Time scale modeling for consensus in sparse directed networks with time-varying topologies

Samuel Martin, Irinel-Constantin Morarescu, Dragan Nesic

▶ **To cite this version:**

Samuel Martin, Irinel-Constantin Morarescu, Dragan Nesic. Time scale modeling for consensus in sparse directed networks with time-varying topologies. 55th IEEE Conference on Decision and Control, CDC 2016, Dec 2016, Las Vegas, NV, United States. hal-01379076

HAL Id: hal-01379076
https://hal.science/hal-01379076

Submitted on 11 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Time scale modeling for consensus in sparse directed networks with time-varying topologies

Samuel Martin*, Irinel-Constantin Morărescu* , Dragan Nešić#

***Abstract*— The paper considers the consensus problem in large networks represented by time-varying directed graphs. A practical way of dealing with large-scale networks is to reduce their dimension by collapsing the states of nodes belonging to densely and intensively connected clusters into aggregate variables. It will be shown that under suitable conditions, the states of the agents in each cluster converge fast toward a local agreement. Local agreements correspond to aggregate variables which slowly converge to consensus. Existing results concerning the time-scale separation in large networks focus on fixed and undirected graphs. The aim of this work is to extend these results to the more general case of time-varying directed topologies. It is noteworthy that in the fixed and undirected graph case the average of the states in each cluster is time-invariant when neglecting the interactions between clusters. Therefore, they are good candidates for the aggregate variables. This is no longer possible here. Instead, we find suitable time-varying weights to compute the aggregate variables as time-invariant weighted averages of the states in each cluster. This allows to deal with the more challenging time-varying directed graph case. We end up with a singularly perturbed system which is analyzed by using the tools of two time-scales averaging which seem appropriate to this system.**

I. Introduction

Large dynamical systems of interacting units such as power grids, transportation networks or the brain can be modeled as multi-agent systems. Unfortunately, these models consist of very large number of coupled differential equations whose analysis becomes intractable. Therefore, it is crucial to reduce the complexity of the system by proposing appropriate reduced-order models that still capture the asymptotic behavior of the network. A natural approach is to merge a number of agents into a single node whose state approximates the states of the agents that generated it. This idea has been successfully applied for networks represented as fixed undirected graphs [1], [2], [3], [4], [5]. An extension of this idea to networks represented as fixed directed graphs was proposed in [6].

The goal of this paper is to provide a methodology to generate appropriate reduced systems for multi-agent systems with a large number of agents interacting through directed time-varying networks. A fundamental assumption is that the interaction network consist of sparsely connected clusters of densely connected agents [3], [4], [7]. An interesting feature of these networks is that the agents first converge fast toward a local agreement inside the clusters and then they slowly converge to a global consensus. This time scale separation has been emphasized and used for the analysis of power grids with fixed and undirected interconnections [3], [4], [5]. The same feature has been used in [8] for clusters detection in large scale networks. However, up to now, only symmetric and time-invariant communication networks have been dealt with.

In the case of fixed and undirected graph the average of the states in each cluster is time invariant when neglecting the interactions between clusters. Therefore, it is natural to use it as an aggregate state. When the interactions are directed and time varying we have to define appropriate aggregate variables. To do so, we introduce time-varying weights defining a time-invariant weighted average (using reasoning similar to [9]) as aggregate state of a cluster isolated from the other clusters. As a result we exhibit a time-scale separation in the network justifying the approximation of the overall network dynamics with the dynamics of the slow aggregate variables. It is important to note that unlike the fixed undirected interconnections graph [3], [4], [5], the resulting singularly perturbed system in this paper is not in standard form and we cannot use the standard tools [10], [11]. Instead it is more natural to employ averaging theory for two time-scales systems [12], [13], [14]. Note that these results are not restricted to periodic systems. The contribution of the paper can be summarized as follows:

- we provide the assumptions guaranteeing that **the convergence is faster inside clusters than between clusters**;
- under these assumptions, we show that **time-scale separation occurs and can be used for analysis of directed and time varying interconnections;**
- **we use averaging theory for two time-scales systems** which seems to be appropriate, under the stated conditions, for the analysis of networks structured in sparsely connected clusters of densely connected agents.

The rest of the paper is organized as follows. Section II introduces the model under study, the assumptions allowing to rewrite the model in singular perturbation form and presents the main contributions related to the singular perturbation modeling and analysis. In Section III we give some technical results concerning the two time-scales modeling. The singular perturbation analysis is provided in Section IV and an illustrative numerical examples is presented in Section V.

Notation. The following standard notation will be used throughout the paper. The set of non-negative integers, real and non-negative real numbers are denoted by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$, respectively. For a matrix A we denote by $\|A\|_\infty$ its infinity norm. The transpose of a matrix A is denoted by $A^\top$. By I_k we denote the $k \times k$ identity matrix. $\mathbf{1}_k$ and $\mathbf{0}_k$ are the column vectors of size k having all the components equal 1 and 0, respectively. A non

The work of S. Martin and I.-C. Morărescu was supported by project PEPS MoMIS MADRES and PEPS INS2I CONAS funded by the CNRS and project *Computation Aware Control Systems (COMPACS)*, ANR-13-BS03-004 funded by ANR. The work of D. Nešić was supported by the Australian Research Council under the Discovery Project DP1094326.

* Université de Lorraine, CRAN, UMR 7039, and CNRS, CRAN, UMR 7039, 2, Avenue de la Forêt de Haye, 54500, Vandœuvre, France. (*samuel.martin,constantin.morarescu*)*@univ-lorraine.fr*.

Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, 3010 VIC, Australia. *dnesic@unimelb.edu.au*

trivial subset S of a set C, denoted as $S \sqsubset C$, is a non-empty set with $S \subsetneq C$. A *directed path of length* p in a given directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is a union of directed edges $\bigcup_{k=1}^{p}(i_k, j_k)$ such that $i_{k+1} = j_k$, $\forall k \in \{1, \dots, p-1\}$. The node j is *connected* to node i in a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ if there exists at least a directed path in $\mathcal{G}$ from i to j (i.e. $i_1 = i$ and $j_p = j$).

II. Preliminaries

A. Model statement

Let $\mathcal{N} \triangleq \{1, \dots, n\}$ be a set of n agents. By abuse of notation we denote both the agent and its index by the same symbol $i \in \mathcal{N}$. Each agent is characterized by a scalar state $x_i \in \mathbb{R}, \forall i \in \mathcal{N}$ that evolves according to the following consensus system

$$\dot{x}_i(t) = \sum_{j=1}^{n} a_{ij}(t)(x_j(t) - x_i(t)), \ \forall i \in \mathcal{N}, \tag{1}$$

where $a_{ij}(t) \geq 0$ are twice continuously differentiable and uniformly upper-bounded functions of time representing the *interaction weights*. Let $x(t) = (x_1(t), \dots, x_n(t))^\top \in \mathbb{R}^n$ be the overall state of the network collecting the states of all the agents. Under these conditions, there exists a unique differentiable function of time $x : \mathbb{R}^+ \to \mathbb{R}^n$ whose components satisfy equation (1) for all $t \in \mathbb{R}^+$ [15], and it is bounded [16]. We call it the *trajectory* of the overall system. We say the trajectory asymptotically reaches a consensus when there exists a common agreement value $\alpha \in \mathbb{R}$ such that

$$\lim_{t \to +\infty} x_i(t) = \alpha, \ \forall i \in \mathcal{N}.$$

In the sequel, agents are assumed to be partitioned in m non-empty clusters: $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_m \subset \mathcal{N}$, that are assumed to be given or can be easily identified. For instance clusters may correspond to groups of agents which are spatially close to each other while different clusters are distant to each other. Let us introduce the following supplementary notation: $\mathcal{M} \triangleq \{1, \dots, m\}$ and n_i denotes the cardinality of cluster $\mathcal{C}_i$ and $\underline{n} = \min_{i \in \mathcal{M}} n_i$. Without loss of generality, we permute the agents' labels according to the partition so that when $j \in \mathcal{C}_i$ and $j' \in \mathcal{C}_{i+1}$, $j < j'$.

Denote $A(t) = (a_{ij}(t))$ the adjacency matrix of communication weights at time t, $D(t) = diag(d_{ii}(t))$ with $d_{ii}(t) = \sum_{j \in \mathcal{N}} a_{ij}(t)$, and $L(t) = D(t) - A(t)$ its associated Laplacian matrix. We denote as $A^I(t) = diag(A^1(t), \dots, A^m(t))$ the block-diagonal intra-cluster adjacency matrix where for $k \in \mathcal{M}$, $A^k(t) = (a_{ij}(t))_{(i,j) \in \mathcal{C}_k^2}$. We also denote by $L^k(t)$ the Laplacian matrix associated to $A^k(t)$ and $L^I(t) = diag(L^1(t), \dots, L^m(t))$ the block-diagonal Laplacian matrix associated to $A^I(t)$. The inter-cluster communications are then described by $A^E(t) = A(t) - A^I(t)$ and its associated Laplacian matrix $L^E(t) = L(t) - L^I(t)$. The inter-cluster adjacency matrix has zero elements $A^E_{ij}(t)$ if i and j belong to the same cluster.

B. Framework assumptions

In the following let us introduce some notation and the main hypotheses of this work. For two subsets of nodes $A, B \subset \mathcal{N}$, the sum of communication weights from B to A is denoted as

$$w_{A \leftarrow B}(t) = \sum_{i \in A, j \in B} a_{ij}(t).$$

Let us also recall some notation in [3], [4], [5] by adapting them to cope with non constant internal communications:

$$\begin{aligned} c^I(t) &= \min_{k \in \mathcal{M}, S \sqsubset \mathcal{C}_k} w_{S \leftarrow (\mathcal{C}_k \setminus S)}(t), \\ \gamma^E(t) &= \max_{k \in \mathcal{M}} w_{\mathcal{C}_k \leftarrow (\mathcal{N} \setminus \mathcal{C}_k)}(t), \\ \delta(t) &= \frac{\gamma^E(t)}{c^I(t)}. \end{aligned} \tag{2}$$

Assumption 1 (Intra-cluster communication): There exist constants $\underline{c} \geq 0, K_I \geq 1$ such that $c^I(t) \geq \underline{c}, \forall t \geq 0$ and for any cluster $k \in \mathcal{M}$ and for all non trivial subsets $S \sqsubset \mathcal{C}_k$,

$$w_{S \leftarrow (\mathcal{C}_k \setminus S)}(t) \leq K_I \cdot w_{(\mathcal{C}_k \setminus S) \leftarrow S}(t), \ \forall t \geq 0.$$

As proven in the sequel, Assumption 1 ensures exponential consensus inside each cluster seen as an independent network (see Lemma 1). This will justify the merging of all the nodes belonging to a cluster into one aggregate node.

Assumption 2 (Inter-cluster communication): There exist constants $K_E \geq 1$ and $\underline{a} \in (0, 1)$ such that for all non trivial subset $S \sqsubset \mathcal{M}$, for all $t \geq 0$, the two following equations hold :

$$\sum_{k \in S} w_{\mathcal{C}_k \leftarrow (\mathcal{N} \setminus \mathcal{C}_k)}(t) \leq K_E \cdot \sum_{k \in S} w_{(\mathcal{N} \setminus \mathcal{C}_k) \leftarrow \mathcal{C}_k}(t), \tag{3}$$

$$\sum_{k \in S} \sum_{h \in \mathcal{M} \setminus S} w_{\mathcal{C}_k \leftarrow \mathcal{C}_h}(t) \geq c^I(t) \varepsilon \underline{a}. \tag{4}$$

Assumption 2 ensures consensus of the aggregate state associated with the clusters, where equation (3) ensures reciprocity of the external interactions while equation (4) guarantees their persistence when Assumption 1 is satisfied (see *e.g.* [17] for definitions). Constant $\underline{a}$ can be chosen arbitrarily small. It is necessary that $K_I \geq 1$ and $K_E \geq 1$ and the equality corresponds to symmetric communication. The next assumption ensures that the total communication weight which a cluster $\mathcal{C}_k$ receives cannot exceed a proportion of the weight received by any non-trivial subset of $\mathcal{C}_k$ from the rest of $\mathcal{C}_k$.

Assumption 3: The ratio between internal and external interaction weights (defined in (2)) is bounded by a constant bound $\varepsilon > 0$: for all $t \geq 0$,

$$\delta(t) \leq \varepsilon.$$

The purpose of Assumption 3 is to ensure that the partition in clusters corresponds to the distribution of communication weights. It prevents cases where two subsets of a cluster are more connected to the outside than with each other. As it will be detailed in section III, Assumption 3 is necessary in order to exhibit a two time-scale separation of the system.

Remark 1: Assumption 3 guarantees that for each cluster $k \in \mathcal{M}$ and for all non trivial subsets $S \sqsubset \mathcal{C}_k$,

$$w_{\mathcal{C}_k \leftarrow (\mathcal{N} \setminus \mathcal{C}_k)}(t) \leq \varepsilon \cdot w_{S \leftarrow (\mathcal{C}_k \setminus S)}(t), \ \forall t \geq 0.$$

Remark 2: Note that Assumption 3 does not follow from Assumptions 1 and 2 hold. For a counter-example see [18].

Remark 3: For instances of classes of communication weights satisfying Assumptions 1 and 2, see [16]. Section V also provides an instance where Assumptions 1, 2 and 3 are all satisfied.

Using the formalism of communication in clusters, one has the following result.

Proposition 1: Supposing that Assumptions 1, 2 and 3 hold, the trajectory of system (1) reaches consensus exponentially fast.

Proof: As shown in [18, Theorem 1], Assumptions 1, 2 and 3 ensure the cut-balance condition in [16]. Moreover, (4) and the lower-bound imposed on $c^I(t)$ guarantee that there exists a strongly connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ such that

$$\int_0^t a_{ij}(s)ds \geq \varepsilon t\underline{a}\underline{c}, \quad \forall (j,i) \in \mathcal{E}$$

and following [17, Proposition 4] one obtains the exponential convergence toward consensus. ∎

Remark 4: Proposition 1 applies in the more general setting of integral equations with non smooth a_{ij} (see [16]) but in the present study, we use only smooth a_{ij} due to the singular perturbation analysis that is presented in the next section.

From Proposition 1, it is known that the trajectory of system (1) will converge to consensus. It remains to characterize this convergence. In particular, we want to show that convergence occurs faster within clusters than between clusters. Therefore aggregating the nodes inside clusters yields a smaller dimension model that approximates the overall dynamics. The analysis is carried out for general time-varying, non-symmetric communication weights.

C. Contributions

Firstly, we show that under Assumptions 1, 2 and 3 the system (1) rewrites in a singular perturbation form:

$$\begin{cases} \dfrac{d\hat{y}}{dt_f}(t_f) = \varepsilon A_{11}(t_f,\varepsilon)\hat{y}(t_f) + \varepsilon A_{12}(t_f,\varepsilon)\hat{z}(t_f), \\ \dfrac{d\hat{z}}{dt_f}(t_f) = \varepsilon A_{21}(t_f,\varepsilon)\hat{y}(t_f) + A_{22}(t_f,\varepsilon)\hat{z}(t_f), \end{cases} \tag{5}$$

where ε is a small parameter, t_f represents the fast time-scale that will be defined as well as all the matrices involved in the dynamics. The matrices $A_{ij}(t_f,\varepsilon)$ will be shown to be continuously differentiable in their arguments and the norms of matrices $A_{11}(t_f,\varepsilon)$, $A_{12}(t_f,\varepsilon)$ and $A_{21}(t_f,\varepsilon)$ will be proven to be bounded above by a constant while the norm of $A_{22}(t_f,\varepsilon)$ is bounded below away from 0 (see equation (18)). The bounds are uniform in time t_f and in ε. Roughly speaking, $\hat{y} \in \mathbb{R}^m$ and $\hat{z} \in \mathbb{R}^{n-m}$ describe the aggregate state of the clusters and the disagreement variables inside clusters, respectively. It is worth mentioning that we will emphasize the presence of a slow time-scale $t_s = \varepsilon t_f$ and show that $\hat{y}$ evolves slowly while $\hat{z}$ evolves fast.

Assumption 4: Suppose that the following limit exists and is independent of t_f

$$A_{av} = \lim_{T\to\infty} \frac{1}{T} \int_{t_f}^{t_f+T} A_{11}(s,0)ds.$$

We define the reduced (slow) system and boundary layer (fast) system:

$$\begin{cases} \dfrac{dy_s}{dt_s}(t_s) = A_{av} y_s(t_s); \\ \dfrac{dz_f}{dt_f}(t_f) = A_{22}\left(t_f, 0\right) z_f(t_f); \end{cases} \tag{6}$$

to approximate the solutions of system (5). The following approximation result is based on Theorem 1 in [12].

Theorem 1: Let $y_s(0) = y_{s,0} \in \mathbb{R}^m$ and $z_f(0) = z_{f,0} \in \mathbb{R}^{n-m}$ some fixed initial conditions to system (6). Set the same initial conditions to system (5), independently of ε. Then, under Assumptions 1, 2, 3 and 4, for any fixed $\varepsilon > 0$, system (5) possesses unique bounded solutions $\hat{y}(t_f,\varepsilon)$ and $\hat{z}(t_f,\varepsilon)$ defined on $\mathbb{R}^+$ (here, the dependence in ε is made explicit), and system (6) possesses unique bounded solutions y_s and z_f defined on $\mathbb{R}^+$. These solutions satisfy

$$\lim_{\varepsilon\to 0} \sup_{t_f\in\mathbb{R}+} \|\hat{y}(t_f,\varepsilon) - y_s(\varepsilon t_f)\| = 0,$$
$$\lim_{\varepsilon\to 0} \sup_{t_f\in\mathbb{R}+} \|\hat{z}(t_f,\varepsilon) - z_f(t_f)\| = 0.$$

Furthermore, for any fixed $\varepsilon > 0$, the trajectories $\hat{y}(t_f,\varepsilon)$ and $y_s(\varepsilon t_f)$ converge exponentially fast to consensus and the trajectories $\hat{z}(t_f,\varepsilon)$ and $z_f(t_f)$ converge exponentially fast to 0, when $t_f \to \infty$.

III. Two time-scale modeling

Using the matrix notation, system (1) can be represented as

$$\dot{x} = -L(t)x = -(L^I(t) + L^E(t))x. \tag{7}$$

A. Aggregate state

When trying to reduce the model and carry out time-scale separation between internal and external dynamics, difficulties arrise when the internal communication are neither constant nor symmetric as it was assumed in [3], [4], [5]. In the present section, we introduce some time-varying weights defining a time-invariant weighted average for stand alone clusters, which defines the aggregate state of the cluster.
Let us consider the isolated internal dynamics :

$$\dot{\tilde{x}}_{\mathcal{C}_k}(t) = -L^k(t)\tilde{x}_{\mathcal{C}_k}(t), \tag{8}$$

which corresponds to the internal dynamics $x_{\mathcal{C}_k}(t)$ when $L^E(t) = 0$, where $x_{\mathcal{C}_k}(t)$ denotes the vector collecting states $x_i(t)$ for $i \in \mathcal{C}_k$. We define, for all $t, \tau \geq 0$ with $t \leq \tau$, the fundamental matrix $\Phi^k(t,\tau)$ of system (8) such that

$$\tilde{x}_{\mathcal{C}_k}(\tau) = \Phi^k(t,\tau)\tilde{x}_{\mathcal{C}_k}(t). \tag{9}$$

As a direct consequence of Lemma 6 in [9], for all $t, \tau \geq 0$ with $\tau \geq t$, for any $i, j \in \mathcal{C}_k$, weight $\Phi^k_{ij}(t,\tau)$ is non-negative and

$$\sum_{j\in\mathcal{C}_k} \Phi^k_{ij}(t,\tau) = 1. \tag{10}$$

Lemma 1: Under Assumption 1 (intra-cluster communication) system (8) converges to consensus at exponential rate.

Lemma 1 is a direct consequence of Proposition 4 [17]. In the sequel, we assume that Assumption 2 is satisfied so that $\tilde{x}_{\mathcal{C}_k}$ converges to a consensus at exponential speed. We denote the limit matrix $\Phi^k(t,\infty) = \lim_{\tau\to\infty} \Phi^k(t,\tau)$, which exists thanks to Lemma 1. Moreover, since $\tilde{x}_{\mathcal{C}_k}$ converges to a consensus independently of the inital conditions, there exists $q^k(t) \in \mathbb{R}^{n_k}$ such that

$$\Phi^k(t,\infty) = \mathbf{1} \cdot (q^k(t))^T. \tag{11}$$

Vector $q^k(t)$ plays an important role in the rest of the study, in particular it serves to define an invariant for the isolated internal dynamics (8) as given in the next lemma.

Lemma 2: The vector $q^k(t)$ satisfies

$$q^k(t)^T \cdot \mathbf{1} = 1$$

and its components $q_i^k(t)$ are positive uniformly bounded *i.e.,* for all $t \geq 0$ and $i \in \mathcal{C}_k$,

$$q_i^k(t) \in [q_{min}, q_{max}] \in (0,1),$$

with

$$q_{min} = (\exp(-K_I)/\bar{n})^{\bar{n}-1}\,,\ q_{max} = 1-(\underline{n}-1)q_{min},$$

where $\underline{n} = \min_{k\in\{1,\dots,m\}} n_k$ and $\bar{n} = \max_{k\in\mathcal{M}} n_k$. Furthermore, quantity $q^k(t)^T \cdot \tilde{x}_{\mathcal{C}_k}(t)$ is invariant in time, *i.e.,*

$$\forall t \geq 0, q^k(t)^T \cdot \tilde{x}_{\mathcal{C}_k}(t) = q^k(0)^T \cdot \tilde{x}_{\mathcal{C}_k}(0).$$

Proof: See the Appendix in [19]. ■

Remark 5: Computing vector $q^k(t)$ is straightforward when the internal communication weights are constant up to a multiplicative factor, *i.e.,* $L^k(t) = l^k(t)L^k(0)$, with $l^k(t) \in \mathbb{R}$ (but not necessarily symmetric). In this case, q^k is constant, obtained as the left eigenvector of $L^k(0)$ associated to the eigenvalue 0 such that $(q^k)^T\mathbf{1}_{n_k} = 1$ (see for instance [20]). More generally, $q^k(t)$ can be computed whenever the internal dynamics is simple enough to be computed analytically. In other cases, the time-scale separation still occurs naturally as the present study shows and Lemma 2 provides bounds for the invariant $q^k(t)^T \cdot \tilde{x}_{\mathcal{C}_k}(t)$.

In the following we introduce variables $y(t) \in \mathbb{R}^m$ which corresponds to the cluster-aggregate state collecting weighted averages of the agents' states in each cluster:

$$y_k(t) = (q^k(t))^\top x_{\mathcal{C}_k}(t) \quad \text{or} \quad y(t) = J(t)x(t) \in \mathbb{R}^m,$$

with $x_{\mathcal{C}_k}$ the vector collecting the states x_i of the nodes $i \in \mathcal{C}_k$ and where $J(t)$ is the following block-diagonal matrix:

$$J(t) = diag\big((q^1(t))^\top,\ (q^2(t))^\top, \dots, (q^m(t))^\top\big) \in \mathbb{R}^{m\times n}. \tag{12}$$

The next lemma states that in each cluster k, the weight vector $q^k(t)$ is taken so that y is invariant through system (7) when no inter-cluster communication takes place. This justifies why $y(t)$ is a good candidate to represent the cluster-aggregate states.

Lemma 3: Suppose Assumption 3 holds. When ε (defined in Assumption 3) equals 0, no inter-cluster communication takes place, and in this case, y is invariant in time *i.e.,*

$$\varepsilon = 0 \Rightarrow L^E(t) = 0 \Rightarrow \dot{y}(t) = 0. \tag{13}$$

In this case, if Assumptions 1 (intra-cluster communication) is also satisfied, the trajectory $x_{\mathcal{C}_k}$ in each cluster k is identical to $\tilde{x}_{\mathcal{C}_k}$ and converges at exponential speed to a local consensus with consensus value $y_k(0)$.

Proof: The first part of statement (13) is a consequence of the second inequality in Assumption 2. The second part of statement (13) is a direct consequence of Lemma 2. The convergence to consensus comes from Lemma 1. Denote $x^*_{\mathcal{C}_k} \in \mathbb{R}$ the consensus value. Then, $\lim_{t\to\infty} x_{\mathcal{C}_k}(t) = x^*_{\mathcal{C}_k}\mathbf{1}$ and

$$y_k(0) = \lim_{t\to\infty} y_k(t) = x^*_{\mathcal{C}_k} \lim q^k(t)^T\mathbf{1}_{n_k} = x^*_{\mathcal{C}_k},$$

where we successively used equation (13), the definition of y_k and Lemma 2. ■

B. Disagreement variables

In the following, we provide a change of variable adapted to the clustered communication with non-symmetric and time-varying weights. First, In order to characterize the disagreement inside each cluster, let us introduce the variable $z(t) \in \mathbb{R}^{n-m}$ which represents for all $k \in \mathcal{M}$ the distances to the first agent in $\mathcal{C}_k$:

$$z(t) = Qx(t),$$

with $Q = diag(Q_1, Q_2, \dots, Q_m)$ where

$$Q_k = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{(n_k-1)\times n_k}.$$

The initial state variable $x(t)$ can be written in function of $y(t)$ and $z(t)$ as follows :

$$x(t) = Hy(t) + \tilde{Q}(t)z(t), \tag{14}$$

where $H = diag\big(\mathbf{1}_{n_1}, \mathbf{1}_{n_2}, \dots, \mathbf{1}_{n_m}\big) \in \mathbb{R}^{n\times m}$, $\tilde{Q}(t) = diag(\tilde{Q}_1(t), \tilde{Q}_2(t), \dots, \tilde{Q}_m(t))$ and $\tilde{Q}_k(t) \in \mathbb{R}^{n_k\times(n_k-1)}$ such that

$$\tilde{Q}_k(t) = \begin{bmatrix} -q_2^k(t) & -q_3^k(t) & \dots & -q_{n_k}^k(t) \\ 1-q_2^k(t) & -q_3^k(t) & \dots & -q_{n_k}^k(t) \\ -q_2^k(t) & 1-q_3^k(t) & \dots & -q_{n_k}^k(t) \\ \vdots & \vdots & \vdots & \vdots \\ -q_2^k(t) & -q_3^k(t) & \dots & 1-q_{n_k}^k(t) \end{bmatrix}.$$

Matrices $Q, \tilde{Q}(t)$ were chosen to verify the following lemma.

Lemma 4: The following properties hold true

- $Q_k\mathbf{1}_{n_k} = 0,\ (q^k(t))^\top\tilde{Q}_k(t) = 0,$
- $Q_k\tilde{Q}_k(t) = I_{n_k-1}$
- $\tilde{Q}_k(t)Q_k = I_{n_k} - \mathbf{1}_{n_k}(q^k(t))^\top$

Proof: The first property is obvious. The second comes from property $(q^k(t))^\top\mathbf{1}_{n_k} = 1$ meaning that $\sum_{i=1}^{n_k} q_i^k(t) = 1$ (see Lemma 2). For the third one, we compute each element of the first column of the product $\tilde{Q}_k(t)Q_k$. The first element is $\sum_{i=2}^{n_k} q_i^k(t)$ and is replaced by $1-q_1^k(t)$ while the rest of them are $\sum_{i=2}^{n_k} q_i^k(t) - 1$ which is replaced by $-q_1^k(t)$. ■

Remark 6: We note that definitions of $y(t)$ and $z(t)$ coincide with those used in [4] if we consider that the interactions are symmetric and constant i.e. $a_{ij}(t) = a_{ji}(t)$. We also note that, for constant symmetric interactions, the variables $y(t)$ and $z(t)$ can be introduced using another metric [3] such that $\tilde{Q} = Q^\top$.

Using equation (14), system (7) can be rewritten to make explicit the dynamics in terms of variables $y(t) = J(t)x(t)$ and $z(t) = Qx(t)$:

$$\begin{cases} \dot{y}(t) = \bar{A}_{11}(t)y(t) + \bar{A}_{12}(t)z(t), \\ \dot{z}(t) = \bar{A}_{21}(t)y(t) + \bar{A}_{22}(t)z(t), \end{cases} \tag{15}$$

where

$$\begin{cases} \bar{A}_{11}(t) = -J(t)L(t)H, \quad \bar{A}_{12}(t) = -J(t)L(t)\tilde{Q}(t), \\ \bar{A}_{21}(t) = -QL(t)H, \quad \bar{A}_{22}(t) = -QL(t)\tilde{Q}(t). \end{cases}$$

C. Time-scale separation

The next lemma shows that the aggregate variable y and the intra-cluster disagreement variable z evolve according to different time-scale, parametrized by factor ε.

Lemma 5: The ∞-norms of the matrices in (15) satisfy

$$\begin{array}{rcll} \|\bar{A}_{11}(t)\|_\infty &=& \|J(t)L(t)H\|_\infty & \leq 2c^I(t)\varepsilon, \\ \|\bar{A}_{12}(t)\|_\infty &=& \|J(t)L(t)\tilde{Q}(t)\|_\infty & \leq 2c^I(t)\varepsilon, \\ \|\bar{A}_{21}(t)\|_\infty &=& \|QL(t)H\|_\infty & \leq 2c^I(t)\varepsilon, \\ \|\bar{A}_{22}(t)\|_\infty &=& \|QL(t)\tilde{Q}(t)\|_\infty & \geq (1-8\varepsilon)c^I(t). \end{array} \tag{16}$$

The proof of the lemma is given in the Appendix in [19].

The only matrix in system (15) which is not $\mathcal{O}(c^I(t)\varepsilon)$ is $\bar{A}_{22}$ which corresponds to the influence of z on its dynamics. As a consequence, Lemma 5 shows that variables y behave as slow variables compared to the z variables. To reveal this fact, we rescale the time with a fast time scale $t_f = \int_0^t c^I(s)ds$ and a slow time scale $t_s = \varepsilon t_f$. This rescaling is well defined since according to Assumption 1, t_f diverges when $t \to \infty$. The transformation $t \mapsto t_f$ is invertible and so is $t \mapsto t_s$. We denote $t = \psi(t_f)$ the reverse transformation. Then $y(t) = y(\psi(t_f)) \triangleq \hat{y}(t_f)$, $z(t) = z(\psi(t_f)) \triangleq \hat{z}(t_f)$. Matrices $\bar{A}_{ij}$ are also rescaled as

$$\left\{ \begin{array}{ll} A_{11}(t_f,\varepsilon) = \frac{\bar{A}_{11}(\psi(t_f))}{c^I(\psi(t_f))\cdot\varepsilon}, & A_{12}(t_f,\varepsilon) = \frac{\bar{A}_{12}(\psi(t_f))}{c^I(\psi(t_f))\cdot\varepsilon}, \\ A_{21}(t_f,\varepsilon) = \frac{\bar{A}_{21}(\psi(t_f))}{c^I(\psi(t_f))\cdot\varepsilon}, & A_{22}(t_f,\varepsilon) = \frac{\bar{A}_{22}(\psi(t_f))}{c^I(\psi(t_f))}. \end{array} \right.$$

Thanks to Assumption 3, the rescaling is well defined. At this point, the dependance in ε has been made explicit to stress that the singular perturbation approximation corresponds to set $\varepsilon = 0$ (see below equation (19)).

Without entering in details, writing the system (15) in the slow time-scale leads to non-standard singularly perturbed system because of the appearance of the term $1/\varepsilon$ inside the matrix arguments. Thus singular perturbation analysis in [10] cannot be applied. Nevertheless, using averaging theory [12], [14] for singularly perturbed systems we can approximate the behavior of system (15) using the fast-time scale dynamics.

Lemma 6: In the fast time-scale system (15) is described by

$$\left\{ \begin{array}{l} \frac{d\hat{y}}{dt_f}(t_f) = \varepsilon A_{11}(t_f,\varepsilon)\hat{y}(t_f) + \varepsilon A_{12}(t_f,\varepsilon)\hat{z}(t_f), \\ \frac{d\hat{z}}{dt_f}(t_f) = \varepsilon A_{21}(t_f,\varepsilon)\hat{y}(t_f) + A_{22}(t_f,\varepsilon)\hat{z}(t_f), \end{array} \right. \tag{17}$$

where the ∞-norms of the matrices in (5) satisfy

$$\begin{array}{ll} \|A_{11}(t_f,\varepsilon)\|_\infty \leq 2, & \|A_{12}(t_f,\varepsilon)\|_\infty \leq 2, \\ \|A_{21}(t_f,\varepsilon)\|_\infty \leq 2, & \|A_{22}(t_f,\varepsilon)\|_\infty \geq (1-8\varepsilon). \end{array} \tag{18}$$

Proof: Using equation (15),

$$\begin{array}{rcl} \frac{d\hat{y}}{dt_f}(t_f) &=& \frac{d(y\circ\psi)}{dt_f}(t_f) = \frac{d\psi}{dt_f}(t_f)\frac{dy}{dt}(\psi(t_f)) \\ &=& \frac{(\bar{A}_{11}(\psi(t_f))y(\psi(t_f)) + \bar{A}_{12}(\psi(t_f))z(\psi(t_f)))}{c^I(\psi(t_f))}. \end{array}$$

Moreover,

$$\bar{A}_{11}(\psi(t_f))/c^I(\psi(t_f)) = \varepsilon A_{11}(t_f,\varepsilon),$$

and analogous equalities hold for $\bar{A}_{12}$ and $\bar{A}_{22}$ while

$$\bar{A}_{22}(\psi(t_f))/c^I(\psi(t_f)) = A_{22}(t_f,\varepsilon).$$

The bounds on the ∞-norms directly come from Lemma 5. ■

IV. Singular perturbation analysis

The dynamics given in Lemma 6 fits in the generic form of systems treated in [12], [14]. Let z_f be the solution of

$$\frac{dz_f}{dt_f}(t_f) = A_{22}(t_f,0)\, z_f(t_f). \tag{19}$$

Note that for $\varepsilon = 0$ the matrix A_{22} and corresponds to intra clusters communication only and $z_f(t_f) = Q\tilde{x}(\psi(t_f))$. So, as given in Lemma 3, system (19) converges to 0 exponentially fast. Denote $A_{11}(s,0) = \lim_{\varepsilon\to 0} A_{11}(s,\varepsilon)$, which exists thanks to equation 18.

Let A_{av} be defined in Assumption 4 and y_s be the solution of

$$\frac{dy_s}{dt_s}(t_s) = A_{av}y_s(t_s). \tag{20}$$

Proposition 2: Denote $y_{s,k}$ for $k \in \mathcal{M}$ the k-th component of y_s. Under Assumption 4, system (20) is a consensus system described element-wise by

$$\frac{dy_{s,k}}{dt_s}(t_s) = \sum_{j\in\mathcal{M}} a^s_{kl}(y_{s,l}(t_s) - y_{s,k}(t_s)), k \in \mathcal{M}, \tag{21}$$

where $a^s_{kl} = \lim_{T\to\infty} \frac{1}{T}\int_{t_f}^{t_f+T} \sum_{j\in\mathcal{C}_l}\sum_{i\in\mathcal{C}_k} \frac{q^k_i(\psi(s))a_{ij}(\psi(s))}{c^I(\psi(s))\varepsilon}$ is independent of t_f. Moreover, under Assumption 2 (Inter-cluster communication) y_s reaches consensus exponentially fast.

Proof: See [19].

■

We are now ready to prove the main result of the study.

Proof of Theorem 1: We want to apply Theorem 1 in [12]. Notice that system (5) rewrites as

$$\left\{ \begin{array}{l} \frac{d\hat{y}}{dt_f}(t_f) = \varepsilon f(t_f,t_s,\hat{y},\hat{z},\varepsilon), \\ \frac{d\hat{z}}{dt_f}(t_f) = g(t_f,t_s,\hat{y},\hat{z},\varepsilon), \end{array} \right. \tag{22}$$

where, following the notation in [12],

$$\begin{array}{l} f(t_f,t_s,\hat{y},\hat{z},\varepsilon) = A_{11}(t_f,\varepsilon)\hat{y}(t_f) + A_{12}(t_f,\varepsilon)\hat{z}(t_f), \\ g(t_f,t_s,\hat{y},\hat{z},\varepsilon) = \varepsilon A_{21}(t_f,\varepsilon)\hat{y}(t_f) + A_{22}(t_f,\varepsilon)\hat{z}(t_f) \end{array} \tag{23}$$

It is worth noting that in our case f and g do not depend explicitly on the slow time t_s. Equation $\dot{z}_f = g(t_f,\tau,y,z_f,0)$ where τ and y are regarded as parameters corresponds to equation (19). According to Lemma 3, z_f converges at exponential speed to 0 so that Hypothesis H1 in [12] is satisfied. Moreover, f_0 in H2 in [12] is

$$\begin{array}{rcl} f_0(\tau,y) &=& \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T} f(s,\tau,y,z_f(s),0)ds \\ &=& \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T} A_{11}(s,0)y + A_{12}(s,0)z_f(s)ds \\ &=& \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T} A_{11}(s,0)ds\cdot y = A_{av}\cdot y \end{array}$$

where we used the exponential decrease of z_f and the bound on $\|A_{12}(s,0)\|_\infty$ in equation (18) to remove the second term under the integral. Assumption 4 shows that H2 and equation (2.4) in [12] are satisfied. The averaged equation (2.5) in [12] corresponds to equation (20), whose trajectory

are bounded and converge to consensus at exponential speed thanks to Proposition 2. As a consequence, Hypothesis H3 and H4 in [12] are satisfied. Finally, H5 is also satisfied because $g(t_f, \tau, y, z_f, 0)$ is independent of y. Under Hypotheses H1 through H5, Theorem 1 in [12] provides the expected result.

V. Numerical illustrations

Consider the multi-agents system (1) with 8 agents and two clusters whose communication pattern is defined in Figure 1. When not null, internal communication weights are set to $a_{ij}(t) = 2+\cos(2t)$ in cluster 1, to $a_{ij}(t) = 1$ in cluster 2 and external communication weights are set to $a_{ij}(t) = \varepsilon(sin(t)+2)/3$ where ε is a constant parameter. As a consequence, Assumption 1, 2, 3 and 4 are all satisfied. In the present case, internal weights vary in time all with the same multiplicative factor so that $q^k(t)$ is time-invariant and correspond to the left eigenvector of $L^k(0)$ associated to eigenvalue 0 such that $q^k \mathbf{1}_{n_k} = 1$ (see Remark 5 for the computation of q^k in the general case). We simulate the system with initial conditions $x(0) = (6, 6.3, 4.4, 5.2, 3, 3.5, 0.4, 2.2)^T$ (see Figure 2). As expected from Proposition 2 and Lemma 3, all aggregate variables converge to consensus and disagreement variables converge to 0, even when ε is not small. When $\varepsilon = 1$ the approximations are not precise, but they approch the real variables uniformly as ε diminishes, as expected from Theorem 1.

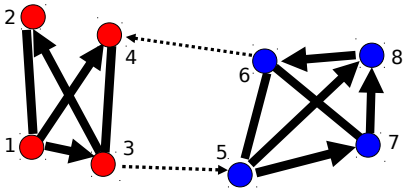

Fig. 1. Communication graph. Red and blue circles are agents in cluster 1 and 2 respectively. Bold and dashed thin arrows correspond to communication inside and between clusters, respectively. Arrows are replaced by lines when the communication is bidirectional. Only communications with non-zero weight are displayed.

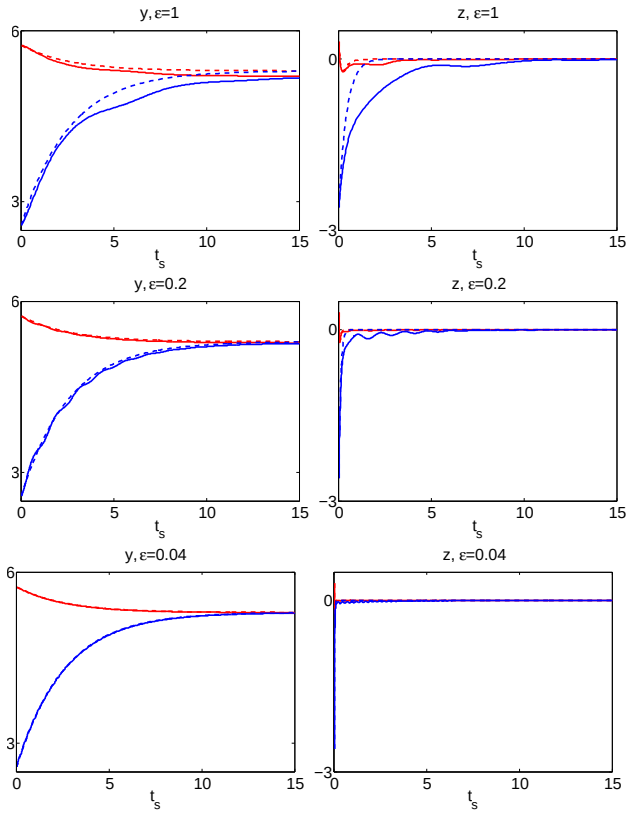

Fig. 2. Trajectories in the slow time scale t_s of the aggregate (left) and disagreement variables (right) along with their approximations, for $\varepsilon \in \{1, 0.2, 0.04\}$ from top to bottom. Strait lines correspond to the original variables while dashed lines are the approximations. Red lines are for cluster 1 and blue lines for cluster 2. Only the first disagreement variable of each cluster is displayed (the two others present a similar behavior).

VI. Conclusion

In the present study, we have extended the singular perturbation analysis for linear consensus systems found in [4] to time-varying and directed networks. We have emphasized that the right tool to study the network at hand is averaging theory for two time-scales systems.

References

[1] I. Hanski, "Metapopulation dynamics," *Nature*, vol. 396, pp. 41–49, 1998.

[2] D. Gfeller and P. D. L. Rios, "Spectral coarse graining and synchronization in oscillator networks," *Phys. Rev. Lett.*, vol. 100, p. 174104, 2008.

[3] E. Bıyık and M. Arcak, "Area aggregation and time-scale modeling for sparse nonlinear networks," *Systems & Control Letters*, vol. 57, no. 2, pp. 142–149, 2007.

[4] J. Chow and P. Kokotović, "Time scale modeling of sparse dynamic networks," *IEEE Transactions on Automatic Control*, vol. 30, no. 8, pp. 714–722, 1985.

[5] D. Romeres, f. Dörfler, and F. Bullo, "Novel results on slow coherency in consensus and power networks," in *European Control Conference, Zürich, Switzerland*, 2013.

[6] I.-C. Morărescu and R. Postoyan, "Dimension reduction for large-scale networked systems," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012.

[7] I.-C. Morărescu, S. Martin, A. Girard, and A. Muller-Gueudin, "Coordination in networks of linear impulsive agents," *IEEE Transactions on Automatic Control*, p. DOI: 10.1109/TAC.2015.2492058, 2016.

[8] I.-C. Morărescu and A. Girard, "Opinion dynamics with decaying confidence: Application to community detection in graphs," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1862 – 1873, 2011.

[9] S. Martin and J. Hendrickx, "Continuous-time consensus under non-instantaneous reciprocity," *IEEE Transactions on Automatic Control*, 2016.

[10] H. Khalil, *Nonlinear Systems (Third Edition)*. Prentice Hall, 2001.

[11] P. Kokotović, H. Khalil, and J. O'Reilly, *Singular perturbation methods in control: analysis and design*. SIAM Series in Classics and Applied Mathematics, 1999.

[12] M. Balachandra and P. R. Sethna, "A generalization of the method of averaging for systems with two time scales," *Archive for Rational Mechanics and Analysis*, vol. 58, no. 3, pp. 261–283, 1975.

[13] D. Nesic and A. Teel, "Input-to-state stability for nonlinear time-varying systems via averaging," *Mathematics of Control, Signals, and Systems*, vol. 14, pp. 257–280, 2001.

[14] A. Teel, L. Moreau, and D. Nesic, "A unified framework for input-to-state stability in systems with two time scales," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1526–1544, 2003.

[15] A. F. Filippov, *Differential equations with discontinuous righthand sides*. Dordrecht: Kluwer Academic Publishers Group, 1988, translated from the Russian.

[16] J. M. Hendrickx and J. N. Tsitsiklis, "Convergence of type-symmetric and cut-balanced consensus seeking systems." *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 214–218, 2013.

[17] S. Martin and A. Girard, "Continuous-time consensus under persistent connectivity and slow divergence of reciprocal interaction weights." *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 2568–2584, 2013.

[18] S. Martin, I.-C. Morărescu, and D.Nešić, "Cut-balance conditions for consensus in networks of clusters," *Automatica*, submitted 2016.

[19] S. Martin, I.-C. Morarescu, and D. Nesic, "Time scale modeling for consensus in sparse directed networks with time-varying topologies," *ArXiv e-prints*, Aug. 2016.

[20] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1520–1533, 2004.