$$

$$\dot{v}_{\{i,j\}} = (\kappa - \lambda_2(\mathbf{w})) v_{\{i,j\}}, \quad v_{\{i,j\}}(0) > 0 \tag{16}$$

and show that the equilibrium point $[\mathbf{w}_{opt}^\top, \boldsymbol{\mu}_{opt}^\top, \mathbf{v}_{opt}^\top]^\top$ satisfies the first-order necessary Karush–Kuhn–Tucker (KKT) [37] conditions for the optimisation problem. We will show that the first-order necessary KKT conditions are not only necessary but also sufficient for the equilibrium point $[\mathbf{w}_{opt}^\top, \boldsymbol{\mu}_{opt}^\top, \mathbf{v}_{opt}^\top]^\top$ to be the optimal solution to our optimisation problem, (10) and (11). Following this, we will discuss conditions for the equilibrium to be exponentially stable. As these conditions are not necessarily always true, we will then suggest a minor modification to the algorithm in (14) to (16). An important condition for the system to work as intended is the following assumption:

Assumption 1

At the optimal edge weights \mathbf{w}_{opt} , which solve the optimisation problem stated in (10) and (11), we assume that the eigenvalue $\lambda_2(\mathbf{w}_{opt}) \neq \lambda_3(\mathbf{w}_{opt})$ and that $\lambda_n(\mathbf{w}_{opt}) \neq \lambda_{n-1}(\mathbf{w}_{opt})$ implying that the spectral radius λ_n is distinct. Furthermore, any feasible algebraic connectivity is bounded away from zero when $\kappa > 0$, and the algebraic connectivity λ_2 is distinct.

We make this assumption to facilitate the analysis of the subsystems and argue that this assumption is required for exponential convergence of edge weights onto the solution in the entire distributed system. In Section 7.1, we discuss the ramifications when this assumption does not hold and demonstrate the performance of the system in this case.

Lemma 1 ([38])

Under Assumption 1, the algebraic connectivity $\lambda_2(\mathbf{w})$, the spectral radius $\lambda_n(\mathbf{w})$ and their associated unit eigenvectors, $\mathbf{v}_2(\mathbf{w})$ and $\mathbf{v}_n(\mathbf{w})$, are analytic functions of the edge weights in the vicinity of the solution, \mathbf{w}_{opt} , and hence, their partial derivatives are well defined in this region.

We may write the collection of (14) to (16) for all edges and nodes in the network in a compact form using Hadamard (elementwise) multiplication $\mathbf{a} = \mathbf{b} \circ \mathbf{c} \Leftrightarrow a_i = b_i c_i \ \forall i$. We also introduce Hadamard exponentiation so that $\mathbf{a}^{\circ 2} = \mathbf{a} \circ \mathbf{a}$.

$$\dot{\mathbf{w}} = -\frac{\partial \lambda_n(\mathbf{w})}{\partial \mathbf{w}} + \frac{1}{2} \left(\boldsymbol{\mu}^{\circ 2} + \frac{\partial \lambda_2(\mathbf{w})}{\partial \mathbf{w}} \circ \mathbf{v}^{\circ 2} \right) \tag{17}$$

$$\dot{\boldsymbol{\mu}} = -\mathbf{w} \circ \boldsymbol{\mu} \tag{18}$$

$$\dot{\mathbf{v}} = (\kappa - \lambda_2(\mathbf{w})) \mathbf{v} \tag{19}$$

Lemma 2

The equilibrium point of the system given by (17) to (19) satisfies the first-order necessary KKT conditions of the optimisation problem (10) with the feasible set (11).

Proof

The optimisation problem is of the form

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimise}} && f(\mathbf{x}) \\ & \text{subject to} && g_i(\mathbf{x}) \leq 0, \ \forall i = 1 \dots m + 1 \end{aligned} \tag{20}$$

where $\mathbf{x} \triangleq \mathbf{w}$, the objective function $f(\mathbf{x}) \triangleq \lambda_n(\mathbf{w})$ is convex, the inequality constraints $g_i(\mathbf{x}) \triangleq -w_{q(i-n)}, \ \forall i = 1 \dots m$ are affine and the final inequality constraint $g_{m+1} \triangleq (\kappa - \lambda_2(\mathbf{w}))$ is convex. The vector of dual variables is simply $\mathbf{y} \triangleq \frac{1}{2} [\boldsymbol{\mu}^{\circ 2\top}, \mathbf{v}^{\circ 2\top}]^\top$

It can be seen that the right-hand sides of (15) and (16) are equal to zero when and only when the complementary slackness condition holds: $\mathbf{g}(\mathbf{x}) \circ \mathbf{y} = \mathbf{0}$. Feasibility of the dual variables is assured by the initial conditions $\mu_{\{i,j\}} = 0, v_i(0) = 0$, so that $\mathbf{y}(t) \geq 0, \ \forall t \geq 0$. Setting the right-hand side of (17) to zero is identical to satisfying the stationarity condition $-\nabla f(\mathbf{x}) = \sum_i \nabla g_i(\mathbf{x}) y_i$. Furthermore, if primal feasibility does not hold, then $\dot{\mathbf{y}} \neq \mathbf{0}$ so long as $\mathbf{y} \neq \mathbf{0}$, and we know that for all finite time $\mathbf{y} > \mathbf{0}$ and may only approach zero exponentially fast (so long as $g_i(\mathbf{x})$ remains bounded from below). \square

Lemma 3

For the optimisation problem defined by (10) and (11), the first-order necessary KKT conditions are also sufficient.

Proof

Proof follows from the fact that the primal problem is convex and strictly feasible, thus Slater's condition holds (see [39] for details). \square

This implies that the equilibrium $[\mathbf{w}_{opt}^\top, \boldsymbol{\mu}_{opt}^\top, \mathbf{v}_{opt}^\top]^\top$ of (14) to (16) is also the solution to the optimisation problem (10), (11). It is now necessary to provide conditions for the system (14) to (16) to be exponentially stable. This will help in the later singular perturbation analysis. Consider the following result:

Theorem 1

The equilibrium point $[\mathbf{w}_{opt}^\top, \boldsymbol{\mu}_{opt}^\top, \mathbf{v}_{opt}^\top]^\top$ is locally exponentially stable if either $\nabla^2 \lambda_n(\mathbf{w}_{opt}) > 0$ or $\nabla^2 \lambda_2(\mathbf{w}_{opt}) < 0$.

Proof

Consider the generic convex minimisation problem (20). If the inequality constraints are such that $g_i(\mathbf{x}) \leq 0, \forall i, \mathbf{x} \in \mathcal{X}$, and the primal feasible set \mathcal{X} is convex and compact, $f(\mathbf{x})$ is strictly convex, there exists a unique optimal point \mathbf{x}_{opt} . At this point, some $r_1 \geq 1$ constraints will be tight (as $f(\mathbf{x})$ is monotonically increasing, we can be sure that at least one inequality will be tight), and some $r_2 \geq 1$ inequality constraints may be slack (there must be at least one slack inequality; by contradiction, if all were tight, then $\mathbf{w}_{opt} = \mathbf{0}$, and thus, $\lambda_2 = 0$, which is infeasible).

We can divide these tight and slack inequalities into strict inequalities $p_j(\mathbf{x}) < 0$ and equality constraints $h_j(\mathbf{x}) = 0$ by redefining

$$\forall i \text{ s.t. } g_i(\mathbf{x}_{opt}) = 0, h_j(\mathbf{x}) \triangleq g_i(\mathbf{x}), j = 1, \dots, r_1 \quad (21)$$

$$\forall i \text{ s.t. } g_i(\mathbf{x}_{opt}) < 0, p_j(\mathbf{x}) \triangleq g_i(\mathbf{x}), j = 1, \dots, r_2 \quad (22)$$

Now, the equivalent optimisation problem can be formulated:

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimise}} & f(\mathbf{x}) \\ \text{subject to} & h_i(\mathbf{x}) = 0, \forall i = 1 \dots r_1 \\ & p_i(\mathbf{x}) < 0, \forall i = 1 \dots r_2, \\ & \text{where } r_1 + r_2 = m + 1 \end{array} \quad (23)$$

The solution to this optimisation problem can be obtained from the following set of ODEs:

$$\dot{\mathbf{x}} = -\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} - \frac{1}{2} \sum_i \frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}} \alpha_i^2 - \frac{1}{2} \sum_i \frac{\partial p_i(\mathbf{x})}{\partial \mathbf{x}} \beta_i^2 \quad (24)$$

$$\dot{\boldsymbol{\alpha}} = \mathbf{h}(\mathbf{x}) \circ \boldsymbol{\alpha} \quad (25)$$

$$\dot{\boldsymbol{\beta}} = \mathbf{p}(\mathbf{x}) \circ \boldsymbol{\beta} \quad (26)$$

Linearisation of the system about the optimal point where $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_{opt}$, $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \boldsymbol{\alpha}_{opt}$, $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta} - \boldsymbol{\beta}_{opt}$, is:

$$\begin{bmatrix} \dot{\tilde{\mathbf{x}}} \\ \dot{\tilde{\boldsymbol{\alpha}}} \\ \dot{\tilde{\boldsymbol{\beta}}} \end{bmatrix} \approx \mathbf{J}_0 \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\boldsymbol{\alpha}} \\ \tilde{\boldsymbol{\beta}} \end{bmatrix} \quad (27)$$

$$\mathbf{J}_0 = \begin{bmatrix} -\nabla^2 f(\mathbf{x}) - \frac{1}{2} \sum_i \alpha_i^2 \nabla^2 h_i(\mathbf{x}) - \frac{1}{2} \sum_i \beta_i^2 \nabla^2 p_i(\mathbf{x}) - (\text{diag}\{\boldsymbol{\alpha}\} \nabla \mathbf{h}(\mathbf{x}))^\top - (\text{diag}\{\boldsymbol{\beta}\} \nabla \mathbf{p}(\mathbf{x}))^\top \\ \text{diag}\{\boldsymbol{\alpha}\} \nabla \mathbf{h}(\mathbf{x}) & \text{diag}\{\mathbf{h}(\mathbf{x})\} & \mathbf{0} \\ \text{diag}\{\boldsymbol{\beta}\} \nabla \mathbf{p}(\mathbf{x}) & \mathbf{0} & \text{diag}\{\mathbf{p}(\mathbf{x})\} \end{bmatrix} \quad (28)$$

evaluated at $\mathbf{x} = \mathbf{x}_{opt}$, $\boldsymbol{\alpha} = \boldsymbol{\alpha}_{opt}$, $\boldsymbol{\beta} = \boldsymbol{\beta}_{opt}$:

$$\mathbf{J}_0 = \begin{bmatrix} -\nabla^2 f(\mathbf{x}_{opt}) - \frac{1}{2} \sum_i \alpha_{opt,i}^2 \nabla^2 h_i(\mathbf{x}_{opt}) - (\text{diag}\{\boldsymbol{\alpha}_{opt}\} \nabla \mathbf{h}(\mathbf{x}_{opt}))^\top & \mathbf{0} \\ \text{diag}\{\boldsymbol{\alpha}_{opt}\} \nabla \mathbf{h}(\mathbf{x}_{opt}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}\{\mathbf{p}(\mathbf{x}_{opt})\} \end{bmatrix} \quad (29)$$

as $\boldsymbol{\beta}_{opt} = \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$.

It is evident that this Jacobian matrix has a block diagonal structure, indicating the decoupling of the dual variables whose associated inequality constraints are slack at the optimal point. In the vicinity of the optimal point, the square root of the (redundant) slack dual variables β_i will decay to zero exponentially fast with rate constants $p_i(\mathbf{x}_{opt}) < 0$. That is, the ‘more slack’ the constraint is, the faster its dual variable will decay away.

The other block is more interesting and has the form

$$\begin{aligned} \mathbf{J}_1 &= \begin{bmatrix} \mathbf{R} & -\mathbf{S} \\ \mathbf{S}^\top & \mathbf{0} \end{bmatrix} \\ \mathbf{R} &= -\nabla^2 f(\mathbf{x}_{opt}) - \frac{1}{2} \sum_i \alpha_{opt,i}^2 \nabla^2 h_i(\mathbf{x}_{opt}) \\ \mathbf{S} &= (\text{diag}\{\boldsymbol{\alpha}_{opt}\} \nabla \mathbf{h}(\mathbf{x}_{opt}))^\top \end{aligned} \quad (30)$$

In the vicinity of the optimal point, the primal variables \mathbf{x} will converge to \mathbf{x}_{opt} exponentially fast provided \mathbf{R} is negative definite. This follows from the following lemma. \square

Lemma 4

The matrix $\mathbf{J}_1 = \begin{bmatrix} \mathbf{R} & -\mathbf{S} \\ \mathbf{S}^\top & \mathbf{0} \end{bmatrix}$ has eigenvalues with negative real part provided the symmetric part of \mathbf{R} is negative definite and \mathbf{S} is full column rank.

Proof

$$\det\{\mathbf{J}_1\} = \det\{\mathbf{R}\} \det\{\mathbf{S}^\top \mathbf{R}^{-1} \mathbf{S}\} \quad (31)$$

It can be seen that $\det\{\mathbf{R}\} \neq 0$ as \mathbf{R} is nonsingular, and thus, \mathbf{R}^{-1} exists. Furthermore, $\det\{\mathbf{S}^\top \mathbf{R}^{-1} \mathbf{S}\}$ is only full rank if \mathbf{S} is full column rank, that is, $\mathbf{S}^\top \mathbf{S}$ is invertible. Therefore, $\det\{\mathbf{J}_1\} \neq 0$; \mathbf{J}_1 has no zero eigenvalues.

Considering the eigenvector equation $\mathbf{J}_1 \mathbf{v}_i = \lambda_i \mathbf{v}_i$, with $\mathbf{v}_i \triangleq [\mathbf{x}_i^\top, \mathbf{y}_i^\top]^\top$, we have

$$\begin{bmatrix} \mathbf{R} & -\mathbf{S} \\ \mathbf{S}^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} \quad (32)$$

$$\mathbf{R} \mathbf{x}_i - \mathbf{S} \mathbf{y}_i = \lambda_i \mathbf{x}_i \quad (33)$$

$$\mathbf{S}^\top \mathbf{x}_i = \lambda_i \mathbf{y}_i \quad (34)$$

Looking solely at (34), we can see that if $\mathbf{x}_i = \mathbf{0}$, then either $\lambda_i = 0$, which we have previously shown not to be true, or $\mathbf{y}_i = \mathbf{0}$, so that $\mathbf{v}_i = \mathbf{0}$, which is simply the trivial solution. Therefore, we can conclude that for all eigenvectors of \mathbf{J}_1 , $\mathbf{x}_i \neq \mathbf{0}$.

Now, we drop the subscripts and assume that \mathbf{v} is a unit eigenvector so that

$$\lambda = \mathbf{v}^* \mathbf{J}_1 \mathbf{v} \quad (35)$$

$$= [\mathbf{x}^*, \mathbf{y}^*] \begin{bmatrix} \mathbf{R} & -\mathbf{S} \\ \mathbf{S}^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \quad (36)$$

$$= \mathbf{x}^* \mathbf{R} \mathbf{x} - \mathbf{x}^* \mathbf{S} \mathbf{y} + \mathbf{y}^* \mathbf{S}^\top \mathbf{x} \quad (37)$$

$$= \mathbf{x}^* \mathbf{R} \mathbf{x} - \mathbf{x}^* \mathbf{S} \mathbf{y} + (\mathbf{y}^* \mathbf{S}^\top \mathbf{x})^\top \quad (38)$$

$$= \mathbf{x}^* \mathbf{R} \mathbf{x} - \mathbf{x}^* \mathbf{S} \mathbf{y} + \mathbf{x}^\top \overline{\mathbf{S} \mathbf{y}} \quad (39)$$

$$= \mathbf{x}^* \mathbf{R} \mathbf{x} - \mathbf{x}^* \mathbf{S} \mathbf{y} + \overline{\mathbf{x}^* \mathbf{S} \mathbf{y}} \quad (40)$$

$$= \mathbf{x}^* \mathbf{R} \mathbf{x} - 2\Im(\mathbf{x}^* \mathbf{S} \mathbf{y}) \cdot i, \quad i = \sqrt{-1} \quad (41)$$

As $\mathbf{x} \neq \mathbf{0}$ and the symmetric part of \mathbf{R} is negative definite, $\Re(\mathbf{x}^* \mathbf{R} \mathbf{x}) < 0$. This is the only contribution to the real part of the eigenvalue; thus, \mathbf{J}_1 is a stable matrix. \square

Note 1

If $f(\mathbf{x})$ and all $h_i(\mathbf{x})$ are convex, then it is apparent that \mathbf{R} is negative definite only if $f(\mathbf{x})$ or any $h_i(\mathbf{x})$ is strictly convex near \mathbf{x}_{opt} . We can be sure that the lower bound on the algebraic connectivity will be tight $h_j(\mathbf{x}) \triangleq \kappa - \lambda_2(\mathbf{w})$ (if it was not, then all weights could be scaled by a common factor less than 1, so that λ_n would also be scaled down further). In our case, any other tight inequality constraints are affine, and so the condition that $f(\mathbf{x})$ or any $h_i(\mathbf{x})$ is strictly convex near \mathbf{x}_{opt} reduces to the sufficient condition that either $\nabla^2 \lambda_n(\mathbf{w}_{opt}) > 0$ or $\nabla^2 \lambda_2(\mathbf{w}_{opt}) < 0$.

3.0.1. Strict convexity: In note 1, it was discussed that we require either $\lambda_n(\mathbf{w})$ to be strictly convex, or $\lambda_2(\mathbf{w})$ to be strictly concave, in the vicinity of the optimal point for exponential convergence. However, neither $\lambda_n(\mathbf{w})$ nor $-\lambda_2(\mathbf{w})$ is necessarily strictly convex. To account for this, we choose to minimise a modified function $f(\mathbf{x}) = (\lambda_n(\mathbf{w}))^2$. It is apparent that minimising $(\lambda_n(\mathbf{w}))^2$ is equivalent to minimising λ_n , so the two optimisation problems are equivalent.

In the following lemma, we show that as $\lambda_n(\mathbf{w})$ is positive and convex (near \mathbf{w}_{opt}), then its square will be strictly convex in the vicinity of the optimal point if that optimal point is distinct.

Lemma 5

If $f(\mathbf{x})$ is (not necessarily strictly) convex and positive, then $f(\mathbf{x})^2$ is strictly convex over non-flat regions.

Proof

$f(\mathbf{x})$ is convex, so Jensen's inequality holds:

$$f(\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1 - \theta) f(\mathbf{x}_2), \quad \forall \theta \in (0, 1) \quad (42)$$

Squaring both sides, so long as $f(\mathbf{x})$ is non-negative, we can be sure that

$$f(\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2)^2 \leq \theta^2 f(\mathbf{x}_1)^2 + 2\theta(1 - \theta) f(\mathbf{x}_1) f(\mathbf{x}_2) + (1 - \theta)^2 f(\mathbf{x}_2)^2, \quad \forall \theta \in (0, 1) \quad (43)$$

But to prove that $f(\mathbf{x})^2$ is strictly convex, we need to show that

$$f(\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2)^2 < \theta f(\mathbf{x}_1)^2 + (1 - \theta) f(\mathbf{x}_2)^2, \quad \forall \theta \in (0, 1) \quad (44)$$

By considering the right-hand side of (43) and (44), strict convexity follows if the following inequality holds

$$\theta^2 f(\mathbf{x}_1)^2 + 2\theta(1 - \theta) f(\mathbf{x}_1) f(\mathbf{x}_2) + (1 - \theta)^2 f(\mathbf{x}_2)^2 < \theta f(\mathbf{x}_1)^2 + (1 - \theta) f(\mathbf{x}_2)^2 \quad (45)$$

$$(\theta^2 - \theta) f(\mathbf{x}_1)^2 + 2\theta(1 - \theta) f(\mathbf{x}_1) f(\mathbf{x}_2) + ((1 - \theta)^2 - (1 - \theta)) f(\mathbf{x}_2)^2 < 0 \quad (46)$$

$$\theta(\theta - 1)(f(\mathbf{x}_1) - f(\mathbf{x}_2))^2 < 0 \tag{47}$$

which clearly holds in the interval so long as $f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$.

Note 2.

If $f(\mathbf{x}_1) = f(\mathbf{x}_2)$, then there may be a manifold of optimal edge weights. In this case, $(\lambda_n(\mathbf{w}))^2$ will not be strictly convex in this region. Instead, the set of optimal points will be a neutrally stable manifold for the set of ODEs, but local convergence to this manifold will be exponentially fast. \square

Substituting in the altered objective function $f(\mathbf{x}) = \lambda_n(\mathbf{w})^2$ in (20), we arrive at the modified weight adaptation law:

$$\dot{w}_{\{i,j\}} = -\frac{\partial(\lambda_n(\mathbf{w}))^2}{\partial w_{\{i,j\}}} + \frac{1}{2} \left(\mu_{\{i,j\}}^2 + \frac{\partial \lambda_2(\mathbf{w})}{\partial w_{\{i,j\}}} v_{\{i,j\}}^2 \right) \tag{48}$$

$$= -2\lambda_n(\mathbf{w}) \frac{\partial \lambda_n(\mathbf{w})}{\partial w_{\{i,j\}}} + \frac{1}{2} \left(\mu_{\{i,j\}}^2 + \frac{\partial \lambda_2(\mathbf{w})}{\partial w_{\{i,j\}}} v_{\{i,j\}}^2 \right) \tag{49}$$

$$\dot{\mu}_{\{i,j\}} = -w_{\{i,j\}} \mu_{\{i,j\}}, \quad \mu_{\{i,j\}}(0) > 0 \tag{50}$$

$$\dot{v}_{\{i,j\}} = (\kappa - \lambda_2(\mathbf{w})) v_{\{i,j\}}, \quad v_{\{i,j\}}(0) > 0 \tag{51}$$

where the collection of all ODEs for each node and edge may be written in the more compact form:

$$\dot{\mathbf{w}} = -2\lambda_n(\mathbf{w}) \frac{\partial \lambda_n(\mathbf{w})}{\partial \mathbf{w}} + \frac{1}{2} \left(\boldsymbol{\mu}^{\circ 2} + \frac{\partial \lambda_2(\mathbf{w})}{\partial \mathbf{w}} \circ \mathbf{v}^{\circ 2} \right) \tag{52}$$

$$\dot{\boldsymbol{\mu}} = -\mathbf{w} \circ \boldsymbol{\mu} \tag{53}$$

$$\dot{\mathbf{v}} = (\kappa - \lambda_2(\mathbf{w})) \mathbf{v} \tag{54}$$

Observing the set of ODEs, (49) to (51), it is clear to see that these are not fully distributed. The functions $\lambda_2(\mathbf{w})$, $\lambda_n(\mathbf{w})$ and their partial derivatives with respect to the edge weights are global functions of all the edge weights. Thus, to achieve a fully distributed strategy, we must make distributed estimates of these functions. This is achieved using the distributed algebraic connectivity estimator from [21] and a distributed spectral radius estimator. Next, we will recall some of the results from [21] in Section 4.1, before continuing in Section 4.2 to explain how the system can be modified for estimating the spectral radius and its partial derivatives with respect to their edge weights.

4. EIGENVALUE ESTIMATION

4.1. Algebraic connectivity estimator

As described in [21], the algebraic connectivity of a weighted undirected network can be estimated using the ODE :

$$\dot{\mathbf{a}} = -\frac{k_1}{n} \mathbf{1}\mathbf{1}^\top \mathbf{a} - k_2 \mathbf{L}\mathbf{a} - k_3 \left(\frac{\mathbf{a}^\top \mathbf{a}}{n} - 1 \right) \mathbf{a} \tag{55}$$

That is, every node follows

$$\dot{a}_i = -k_1 \langle \mathbf{a} \rangle + k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (a_j - a_i) - k_3 (\langle \mathbf{a}^{\circ 2} \rangle - 1) a_i \tag{56}$$

with $\langle \mathbf{a} \rangle = \frac{1}{n} \sum_i a_i$ being the arithmetic mean of the components of \mathbf{a} and $\langle \mathbf{a}^{\circ 2} \rangle = \frac{1}{n} \sum_i a_i^2$ being the mean of the sum of squared components. It can be shown that if the algebraic connectivity is distinct, the stationary point \mathbf{a}^* is defined as

$$\mathbf{a}^* = \pm \mathbf{v}_2(\mathbf{w}) \sqrt{\frac{n(k_3 - k_2 \lambda_2(\mathbf{w}))}{k_3}} \quad (57)$$

where \mathbf{v}_2 is the unit eigenvector associated with $\lambda_2(\mathbf{w})$. This allows the following estimates to be performed [21]:

$$\widehat{\lambda_2(\mathbf{w})} = \frac{k_3}{k_2} \left(1 - \frac{\|\mathbf{a}\|_2^2}{n} \right) \quad (58)$$

$$\frac{\partial \widehat{\lambda_2(\mathbf{w})}}{\partial w_{\{i,j\}}} = \frac{(a_i - a_j)^2}{\|\mathbf{a}\|_2^2} \quad (59)$$

The equilibrium point \mathbf{a}^* is locally exponentially stable, provided that $k_1 > k_2 \lambda_2$ and $k_3 > k_2 \lambda_2$, $k_2 > 0$. Moreover, the rate constant of this local exponential convergence is given by

$$c_1 = -\min\{k_1 - k_2 \lambda_2, 2(k_3 - k_2 \lambda_2), k_2(\lambda_3 - \lambda_2)\} \quad (60)$$

Again, observing (56), it is apparent that each node requires access to the global functions $\langle \mathbf{a} \rangle$ and $\langle \mathbf{a}^{\circ 2} \rangle$, and so the system is not fully distributed. As in [21], a solution is to use further subsystems to make distributed estimates of these quantities, using PI average consensus estimators [36]. These estimators will be discussed in Section 5.

4.2. Spectral radius estimator

The estimation approach used for the algebraic connectivity can be now extended to obtain an estimate of the spectral radius λ_n and its partial derivatives with respect to each edge weight. However, instead of the action from \mathbf{L} being used to drive the system towards consensus so that the slowest mode dominates (which associated with λ_2), \mathbf{L} now needs to provide a diverging force so that the fastest mode (associated with λ_n) dominates. Again, the consensus mode is deflated through the action of the first term, and the third term provides a normalising force so that the vector \mathbf{b} remains bounded and is driven away from the origin. Specifically, we propose to use the dynamics:

$$\dot{\mathbf{b}} = -\frac{k_1}{n} \mathbf{1} \mathbf{1}^\top \mathbf{b} + k_2 \mathbf{L} \mathbf{b} - k_3 \left(\frac{\mathbf{b}^\top \mathbf{b}}{n} - 1 \right) \mathbf{b} \quad (61)$$

Taking a local view of the system, every node follows,

$$\dot{b}_i = -k_1 \langle \mathbf{b} \rangle - k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (b_j - b_i) - k_3 (\langle \mathbf{b}^{\circ 2} \rangle - 1) b_i \quad (62)$$

using the global functions $\langle \mathbf{b} \rangle = \frac{1}{n} \sum_i b_i$ and $\langle \mathbf{b}^{\circ 2} \rangle = \frac{1}{n} \sum_i b_i^2$.

Taking the orthogonal matrix $\mathbf{V} = [\frac{1}{\sqrt{n}} \mathbf{1}, \mathbf{v}_2, \dots, \mathbf{v}_n]$, which diagonalises the graph Laplacian $\mathbf{\Lambda} = \mathbf{V}^\top \mathbf{L} \mathbf{V}$, as a change of basis, the system may be easily diagonalised using the transformation $\mathbf{b}_d = \mathbf{V}^\top \mathbf{b}$ yielding

$$\dot{\mathbf{b}}_d = \left(-k_1 \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + k_2 \mathbf{\Lambda} + k_3 \mathbf{I} \right) \mathbf{b}_d - \frac{k_3}{n} \mathbf{b}_d \mathbf{b}_d^\top \mathbf{b}_d \quad (63)$$

Each mode can be taken in turn, so that for the consensus mode $b_{d,1}$

$$\dot{b}_{d,1} = \left(k_3 - k_1 - \frac{k_3}{n} \mathbf{b}_d^\top \mathbf{b}_d \right) b_{d,1} \quad (64)$$

which has stationary points at $b_{d,1}^* = 0$ or when $\mathbf{b}_d^{*\top} \mathbf{b}_d^* = \frac{n(k_3 - k_1)}{k_3}$. Clearly, if $k_1 > k_3$, then this second equilibrium point does not exist, and the first equilibrium point is globally exponentially stable; the consensus mode will deflate.

Other modes follow equations of the form

$$\dot{b}_{d,i} = \left(k_3 + k_2 \lambda_i - \frac{k_3}{n} \mathbf{b}_d^{*\top} \mathbf{b}_d^* \right) b_{d,i} \tag{65}$$

for $i = 2, \dots, n$

4.2.1. Location of equilibria: For each of the non-consensus modes, either $b_{d,i}^* = 0$ or $\mathbf{b}_d^{*\top} \mathbf{b}_d^* = \frac{n(k_3 + k_2 \lambda_i)}{k_3}$. Thus, there is one equilibrium point located at the origin $\mathbf{b}_d^* = \mathbf{0}$, and for each unique eigenvalue λ_i , there is an associated equilibrium space such that $\mathbf{b}_d^{*\top} \mathbf{b}_d^* = \frac{n(k_3 + k_2 \lambda_i)}{k_3}$, with zero elements $b_{d,j} = 0$ for all $\lambda_j \neq \lambda_i$.

In the case that λ_i has unitary multiplicity, then the associated equilibrium is a set of two points equidistant from the origin (a 0-sphere) given by

$$b_{d,i}^* = \pm \sqrt{\frac{n(k_3 + k_2 \lambda_i)}{k_3}} \tag{66}$$

$$b_{d,j}^* = 0, \forall j \neq i \tag{67}$$

and in the general case that the eigenvalue is not distinct $\lambda_i = \dots = \lambda_{i+p}$, having algebraic multiplicity $p + 1$, then the associated equilibrium set is a p -sphere of radius $\sqrt{\frac{n(k_3 + k_2 \lambda_i)}{k_3}}$ defined by the set of solutions:

$$\sum_{k \in \{i, \dots, i+p\}} b_{d,k}^{*2} = \frac{n(k_3 + k_2 \lambda_i)}{k_3} \tag{68}$$

$$b_{d,j}^* = 0, \forall j \notin \{i, \dots, i + p\} \tag{69}$$

Lemma 6

The system defined in (61) is bounded, contains no periodic orbits and there exists at least one stable equilibrium.

Proof

We notice that the diagonalised system, (63), is a gradient system $\dot{\mathbf{b}}_d = -\nabla V(\mathbf{b}_d)$, where

$$V(\mathbf{b}_d) = \frac{1}{2} k_1 b_{d,1}^2 - \frac{1}{2} k_2 \sum_{i=2}^n \lambda_i b_{d,i}^2 - \frac{1}{2} k_3 \sum_{i=1}^n b_{d,i}^2 + \frac{k_3}{4n} \sum_{i=1}^n \sum_{j=1}^n b_{d,i}^2 b_{d,j}^2 \tag{70}$$

As such, we can be sure that the system contains no periodic orbits (see [40] for further details). Further to this, we can immediately see that the system is bounded, as for large \mathbf{b}_d , the positive quartic terms will dominate. The Lyapunov-like potential function $V(\mathbf{b}_d)$ is quartic; thus, by continuity, there must exist a global minimum of the function, and this critical point will be a stable equilibrium. \square

Theorem 2

Under Assumption 1, with control parameters satisfying $k_1 > k_3 > 0$, $k_2 > 0$, the equilibrium set associated with λ_n , which consists of two distinct points,

$$\mathbf{b}^* = \pm \mathbf{v}_n \sqrt{\frac{n(k_3 + k_2 \lambda_n)}{k_3}} \tag{71}$$

is the sole stable equilibrium set. Moreover, any trajectory that does not originate at an equilibrium will locally exponentially tend to one of these two points with a constant rate of convergence given by

$$c_2 = -k_2(\lambda_n - \lambda_{n-1}) \tag{72}$$

Proof

Taking a small perturbation $\tilde{\mathbf{b}}_d = \mathbf{b}_d^* - \mathbf{b}_d$, the system defined in (63) can be linearised about the equilibrium \mathbf{b}_d^* :

$$\tilde{\mathbf{b}}_d \approx \left(-k_1 \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + k_2 \mathbf{\Lambda} + k_3 \mathbf{I} - \frac{k_3}{n} (2\mathbf{b}_d^* \mathbf{b}_d^{*\top} + (\mathbf{b}_d^{*\top} \mathbf{b}_d^*) \mathbf{I}) \right) \tilde{\mathbf{b}}_d \tag{73}$$

In the simplest case where all eigenvalues of the graph Laplacian are distinct, and thus, the equilibria associated with each eigenvalue is a set of two distinct points (a 0-sphere), each i th equilibrium $\mathbf{b}_{d,i}$ will have one non-zero element only at the i th position. Thus, the only non-diagonal term in (73), $2\mathbf{b}_d^* \mathbf{b}_d^{*\top}$, will simplify to a diagonal matrix with the value $2b_{d,i}^{*2}$ located at the (i, i) element and zeros elsewhere. Therefore, the linearised system is diagonal, and eigenvalues may be simply read off, noting that $\mathbf{b}_d^{*\top} \mathbf{b}_d^* = b_{d,i}^{*2} = \frac{n(k_3 + k_2 \lambda_i)}{k_3}$.

At the 1st equilibrium point (at the origin, $\mathbf{b}_d = \mathbf{0}$), the eigenvalues are

$$\{-k_1 + k_3, k_3 + k_2 \lambda_2, \dots, k_3 + k_2 \lambda_n\} \tag{74}$$

The eigenvalues at the i th equilibria ($i \in \{2, \dots, n\}$) are

$$\{-k_1 - k_2 \lambda_i, k_2(\lambda_j - \lambda_i) \forall j \neq i, -2(k_3 + k_2 \lambda_i)\} \tag{75}$$

It can be seen that if $k_1 > k_3$, with $k_1, k_2, k_3 > 0$, then all eigenvalues are negative only at the n th equilibrium. Again, we stress that this corresponds to an equilibrium set of two distinct points under Assumption 1 (when λ_n is distinct), and thus, each of these points is going to be an exponentially stable equilibrium point. For all other equilibria, there is at least one positive eigenvalue, and the points will be unstable. Therefore, in the case that all eigenvalues are distinct, almost any trajectory will converge exponentially on one of the two points in the n th equilibrium set.

When eigenvalues are not distinct, the analysis becomes somewhat more complicated. Consider the non-trivial eigenvalue $\lambda_i = \dots = \lambda_{i+p} \neq 0$ having multiplicity $p + 1$, then the matrix $2\mathbf{b}_d^* \mathbf{b}_d^{*\top}$ in the linearised system, about any point on the p -sphere equilibrium, will not be diagonal. It will, however, be block diagonal, with all elements being zero except for the square block between the (i, i) and $(i + p, i + p)$ elements. Therefore, it is still simple to determine the $n - (p + 1)$ eigenvalues, which are not affected by this block:

$$\{k_3 - k_1, k_2(\lambda_j - \lambda_i) \forall j \notin \{i, \dots, i + p\}\} \tag{76}$$

Thus, if there is any eigenvalue λ_j larger than λ_i , the associated p -sphere equilibrium will be unstable in at least one direction. By the fact that the system is a bounded gradient system, there must be at least one stable equilibrium, and through elimination, we deduce that this is the equilibrium associated with the spectral radius λ_n .

We therefore conclude that for a Laplacian where λ_n has multiplicity $p + 1$, from any initial condition that is not in the equilibrium set, the trajectory will converge to the p -sphere equilibrium defined by

$$\sum_{i \in \{n-p, \dots, n\}} b_{d,i}^{*2} = \frac{n(k_3 + k_2 \lambda_n)}{k_3} \tag{77}$$

$$b_{d,j}^* = 0, \forall j \notin \{n - p, \dots, n\} \tag{78}$$

Reversing the diagonal transformation, we find the stable equilibrium set for the system in (61):

$$\mathbf{b}^* = \sum_{i=n-p}^n \alpha_i \mathbf{v}_i, \quad \sum \alpha_i^2 = \frac{n(k_3 + k_2 \lambda_n)}{k_3} \quad (79)$$

In the special case where λ_n is distinct, then \mathbf{b}^* simplifies to

$$\mathbf{b}^* = \pm \mathbf{v}_n \sqrt{\frac{n(k_3 + k_2 \lambda_n)}{k_3}} \quad (80)$$

□

4.2.2. *Use as an estimator.*: From this stable equilibrium, we see that \mathbf{b} is an estimate for the eigenvector associated with λ_n . In the case that λ_n is distinct (Assumption 1 holds), then \mathbf{b} almost surely converges to the associated eigenvector, whilst when λ_n is not distinct, then \mathbf{b} almost surely converges to a vector in the associated eigenspace.

Whether distinct or not, the magnitude of \mathbf{b}^* is dependent on λ_n , so by rearranging (79), we have

$$\|\mathbf{b}^*\|_2^2 = \frac{n(k_3 + k_2 \lambda_n)}{k_3} \quad (81)$$

$$\lambda_n = \frac{k_3}{k_2} \left(\frac{\|\mathbf{b}^*\|_2^2}{n} - 1 \right) \quad (82)$$

and an estimate of λ_n can be obtained as

$$\widehat{\lambda}_n = \frac{k_3}{k_2} \left(\frac{\|\mathbf{b}\|_2^2}{n} - 1 \right) \quad (83)$$

Estimates of the partial derivatives with respect to the edge weights $\frac{\partial \lambda_n}{\partial w_{\{i,j\}}}$ may also be obtained using this estimator. Taking the eigenvalue equation for eigenvalue $\lambda(\mathbf{w})$, with associated unit eigenvector $\hat{\mathbf{v}}(\mathbf{w})$, and taking the derivative with respect to the edge weight yield

$$\lambda \hat{\mathbf{v}}(\mathbf{w}) = \mathbf{L}(\mathbf{w}) \hat{\mathbf{v}}(\mathbf{w}) \quad (84)$$

$$\lambda = \hat{\mathbf{v}}^T \mathbf{L} \hat{\mathbf{v}} \quad (85)$$

$$\frac{\partial \lambda}{\partial w_{\{i,j\}}} = \frac{\partial \hat{\mathbf{v}}^T}{\partial w_{\{i,j\}}} \mathbf{L} \hat{\mathbf{v}} + \hat{\mathbf{v}}^T \frac{\partial \mathbf{L}}{\partial w_{\{i,j\}}} \hat{\mathbf{v}} + \hat{\mathbf{v}}^T \mathbf{L} \frac{\partial \hat{\mathbf{v}}}{\partial w_{\{i,j\}}} \quad (86)$$

As \mathbf{L} is symmetric, this ensures that, for the unit eigenvector $\hat{\mathbf{v}}$,

$$\frac{\partial \hat{\mathbf{v}}^T}{\partial w_{\{i,j\}}} \mathbf{L} \hat{\mathbf{v}} + \hat{\mathbf{v}}^T \mathbf{L} \frac{\partial \hat{\mathbf{v}}}{\partial w_{\{i,j\}}} = \lambda \frac{\partial (\hat{\mathbf{v}}^T \hat{\mathbf{v}})}{\partial w_{\{i,j\}}} = 0 \quad (87)$$

Therefore,

$$\frac{\partial \lambda}{\partial w_{\{i,j\}}} = \hat{\mathbf{v}}^T \frac{\partial \mathbf{L}}{\partial w_{\{i,j\}}} \hat{\mathbf{v}} = (\hat{v}_i - \hat{v}_j)^2 \quad (88)$$

Using 88 and $\mathbf{b} / \|\mathbf{b}\|_2$ as an estimate of the unit eigenvector \mathbf{v}_n , an estimate of the partial derivative can be obtained as

$$\widehat{\frac{\partial \lambda_n}{\partial w_{i,j}}} = \frac{(b_i - b_j)^2}{\|\mathbf{b}\|_2^2} \quad (89)$$

just as for the algebraic connectivity sensitivities in 59.

4.3. Distributed estimation

Both the λ_2 and λ_n estimators require node dynamics, (56) and (62), to be dictated by global variables, specifically through the quantities $\langle \mathbf{a} \rangle$, $\langle \mathbf{a}^{\circ 2} \rangle$ and $\langle \mathbf{b} \rangle$, $\langle \mathbf{b}^{\circ 2} \rangle$, respectively. These four functions are arithmetic means of the components of \mathbf{a} and \mathbf{b} and of the squares of the components. As such, they require information from every node and not just those in the neighbourhood of node i . As in [21], a solution is to use PI average consensus estimators [36] to obtain distributed estimates of these means and use these local estimates in the eigenvector estimators.

5. PROPORTIONAL–INTEGRAL AVERAGE CONSENSUS

In [36], it is shown that the arithmetic mean of a vector, say \mathbf{u} , can be estimated in a distributed way on a network, using the linear system:

$$\begin{aligned}\dot{\pi}_i^{\mathbf{u}} &= k_{\gamma}(u_i - \pi_i^{\mathbf{u}}) + k_P \sum_{j \in \mathcal{N}_i} w_{\{i,j\}}(\pi_j^{\mathbf{u}} - \pi_i^{\mathbf{u}}) + k_I \sum_{j \in \mathcal{N}_i} w_{\{i,j\}}(\zeta_j^{\mathbf{u}} - \zeta_i^{\mathbf{u}}) \\ \dot{\zeta}_i^{\mathbf{u}} &= -k_I \sum_{j \in \mathcal{N}_i} w_{\{i,j\}}(\pi_j^{\mathbf{u}} - \pi_i^{\mathbf{u}})\end{aligned}\quad (90)$$

Here, we use the notation $\pi_i^{\mathbf{u}}$ to designate the i th node's proportional variable for estimating the mean of the vector \mathbf{u} , with $\zeta_i^{\mathbf{u}}$ being the corresponding integrator variable. This notation is introduced as we will need to use four PI average consensus systems for estimating the means:

$$\langle \mathbf{a} \rangle = \frac{\mathbf{1}^{\top} \mathbf{a}}{n}, \quad \langle \mathbf{a}^{\circ 2} \rangle = \frac{\mathbf{a}^{\top} \mathbf{a}}{n}, \quad \langle \mathbf{b} \rangle = \frac{\mathbf{1}^{\top} \mathbf{b}}{n}, \quad \langle \mathbf{b}^{\circ 2} \rangle = \frac{\mathbf{b}^{\top} \mathbf{b}}{n}\quad (91)$$

As such, the superscripts for the four PI consensus systems will be: \mathbf{a} , $\mathbf{a}^{\circ 2}$, \mathbf{b} , and $\mathbf{b}^{\circ 2}$. However, for the moment, we will recall some results about the generic PI system (90).

The collection of all agents proportional and integrator variable dynamics can be concisely written as the matrix differential equation:

$$\begin{bmatrix} \dot{\pi}^{\mathbf{u}} \\ \dot{\zeta}^{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} -k_P \mathbf{L}(\mathbf{w}) - k_{\gamma} \mathbf{I} & -k_I \mathbf{L}(\mathbf{w}) \\ k_I \mathbf{L}(\mathbf{w}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \pi^{\mathbf{u}} \\ \zeta^{\mathbf{u}} \end{bmatrix} + k_{\gamma} \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix}\quad (92)$$

It can be shown that the stable equilibrium manifolds (provided $k_{\gamma}, k_P, k_I > 0$) for these systems are

$$\pi^{\mathbf{u},*} = \frac{\mathbf{1} \mathbf{1}^{\top} \mathbf{u}}{n}\quad (93)$$

$$\zeta^{\mathbf{u},*} = -\frac{k_{\gamma}}{k_I} \mathbf{L}^{\dagger}(\mathbf{w}) \mathbf{u} + \frac{\mathbf{1} \mathbf{1}^{\top} \zeta^{\mathbf{u}}(0)}{n}\quad (94)$$

where $\mathbf{L}^{\dagger}(\mathbf{w})$ is the Moore–Penrose inverse of the graph Laplacian:

$$\mathbf{L}^{\dagger}(\mathbf{w}) \triangleq \left(\mathbf{L}(\mathbf{w}) - \frac{\mathbf{1} \mathbf{1}^{\top}}{n} \right)^{-1} - \frac{\mathbf{1} \mathbf{1}^{\top}}{n}\quad (95)$$

Considering as input \mathbf{u} and as output $\pi^{\mathbf{u}}$, there is an uncontrollable and unobservable mode in the integrator variables for each PI average consensus estimator associated with the consensus mode, $\mathbf{1}^{\top} \zeta^{\mathbf{u}}(0)$, which does not have any effect on any other variables. Deflating the systems in this mode through simultaneous diagonalisation, and removal of the zero rows/columns, then results in a distinct equilibrium point. Let $\mathbf{V}(\mathbf{w})$ be the matrix of right eigenvectors of $\mathbf{L}(\mathbf{w})$ so that $\mathbf{L} = \mathbf{V}^{\top}(\mathbf{w}) \mathbf{\Lambda}(\mathbf{w}) \mathbf{V}(\mathbf{w})$. Because we know that the unit eigenvector associated with the consensus mode is simply $\mathbf{v}_1 = \frac{1}{\sqrt{n}}$, we can define an $n \times (n - 1)$ matrix of the unknown unit eigenvectors $\mathbf{W}(\mathbf{w})$

$$\mathbf{V}(\mathbf{w}) = \left[\frac{1}{\sqrt{n}} \mathbf{W}(\mathbf{w}) \right]\quad (96)$$

and consequently perform the diagonalisation and truncation of the uncontrollable/unobservable mode

$$\boldsymbol{\pi}_d^u = \mathbf{V}^\top(\mathbf{w})\boldsymbol{\pi}^u, \boldsymbol{\xi}_d^u = \mathbf{W}^\top(\mathbf{w})\boldsymbol{\xi}^u \quad (97)$$

so that

$$\begin{bmatrix} \dot{\boldsymbol{\pi}}_d^u \\ \dot{\boldsymbol{\xi}}_d^u \end{bmatrix} = \begin{bmatrix} -k_P \boldsymbol{\Lambda}(\mathbf{w}) - k_\gamma \mathbf{I} & k_I \boldsymbol{\Lambda}_r^\top(\mathbf{w}) \\ -k_I \boldsymbol{\Lambda}_r(\mathbf{w}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\pi}_d^u \\ \boldsymbol{\xi}_d^u \end{bmatrix} + k_\gamma \begin{bmatrix} \mathbf{V}^\top(\mathbf{w})\mathbf{u} \\ \mathbf{0} \end{bmatrix} \quad (98)$$

with $\boldsymbol{\Lambda}_r(\mathbf{w})$ being the truncated matrix, with the first row removed:

$$\boldsymbol{\Lambda}_r(\mathbf{w}) = \begin{bmatrix} 0 & \lambda_2(\mathbf{w}) & 0 & 0 & \dots \\ 0 & 0 & \lambda_3(\mathbf{w}) & 0 & \dots \\ 0 & 0 & 0 & \ddots & \\ \vdots & \vdots & & & \lambda_n(\mathbf{w}) \end{bmatrix} \quad (99)$$

By Lemma 4, the system matrix of (98) is Hurwitz. Thus, the distinct, stable equilibrium point in the transformed variables is

$$\boldsymbol{\pi}_d^{u,*} = \begin{bmatrix} \frac{\mathbf{1}^\top \mathbf{u}}{\sqrt{n}} \\ \mathbf{0} \end{bmatrix} \quad (100)$$

$$\boldsymbol{\xi}_d^{u,*} = -\frac{k_\gamma}{k_I} \boldsymbol{\Lambda}_{c,r}^{-1}(\mathbf{w}) \mathbf{W}^\top(\mathbf{w})\mathbf{u} \quad (101)$$

where the square matrix $\boldsymbol{\Lambda}_{c,r}(\mathbf{w})$ is the further truncated matrix (the first row and the first column are removed):

$$\boldsymbol{\Lambda}_{c,r}(\mathbf{w}) = \begin{bmatrix} \lambda_2(\mathbf{w}) & 0 & 0 & \dots \\ 0 & \lambda_3(\mathbf{w}) & 0 & \dots \\ 0 & 0 & \ddots & \\ \vdots & & & \lambda_n(\mathbf{w}) \end{bmatrix} \quad (102)$$

6. SINGULAR PERTURBATION FORMULATION

We are now ready to substitute the distributed estimates from the faster subsystems into the slower subsystems, formulating an entirely distributed system. Starting with the base subsystems, the four PI average consensus estimators, each node i maintains an estimate of the four means:

$$\widehat{\langle \mathbf{a} \rangle}^{(i)} = \pi_i^{\mathbf{a}} \quad (103)$$

$$\widehat{\langle \mathbf{a}^{\circ 2} \rangle}^{(i)} = \pi_i^{\mathbf{a}^{\circ 2}} \quad (104)$$

$$\widehat{\langle \mathbf{b} \rangle}^{(i)} = \pi_i^{\mathbf{b}} \quad (105)$$

$$\widehat{\langle \mathbf{b}^{\circ 2} \rangle}^{(i)} = \pi_i^{\mathbf{b}^{\circ 2}} \quad (106)$$

where we indicate the node i 's estimate of a function using a hat and superscript (i) . These estimates are substituted into the eigenvalue estimators, (56) and (62), to yield the distributed estimators:

$$\begin{aligned} \dot{a}_i &= -k_1 \widehat{\langle \mathbf{a} \rangle}^{(i)} + k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (a_j - a_i) - k_3 (\widehat{\langle \mathbf{a}^{\circ 2} \rangle}^{(i)} - 1) a_i \\ &= -k_1 \pi_i^{\mathbf{a}} + k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (a_j - a_i) - k_3 (\pi_i^{\mathbf{a}^{\circ 2}} - 1) a_i \end{aligned} \quad (107)$$

$$\begin{aligned} \dot{b}_i &= -k_1 \widehat{\langle \mathbf{b} \rangle}^{(i)} - k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (b_j - b_i) - k_3 (\widehat{\langle \mathbf{b}^{o2} \rangle}^{(i)} - 1) b_i \\ &= -k_1 \pi_i^{\mathbf{b}} - k_2 \sum_{j \in \mathcal{N}_i} w_{\{i,j\}} (b_j - b_i) - k_3 (\pi_i^{\mathbf{b}^{o2}} - 1) b_i \end{aligned} \quad (108)$$

From (58), (59), (83) and (89), these distributed estimators allow the following local estimates to be made at each node i :

$$\widehat{\lambda}_2^{(i)} = \frac{k_3}{k_2} \left(1 - \pi_i^{\mathbf{a}^{o2}} \right) \quad (109)$$

$$\frac{\partial \widehat{\lambda}_2(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i)} = \frac{(a_i - a_j)^2}{n \pi_i^{\mathbf{a}^{o2}}} \quad (110)$$

$$\widehat{\lambda}_n^{(i)} = \frac{k_3}{k_2} \left(\pi_i^{\mathbf{b}^{o2}} - 1 \right) \quad (111)$$

$$\frac{\partial \widehat{\lambda}_n(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i)} = \frac{(b_i - b_j)^2}{n \pi_i^{\mathbf{b}^{o2}}} \quad (112)$$

Each edge $\{i, j\}$ has access to both the local variables at node i and at node j ; thus, it can also make local estimates of these functions. So as not to bias one parent node over the other, we choose to use the mean of the consensus variables from i and j when either would do, as we expect consensus to be reached.

$$\widehat{\lambda}_2^{(i,j)} = \frac{k_3}{k_2} \left(1 - \frac{\pi_i^{\mathbf{a}^{o2}} + \pi_j^{\mathbf{a}^{o2}}}{2} \right) \quad (113)$$

$$\frac{\partial \widehat{\lambda}_2(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i,j)} = \frac{2(a_i - a_j)^2}{n(\pi_i^{\mathbf{a}^{o2}} + \pi_j^{\mathbf{a}^{o2}})} \quad (114)$$

$$\widehat{\lambda}_n^{(i,j)} = \frac{k_3}{k_2} \left(\frac{\pi_i^{\mathbf{b}^{o2}} + \pi_j^{\mathbf{b}^{o2}}}{2} - 1 \right) \quad (115)$$

$$\frac{\partial \widehat{\lambda}_n(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i,j)} = \frac{2(b_i - b_j)^2}{n(\pi_i^{\mathbf{b}^{o2}} + \pi_j^{\mathbf{b}^{o2}})} \quad (116)$$

Substituting these local estimates into the weight optimisation layer (49) and (51), we finally accomplish the goal of formulating an entirely distributed adaptive solution to the optimisation problem described in (10) and (11).

$$\dot{w}_{\{i,j\}} = -2 \widehat{\lambda}_n(\mathbf{w}) \frac{\partial \widehat{\lambda}_n(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i,j)} + \frac{1}{2} \frac{\widehat{\lambda}_2(\mathbf{w})}{\partial w_{\{i,j\}}}^{(i,j)} \left(\frac{v_i + v_j}{2} \right)^2 + \frac{1}{2} \mu_{\{i,j\}}^2 \quad (117)$$

$$= -\frac{4k_3}{nk_2} \left(\frac{\pi_i^{\mathbf{b}^{o2}} + \pi_j^{\mathbf{b}^{o2}}}{2} - 1 \right) \frac{(b_i - b_j)^2}{(\pi_i^{\mathbf{b}^{o2}} + \pi_j^{\mathbf{b}^{o2}})} + \frac{(a_i - a_j)^2}{n(\pi_i^{\mathbf{a}^{o2}} + \pi_j^{\mathbf{a}^{o2}})} \left(\frac{v_i + v_j}{2} \right)^2 + \frac{1}{2} \mu_{\{i,j\}}^2$$

$$\dot{\mu}_{\{i,j\}} = -w_{\{i,j\}} \mu_{\{i,j\}}, \quad \mu_{\{i,j\}}(0) > 0 \quad (118)$$

$$\begin{aligned} \dot{v}_i &= \left(\kappa - \widehat{\lambda}_2(\mathbf{w})^{(i)} \right) v_i \\ &= \left(\kappa - \frac{k_3}{k_2} \left(1 - \pi_i^{a \circ 2} \right) \right) v_i, \quad v_i(0) > 0 \end{aligned} \tag{119}$$

Note 2

Note that we are now locating the variables v_i , associated with the inequality constraint $\lambda_2(\mathbf{w}) \geq \kappa$, at each node, rather than at each edge in (51), to reduce the amount of duplication (n is typically less than m). It is also more common that computation be performed in each node, and to this end, for instance, the $\mu_{\{i,j\}}$ variables could also be distributed over the nodes by duplication in the incident nodes. Let the dual variable located at the i th node and responsible for guaranteeing non-negativity for the edge $\{i, j\}$ be denoted $\mu_i^{\{i,j\}}$, then these variables can adapt according to

$$\dot{\mu}_i^{\{i,j\}} = -w_{\{i,j\}} \mu_i^{\{i,j\}}, \quad \mu_i^{\{i,j\}}(0) > 0 \tag{120}$$

If $\mu_i^{\{i,j\}}(0) = \mu_j^{\{i,j\}}(0)$, then both incident nodes of the edge $\{i, j\}$ maintain a duplicate of the dual variable $\mu_{\{i,j\}}$ as before in (118), with $\mu_i^{\{i,j\}}(0) = \mu_j^{\{i,j\}}(0) = \mu_{\{i,j\}}(0)$. It can be seen that any edge-based dynamics can be duplicated on the incident nodes in this fashion. However, for the sake of clarity and conciseness of notation, we continue to treat these variables as if they were located at edges to avoid the need for duplication.

The system is now entirely decentralised, and for clarity, the interactions between the various subsystems within an individual node are graphically represented in Figure 2. To clarify the separation

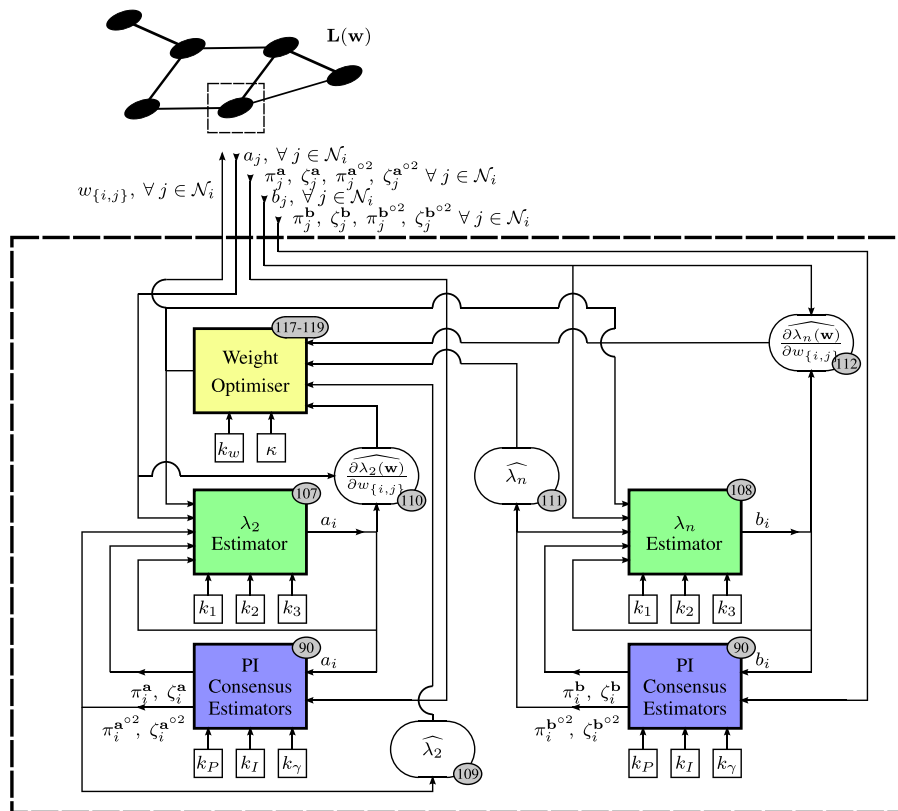


Figure 2. A simplified flowchart of information flow within a node. Each block is labelled with a number pointing to which equation in the paper the block represents. There are three timescales for the ordinary differential equations indicated by the colours: yellow (slow), green (faster) and blue (fastest). White rounded blocks are simply algebraic equations, and white squares signify control constants. [Colour figure can be viewed at wileyonlinelibrary.com]

in timescales between the different layers, we normalise the eigenvalue estimators by $\epsilon_1 = \frac{1}{k_2}$ and for the PI average consensus layers, by $\epsilon_2 = \frac{1}{k_P}$. We are interested in the behaviour of this system as $\epsilon_1 \rightarrow 0$ and $\frac{\epsilon_2}{\epsilon_1} \rightarrow 0$, when the control parameter ratios $\frac{k_1}{k_2}$, $\frac{k_3}{k_2}$, $\frac{k_\gamma}{k_P}$ and $\frac{k_I}{k_P}$ remain fixed. Again, the complete set of equations can be written in a more compact form, using the oriented and unoriented incidence matrices \mathbf{P} and \mathbf{M} :

$$\dot{\mathbf{w}} = -\frac{4k_3}{nk_2} \left(\frac{1}{2} \mathbf{M}\boldsymbol{\pi}^{\mathbf{b}^{\circ 2}} - \mathbf{1} \right) \circ (\mathbf{P}\mathbf{b})^{\circ 2} \circ (\mathbf{M}\boldsymbol{\pi}^{\mathbf{b}^{\circ 2}})^{\circ -1} + \frac{1}{4n} (\mathbf{P}\mathbf{a})^{\circ 2} (\mathbf{M}\boldsymbol{\nu})^{\circ 2} \circ (\mathbf{M}\boldsymbol{\pi}^{\mathbf{a}^{\circ 2}})^{\circ -1} + \frac{1}{2} \boldsymbol{\mu}^{\circ 2} \quad (121)$$

$$\dot{\boldsymbol{\mu}} = -\mathbf{w} \circ \boldsymbol{\mu} \quad (122)$$

$$\dot{\boldsymbol{\nu}} = \left(\kappa \mathbf{1} - \frac{k_3}{k_2} (\mathbf{1} - \boldsymbol{\pi}^{\mathbf{a}^{\circ 2}}) \right) \circ \boldsymbol{\nu} \quad (123)$$

$$\epsilon_1 \dot{\mathbf{a}} = -\frac{k_1}{k_2} \boldsymbol{\pi}^{\mathbf{a}} - \mathbf{L}(\mathbf{w})\mathbf{a} - \frac{k_3}{k_2} (\boldsymbol{\pi}^{\mathbf{a}^{\circ 2}} - \mathbf{1}) \circ \mathbf{a} \quad (124)$$

$$\epsilon_1 \dot{\mathbf{b}} = -\frac{k_1}{k_2} \boldsymbol{\pi}^{\mathbf{b}} + \mathbf{L}(\mathbf{w})\mathbf{b} - \frac{k_3}{k_2} (\boldsymbol{\pi}^{\mathbf{b}^{\circ 2}} - \mathbf{1}) \circ \mathbf{b} \quad (125)$$

$$\epsilon_2 \dot{\boldsymbol{\pi}}^{\mathbf{a}} = \left(-\mathbf{L}(\mathbf{w}) - \frac{k_\gamma}{k_P} \mathbf{I} \right) \boldsymbol{\pi}^{\mathbf{a}} - \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\zeta}^{\mathbf{a}} + \frac{k_\gamma}{k_P} \mathbf{a} \quad (126)$$

$$\epsilon_2 \dot{\boldsymbol{\zeta}}^{\mathbf{a}} = \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\pi}^{\mathbf{a}} \quad (127)$$

$$\epsilon_2 \dot{\boldsymbol{\pi}}^{\mathbf{a}^{\circ 2}} = \left(-\mathbf{L}(\mathbf{w}) - \frac{k_\gamma}{k_P} \mathbf{I} \right) \boldsymbol{\pi}^{\mathbf{a}^{\circ 2}} - \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\zeta}^{\mathbf{a}^{\circ 2}} + \frac{k_\gamma}{k_P} \mathbf{a}^{\circ 2} \quad (128)$$

$$\epsilon_2 \dot{\boldsymbol{\zeta}}^{\mathbf{a}^{\circ 2}} = \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\pi}^{\mathbf{a}^{\circ 2}} \quad (129)$$

$$\epsilon_2 \dot{\boldsymbol{\pi}}^{\mathbf{b}} = \left(-\mathbf{L}(\mathbf{w}) - \frac{k_\gamma}{k_P} \mathbf{I} \right) \boldsymbol{\pi}^{\mathbf{b}} - \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\zeta}^{\mathbf{b}} + \frac{k_\gamma}{k_P} \mathbf{b} \quad (130)$$

$$\epsilon_2 \dot{\boldsymbol{\zeta}}^{\mathbf{b}} = \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\pi}^{\mathbf{b}} \quad (131)$$

$$\epsilon_2 \dot{\boldsymbol{\pi}}^{\mathbf{b}^{\circ 2}} = \left(-\mathbf{L}(\mathbf{w}) - \frac{k_\gamma}{k_P} \mathbf{I} \right) \boldsymbol{\pi}^{\mathbf{b}^{\circ 2}} - \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\zeta}^{\mathbf{b}^{\circ 2}} + \frac{k_\gamma}{k_P} \mathbf{b}^{\circ 2} \quad (132)$$

$$\epsilon_2 \dot{\boldsymbol{\zeta}}^{\mathbf{b}^{\circ 2}} = \frac{k_I}{k_P} \mathbf{L}(\mathbf{w})\boldsymbol{\pi}^{\mathbf{b}^{\circ 2}} \quad (133)$$

The system exhibits a three timescale structure of the form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \\ \epsilon_1 \dot{\mathbf{y}}_1 &= \mathbf{g}_1(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \\ \epsilon_2 \dot{\mathbf{y}}_2 &= \mathbf{g}_2(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \end{aligned} \quad (134)$$

with column vectors $\mathbf{x} \triangleq [\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\nu}]$, $\mathbf{y}_1 \triangleq [\mathbf{a}, \mathbf{b}]$ and $\mathbf{y}_2 \triangleq [\boldsymbol{\pi}^a, \boldsymbol{\zeta}^a, \boldsymbol{\pi}^{a^{o2}}, \boldsymbol{\zeta}^{a^{o2}}, \boldsymbol{\pi}^b, \boldsymbol{\zeta}^b, \boldsymbol{\pi}^{b^{o2}}, \boldsymbol{\zeta}^{b^{o2}}]$.

Theorem 1 from Hoppensteadt [30] deals specifically with singularly perturbed systems of this type, and conditions are given under which the solution of system described by (134) converges to that of the reduced system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \\ \mathbf{0} &= \mathbf{g}_1(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \\ \mathbf{0} &= \mathbf{g}_2(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2) \end{aligned} \tag{135}$$

as $\epsilon_1 \rightarrow 0$ and $\frac{\epsilon_2}{\epsilon_1} \rightarrow 0$. We now proceed to go through these conditions and discuss under which assumptions they hold.

6.1. Transformation to place the stable equilibrium point at the origin

Condition 1 from [30] requires that there exists an equilibrium point for each of the fast subsystems, that this equilibrium point is located at the origin of systems faster than it and also that this equilibrium point is isolated. That is, it is required that

$$\mathbf{g}_1(\mathbf{x}, \mathbf{0}, \mathbf{0}) = \mathbf{0} \tag{136}$$

$$\mathbf{g}_2(\mathbf{x}, \mathbf{y}_1, \mathbf{0}) = \mathbf{0} \tag{137}$$

To achieve this condition, a transformation of variables is required. Observing the equilibrium point of the PI average consensus estimators ((93) and (94)), we define

$$\mathbf{Y}_2(\mathbf{x}, \mathbf{y}_1) = \begin{bmatrix} \boldsymbol{\pi}^{a,*}(\mathbf{y}_1) \\ \boldsymbol{\zeta}^{a,*}(\mathbf{x}, \mathbf{y}_1) \\ \boldsymbol{\pi}^{a^{o2},*}(\mathbf{y}_1) \\ \boldsymbol{\zeta}^{a^{o2},*}(\mathbf{x}, \mathbf{y}_1) \\ \boldsymbol{\pi}^{b,*}(\mathbf{y}_1) \\ \boldsymbol{\zeta}^{b,*}(\mathbf{x}, \mathbf{y}_1) \\ \boldsymbol{\pi}^{b^{o2},*}(\mathbf{y}_1) \\ \boldsymbol{\zeta}^{b^{o2},*}(\mathbf{x}, \mathbf{y}_1) \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{1}\mathbf{1}^\top \mathbf{a}}{n} \\ -\frac{k_\gamma}{k_I} \mathbf{L}^\dagger(\mathbf{w}) \mathbf{a} + \frac{\mathbf{1}\mathbf{1}^\top \boldsymbol{\zeta}^a(0)}{n} \\ \frac{\mathbf{1}\mathbf{a}^\top \mathbf{a}}{n} \\ -\frac{k_\gamma}{k_I} \mathbf{L}^\dagger(\mathbf{w}) \mathbf{a}^{o2} + \frac{\mathbf{1}\mathbf{1}^\top \boldsymbol{\zeta}^{a^{o2}}(0)}{n} \\ \frac{\mathbf{1}\mathbf{1}^\top \mathbf{b}}{n} \\ -\frac{k_\gamma}{k_I} \mathbf{L}^\dagger(\mathbf{w}) \mathbf{b} + \frac{\mathbf{1}\mathbf{1}^\top \boldsymbol{\zeta}^b(0)}{n} \\ \frac{\mathbf{1}\mathbf{b}^\top \mathbf{b}}{n} \\ -\frac{k_\gamma}{k_I} \mathbf{L}^\dagger(\mathbf{w}) \mathbf{b}^{o2} + \frac{\mathbf{1}\mathbf{1}^\top \boldsymbol{\zeta}^{b^{o2}}(0)}{n} \end{bmatrix} \tag{138}$$

Lemma 7

The equilibrium point of the PI average consensus estimators, $\mathbf{Y}_2(\mathbf{x}, \mathbf{y}_1)$, is an analytic function of \mathbf{x} and \mathbf{y}_1 .

Proof

It is evident that the equilibria of the proportional variables are analytic, as they are either linear or quadratic functions in \mathbf{y}_1 . For analyticity of the integrator equilibria, notice that the Moore–Penrose inverse of the graph Laplacian $\mathbf{L}^\dagger(\mathbf{w}) = (\mathbf{L}(\mathbf{w}) - \frac{\mathbf{1}\mathbf{1}^\top}{n})^{-1} - \frac{\mathbf{1}\mathbf{1}^\top}{n}$ has elements that are ratios of polynomials in \mathbf{w} , with non-zero denominator in the feasible set of \mathbf{w} : numerators arise from the cofactors of $\mathbf{L}(\mathbf{w}) - \frac{\mathbf{1}\mathbf{1}^\top}{n}$, where elements of $\mathbf{L}(\mathbf{w})$ are linear in \mathbf{w} , and the denominator is $\det\{\mathbf{L}(\mathbf{w}) - \frac{\mathbf{1}\mathbf{1}^\top}{n}\} \neq 0$ if $\lambda_2 > 0$. \square

Looking at (57) and (80), we define the equilibrium of the eigenvalue estimation systems

$$\mathbf{Y}_1(\mathbf{x}) = \begin{bmatrix} \mathbf{a}^*(\mathbf{x}) \\ \mathbf{b}^*(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} +\mathbf{v}_2(\mathbf{w}) \sqrt{\frac{n(k_3 - k_2 \lambda_2(\mathbf{w}))}{k_3}} \\ +\mathbf{v}_n(\mathbf{w}) \sqrt{\frac{n(k_3 + k_2 \lambda_n(\mathbf{w}))}{k_3}} \end{bmatrix} \tag{139}$$

which by Lemma 1 is a real analytic function in \mathbf{x} in the vicinity of $\mathbf{x} = \mathbf{w}_{opt}$ under Assumption 1, provided that $k_3 \neq 0$ and $k_3 > k_2 \lambda_2(\mathbf{w})$. Note that here, we are restricting our view to just

one equilibrium point of the four possible combinations of stable equilibria, in accordance with condition 1 of [30]. However, the same analysis could be performed on each of the distinct points, yielding the same result for each combination of equilibria.

The equilibrium point of the reduced system is also known and is the constant point

$$\mathbf{x}^* = \begin{bmatrix} \mathbf{w}_{opt} \\ \boldsymbol{\mu}^* = +\sqrt{2}\mathbf{q}^{\circ\frac{1}{2}} \\ \mathbf{v}^* = +\mathbf{1}\sqrt{2r} \end{bmatrix} \quad (140)$$

where \mathbf{q} and r are, respectively, the optimal dual variables corresponding to the inequalities $\mathbf{w} \geq 0$, and the optimal dual variable corresponding to the inequality constraint $\lambda_2(\mathbf{w}) \geq \kappa$, which solve the first-order necessary KKT conditions of the modified optimisation problem.

As such, we can apply the transformations:

$$\mathbf{F}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) = \mathbf{f}(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}), \tilde{\mathbf{y}}_2 + \mathbf{Y}_2(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}))) \quad (141)$$

$$\mathbf{G}_1(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) = \mathbf{g}_1(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}), \tilde{\mathbf{y}}_2 + \mathbf{Y}_2(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}))) \quad (142)$$

$$\begin{aligned} & -\epsilon_1 \left(\frac{\partial \mathbf{Y}_1}{\partial \mathbf{x}} \cdot \mathbf{F}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) \right) \\ \mathbf{G}_2(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) &= \mathbf{g}_2(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}), \tilde{\mathbf{y}}_2 + \mathbf{Y}_2(\mathbf{x}^* + \tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1 + \mathbf{Y}_1(\mathbf{x}^* + \tilde{\mathbf{x}}))) \quad (143) \\ & -\epsilon_2 \left(\frac{\partial \mathbf{Y}_2}{\partial \mathbf{x}} \cdot \mathbf{F}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) + \frac{\partial \mathbf{Y}_2}{\partial \mathbf{y}_1} \cdot \mathbf{G}_1(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) \right) \end{aligned}$$

So that

$$\mathbf{G}_1(\tilde{\mathbf{x}}, \mathbf{0}, \mathbf{0}) = \mathbf{0} \quad (144)$$

$$\mathbf{G}_2(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \mathbf{0}) = \mathbf{0} \quad (145)$$

and also that condition 2 in [30], which the equilibrium point of the slow system lies at the origin, is satisfied:

$$\mathbf{F}(\mathbf{0}, \mathbf{0}, \mathbf{0}) = \mathbf{0} \quad (146)$$

6.2. Analyticity of the right-hand sides

- Condition 3 in [30] requires that the functions $\mathbf{F}(\cdot)$, $\mathbf{G}_1(\cdot)$ and $\mathbf{G}_2(\cdot)$ and their derivatives with respect to the components of $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}_1$, $\tilde{\mathbf{y}}_2$ are continuous in a ball of radius R around the origin.
- Condition 4 is a requirement that the elements of $\mathbf{F}(\cdot)$, $\mathbf{G}_1(\cdot)$ and $\mathbf{G}_2(\cdot)$ and the elements of $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$, $\frac{\partial \mathbf{G}_i}{\partial \mathbf{x}}$, $\frac{\partial \mathbf{G}_i}{\partial \mathbf{y}_j}$ are bounded within the ball around the origin.
- Further to continuity and boundedness, conditions 5 and 6 specify uniform smoothness conditions on $\mathbf{F}(\cdot)$ and $\mathbf{G}_2(\cdot)$. Specifically, it is required that there is a continuous non-negative function $v(|\tilde{\mathbf{y}}|)$ with $v(0) = 0$, defining $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}_1^\top, \tilde{\mathbf{y}}_2^\top]^\top$ such that

$$|\mathbf{F}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) - \mathbf{F}(\tilde{\mathbf{x}}, \mathbf{0})| \leq v(|\tilde{\mathbf{y}}|) \quad (147)$$

Likewise, there exists a continuous non-negative function $u(|\tilde{\mathbf{y}}_2|)$ with $u(0) = 0$ such that

$$|\mathbf{G}_1(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) - \mathbf{G}_1(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \mathbf{0})| \leq u(|\tilde{\mathbf{y}}_2|) \quad (148)$$

All of these conditions are immediately satisfied if $\mathbf{F}(\cdot)$, $\mathbf{G}_1(\cdot)$ and $\mathbf{G}_2(\cdot)$ are analytic functions of $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}_1$, $\tilde{\mathbf{y}}_2$.

Lemma 8

$\mathbf{F}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2)$ is an analytic function in the vicinity of the equilibrium point, provided Assumption 1 holds.

Proof

$\mathbf{f}(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2)$ is an analytic function of its variables, provided $\pi_i^{a^{o2}} \neq -\pi_j^{a^{o2}}, \forall i, j$ and $\pi_i^{b^{o2}} \neq -\pi_j^{b^{o2}}, \forall i, j$. This is apparent as $\mathbf{f}(\cdot)$ is a sum of ratios of polynomials in its variables, and the denominators may only be zero if $\pi_i^{a^{o2}} = -\pi_j^{a^{o2}}$ or $\pi_i^{b^{o2}} = -\pi_j^{b^{o2}}$. At the equilibrium point, $\boldsymbol{\pi}^{a^{o2},*} = \frac{\mathbf{1}\mathbf{a}^{*\top}\mathbf{a}^*}{\mathbf{1}^{\top}\mathbf{a}^*}, \mathbf{a}^* \neq \mathbf{0}$, so $\pi_i^{a^{o2},*} = \pi_j^{a^{o2},*} \neq 0$. Likewise, for the spectral radius estimator $\pi_i^{b^{o2},*} = \pi_j^{b^{o2},*} \neq 0$

Further to this, both $\mathbf{Y}_1(\mathbf{x})$ and $\mathbf{Y}_2(\mathbf{x}, \mathbf{y}_1)$ are analytic functions in the vicinity of the equilibrium point under Assumption 1. Thus, $\mathbf{F}(\cdot)$, which is a composition of analytic functions, is analytic. \square

Lemma 9

$\mathbf{G}_1(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2)$ and $\mathbf{G}_2(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2)$ are analytic functions in the vicinity of the equilibrium point, provided Assumption 1 holds.

Proof

It can be seen that both $\mathbf{g}_1(\cdot)$ and $\mathbf{g}_2(\cdot)$ are quadratic functions of their variables and are thus analytic. Further to this, both stationary manifolds $\mathbf{Y}_1(\cdot)$ and $\mathbf{Y}_2(\cdot)$ are analytic provided Assumption 1 holds. Hence, the function $\mathbf{G}_1(\cdot)$, which is a composition of the analytic functions $\mathbf{G}_1(\cdot), \mathbf{Y}_1(\cdot), \mathbf{Y}_2(\cdot)$ and $\mathbf{F}(\cdot)$, is analytic. The same argument can be made for $\mathbf{G}_2(\cdot)$, which is a composition of the aforementioned analytic functions and $\mathbf{G}_1(\cdot)$. \square

6.3. Stability of the boundary layer systems and reduced system

The final two conditions in [30], conditions 7 and 8, concern the uniform asymptotic stability of the boundary layer systems and the reduced system. That is, it is required that the reduced system

$$\dot{\tilde{\mathbf{x}}} = \mathbf{F}(\tilde{\mathbf{x}}, \mathbf{0}, \mathbf{0}) \tag{149}$$

is uniformly asymptotically stable on the origin, as are

$$\frac{\partial \tilde{\mathbf{y}}_1}{\partial \tau_1} = \mathbf{G}_1(\mathbf{x}^*, \tilde{\mathbf{y}}_1, \mathbf{0}) \tag{150}$$

and

$$\frac{\partial \tilde{\mathbf{y}}_2}{\partial \tau_2} = \mathbf{G}_1(\mathbf{x}^*, \mathbf{y}_1^*, \tilde{\mathbf{y}}_2) \tag{151}$$

when taking \mathbf{x}^* and \mathbf{y}_1^* as any constant values (in a ball around the origin) and defining $\tau_i \triangleq \frac{t}{\epsilon_i}$.

The fully distributed system has been designed so that the reduced system (17) is locally exponentially stable (Theorem 1) implying uniform asymptotic stability. Likewise, each boundary layer system, found by taking slower variables as constant and faster variables as having already equilibrated, is simply the ideal estimators (55), (61) and (92) and is thus exponentially stable as proven in Theorem 2 and elsewhere [21, 36].

All the conditions from Theorem 1 in [30] hold, and so we may draw the conclusion that there exists a sufficiently small $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2]^\top$ such that $\epsilon_1 \rightarrow 0$ and $\frac{\epsilon_2}{\epsilon_1} \rightarrow 0$ as $|\boldsymbol{\epsilon}| \rightarrow 0$, so that the equilibrium point corresponding to the solution of the optimisation problem is locally exponentially stable. A suitable choice for $\boldsymbol{\epsilon}$ is $[\frac{1}{k_2}, \frac{1}{k_P}]^\top$, where $k_P = k_2^2$. There exists a $k_P > \frac{1}{\epsilon^*}$ sufficiently large so that the equilibrium point of the entire distributed system is locally exponentially stable.

7. NUMERICAL VALIDATION

In this example, we start with a small network of $n = 8$ nodes and $m = 11$ edges (Figure 3). Each edge is assigned an initial weight of one, $w_{\{i,j\}}(0) = 1$, so that the initial algebraic connectivity $\lambda_2(\mathbf{w}(0)) \approx 0.7611$ and initial spectral radius of $\lambda_n(\mathbf{w}(0)) \approx 5.8635$. Edge weights are then controlled according to the distributed multilayer system, with the objective of minimising λ_n , whilst

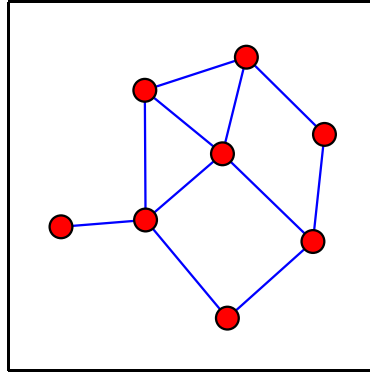
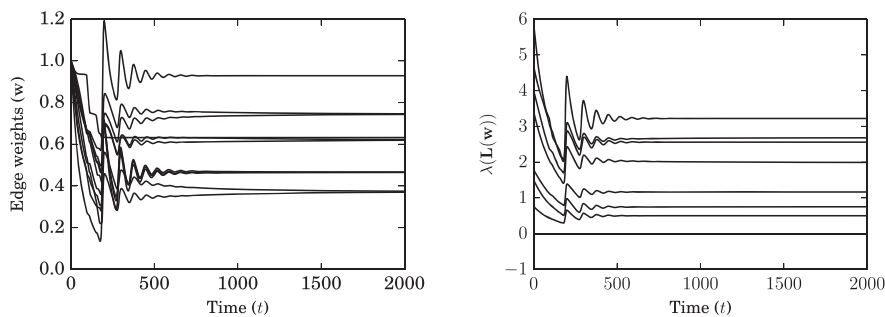


Figure 3. The initial network, at time zero, with all edge weights of $w_{\{i,j\}}(0) = 1$. [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Edge weights change over time from their initial value of $\mathbf{w}(0) = 1$, and converge onto the stationary point \mathbf{w}_{opt} . (b) As edge weights in the network change, the spectrum of the graph Laplacian also changes. As required by the optimisation problem, the algebraic connectivity maintains the bound $\lambda_2 \geq 0.5$, whilst the spectral radius λ_n is minimised.

Figure 4. Trajectories of the edge weights and the eigenvalues of the graph Laplacian.

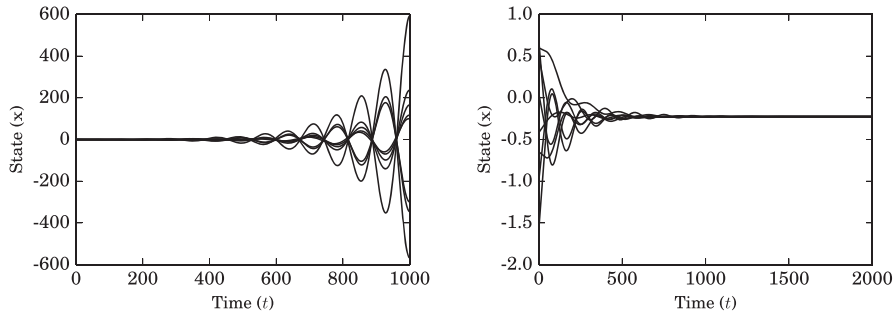
holding the constraints that all weights are non-negative and a minimum algebraic connectivity of $\lambda_2 \geq 0.5$ is maintained.

As the simulation progresses, edge weights evolve over time and eventually settle to a stationary value (Figure 4(a)). The edge weights in the network directly affect the graph Laplacian, resulting in the eigenvalues changing over time (Figure 4(b)) and for this particular example, eigenvalues at the solution $\mathbf{w} = \mathbf{w}_{opt}$ are distinct, so that edge weights converge as expected. At the end of the simulation, $t = 2000$, the algebraic connectivity has settled to its lower bound as expected $\lambda_2(\mathbf{w}(2000)) \approx 0.5000$, and the spectral radius has decreased to $\lambda_n(\mathbf{w}(2000)) \approx 3.2211$, shrinking to approximately 55% of its initial value. This result agrees with the optimal value found using a centralised SDP solver, with the relative error in the minimal spectral radius being less than 0.001%.

The simple linear consensus protocol $\dot{\mathbf{x}}(t) = -0.01 \times \mathbf{L}(\mathbf{w})\mathbf{x}(t - \tau)$ with homogeneous time delay $\tau = 40$ is run both on the initial and final states of the network (Figure 5). From the bound given in [1], we know that the system is only stable on the consensus mode if $\lambda_n \leq 100 \times \frac{\pi}{2 \times 40} \approx 3.9270$. We see that in the initial network (at time $t = 0$), the spectral radius exceeds this bound and as a consequence trajectories in Figure 5(a), diverge. When the same system is realised on the near-optimal network (Figure 5(b)), trajectories now converge as the consensus mode is stable.

7.1. Ramifications of Assumption 1 not holding

To illustrate the scenario where Assumption 1 does not hold, we use a larger network of $n = 50$, $m = 118$, so that there are $n - 1 = 49$ non-trivial eigenvalues in the connected graph, see Figure 6.



(a) The spectral radius of the initial network exceeds the calculated bound $\lambda_n \approx 5.8635 > 3.9270$, and so the consensus mode is unstable; trajectories diverge. (b) At time $t = 2000$, edge weights in the network have converged onto values reducing the spectral radius of the graph Laplacian below the bound $\lambda_n \approx 3.2211 < 3.9270$; consensus is stable.

Figure 5. The simple linear consensus protocol with homogeneous delay of $\tau = 40$ is run, both on the initial network with homogenous weights $\mathbf{L}(\mathbf{w}(0) = \mathbf{1})$ and on the final state of the network $\mathbf{L}(\mathbf{w}(2000) \approx \mathbf{w}_{opt})$.

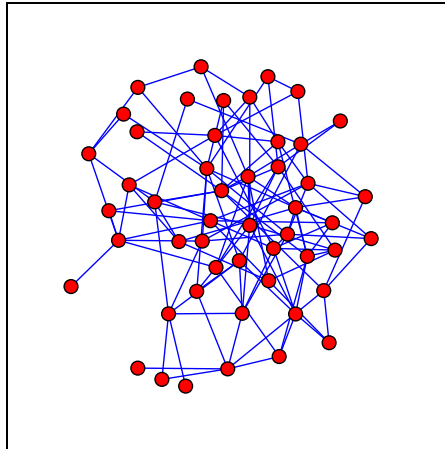


Figure 6. The larger network, $n = 50, m = 118$, at time zero, with edge weights each of $w_{\{i,j\}}(0) = 1$. [Colour figure can be viewed at wileyonlinelibrary.com]

As $\lambda_n(\mathbf{w})$ is decreased, the spectrum of the graph Laplacian is compressed into a smaller region, increasing the likelihood that at the optimal edge weights $\lambda_2(\mathbf{w})$ and $\lambda_n(\mathbf{w})$ will not be distinct (Figure 7(b)).

In this example, the spectral radius of the graph Laplacian is again decreased over time, from an initial value of $\lambda_n(\mathbf{w}(0)) \approx 11.1450$ to $\lambda_n(\mathbf{w}(4000)) \approx 5.2514$, whilst the algebraic connectivity falls from $\lambda_2(\mathbf{w}(0)) \approx 0.6589$ to $\lambda_2(\mathbf{w}(4000)) \approx 0.4988$. Note that this time at $t = 4000$, the algebraic connectivity lies slightly outside of its constraint. This is a result of Assumption 1 failing. At the optimal edge weights \mathbf{w}_{opt} , the algebraic connectivity has a multiplicity greater than unity. Therefore, the algebraic connectivity estimator subsystem does not converge to a specific eigenvector associated with λ_2 but converges instead onto its associated eigenspace. The effect of this is that the algebraic connectivity derivative estimates, $\frac{\partial \lambda_2}{\partial \mathbf{w}}$, will be incorrect as the value is not well defined for non-distinct eigenvalues. This error forces the edge weights off of the optimal values, and a persistent oscillation in the neighbourhood of the optimum is set up. The magnitude and frequency of these oscillations can be controlled by the size of the separation in timescales between the weight dynamics and the eigenvalue estimation, as this controls how far the edge weights can overshoot before the eigenvalue estimator re-converges on the correct value for $\frac{\partial \lambda_2}{\partial \mathbf{w}}$. We would however like to compare the values of the extremal eigenvalues with those found using a centralised solver, specifically using an SDP formulation presented in [15]. The centralised solver finds optimal eigenvalues of $\lambda_{2,opt} \approx 0.5000$ and $\lambda_{n,opt} \approx 5.2178$. Comparing these values with the extremal eigenvalues at

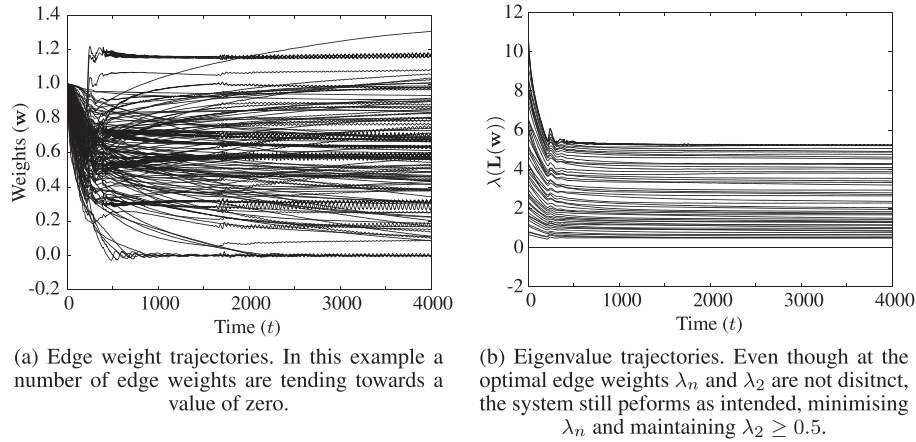


Figure 7. Edge weight trajectories and eigenvalue trajectories.

$t = 4000$, of $\lambda_2(\mathbf{w}(4000)) \approx 0.4988$ and $\lambda_n(\mathbf{w}(4000)) \approx 5.2514$, we can see that the decentralised method performs admirably despite Assumption 1 not holding, finding edge weights that result in a cost just 0.64% greater than those found using a centralised solver. In contrast, the decentralised method presented in [25] can only adapt edge weights up to the point when the extremal eigenvalues become indistinct and the algorithm breaks down, resulting in a greater disparity between the solution to the optimisation problem and the result from the decentralised method. It is foreseen that this issue of non-distinct extremal eigenvalues will grow with increasing network size.

8. CONCLUSION

In this paper, we have formulated a multilayer continuous control law for optimising edge weights in an undirected network, so as to minimise the spectral radius of the weighted graph Laplacian matrix, whilst maintaining a minimum bound on the algebraic connectivity, and enforcing the constraint that no edge weight may be negative. Moreover, this control law is completely distributed over the network; both edges and nodes require only local variables that are stored in their neighbours local memory. A notable application of this optimisation is that the robustness of the network to time delays in the simple linear consensus protocol is increased, and the constraint on the lower bound of the algebraic connectivity maintains connectivity in the network, so that consensus can be achieved. Moreover, the algebraic connectivity is strongly related to the performance of consensus protocols on networks. It should also be noticed that minimising λ_n whilst maintaining a minimum λ_2 is equivalent to minimising the synchronisability ratio $\frac{\lambda_n}{\lambda_2}$ whilst maintaining a desired algebraic connectivity, and so the method presented here may be readily adapted for a number of synchronisation applications.

Local exponential convergence has been proven for sufficiently large timescale separation between the layers, under the assumption that both $\lambda_2(\mathbf{w}_{opt})$ and $\lambda_n(\mathbf{w}_{opt})$ are distinct at the optimal edge weights. This assumption is relatively strong, and for large networks, with $n - 1$ non-trivial eigenvalues, in which the extremal eigenvalues are being driven towards each other, this assumption will not necessarily hold. When the assumption does not hold, the system is still shown to perform well, but edge weights do not converge as in the distinct eigenvalue case. Instead, small oscillations about the optimal values persist, and a limit cycle is set up. This effect is due to the extremal eigenvalues crossing and discontinuities in the eigenvectors associated with the extremal eigenvalues: as the spectral radius is decreased, the trajectories of λ_n and λ_{n-1} may cross over, and in the time it takes for the eigenvalue estimation to re-converge on the new eigenvector associated with λ_n , the weight dynamics, using incorrect estimates, will overshoot, setting up the persistent oscillation.

Future work will focus on relaxing Assumption 1, so that convergence can be guaranteed when the extremal eigenvalues $\lambda_2(\mathbf{w}_{opt})$ and $\lambda_n(\mathbf{w}_{opt})$ are not distinct. Also, a constructive method shall be sought for determining a sufficient timescale separation between the layers.

ACKNOWLEDGEMENT

This work was supported by the Engineering and Physical Sciences Research Council (1249225).

REFERENCES

1. Olfati-Saber R, Murray RM. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control* 2004; **49**(9):1520–1533.
2. Olfati-Saber R, Fax A, Murray RM. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE* 2007; **95**(1):215–233.
3. Wen G, Duan Z, Yu W, Chen G. Consensus in multi-agent systems with communication constraints. *International Journal of Robust and Nonlinear Control* 2012; **22**(2):170–182.
4. Wan Y, Namuduri K, Akula S, Varanasi M. The impact of multi-group multi-layer network structure on the performance of distributed consensus building strategies. *International Journal of Robust and Nonlinear Control* 2013; **23**(6):653–662.
5. Huang L, Chen Q, Lai Y-C, Pecora LM. Generic behavior of master-stability functions in coupled nonlinear dynamical systems. *Physical Review E* 2009; **80**(3):036204-1–036204-11.
6. Pecora LM, Carroll TL. Master stability functions for synchronized coupled systems. *Physical Review Letters* 1998; **80**(10):2109–2112.
7. Comellas F, Gago S. Synchronizability of complex networks. *Journal of Physics A: Mathematical and Theoretical* 2007; **40**(17):4483–4492.
8. Sundararaman B, Buy U, Kshemkalyani AD. Clock synchronization for wireless sensor networks: a survey. *Ad hoc networks* 2005; **3**(3):281–323.
9. He J, Cheng P, Shi L, Chen J. Sats: secure average-consensus-based time synchronization in wireless sensor networks. *IEEE Transactions on Signal Processing* 2013; **61**(24):6387–6400.
10. Dörfler F, Chertkov M, Bullo F. Synchronization in complex oscillator networks and smart grids. *Proceedings of the National Academy of Sciences* 2013; **110**(6):2005–2010.
11. Zhao C, He J, Cheng P, Chen J. Consensus-based energy management in smart grid with transmission losses and directed communication. *IEEE Transactions on Smart Grid*; **PP**(99):1–13.
12. De Gennaro MC, Jadbabaie A. Decentralized control of connectivity for multi-agent systems. *45th IEEE Conference on Decision and Control*, IEEE, San Diego, USA, 2006; 3628–3633.
13. Zavlanos MM, Pappas GJ. Distributed connectivity control of mobile networks. *IEEE Transactions on Robotics* 2008; **24**(6):1416–1428.
14. Fiedler M. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal* 1973; **23**(2):298–305.
15. Boyd S. Convex optimization of graph Laplacian eigenvalues. *Proceedings of the International Congress of Mathematicians* 2006; **3**:1311–1319.
16. Kim Y, Mesbahi M. On maximizing the second smallest eigenvalue of a state-dependent graph Laplacian. *IEEE Transactions on Automatic Control* 2006; **51**(1):116–120.
17. Simonetto A, Keviczky T, Babuška R. Constrained distributed algebraic connectivity maximization in robotic networks. *Automatica* 2013; **49**(5):1348–1357.
18. Zavlanos BMM, Egerstedt MB, Pappas GJ. Graph-theoretic connectivity control of mobile robot networks. *Proceedings of the IEEE* 2011; **99**(9):1525–1540.
19. Schuresko M, Cortés J. Distributed motion constraints for algebraic connectivity of robotic networks. *Journal of Intelligent and Robotic Systems* 2009; **56**(1-2):99–126.
20. Wang L, Wang X, Hu X. Connectivity maintenance and distributed tracking for double-integrator agents with bounded potential functions. *International Journal of Robust and Nonlinear Control* 2015; **25**(4):542–558.
21. Yang P, Freeman RA, Gordon GJ, Lynch KM, Srinivasa SS, Sukthankar R. Decentralized estimation and control of graph connectivity for mobile sensor networks. *Automatica* 2010; **46**(2):390–396.
22. Satici AC, Poonawala HA, Eckert H, Spong MW. Connectivity preserving formation control with collision avoidance for nonholonomic wheeled mobile robots. *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Tokyo, Japan, 2013; 5080–5086.
23. Kempton L, Herrmann G, di Bernardo M. Adaptive weight selection for optimal consensus performance. *53rd Annual Conference on Decision and Control*, IEEE, Los Angeles, USA, 2014; 2234–2239.
24. Qu Z, Li C, Lewis F. Cooperative control with distributed gain adaptation and connectivity estimation for directed networks. *International Journal of Robust and Nonlinear Control* 2014; **24**(3):450–476.
25. Bertrand A, Moonen M. Topology-aware distributed adaptation of Laplacian weights for in-network averaging. *21st European Signal Processing Conference (EUSIPCO 2013)*, IEEE, 2013; 1–5.
26. Donetti L, Neri F, Munoz MA. Optimal network topologies: expanders, cages, ramanujan graphs, entangled networks and all that. *Journal of Statistical Mechanics: Theory and Experiment* 2006; **2006**(08):P08007, 1–24.
27. Gorochoowski TE, di Bernardo M, Grierson CS. Evolving enhanced topologies for the synchronization of dynamical complex networks. *Physical Review E* 2010; **81**(5):056212, 1–15.
28. Liu T, Hill DJ, Zhao J. Synchronization of dynamical networks by network control. *IEEE Transactions on Automatic Control* 2012; **57**(6):1574–1580.

29. Jalili M, Rad AA. Rewirings based on the eigenvectors of the Laplacian matrix for enhancing synchronizability of dynamical networks. *7th IEEE International Conference on Industrial Informatics*, IEEE, 2009; 588–593.
30. Hoppensteadt F. On systems of ordinary differential equations with several parameters multiplying the derivatives. *Journal of Differential Equations* 1969; **5**(1):106–116.
31. Horn RA, Johnson CR. *Matrix Analysis*. Cambridge University Press, 2012.
32. De Abreu NMM. Old and new results on algebraic connectivity of graphs. *Linear algebra and its applications* 2007; **423**(1):53–73.
33. Mesbahi M, Egerstedt M. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.
34. Münz U, Papachristodoulou A, Allgöwer F. Delay robustness in consensus problems. *Automatica* 2010; **46**(8): 1252–1265.
35. Lewis AS. Convex analysis on the hermitian matrices. *SIAM Journal on Optimization* 1996; **6**(1):164–177.
36. Freeman RA, Yang P, Lynch KM. Stability and convergence properties of dynamic average consensus estimators. *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, 2006.
37. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press, 2004.
38. Lax PD. *Linear Algebra and Its Applications* (2nd edn). Wiley-Interscience: New York, 2007.
39. Bertsekas DP. *Nonlinear Programming*. Athena Scientific, 1999.
40. Guckenheimer J, Holmes P. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Vol. 42. Springer Science & Business Media, 2013.