Convergence Rates of Distributed Consensus over Cluster Networks: 
A Two-Time-Scale Approach

Amit Dutta, Almuatazbellah M. Boker, Thinh T. Doan*

Abstract—We study the popular distributed consensus method over networks composed of a number of densely connected clusters with a sparse connection between them. In these cluster networks, the method often constitutes two-time-scale dynamics, where the internal nodes within each cluster reach consensus quickly relative to the aggregate nodes across clusters. Our main contribution is to provide the rate of convergence of the distributed consensus method in a way that characterizes explicitly the impacts of the internal and external graphs on its performance. Our main result shows that the consensus method converges exponentially and only scales with a few number of nodes, which is relatively small to the size of the network. The key technique in our analysis is to consider a Lyapunov function that captures the impact of different time-scale dynamics on the convergence of the method. Our approach avoids using model reduction, which is the typical way according to singular perturbation theory and relies on relatively simple definitions of the slow and fast variables. We illustrate our theoretical results by a number of numerical simulations over different cluster networks.

Index Terms—consensus control; cluster networks; singular perturbation; two-time scale dynamics

I. INTRODUCTION

Cluster networks are omnipresent nowadays. They can be seen, for example, in small-world networks [1], wireless sensor networks [2] and power systems networks [3]. Clustering in this context refers to the network topology when there is dense communication inside the clusters and sparse communication among the different clusters. It was shown in [4] that for these types of networks, the dynamics evolve according to two-time scales. More specifically, the dynamics of the local interactions of the nodes within clusters evolve at faster time-scale relative to the aggregate dynamics of the clusters. This motivates exploiting this intrinsic feature of time scale separation to analyse the performance of these types of networks. This paper is a step in this direction, where our focus is to characterize the convergence rates of the popular distributed control methods over cluster networks, namely distributed consensus averaging.

Most of the results that exploit the two-time-scale phenomenon have been focused on model-reduction, for time-invariant linear dynamics in [4], for time-varying linear dynamics in [5] and for nonlinear dynamics in [6], and control design; see for example [7], [8], [9], [10], [11] and the references therein. The two-time-scale observation was also leveraged to solve different networked control systems problems. For example, the consensus control problem was solved in [7]. The leader-follower problem was solved in [12] in the context of discrete-time systems.

It is known that for (undirected) connected networks, all the nodes will reach consensus regardless of the network topology [13]. Furthermore, it is also known that the rate of change of convergence of all nodes to the consensus equilibrium is dependent on how well the network is connected. That is, the more connected the network, with more nodes communicating among each other, the faster the network reaches consensus. However, little attention has been devoted to investigating the inter-dependency of the convergence rates of the different dynamics when the network has a cluster structure. Depending on the application scenario, this can prove to be useful. For example, knowing how each cluster converges to the consensus equilibrium motivates local network design considerations. This, in turn, can lead to distributed feedback strategies to achieve local and global needs simultaneously. Other applications can include distributed learning which involves designing a local learning framework for each cluster which can be shared across the network periodically, thus learning different parts of a global problem at the same time.

Main contributions. In summary, our main contribution is to provide the convergence rate of the distributed consensus method over cluster networks, which is missing in the literature. Our main result shows that this rate converges exponentially and only scales with a few number of nodes, which is relatively small to the size of the network. The key technique in our analysis is to consider a Lyapunov function, which is fundamentally different from the one based on model reduction in the existing literature. Finally, we illustrate our theoretical results by a number of numerical simulations over different cluster networks.

Technical Approach. Inspired by singular perturbation theory [14] and recent analysis for the centralized two-time-scale stochastic approximation [15], [16], we analyze the convergence properties of cluster consensus networks in a way that respects time-scale separation. More specifically, we show the convergence rates of both the slow and fast dynamics through the use of a composite Lyapunov function. This function is scaled by a singular-perturbation parameter that represents the clustering strength and simultaneously captures the time separation. This approach is different than the typical singular perturbation approach reported in [14] and [6] in that it is not based on reducing the system model into two smaller models. This means that our approach does not require the system to be modeled in the standard singular perturbation form, which requires the fast subsystem...
Fig. 1: A 42-nodes network is partitioned into three densely connected clusters with sparse connections between them.

Novelty. Existing literature for general network shows that the convergence time depends on the number of nodes, which is much larger than the number of clusters. Hence our analysis provides a more compact result than the existing works. In particular, our analysis highlights how the nodes interact within their cluster versus their external communication with other nodes in other clusters. Knowing how the dynamics within and across clusters evolve provides an useful approach to design better distributed control strategies in different applications including robotics and power networks.

II. PROBLEM FORMULATION

We consider an averaging-consensus problem over a network of $N$ nodes with $M$ number of links where the states of the nodes $i = 1, 2, \ldots, N$ are represented by $x_i \in \mathbb{R}$. In this problem, each node $i$ initializes with a constant $c_i$ and their goal is to cooperatively compute the average $\bar{c}$ of their constants, i.e., $\bar{c} = \frac{1}{N} \sum_{i=1}^{N} c_i$.

We assume that the nodes can communicate with each other, where the communication structure is described by an undirected and connected graph $G = (V, E)$, where $V = \{1, \ldots, N\}$ and $E = V \times V$ are the vertex and edge sets, respectively. To find $\bar{c}$, distributed averaging-consensus methods have been studied extensively in the literature [17]. In particular, each node $i$ maintains a variable $x_i$ and initializes at its constant $c_i$, i.e., $x_i(0) = c_i$.

A standard way to characterize the consensus dynamics for undirected graph is by knowing if any of two nodes are connected. For this purpose, we introduce the graph Laplacian matrix $L \in \mathbb{R}^{N \times N}$, whose $ij$-th entry is

$$\ell_{ij} = \begin{cases} \text{deg}(i) & \text{if } i = j, \\ -1 & \text{if } i \neq j, \\ 0 & \text{otherwise}, \end{cases}$$

where $\text{deg}(i)$ is the degree of $i$th which represents the total number of edges that are incident on the node. The entries are zero if nodes are not connected. This implies that $L = DD^T$.

Considering $x = (x_1, x_2, \ldots, x_N)^T$, we have

$$\dot{x} = -Lx.$$  \hspace{1cm} (2)

With this dynamics one can show that from [18] $\lim_{t \to \infty} x_i(t) = \bar{c}, \text{ } \forall i \in V$.

Note that $L$ is a symmetric matrix. Moreover, when $G$ is connected, the sum of every row of $L$ is equal to zero. We also know from [13] that $L$ is positive semidefinite; hence its real eigenvalues can be ordered as $\sigma_1(L) \leq \sigma_2(L) \leq \ldots \leq \sigma_N(L)$, with $\sigma_1(L) = 0$, where $\sigma_i(L)$ is the $i$th eigenvalue of $L$. In this paper, our focus is to study the performance of the distributed consensus-averaging method when the communication graph $G$ has a cluster structure which exhibits two-time-scale dynamics. In the next section we model the cluster network by characterizing the slow, fast and inter-cluster variables which play a significant role in the main analysis of this paper.

III. MODELING OF THE CLUSTER NETWORK DYNAMICS

In this section, we characterize the network properties and define the slow and fast states. Towards this goal, we consider the case when $G$ is divided into $r$ densely connected clusters $C_1, C_2, \ldots, C_r$ while there are sparse connections between clusters. Fig. 1 illustrates such a graph where $r = 3$ and communications between nodes within any cluster are dense while they are sparse across different clusters. Next, we denote by $N_\alpha = |C_\alpha|$ the number of nodes in cluster $\alpha$ and $\sum_{\alpha=1}^{r} N_\alpha = N$. We argue that our formulation makes it easier in studying the convergence rates of the two-time-scale behavior of distributed methods over cluster networks compared to the existing singular perturbation approach. We next define the consensus dynamics of a cluster network, which will be the basis for defining the slow and fast states.

A. Consensus Dynamics

In this section, we present the formulation of the distributed consensus averaging method for a cluster network. For this we consider $G_\alpha = (V_\alpha, E_\alpha)$ be the graph representing the connection between nodes in cluster $C_\alpha$, for $\alpha = 1, \ldots, r$. In addition, let $G_E$ be the graph describing the external connections between clusters. Associated with each $C_\alpha$ we have a Laplacian matrix denoted by $L_\alpha$ corresponding to $G_\alpha$ and let $L^f = \text{diag}\{L_1, L_2, \ldots, L_r\}$ which is regarded as internal Laplacian matrix. In addition, let $L^E$ be the external Laplacian matrix corresponding to $G_E$. Thus for the cluster network we have $L = L^f + L^E$, which by (2) gives,

$$\dot{x} = -(L^f + L^E)x.$$  \hspace{1cm} (3)

Assumption 1. We assume that each $G_\alpha$, for all $\alpha = 1, \ldots, r$, and $G_E$ are connected and undirected.

For the internal graph we have $L^f = \text{diag}\{L^f_1, L^f_2, \ldots, L^f_r\}$. Since all the clusters have similar topology, they have the
same internal Laplacian matrix given by
\[
L^I_1 = L^I_2 = L^I_3 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.
\]
The external Laplacian matrix \(L^E\) is given by
\[
L^E = \begin{bmatrix}
L^E_{11} & L^E_{12} & L^E_{13} \\
L^E_{21} & L^E_{22} & L^E_{23} \\
L^E_{31} & L^E_{32} & L^E_{33}
\end{bmatrix}.
\]

Since each node in the external graph has degree 2, \(L^E_{11} = L^E_{22} = L^E_{33} = \text{diag}(2, 0, 0)\). Only the first node of each cluster communicates with other external nodes in this graph. Hence, we have
\[
L^E_{12} = L^E_{13} = L^E_{21} = L^E_{23} = L^E_{31} = L^E_{32} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\]

**B. Fast and slow variables within clusters**

As mentioned, distributed consensus methods over cluster networks often constitute two-time-scale dynamics. We note our model of fast and slow variables introduced in this paper is fundamentally different from the ones used in [6] and [4].

1) Slow variable

We define the slow variable in each cluster \(\alpha\) as the average of its nodes’ variables, i.e., let \(\bar{x}_\alpha\) be the slow variable of \(C_\alpha\) defined as \(\bar{x}_\alpha = \frac{1}{N_{\alpha}} \sum_{i \in N_{\alpha}} x^\alpha_i\), where \(N_{\alpha} = [1, 1, ..., 1]^T \in \mathbb{R}^{N_{\alpha}}\) and \(N_{\alpha}\) is the number of nodes in \(C_\alpha\). We denote by \(U\) and \(P\), respectively, as
\[
U = \text{diag}(1_{N_1}, 1_{N_2}, ..., 1_{N_{\alpha}}) \in \mathbb{R}^{N \times r},
\]
\[
P = \text{diag}(1_{N_1}^T, 1_{N_2}^T, ..., 1_{N_{\alpha}}^T) \in \mathbb{R}^{r \times r}.
\]

Note that \(P = U^T U\). Moreover, let \(y\) be defined as \(y = [y_1, y_2, ..., y_r]^T = [x_1, x_2, ..., x_r]^T \in \mathbb{R}^r\), where \(y_\alpha = \bar{x}_\alpha\). Then we denote by \(y\) the slow variable for the entire network
\[
y = P^{-1} U^T x.
\]

It follows from the formulation that the slow variable represents the states of the nodes of the aggregate network derived from the cluster network. Let the aggregate network be characterized by \(G_E\). We denote the Laplacian corresponding to the external aggregate graph as \(L^E \in \mathbb{R}^{r \times r}\). Note that \(L^E\) is a sparse matrix due to the cluster structure of our graph. We consider the following useful result about the relationship between \(L^E\) and \(L^E\).

**Lemma 1.** The Laplacian \(L^E\) satisfies
\[
L^E = U^T L^E U.
\]

2) Fast variable

Since \(G_\alpha\) is a dense graph, we expect that each \(x^\alpha_i\) converges quickly to \(\bar{x}_\alpha\). Thus, we define the fast variables within any cluster as the relative difference of its states to its slow variable. It can also be viewed as the residue or error of the state of a node with respect to the nodes’ average in a cluster. Hence, we define the fast variable \(e^\alpha_{x_i}\) in \(C_\alpha\) as
\[
e^\alpha_{x_i} = x^\alpha_i - \bar{x}_\alpha.
\]

**Remark 1.** It is worth noting that the definition of the fast state (7) is different than that in [4] and [6], where the fast state is defined as the relative difference in state values with respect to a reference node (typically the first node). This results in a simpler representation of the system dynamics as there is no need to use a complex similarity transformation. In addition, the formulation (7) will help to characterize explicitly the rates of the algorithm, which may not be obvious to derive from the ones in [4], [6].

For each cluster \(\alpha\), we denote by the diagonally dominant centering matrix as \(W_{\alpha} = (I_{N_{\alpha}} - \frac{1}{N_{\alpha}} 1_{N_{\alpha}} 1_{N_{\alpha}}^T) \in \mathbb{R}^{N_{\alpha} \times N_{\alpha}}\), and let \(e^\alpha_{x_i}\) be a vector in \(\mathbb{R}^{N_{\alpha}}\), whose \(i\)-th entry is \(e^\alpha_{x_i}\). In view of (7), we have \(e^\alpha_{x_i} = x^\alpha_i - \bar{x}_\alpha 1_{N_{\alpha}} = x^\alpha_i - \frac{1}{N_{\alpha}} \sum_{\alpha} 1_{N_{\alpha}}^T x^\alpha_i 1_{N_{\alpha}} = (I_{N_{\alpha}} - \frac{1}{N_{\alpha}} 1_{N_{\alpha}} 1_{N_{\alpha}}^T) x^\alpha_i = W_{\alpha} e^\alpha_{x_i}\). Finally, we denote by \(W = \text{diag}(W_{\alpha}) \in \mathbb{R}^{N \times N}\) and \(e_x\) as \(e_x = [(e^1_{x_1})^T, ..., (e^r_{x_r})^T]^T \in \mathbb{R}^N\). Thus, the fast variable for the entire network \(G\) is given by
\[
e_x = x - U y = W x.
\]

3) Fast and Slow Dynamics

We next present the dynamics for the fast and slow variables based on (5) and (8).

**Lemma 2.** The fast and slow variables \(e_x\) and \(y\) satisfy
\[
\dot{e}_x = \frac{de_x}{dt} = -WL^I e_x - WLE e_x - WL^E y,
\]
\[
\dot{y} = \frac{dy}{dt} = -P^{-1} L^E y - P^{-1} U^T L^E e_x.
\]

We observe that the dynamics of the slow variable depends on \(\hat{L}^E\). Since the slow variable can also be viewed as an aggregate variable, its states are expected to evolve according to the external aggregate communication graph \(G_E\). The second term in (10) can be seen as form of noise which depends on the fast variable, whose contribution will be prominent till the nodes inside the area synchronize. After local synchronization, the clusters behave as aggregate nodes, which eventually converge to the average of their initial conditions. This average coincides with the average of the initial conditions of all the nodes in the network. This motivates us to define a variable for each area that represents an error in consensus, and is expressed as \(e^\alpha_y = y_\alpha - \bar{y}\), where \(\bar{y} = \frac{1}{r} \sum_{\alpha = 1}^r y_\alpha\) is the average of the slow variables of each clusters which gives the average of all the states in the network. For the entire network, define \(e_y = [e^1_y, ..., e^r_y]^T \in \mathbb{R}^{r \times 1}\), which can also be expressed as
\[
e_y = y - \bar{y} 1_r = (I_r - \frac{1}{r} 1_r 1_r^T) y = W_r y,
\]
where \(1_r = [1, 1, ..., 1]^T \in \mathbb{R}^r\). Recall that \(r\) is the number of areas. Here \(W_r\) is the centering matrix of dimension \(r \times r\) associated with the external communication graph. The dynamics of \(e_y\) is given in the following lemma.

**Lemma 3.** The inter-area state \(e_y\) satisfies
\[
\dot{e}_y = -W_r P^{-1} \hat{L}^E e_y - W_r P^{-1} U^T L^E e_x.
\]
IV. MAIN RESULTS

In this section, we present the main results for our analysis of the convergence rate of the consensus problem over a cluster network. For our analysis we consider quadratic Lyapunov functions for the inter-area slow variable and the fast variable respectively given by $V_{e_y} = \frac{1}{2}\|e_y\|^2$ and $V_{e_x} = \frac{1}{2}\|e_x\|^2$. The convergence rate for the ensemble of the entire network can be derived by considering an aggregate Lyapunov function given as

$$V = V_{e_y} + \epsilon V_{e_x}. \quad (13)$$

Here $\epsilon > 0$ can be chosen conveniently depending upon the network structure as shown in Theorem 1. This type of Lyapunov functions has been used recently in studying the convergence rates of the classic two-time-scale stochastic approximation; see for example [16], [19]. Also, we denote $N_{\text{min}}$ and $N_{\text{max}}$, respectively, as $N_{\text{min}} = \min_{\alpha} N_{\alpha}$, $N_{\text{max}} = \max_{\alpha} N_{\alpha}$. We further denote $\sigma^2_{L} = \min_{\alpha} \sigma^2(L_{\alpha})$, which is strictly positive since the cluster $C_{\alpha}$ is connected. Also, let $\sigma^2(L) \equiv \sigma^2(L_{\alpha})$ be the second smallest eigenvalue of $L$, which is also strictly positive since $G_E$ is connected. Finally, since $P^{-1}$ is a diagonal matrix with entries $\frac{1}{N_{\alpha}}$ as diagonal elements, we have

$$\frac{1}{N_{\text{max}}} \leq P^{-1} \leq \frac{1}{N_{\text{min}}} = \frac{1}{N_{\text{max}}} I \leq P \leq N_{\text{max}} I \quad (14)$$

We next consider the following useful lemmas about the Lyapunov functions $V_{e_y}$ and $V_{e_x}$.

**Lemma 4.** The Lyapunov function $V_{e_y}$ satisfies

$$\frac{dV_{e_y}}{dt} \leq -\frac{1}{2} e^T y P^{-1} \bar{L} e_y + \frac{1}{2N_{\text{min}}} e^T x L e_x. \quad (15)$$

**Lemma 5.** The Lyapunov function $V_{e_x}$ satisfies

$$\frac{dV_{e_x}}{dt} \leq -e^T x W L e_x - \frac{1}{2} e^T x L W L e_x + 1 e^T x L e_x. \quad (16)$$

The proofs of Lemma 1-5 are omitted due to space limit and can be found in the extended version of this paper at [20].

**Assumption 2.** $G$ has a cluster structure that satisfies

$$\sigma^2_{L} \geq \frac{2\|L\|^2 N_{\text{max}}}{\sigma^2(L^E) N_{\text{min}}}. \quad (17)$$

**Remark 3.** 1. Our result in (19) implies that the distributed consensus method converges exponentially. In addition, the rate of convergence only depends on the external graph $G_E$ through $\sigma^2_L$, which only scales with a small number of nodes since there are a few number of communications between clusters. On the other hand, the existing analysis for the rate of distributed consensus method depends on the total number of nodes $N$ when applying to the entire network.

2. The parameter $\epsilon$ can be viewed as a singular-perturbation parameter representing the time-scale difference in the fast and slow dynamics, similar to the one used in [6].

**Proof.** We begin our proof by recalling the aggregate Lyapunov function from (13) given by $V = V_{e_y} + \epsilon V_{e_x}$, which by using (15) and (16) we have

$$\dot{V} = \dot{V}_{e_y} + \epsilon \dot{V}_{e_x} \leq -\frac{1}{2} e^T y P^{-1} \bar{L} e_y + \frac{1}{2N_{\text{min}}} e^T x L e_x - e^T x W L e_x - \frac{1}{2} e^T x L W L e_x + \epsilon e^T y L e_y. \quad (20)$$

We simplify the above inequality by bounding each term using the inequalities that we have established before. First, using the definition of $P$ in (4) and $N_{\text{max}}$ in (IV) yields $-\frac{1}{2} e^T y P^{-1} \bar{L} e_y \leq \frac{1}{2N_{\text{min}}} e^T y L e_y$. Here we observe that the Laplacian $L^E$ has a zero eigen value corresponding to the eigen vector $1_r$ while the other eigen values are strictly positive. Recall that $1_r$ is in the null space of $W_r$, i.e., $1^T_r W_r = W_r 1^T_r = 0$. Also, since $e_y = W_r e_y$ we have that $e_y \in \text{span}\{1_r^T\}$. Thus we have $-e^T x L e_y \leq -\sigma^2_L ||e_y||^2$, which in view of (14) implies that

$$\frac{1}{2} e^T y P^{-1} \bar{L} e_y \leq -\frac{\sigma^2_L}{2N_{\text{max}}} ||e_y||^2. \quad (21)$$

Since $L^E$ is positive semi-definite and $e_r^T L = e_r^T W L W = e_r^T W = \epsilon_r^T e_r$, hence we have $-\frac{1}{2} e^T x W L e_x = -\frac{1}{2} e^T x L e_x \leq 0$. Next considering the last term in the inequality (20) we have $-e^T x W L e_x = -e^T x L e_x$. We note that the fast variable can be expressed as sub-vector corresponding to each cluster as $e_x = [(e_x^T_{1})\ldots(e_x^T_{r})]^T$. Following this we have $-e^T x L e_x = -e^T \sum_{\alpha=1}^{r} e_x^T_{\alpha} L e_x^T_{\alpha} \leq -\epsilon \sum_{\alpha=1}^{r} \sigma^2_{L} ||e_x^T||^2 \leq -\epsilon \sum_{\alpha=1}^{r} \sigma^2_{L} ||e_x^T||^2 = -\epsilon \sigma^2 L ||e_x||^2$. Next we observe that since the external Laplacian matrix is sparse then we have $||L^E|| \leq ||L||$. Next we simplify the second and last term from (20), so that we have

$$\frac{1}{2N_{\text{min}}} e^T x L e_x \leq \frac{||L||}{2N_{\text{min}}} \|e_x\|^2; \quad \epsilon \frac{1}{2} e^T y T U T L U e_y \leq \frac{||L^E||}{2} \|e_y\|^2. \quad (22)$$

By using (21)–(22) into (20) we have

$$\dot{V} \leq -\frac{\sigma^2_{L}}{2N_{\text{max}}} \|e_y\|^2 + \frac{||L||}{2N_{\text{min}}} \|e_x\|^2 - \epsilon \sigma^2_{L} \|e_x\|^2 + \frac{||L^E||}{2} \|e_y\|^2 - \left(\epsilon \sigma^2_{L} - \frac{||L^E||}{2N_{\text{min}}} \|e_x\|^2 \right) \|e_y\|^2. \quad (23)$$
We note that to guarantee stability we need $\dot{V} < 0$. Accordingly, choosing $\epsilon = \sigma^2 E \frac{2}{N_{\text{max}} \|E\|}$ leads to

$$\dot{V} \leq -\frac{\sigma^2 E_\eta}{2N_{\text{max}}} V_{\eta} - \left( \frac{\sigma^2 \sigma_x E_\eta}{N_{\text{max}} \|E\|} - \frac{\|E\|}{N_{\text{min}}} \right) V_{x}. \quad (24)$$

Using Assumption 2 we have $\frac{\sigma^2 \sigma_x E_\eta}{N_{\text{max}} \|E\|} - \frac{\|E\|}{N_{\text{min}}} \geq \frac{\|E\|}{N_{\text{min}}}$, which when substituting (24) in yields

$$\dot{V} \leq \frac{\sigma^2 (L^E)}{2N_{\text{max}}} V_{\eta} - \left( \frac{\|E\|}{N_{\text{min}}} \right) V_{x} + \frac{\sigma^2 (L^E)}{2N_{\text{max}}} \epsilon V_{x}$$

$$= -\frac{\sigma^2 (L^E)}{2N_{\text{max}}} (V_{\eta} + \epsilon V_{x}) - \left( \frac{\|E\|}{N_{\text{min}}} - \frac{\sigma^2 (L^E)}{4N_{\text{max}}^2 \|E\|^2} \right) V_{x}. \quad (25)$$

Here we note have $\|E\| > \sigma^2 (L^E)$, hence we have a non-positive coefficient of $V_{x}$. Thus using (13) leads us $\dot{V} \leq -\frac{\sigma^2 (L^E)}{2N_{\text{max}}} V$. This immediately gives (19) which concludes our proof.

V. Simulation results

In this section, we illustrate our theoretical results by a number of numerical experiments on distributed consensus methods over different cluster networks. Here we investigate how the convergence of the consensus dynamics depends upon changing the connectivity between the clusters and further changing the connectivity between the clusters.

Fig. 3: A network with 3 clusters composed of 15 (blue), 30 (black), and 20 (red) node.

A. Small Network

For our simulation we consider a cluster network of 65 nodes divided into 3 distinct clusters and presented in Fig. 3. We simulate two different networks, one with only three external communication edges as shown by green dashed edges, and the other with a total of 6 communication edge as shown by both the dashed and solid edges.

We first verify the condition in Assumption 2. The network in Fig. 3 has $N_{\text{min}} = 15$ and $N_{\text{max}} = 30$. The eigenvalues of the different Laplacian matrices are $\sigma^2 (L^I) = 6.79$, $\sigma^2 (L^I_\eta) = 16.48$, $\sigma^2 (L^I_\eta) = 6.52$ and $\sigma^2 (L^E) = 3$. It is straight forward to verify that this network satisfies condition (17) in given in Assumption (2).

Fig. 4: Simulation 1: evolution of different variables for the network in Fig. 3 using 3 external communication links.

Fig. 5: Simulation 2: evolution of different variables for the network in Fig. 3 using 6 external communication links.

We next present our simulations in Fig. 4 (where we consider 3 external communication edges) and in Fig. 5 (where there are 6 external communication edges). For our simulations, we have $\epsilon = 0.017$. In addition, the nodes’ initial values are randomly initialized. We use (1) to generate $L^I$ and $L^E$. The two time-scale nature of the network can be observed in Fig. 4, which shows the time response of the states $x$, the inter-area variable $e_y$, and the fast variable $e_x$. It can be seen from Fig. 4 (b) and (c) that the nodes perform first a local synchronization, then they slowly drive the entire network to the average of their initial conditions, which agrees with our theoretical results. Fig. 5 shows a similar result for the same network but using 6 external communication edges. Moreover, it also shows that the convergence of the variables is faster when the number of edges of the external graph increases as compared to the ones in Fig. 4. Our simulations show that we can study how the states of all the 65 nodes evolve by considering the dynamics of the inter-area variables, which has a much smaller dimension given by the number of clusters. This agrees with our theoretical result in Theorem 1.
B. Large Network

We now provide simulations on the performance of distributed consensus over large cluster networks where we intend to observe the convergence of the different network variables as we increase the number of clusters. Here we consider 6 clusters with 50 nodes in each cluster as shown in Fig. 6. The individual clusters are generated in a similar manner to the simulations above. We consider that any two clusters communicate with each other using only two links. The simulation results of this experiment is shown in Fig. 7, where we have the same observation as the ones in small networks.

VI. CONCLUSION AND FUTURE WORK

In this paper, we propose a distributed two-time-scale consensus algorithm for cluster networks where the time scales arise both from the cluster structure. We provided an explicit formula for the convergence rate of such algorithms. Our theoretical results are validated by few numerical simulations. As future work we will address more practical issues like communication constraints and privacy over cluster network.

VII. ACKNOWLEDGEMENT

The work of Amit Dutta and Thinh T. Doan was partially supported by the Commonwealth Cyber Initiative, an investment in the advancement of cyber R&D, innovation, and workforce development. For more information about CCI, visit www.cyberinitiative.org.

REFERENCES

[1] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” nature, vol. 393, no. 6684, pp. 440–442, 1998.

[2] F. Zhao and L. J. Guibas, Wireless sensor networks: an information processing approach. Morgan Kaufmann, 2004.

[3] J. H. Chow, Power System Coherency and Model Reduction. Springer, 2013.

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

[5] S. Martin, I.-C. Morărescu, and D. Nešić, “Time scale modeling for consensus in sparse directed networks with time-varying topologies,” in 2016 IEEE 55th Conference on Decision and Control (CDC). IEEE, 2016, pp. 7–12.

[6] E. Byzik and M. Arcak, “Area aggregation and time-scale modeling for sparse nonlinear networks,” Systems & Control Letters, vol. 57, no. 2, pp. 142–149, 2008.

[7] A. M. Boker, T. R. Nudell, and A. Chakrabortty, “On aggregate control of clustered consensus networks,” in 2015 American Control Conference (ACC), 2015, pp. 5527–5532.

[8] S. Mukherjee, H. Bai, and A. Chakrabortty, “Reduced-dimensional reinforcement learning control using singular perturbation approximations,” Automatica, vol. 126, p. 109451, 2021.

[9] A. M. Boker, C. Yuan, F. Wu, and A. Chakrabortty, “Aggregate control of clustered networks with inter-cluster time delays,” in 2016 American Control Conference (ACC). IEEE, 2016, pp. 5340–5345.

[10] I.-c. Morarescu, S. Martin, A. Girard, and A. Muller-gueudin, “Coordination in networks of linear impulsive agents,” IEEE Transactions on Automatic Control, vol. 61, no. 9, pp. 2402–2415, 2016.

[11] V. T. Pham, N. Messai, D. Hoa Nguyen, and N. Manamanni, “Robust formation control under state constraints of multi-agent systems in clustered networks,” Systems and Control Letters, vol. 140, p. 104689, 2020.

[12] T. V. Pham, T. T. Doan, and D. H. Nguyen, “Delivered twos-time-scale methods over clustered networks,” in American Control Conference (ACC), 2020.

[13] M. Mesbahi and M. Egerstedt, Graph theoretic methods in multiagent networks. Princeton University Press, 2010.

[14] P. Kokotovic, H. K. Khalil, and J. O’Reilly. Singular Perturbation Methods in Control Analysis. SIAM, 1999.

[15] T. T. Doan, “Finite-time analysis and restarting scheme for linear two-time-scale stochastic approximation,” SIAM Journal on Control and Optimization, vol. 59, no. 4, pp. 2798–2819, 2021.

[16] ——, “Nonlinear two-time-scale stochastic approximation: Convergence and finite-time performance,” Available at: https://arxiv.org/abs/2011.01868, 2020.

[17] S. S. Kia, B. Van Scoy, J. Cortés, R. A. Freeman, K. M. Lynch, and S. Martínez, “Tertorial on dynamic average consensus: The problem, its applications, and the algorithms,” vol. 39, 2019, pp. 40–72, jologs.

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

[19] T. T. Doan, “Finite-time convergence rates of nonlinear two-time-scale stochastic approximation under Markovian noise,” Available at: https://arxiv.org/abs/2104.01627, 2021.

[20] A. Dutta, A. M. Boker, and T. T. Doan, “Convergence rates of distributed consensus over cluster networks: A two-time-scale approach,” arXiv preprint arXiv:2104.07781, 2021.