

Distributed Solver for Discrete-Time Lyapunov Equations Over Dynamic Networks with Linear Convergence Rate

Xia Jiang, Xianlin Zeng, Jian Sun and Jie Chen, *Fellow, IEEE*

Abstract

This paper investigates the problem of solving discrete-time Lyapunov equations (DTLE) over a multi-agent system, where every agent has access to its local information and communicates with its neighbors. To obtain a solution to DTLE, a distributed algorithm with uncoordinated constant step sizes is proposed over time-varying topologies. The convergence properties and the range of constant step sizes of the proposed algorithm are analyzed. Moreover, a linear convergence rate is proved and the convergence performances over dynamic networks are verified by numerical simulations.

Index Terms

distributed algorithm, linear convergence rate, discrete-time Lyapunov equation, convex optimization, dynamic network

I. INTRODUCTION

Distributed algorithms are vital for control and optimization over large-scale systems, such as sensor networks [26], wireless communication networks [28] and smart grids [5]. The research of distributed optimization algorithms has attracted considerable interest. Distributed optimization problems, which search for an optimal solution of a sum of individual objective functions with local or coupled constraints, have been widely investigated. A large number of effective discrete-time algorithms have been developed to solve distributed optimization problems in multi-agent networks, such as distributed gradient/subgradient methods [14], [17], [20], [21], [27].

In recent years, many distributed optimization algorithms have been proposed for the computational problems of solving large-scale linear algebraic equations of the form $Ax = b$, where A is a matrix and x, b are vectors of appropriate dimensions [1], [15], [16], [19], [23]. [19] provided a

distributed computing algorithm and detailed discussions for solving linear algebraic equations, where every agent only knows a subset of the partitioned matrix $\begin{bmatrix} A & b \end{bmatrix}$ and exchanges the estimates of solution with its neighbors. Based on the assumption of the existence of at least one exact solution to linear algebraic equations, [15] proposed a distributed discrete-time algorithm with a linear convergence rate. If linear equations have no exact solutions, distributed algorithms to solve the least-squares solution are investigated [16], [23]. Linear matrix equations are more general than linear algebraic equations and the above distributed algorithms cannot be applied directly to solving linear matrix equations.

Linear matrix equations have wide applications in the design of modern complex systems. In particular, the discrete-time Lyapunov equation (DTLE), which is one of the most common linear matrix equations, is particularly important in the analysis of linear dynamic systems, such as stability analysis, controllability analysis and optimal control [7], [10]. There is a large amount of literature on centralized numerical methods to solve Lyapunov equations [2], [8], [9], [25]. With the growing data in systems and the expansion of the system scale, many excellent works have studied efficient methods for solving large-scale Lyapunov equations [2], [8]. However, because system information is distributed among separated equipments and computing capability of one agent is limited, the need to solve Lyapunov equations in a distributed way arises. Recently, there are several works studying distributed algorithms for computations of linear matrix equations [13], [30]. [30] has investigated the problem of solving general linear matrix equations and proposed distributed continuous-time algorithms over time-invariant undirected topologies. The distributed algorithm for solving DTLE proposed in [13] is a gradient-based method over time-invariant undirected graphs with diminishing step sizes, whose convergence rate is not fast enough.

This paper proposes a discrete-time distributed algorithm for solving DTLE of the form $AXA' - X + Q = 0_{n \times n}$ over time-varying topologies, which has a linear convergence rate. The contributions are summarized below.

- The paper studies the problem of solving discrete-time Lyapunov equation in a distributed way, which is vital in the analysis of linear dynamical systems. [30] proposed a distributed continuous-time algorithm over time-invariant undirected graphs, which cannot be applied over time-varying graphs. To our best knowledge, there is no work studying how to solve the discrete-time Lyapunov matrix equation over time-varying topologies in a distributed way.

- This paper proposes a distributed discrete-time algorithm with uncoordinated constant step sizes over time-varying graphs. The proposed method overcomes the shortage of algorithms with diminishing step sizes that fail to achieve linear convergence rates [6], [13], [18]. [24], [27] have developed algorithms with identical constant step sizes over multi-agent topologies. However, the methods can only achieve linear convergence rates under strongly convex assumptions. In this paper, the proposed method achieves a linear convergence rate without the assumption of strong convexity. What's more, the proposed algorithm has uncoordinated step sizes.
- Using the theories of optimization and control, we provide sufficient conditions for the linear convergence rate of the proposed algorithm under time-varying topologies and the range of constant step sizes of the proposed algorithm.

The remainder of the paper is organized as follows.

The mathematical preliminaries and the problem description of solving discrete-time Lyapunov equation in a distributed manner are given in section II. A distributed discrete-time algorithm for the optimization problem and its convergent properties are provided in section III. We provide theoretical proofs for our main results in section IV. The convergence performances of the proposed algorithm are verified by numerical simulations in Section V and the conclusion is made in section VI.

II. PRILIMINARIES & PROBLEM DESCRIPTION

A. Mathematical Priliminaries

We denote \mathbb{R} as the set of real numbers, \mathbb{R}^n as the set of n -dimensional real column vectors, $\mathbb{R}^{n \times m}$ as the set of n -by- m real matrices, $\mathbf{1}_m$ as an m -dimensional column vector with all elements being one and E_n as the $n \times n$ identity matrix, respectively. We denote A' as the transpose of matrix A , $0_{n \times q}$ as the $n \times q$ matrix with all elements of 0, $\lambda_{min}(A)$ as the minimum eigenvalue of the matrix A , $\lambda_{max}(A)$ as the maximum eigenvalue of the matrix A and $A \otimes B$ as the Kronecker product of matrices A and B . For a matrix A , the formula $A > 0$ ($A \geq 0$) denotes that A is positive definite (positive semi-definite) and $A < 0$ ($A \leq 0$) denotes that A is negative definite (negative semi-definite). In addition, for a real vector v , $\|v\|$ is the Euclidean norm. For a real matrix A , $\|A\|_F$ denotes the Frobenius norm of the real matrix defined by $\|A\|_F = \sqrt{tr(A'A)} = \sqrt{\sum_{i,j} A_{ij}^2}$. Let $\langle \cdot, \cdot \rangle_F$ be the Frobenius inner product of real matrices defined by $\langle B_1, B_2 \rangle_F = tr(B_1'B_2) = \sum_{i,j} (B_1)_{i,j} (B_2)_{i,j}$ with $B_1, B_2 \in \mathbb{R}^{m \times n}$. If

$f : \mathbb{R}^{m_1 \times m_2} \times \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}$, $\nabla_X f(X, Y) = [\nabla_{X_{ij}} f(X, Y)] \in \mathbb{R}^{m_1 \times m_2}$ denotes the partial gradient of function f with respect to X . For a matrix $A \in \mathbb{R}^{n \times m}$, $\text{vec}(A)$ denotes the vectorization of A which is an $mn \times 1$ column vector obtained by juxtaposing the consecutive rows of the matrix next to each other and taking the transpose of the obtained long multi-row: $\text{vec}(A) = [a_{11}, \dots, a_{1m}, a_{21}, \dots, a_{2m}, \dots, a_{n1}, \dots, a_{nm}]'$.

B. Graph Theory

The dynamic communication between m agents at time k is denoted by $\mathcal{G}_k(\nu, \varepsilon_k, \mathcal{A}_k)$, where $\nu = \{1, \dots, m\}$ is a finite nonempty node set, $\varepsilon_k \subset \nu \times \nu$ is the edge set of time k and $\mathcal{A}_k = [a_{ij,k}] \in \mathbb{R}^{m \times m}$ is the adjacency matrix at time k . For undirected graphs, all the adjacency matrices $\{\mathcal{A}_k\}$ are symmetric matrices such that $a_{ij,k} = a_{ji,k}$ and diagonal element $a_{ii,k} = 0$. If an edge $(i, j) \in \varepsilon_k$, then node j is called a neighbor of agent i at time k and $a_{ij,k} > 0$, otherwise, $a_{ij,k} = 0$. Let $\mathcal{N}_{i,k}$ denotes the set of neighbor nodes that connect to node i at time k , i.e., $\mathcal{N}_{i,k} = \{j : (i, j) \in \varepsilon_k\}$. The Laplacian matrix \mathcal{L}_k at time k is defined as $l_{ij,k} = -a_{ij,k}$, for $i \neq j$ and $l_{ii,k} = \sum_{j \in \mathcal{N}_{i,k}} a_{ij,k}$, which ensures that $\sum_{j=1}^m l_{ij,k} = 0$. What's more, $\mathcal{L}_k = \mathcal{L}'_k$ for every undirected graph \mathcal{G}_k . The dynamic undirected graphs are assumed to be simple which implies that there is no repeated edge or self-loop.

C. DTLE Problem Description

The discrete-time Lyapunov equation (DTLE) for stability/controllability [10] is

$$AXA' - X + Q = 0_{n \times n}, \quad (1)$$

where $X, A, Q \in \mathbb{R}^{n \times n}$ and X is the unknown variable. In addition, X and Q are symmetric matrices. If the Lyapunov equation (1) is used to analyze one system's stability, Q is a symmetric positive definite matrix.

The matrices A and Q are decomposed as follows:

$$A = \begin{bmatrix} A_{r1} \\ \cdot \\ \cdot \\ \cdot \\ A_{rm} \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad Q = [Q_{l1} \quad \cdot \quad \cdot \quad \cdot \quad Q_{lm}] \in \mathbb{R}^{n \times n}, \quad (2)$$

where $A_{ri} \in \mathbb{R}^{n_i \times n}$, $Q_{li} \in \mathbb{R}^{n \times n_i}$, $\sum_{i=1}^m n_i = n$.

In this structure, every agent i only knows A_{ri} , Q_{li} , and communicates with its neighbors $\{j : j \in \mathcal{N}_i\}$ to obtain a same solution to (1). To distinguish the row-blocks or column-blocks of a matrix, we use subscript ri to denote its i th row-block and subscript li to denote its i th column-block.

The objective of this paper is to propose a fully distributed discrete-time algorithm over time-varying undirected topologies to solve DTLE (1) with a linear convergence rate. We need the following assumptions.

Assumption 1: Equation (1) has at least one solution.

Assumption 2: Graph \mathcal{G}_k is undirected for all $k \geq 0$ and the adjacency matrix \mathcal{A}_k of graph \mathcal{G}_k is doubly stochastic.

Remark 1: Assumption 1 makes the problem well-defined, which also implies that the solution set is non-empty. Assumption 2 is common for the design of distributed algorithms over multi-agent topologies.

III. DISTRIBUTED SOLVER DESIGN

In this section, we reformulate the problem into a separable one and propose a distributed discrete-time algorithm with a linear convergence rate.

A. Problem Reformulation

To deal with the coupling in AXA' , we introduce an auxiliary variable $Y \in \mathbb{R}^{n \times n}$ to transform the term into a separable structure. Specifically, equation (1) is equivalent to

$$\begin{aligned} Y &= AX, \\ YA' &= XE_n - Q, \end{aligned}$$

where E_n is an $n \times n$ identity matrix.

Define $X_i \in \mathbb{R}^{n \times n}$ and $Y_i \in \mathbb{R}^{n \times n}$ as estimates of X and Y of agent $i \in \{1, \dots, m\}$ respectively. With the matrix decomposition (2), $Y = AX$ can be transformed into $Y_i^{ri} = A_{ri}X_i$ and $Y_i = [Y_{r1}^i, \dots, Y_{rm}^i]'$; By multiplying X with an identity matrix E_n , $YA' = XE_n - Q$ can be

transformed into $Y_i A'_{ri} = X_i E_{li} - Q_{li}$ and $E_n = [E_{l1}, \dots, E_{lm}]$. With the above decompositions, we decentralize the DTLE (1) into

$$\begin{aligned} Y_i^{ri} &= A_{ri} X_i, \\ Y_i A'_{ri} &= X_i E_{li} - Q_{li}, \\ X_i &= X_j, \\ Y_i &= Y_j, \end{aligned} \tag{3}$$

for all $i, j \in \{1, \dots, m\}$.

Then, we further transform the distributed computation of (3) into a distributed optimization problem.

$$\begin{aligned} \text{Min}_{(\hat{X}, \hat{Y})} \quad & F(\hat{X}, \hat{Y}) = \sum_{i=1}^m f_i(X_i, Y_i), \\ \text{s.t.} \quad & X_i = X_j, \quad Y_i = Y_j, \quad i, j \in \{1, \dots, m\}, \end{aligned} \tag{4}$$

where $f_i(X_i, Y_i) = \frac{1}{2} \|Y_i^{ri} - A_{ri} X_i\|_F^2 + \frac{1}{2} \|Y_i A'_{ri} - X_i E_{li} + Q_{li}\|_F^2$, $\hat{Y} = [Y'_1, \dots, Y'_n]' \in \mathbb{R}^{mn \times n}$, $\hat{X} = [X'_1, \dots, X'_n]' \in \mathbb{R}^{mn \times n}$, agent i only knows the information of A_{ri} , E_{li} and Q_{li} . Throughout this paper, we use (\hat{X}^*, \hat{Y}^*) to denote an optimal solution to problem (4).

Remark 2: Because the optimal function value of (4) $F^* = 0$, it implies that four equations in (3) are satisfied. Then an optimal solution of problem (4) is \hat{X}^* if and only if $\hat{X}^* = \mathbf{1}_m \otimes X^*$ and X^* is a solution of (1).

B. Distributed algorithm

From problem reformulation, every agent i has the information state $X_{i,k} \in \mathbb{R}^{n \times n}$ and the auxiliary variable $Y_{i,k} \in \mathbb{R}^{n \times n}$, which are the estimates of the solutions of the original optimization problem (4) at time k .

Agent i updates its estimates $X_{i,k+1}$ according to the following rules. For all $k \geq 0$ and all $i \in \{1, \dots, m\}$:

$$\begin{aligned} X_{i,k+1} &= X_{i,k} - \alpha_i d_{X_{i,k}} - \frac{\alpha_i}{2} \sum_{j \in \mathcal{N}_{i,k}} a_{ij,k} (X_{i,k} - X_{j,k}), \\ Y_{i,k+1} &= Y_{i,k} - \alpha_i d_{Y_{i,k}} - \frac{\alpha_i}{2} \sum_{j \in \mathcal{N}_{i,k}} a_{ij,k} (Y_{i,k} - Y_{j,k}), \\ d_{X_{i,k}} &= -A'_{ri} (Y_{i,k}^{ri} - A_{ri} X_{i,k}) - (Y_{i,k} A'_{ri} - X_{i,k} E_{li} + Q_{li}) E'_{li}, \\ d_{Y_{i,k}} &= E_{li} (Y_{i,k}^{ri} - A_{ri} X_{i,k}) + (Y_{i,k} A'_{ri} - X_{i,k} E_{li} + Q_{li}) A_{ri}, \end{aligned} \tag{5}$$

where $a_{ij,k}$ is the weight of time k between agent i and j , $d_{X_i,k} = \nabla_{X_i,k} f_i(X_{i,k}, Y_{i,k})$, which is a gradient of the objective function $f_i(X_i, Y_i)$ with respect to X_i at time k and similarly, $d_{Y_i,k} = \nabla_{Y_i,k} f_i(X_{i,k}, Y_{i,k})$.

For ease of presentation, we introduce an assumption to describe the range of step sizes.

Assumption 3: In the proposed algorithm (5), the step size of every agent i , α_i satisfies that $\min\{1, 1/\xi_i\} > \alpha_i > 0$, where ξ_i satisfies the following inequality

$$\begin{aligned} & \| -A'_{ri}T_{i1} - T_{i2}E'_{li} \|_F^2 + \| E_{li}T_{i1} + T_{i2}A_{ri} \|_F^2 \\ & \leq \xi_i \| T_{i1} \|_F^2 + \xi_i \| T_{i2} \|_F^2, \end{aligned}$$

for all $T_{i1} \in \mathbb{R}^{n_i \times n}$, $T_{i2} \in \mathbb{R}^{n \times n_i}$.

Remark 3: Because of the special form of the above inequality, it is easy to find a proper variable ξ_i . For instance, by transformation, every ξ_i satisfying $\xi_i \geq 2(\|A'_{ri}\|^2 + \|E_{li}\|^2)$ also satisfies the above inequality. So, we can find a proper step size α_i satisfying Assumption 3 easily.

For the simplicity of analysis, we define $Z_{i,k} = \begin{bmatrix} Y_{i,k} \\ X_{i,k} \end{bmatrix}$, $\hat{Z}_k = \begin{bmatrix} Z_{1,k} \\ \vdots \\ Z_{m,k} \end{bmatrix}$, and $f_i(Z_i) = f_i(X_i, Y_i)$

without causing ambiguity. Rewrite the algorithm (5) compactly in terms of the matrix $Z_{i,k}$ as

$$\begin{aligned} Z_{i,k+1} &= Z_{i,k} - \alpha_i d_{Z_{i,k}} - \frac{\alpha_i}{2} \sum_{j \in \mathcal{N}_{i,k}} a_{ij,k} (Z_{i,k} - Z_{j,k}) \\ &= (1 - \frac{\alpha_i}{2} l_{ii,k}) Z_{i,k} - \frac{\alpha_i}{2} \sum_{j \in \mathcal{N}_{i,k}} l_{ij,k} Z_{j,k} - \alpha_i d_{Z_{i,k}}, \end{aligned}$$

where $i \in \{1, \dots, m\}$, $k \geq 0$ and $d_{Z_{i,k}} = \nabla_{Z_{i,k}} f_i(Z_{i,k})$.

We introduce a matrix $W_k = [w_{ij,k}] \in \mathbb{R}^{m \times m}$ at time k and let $w_{ii,k} = 1 - \frac{\alpha_i}{2} l_{ii,k}$ and $w_{ij,k} = -\frac{\alpha_i}{2} l_{ij,k}$. Then the above equation can be further written as

$$Z_{i,k+1} = \sum_{j=1}^m w_{ij,k} Z_{j,k} - \alpha_i \nabla f_i(Z_{i,k}), \quad (6)$$

where $\sum_{j=1}^m w_{ij,k} = 1$ and the matrix $W_k \in \mathbb{R}^{m \times m}$ is doubly stochastic if Assumption 2 holds.

Remark 4: There are a lot of excellent works studying distributed optimization algorithms with diminishing/constant step sizes in detail. In comparison, the algorithm proposed in this paper has the following differences and advantages.

- The step sizes of many algorithms are diminishing, hindering the possibility of linear convergence rates [13], [18]. The proposed algorithm has constant step sizes and a linear convergence rate.
- Some studies with constant step sizes have achieved linear convergence rates for strongly convex functions [24], [27]. In addition, these methods require an identical step size for all agents. Nevertheless, the constant step sizes in the proposed algorithm are not necessarily identical and the method achieves a linear convergence rate without the strongly convex assumptions, making the algorithm more practical.
- In [29], [30], they studied distributed algorithms over time-invariant undirected graphs. In comparison, the proposed algorithm of this paper can be applied to time-varying undirected topologies. Furthermore, the proposed algorithm uses no dual variable, making it more straightforward than primal-dual methods in [30].

C. Convergence properties

In this subsection, convergence properties and the convergence rate of the proposed algorithm are presented.

To derive convergence properties of the proposed algorithm (5), we need some common assumptions.

Assumption 4: The graph sequence $\{\mathcal{G}_k\}$ is uniformly connected. That is, there exists an integer $B \geq 1$ such that agent i sends its information to all other agents at least once every B consecutive time slots.

Assumption 5: There exists a scalar η with $0 < \eta < 1$ such that for all $i \in \{1, \dots, m\}$, $w_{ii,k} \geq \eta$ for all $k \geq 0$ and $w_{ij,k} \geq \eta$ if agent j communicates directly with agent i at time k . Otherwise, $w_{ij,k} = 0$.

Then the following theorem shows convergence results of the the proposed method.

Theorem 1: Suppose Assumptions 1, 2, 3 – 5 hold.

- (i) The variable state of every agent i generated by algorithm (5) converges to one same solution, which means $\hat{Z}_k \rightarrow \hat{Z}^*$ as $k \rightarrow \infty$ and \hat{Z}^* is an optimal solution to problem (4).
- (ii) $\sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) = O(\frac{1}{k})$, where $\bar{Z}_{i,k}$ is time average state variable of agent i , $\bar{Z}_{i,k} = \frac{\sum_{l=1}^k Z_{i,l}}{k}$.

Theorem 1 shows that over uniformly strongly connected graphs, all agents' variables converge to one same optimal solution. Next, we give sufficient conditions that the proposed algorithm has a linear convergence rate.

Assumption 6: The topological sequence has finite graphs and every graph is connected.

Remark 5: In Assumption 6, the graph at every iteration k needs to be connected and the number of graphs in topological sequence is finite, which is stronger than the conditions of Assumptions 4 and 5. One topological sequence satisfying Assumption 6 must satisfy Assumptions 4 and 5, but the converse is not true.

The linear convergence rate of the proposed algorithm is presented in the following theorem.

Theorem 2: Suppose Assumptions 1, 2, 3 and 6 hold. The variable state \hat{Z}_k generated by algorithm (6) converges to an optimal solution \hat{Z}^* of problem (4) at a linear convergence rate.

IV. THEORETICAL PROOF

In this section, we present theoretical proofs for the main results on convergence properties and convergence rate of algorithm (5).

A. Proof of Theorem 1

Before we give the proof of Theorem 1, we introduce a few lemmas.

Lemma 1:

Weyl's inequality: For matrix formula $M = H + P$, if M, H and P are all n by n Hermitian matrices (symmetric matrices for real matrices), the i th eigenvalue of M satisfies that $\lambda_i(H) + \lambda_{\min}(P) \leq \lambda_i(M) \leq \lambda_i(H) + \lambda_{\max}(P)$.

Lemma 2: Suppose Assumptions 1, 2 and 3 hold.

- (i) the sequence $\{\hat{Z}_k\}$ generated by algorithm (6) is bounded and every equilibrium of the sequence is Lyapunov stable.
- (ii) $\lim_{k \rightarrow \infty} f_i(Z_{i,k}) = 0$ for all agent i and $\sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) = O(\frac{1}{k})$, where $\bar{Z}_{i,k}$ is time average state variable of agent i , $\bar{Z}_{i,k} = \frac{\sum_{l=1}^k Z_{i,l}}{k}$.
- (iii) $\lim_{k \rightarrow \infty} e_{i,k} = 0$ where $e_{i,k} = \alpha_i \nabla_{Z_{i,k}} f_i(Z_{i,k})$ for all agent i .

Proof: (i) Let \hat{Z}^* be a solution to problem (4), which is also an equilibrium of the algorithm (6). Define

$$D(k) = \|\Lambda^{-\frac{1}{2}}(\hat{Z}_{k+1} - \hat{Z}^*)\|_F^2 - \|\Lambda^{-\frac{1}{2}}(\hat{Z}_k - \hat{Z}^*)\|_F^2,$$

where $\Lambda = \text{diag}\{\alpha_1 \otimes E_{2n}, \dots, \alpha_m \otimes E_{2n}\}$. Recall that $L_k = \mathcal{L}_k \otimes E_{2n}$. Since $\hat{Z}_{k+1} = \hat{Z}_k - \frac{1}{2}\Lambda L_k \hat{Z}_k - \Lambda \nabla_{\hat{Z}_k} F(\hat{Z}_k)$ by algorithm (6), we plug \hat{Z}_{k+1} in $D(k)$ and get the following equation

$$\begin{aligned} D(k) &= 2 \left\langle -\frac{1}{2} L_k \hat{Z}_k - \nabla_{\hat{Z}_k} F(\hat{Z}_k), \hat{Z}_k - \hat{Z}^* \right\rangle_F \\ &\quad + \left\| \Lambda^{-\frac{1}{2}} \left(\frac{1}{2} \Lambda L_k \hat{Z}_k + \Lambda \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right) \right\|_F^2 \\ &= - \left\langle \hat{Z}_k - \hat{Z}^*, L_k \hat{Z}_k \right\rangle_F - 4F(\hat{Z}_k) \\ &\quad + \left\| \frac{1}{2} \Lambda^{\frac{1}{2}} L_k \hat{Z}_k + \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2. \end{aligned}$$

What's more, by triangle inequality, the last term satisfies that $\left\| \frac{1}{2} \Lambda^{\frac{1}{2}} L_k \hat{Z}_k + \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2 \leq \frac{1}{2} \left\| \Lambda^{\frac{1}{2}} L_k \hat{Z}_k \right\|_F^2 + 2 \left\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2$. We can obtain

$$\begin{aligned} D(k) &\leq - \left\langle \hat{Z}_k - \hat{Z}^*, L_k \hat{Z}_k \right\rangle_F - 4F(\hat{Z}_k) \\ &\quad + \frac{1}{2} \left\| \Lambda^{\frac{1}{2}} L_k \hat{Z}_k \right\|_F^2 + 2 \left\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2 \\ &= - \left\langle \hat{Z}_k - \hat{Z}^*, L_k (\hat{Z}_k - \hat{Z}^*) \right\rangle_F \\ &\quad + \frac{1}{2} \left\langle \hat{Z}_k - \hat{Z}^*, L'_k \Lambda L_k (\hat{Z}_k - \hat{Z}^*) \right\rangle_F \\ &\quad - 4F(\hat{Z}_k) + 2 \left\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2 \\ &= \left\langle \hat{Z}_k - \hat{Z}^*, \left(\frac{1}{2} L'_k \Lambda L_k - L_k \right) (\hat{Z}_k - \hat{Z}^*) \right\rangle_F \\ &\quad - 4F(\hat{Z}_k) + 2 \left\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \right\|_F^2. \end{aligned} \tag{7}$$

Firstly, as we know, L_k is symmetric positive semi-definite for undirected graphs. So, there is a symmetric matrix $L_k^{\frac{1}{2}}$, which satisfies $L_k^{\frac{1}{2}} L_k^{\frac{1}{2}} = L_k$ [3]. $L_k^{\frac{1}{2}}$ is unique and called the square root of L_k . Then, we have

$$\begin{aligned} &L_k - \frac{1}{2} L'_k \Lambda L_k \\ &= L_k^{\frac{1}{2}} \left(E - \frac{L_k^{\frac{1}{2}} \Lambda L_k^{\frac{1}{2}}}{2} \right) L_k^{\frac{1}{2}}. \end{aligned} \tag{8}$$

For the diagonal step-size matrix Λ , since that $(E - \Lambda) \geq 0$, then $L_k^{\frac{1}{2}} L_k^{\frac{1}{2}} - L_k^{\frac{1}{2}} \Lambda L_k^{\frac{1}{2}} \geq 0$. Furthermore, $(E - \frac{L_k^{\frac{1}{2}} \Lambda L_k^{\frac{1}{2}}}{2}) - (E - \frac{L_k}{2}) \geq 0$. The absolute value of eigenvalues of the doubly stochastic matrix is no larger than 1 and $L_k = (E - \mathcal{A}_k) \otimes E_{2n}$. By Lemma 1, the maximum eigenvalue of the Laplacian matrix L_k is 2, then $\lambda_{max}(\frac{1}{2} L_k) = 1$. It is obvious that $E - \frac{L_k}{2} \geq 0$ by Lemma 1.

What's more, because $(E - \frac{L_k \Lambda L_k}{2}) - (E - \frac{L_k}{2}) \geq 0$ and $E - \frac{L_k}{2} \geq 0$, we can also obtain that $(E - \frac{L_k \Lambda L_k}{2}) \geq 0$ by Lemma 1. Therefore, the matrix formula (8) is positive semi-definite and $(\frac{1}{2}L'_k \Lambda L_k - L_k) \leq 0$, which implies that the first term of equation (7) is non-positive.

Secondly, recall the definition of step size limitation ξ_i in Assumption 3. Assume that $T_{i1} = Y_{i,k}^{ri} - A_{ri} X_{i,k}$ and $T_{i2} = Y_{i,k} A'_{ri} - X_{i,k} E_{li} + Q_{li}$. Then, by the definition of $f_i(X_i, Y_i)$ and Assumption 3, we have

$$\begin{aligned} & \| -A'_{ri} T_{i1} - T_{i2} E'_{li} \|_F^2 + \| E_{li} T_{i1} + T_{i2} A_{ri} \|_F^2 \leq 2\xi_i f_i(X_i, Y_i), \\ & \| \nabla_{Z_{i,k}} f_i(Z_{i,k}) \|_F^2 \leq 2\xi_i f_i(Z_{i,k}), \\ & \| \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 \leq 2 \sum_{i=1}^m \xi_i f_i(Z_{i,k}), \\ & \| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 \leq 2 \sum_{i=1}^m \alpha_i \xi_i f_i(Z_{i,k}) \leq 2F(\hat{Z}_k). \end{aligned} \quad (9)$$

Therefore, $2\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 \leq 4F(\hat{Z}_k)$.

In summary, due to the choice of α_i , it follows that $\left\langle \hat{Z}_k - \hat{Z}^*, (\frac{1}{2}L'_k \Lambda L_k - L_k)(\hat{Z}_k - \hat{Z}^*) \right\rangle_F \leq 0$ and $-4F(\hat{Z}_k) + 2\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 \leq 0$. Therefore, $D(k) \leq 0$, which implies that the sequence $\{\hat{Z}_k\}$ is bounded and \hat{Z}^* is a Lyapunov stable equilibrium of (4).

(ii) From $\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 \leq 2 \sum_{i=1}^m \alpha_i \xi_i f_i(Z_{i,k})$ in (9), we deduce that $2\| \Lambda^{\frac{1}{2}} \nabla_{\hat{Z}_k} F(\hat{Z}_k) \|_F^2 - 4F(\hat{Z}_k) \leq -4 \sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(Z_{i,k}) \leq 0$. Denote $4 \sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(Z_{i,k})$ as a_k such that $a_k \geq 0$. Summing (7) over k , we have

$$\| \Lambda^{-\frac{1}{2}} (\hat{Z}_{k+1} - \hat{Z}^*) \|_F^2 - \| \Lambda^{-\frac{1}{2}} (\hat{Z}_0 - \hat{Z}^*) \|_F^2 \leq - \sum_{s=1}^k a_s. \quad (10)$$

Since the variable sequence is bounded in (i), then we have $\sum_{s=1}^{\infty} a_s < \infty$ as $k \rightarrow \infty$, which implies $\lim_{k \rightarrow \infty} a_k = 0$. By the definition of a_k , we have that $\lim_{k \rightarrow \infty} f_i(Z_{i,k}) = 0$ for all agent i .

Because the function f_i is convex, it follows from Jensen's inequality [4] that $4k \sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) \leq \sum_{s=1}^k a_s$. By equation (10),

$$\sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) \leq \frac{\| \Lambda^{-\frac{1}{2}} (\hat{Z}_0 - \hat{Z}^*) \|_F^2}{4k},$$

and hence, $\sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) = O(\frac{1}{k})$.

(iii) By part (ii), $\lim_{k \rightarrow \infty} f_i(Z_{i,k}) = 0$ for all agent i . Since $e_{i,k} = \alpha_i \nabla_{Z_{i,k}} f_i(Z_{i,k})$, it is straightforward that $\lim_{k \rightarrow \infty} e_{i,k} = 0$ for all agent i . \square

Recall that $Z_{i,k+1} = \sum_{j=1}^m w_{ij,k} Z_{j,k} - e_{i,k}$, where $e_{i,k} = \alpha_i \nabla_{Z_{i,k}} f_i(Z_{i,k})$. Define a transition matrix $\Phi(k, s)$ from time s to time k such that

$$\Phi(k, s) = W_s W_{s+1} \cdots W_{k-1} W_k, \quad (11)$$

where $\Phi(k, k) = W_k$ for all k . It follows from the transition matrix and the proposed algorithm (6) that the relation between $Z_{i,k+1}$ and the estimates $Z_{1,s}, \dots, Z_{m,s}$ at time $s \leq k$ is given by

$$\begin{aligned} Z_{i,k+1} &= \sum_{j=1}^m \Phi(k, s)_{ij} Z_{j,s} \\ &+ \sum_{r=s+1}^k \left(\sum_{j=1}^m \Phi(k, s)_{ij} e_{j,r-1} \right) + e_{i,k}. \end{aligned} \quad (12)$$

We define an auxiliary sequence $\{H_k\}$, where H_k is given by $H_k = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^m w_{ij,k} Z_{j,k}$. Since the matrix $W(k)$ is doubly stochastic, it follows that:

$$H_k = \frac{1}{m} \sum_{j=1}^m Z_{j,k}.$$

Furthermore, by the doubly stochasticity of the transition matrix, it follows from the relations equation (12) that

$$H_k = \frac{1}{m} \sum_{j=1}^m Z_{j,s} + \frac{1}{m} \sum_{r=s+1}^k \left(\sum_{j=1}^m e_{j,r-1} \right), \quad \forall k \geq s. \quad (13)$$

The next lemma shows that the limiting behavior of the agent's estimate $Z_{i,k}$ is the same as the limiting behavior of H_k as $k \rightarrow \infty$.

Lemma 3: Let Assumptions 1, 2, 4 and 5 hold. For all agent i , $\lim_{k \rightarrow \infty} \|Z_{i,k} - H_k\|_F = 0$.

Proof: Because the matrix $W(k)$ is a doubly stochastic matrix, the transition matrix $\Phi(k, s)$ is also a doubly stochastic matrix. By Proposition 1 in [21], the entries $[\Phi(k, s)]_{ij}$ of the transition matrix converge to $\frac{1}{m}$ as $k \rightarrow \infty$ with a linear rate for all $i, j \in \{1 \cdots m\}$,

$$\left| [\Phi(k, s)]_{ij} - \frac{1}{m} \right| \leq 2 \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} (1 - \eta^{B_0})^{\frac{k-s}{B_0}}, \quad (14)$$

for all $k \geq s$, where η is the lower bound of Assumption 5, $B_0 = (m-1)B$ and B is the intercommunication interval bound of Assumption 4. By the Lemma 2 (iii), we have $\lim_{k \rightarrow \infty} e_{i,k} = 0$ for all agent i . By the relationship (14) and Lemma 4 in [21], we get that $\lim_{k \rightarrow \infty} \|Z_{i,k} - H_k\|_F = 0$.

□

We have showed that the sequence $\{\hat{Z}_k\}$ is bounded and any optimal solution in the optimal solution set is Lyapunov stable. What's more, every agent's state $Z_{i,k}$ converges to one same

individual average H_k as $k \rightarrow \infty$. Next, we will prove that the sequence $\{\hat{Z}_k\}$ converges to one same optimal solution in the optimal solution set. Before that, we need a related lemma.

Lemma 4: Let the sequence $\{\hat{Z}_k\}$ be generated by algorithm (5). Define $\Omega(\hat{Z}(\cdot))$ as the positive limit set. If $\hat{Z} \in \Omega(\hat{Z}(\cdot))$ is a Lyapunov stable equilibrium point of algorithm (5), then $\hat{Z} = \lim_{k \rightarrow \infty} \hat{Z}_k$ and $\Omega(\hat{Z}(\cdot)) = \{\hat{Z}\}$.

Proof: The proof is similar to that of Proposition 3.1 in [12] and is omitted. \square

Then we are ready to give the proof of Theorem 1.

Proof of Theorem 1: From Lemmas 2 (ii), (iii) and 3, $\{\hat{Z}_k\}$ converges to the set of equilibria of algorithm (5). Since every equilibrium point of algorithm (5) is a Lyapunov stable equilibrium followed by Lemma 2 (i). It follows from Lemma 4 that $\{\hat{Z}_k\}$ converges to one equilibrium of algorithm (5). Equivalently, the sequence $\{\hat{Z}_{i,k}\}$ converges to one same optimal solution for all agent i . From Lemma 2 (ii), $\sum_{i=1}^m (1 - \alpha_i \xi_i) f_i(\bar{Z}_{i,k}) = O(\frac{1}{k})$, where $\bar{Z}_{i,k}$ is time average state variable of agent i . \square

B. Proof of Theorem 2

To begin with, the function $f_i(X_i, Y_i)$ can be rewritten as a quadratic form by matrix vectorization. In particular, $vec(ABC) = (A \otimes C')vec(B)$. Then,

$$\begin{aligned}
f_i &= \frac{1}{2} \|Y_i^{ri} - A_{ri} X_i\|_F^2 + \frac{1}{2} \|Y_i A'_{ri} - X_i E_{li} + Q_{li}\|_F^2 \\
&= \frac{1}{2} \left\| \begin{bmatrix} E_{ri} & -A_{ri} \end{bmatrix} \begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right\|_F^2 \\
&\quad + \frac{1}{2} \left\| \begin{bmatrix} Y_i - Y_i^* & X_i - X_i^* \end{bmatrix} \begin{bmatrix} A'_{ri} \\ -E_{li} \end{bmatrix} \right\|_F^2 \\
&= \frac{1}{2} \left\| \text{vec} \left(\begin{bmatrix} E_{ri} & -A_{ri} \end{bmatrix} \begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right) \right\|^2 \\
&\quad + \frac{1}{2} \left\| \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* & X_i - X_i^* \end{bmatrix} \begin{bmatrix} A'_{ri} \\ -E_{li} \end{bmatrix} \right) \right\|^2 \\
&= \frac{1}{2} \left\| \begin{bmatrix} E_{ri} & -A_{ri} \end{bmatrix} \otimes E_n \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right) \right\|^2 \\
&\quad + \frac{1}{2} \left\| E_n \otimes \begin{bmatrix} A_{ri} & -E'_{li} \end{bmatrix} \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* & X_i - X_i^* \end{bmatrix} \right) \right\|^2 \\
&= \frac{1}{2} \|C_{i1} \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right)\|^2 \\
&\quad + \frac{1}{2} \|C_{i0} \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* & X_i - X_i^* \end{bmatrix} \right)\|^2,
\end{aligned}$$

where $C_{i1} = \begin{bmatrix} E_{ri} & -A_{ri} \end{bmatrix} \otimes E_n \in \mathbb{R}^{n n_i \times 2n^2}$ and $C_{i0} = E_n \otimes \begin{bmatrix} A_{ri} & -E'_{li} \end{bmatrix} \in \mathbb{R}^{n n_i \times 2n^2}$. Through

finite elementary transformations of C_{i0} , we get $f_i(X_i, Y_i) = \frac{1}{2} \left\| C_{i1} \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right) \right\|_F^2 +$

$$\frac{1}{2} \left\| C_{i2} \text{vec} \left(\begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \right) \right\|_F^2 = \frac{1}{2} \langle \mathcal{Z}_{vi}, P_i \mathcal{Z}_{vi} \rangle_F, \text{ where } \mathcal{Z}_{vi} = \text{vec} \begin{bmatrix} Y_i - Y_i^* \\ X_i - X_i^* \end{bmatrix} \in \mathbb{R}^{2n^2} \text{ and}$$

$$P_i = (C'_{i1} C_{i1} + C'_{i2} C_{i2}). \quad (15)$$

In the equation (15), C_{i2} is obtained from C_{i0} through finite elementary transformations.

$$\text{Define } \mathcal{Z} = \begin{bmatrix} \mathcal{Z}_{v1} \\ \vdots \\ \mathcal{Z}_{vm} \end{bmatrix} = \text{vec}(\hat{Z} - \hat{Z}^*) \in \mathbb{R}^{2mn^2}. \text{ With the formula } f_i(X_i, Y_i) = \frac{1}{2} \langle \mathcal{Z}_{vi}, P_i \mathcal{Z}_{vi} \rangle_F,$$

we rewrite $F(\hat{X}, \hat{Y})$ in problem (4) as

$$F(\hat{X}, \hat{Y}) = \frac{1}{2} \langle \mathcal{Z}, \mathbf{P} \mathcal{Z} \rangle_F, \quad (16)$$

where $\mathbf{P} = \text{diag}\{P_1, \dots, P_m\} \in \mathbb{R}^{2mn^2 \times 2mn^2}$.

Since $\hat{Z}_{k+1} - \hat{Z}^* = (\hat{Z}_k - \hat{Z}^*) - \frac{1}{2} \Lambda L_k (\hat{Z}_k - \hat{Z}^*) - \Lambda \nabla_{\hat{Z}_k} F(\hat{Z}_k)$, using the matrix vectorization idea, we obtain

$$\mathcal{Z}_{k+1} = \mathcal{Z}_k - \left(\frac{\Lambda L_k}{2} \otimes E_n \right) \mathcal{Z}_k - (\Lambda \otimes E_n) \mathbf{P} \mathcal{Z}_k. \quad (17)$$

where $L_k = \mathcal{L}_k \otimes E_{2n} \in \mathbb{R}^{2mn \times 2mn}$ and $\Lambda = \text{diag}\{\alpha_1 \otimes E_{2n}, \dots, \alpha_m \otimes E_{2n}\} \in \mathbb{R}^{2mn \times 2mn}$. Let \mathbf{L}_k denote $(L_k \otimes E_n)$ and let $\mathbf{\Lambda}$ denote $\Lambda \otimes E_n$. The equation (17) can be further rewritten as

$$\mathcal{Z}_{k+1} = \mathcal{Z}_k - \left(\frac{\mathbf{\Lambda} \mathbf{L}_k}{2} + \mathbf{\Lambda} \mathbf{P} \right) \mathcal{Z}_k. \quad (18)$$

To proceed, we recall some results on *semistability* of dynamic systems. The notion of *semistability* pertains to a continuum of initial state dependent equilibria, implying that any system trajectory converges to a equilibrium dependent on its initial state [11]. *Semistability* is the property where every trajectory that starts in a neighborhood of a Lyapunov stable equilibrium converges to a (possibly different) Lyapunov stable equilibrium.

A discrete-time switched linear system on \mathbb{R}^n is

$$x(k+1) = A(k)x(k), \quad (19)$$

where variable state $x(k)$ evolves by switching over a finite dynamic coefficient matrix sequence $\{A(k)\}$.

By equation (18), the proposed algorithm can be considered as a discrete-time switched linear system whose variable is $\hat{Z}_k - \hat{Z}^*$. For the problem (4), if Assumption 1 holds, then the algorithm's corresponding switched linear system has a continuum of equilibria.

Based on *semistability* theories of switched linear systems, we get the desired result. Before giving the proof of convergence rate, we present a related lemma, which has been well studied in [22].

Lemma 5: A switched linear system (19) is linearly semistable under arbitrary switching if and only if it is semistable under arbitrary switching.

In Lemma 5, the linear semistability of a system means that the system's states converge to one stable point at a linear rate.

Proof of Theorem 2: Suppose Assumption 6 holds, \mathcal{G}_k is connected for any iteration time k . Then, the variable states converge to one optimal solution under arbitrary switching topologies by Theorem 1.

Note that the proposed algorithm can be considered as a discrete-time switched linear system. Because variable states converge to one Lyapunov stable equilibrium of the solution set, the switched linear system is semistable under arbitrary switching topologies. By Lemma 5, the switched linear system corresponding to the proposed algorithm is linearly semistable under arbitrary switching topologies. If Assumption 1 holds, the stable point set of the switched linear system is the set where $\hat{Z}_k = \hat{Z}^*$ and $Z_i^* = Z_j^*$. As a result, we conclude that $\{\hat{Z}_k\}$ converges to one optimal solution \hat{Z}^* at a linear rate. \square

V. SIMULATION

In this section, some numerical simulations are presented to show the efficacy of the proposed algorithm.

Example: For linear time-invariant systems, the Lyapunov test for controllability is $AXA' - X = -BB'$ [10]. Compared to problem (1), $Q = BB'$, which is symmetric. Suppose that the system information is distributed among undirected graphs containing 5 agents. The matrices

are decomposed as $A = \begin{bmatrix} A_{r1} \\ \vdots \\ A_{r5} \end{bmatrix}$ and $Q = \begin{bmatrix} Q_{l1} \\ \vdots \\ Q_{l5} \end{bmatrix}$.

Every agent i only knows local information of system A_{ri} , Q_{li} and communicates with neighbors over undirected graphs. We consider two types of time-varying undirected topologies, which are illustrated in Figs. 1,2. For each type of topological graphs, we select one topology randomly from them at every iteration time k . Note that there can be many time-varying topologies. Here, we only give three of them for simplicity. Every agent updates its states according to the proposed algorithm (5). The constant step sizes for different agents are listed as a vector sp .

We provide a linear time-invariant system for example, which is controllable by the rank test of the controllable matrix. The coefficient matrices of the selected system A and B are listed in table I and the matrix dimension $n = 10$. Using algorithm (5) with iteration $k = 6000$, we solve the Lyapunov matrix equation to verify the system's controllability. The simulation results are shown in Figs. 3-5. Furthermore, all eigenvalues of the convergence matrix are positive, which

$A =$	0.0061	0.1355	0.0998	0.1051	0.1085	0.0007	0.1095	0.0817	0.0036	0.0625	$B =$	1	1
	0.1492	0.1171	0.0385	0.0708	0.0791	0.0387	0.0523	0.0311	0.0163	0.1427		1	0
	0.1014	0.1360	0.0263	0.0603	0.0087	0.0481	0.0586	0.0165	0.1036	0.0670		0	0
	0.1255	0.0109	0.0290	0.1362	0.1128	0.0086	0.0244	0.1072	0.1397	0.0689		0	1
	0.1439	0.1434	0.0378	0.0980	0.0884	0.0410	0.1360	0.1231	0.1216	0.0058		1	0
	0.0268	0.1180	0.1009	0.0915	0.0303	0.1404	0.0200	0.1167	0.0305	0.1475		0	0
	0.0944	0.1159	0.0598	0.1209	0.1100	0.0749	0.0108	0.1298	0.0572	0.0684		1	0
	0.1189	0.0558	0.0391	0.1120	0.0946	0.0952	0.1077	0.0560	0.0840	0.0900		1	1
	0.0958	0.0067	0.0693	0.0899	0.1106	0.0007	0.0687	0.0153	0.0499	0.0388		0	1
	0.0828	0.0712	0.0463	0.1241	0.0535	0.1397	0.0728	0.1038	0.1492	0.0129		0	0

TABLE I

COEFFICIENT MATRICES A AND B

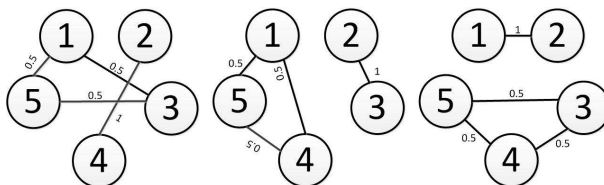


Fig. 1. uniformly connected graphs

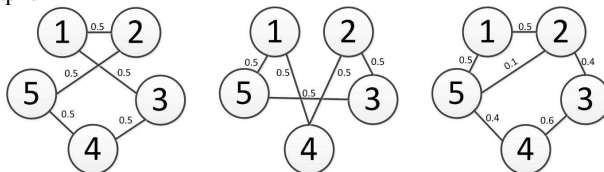


Fig. 2. connected graphs

implies that the matrix is positive definite and the linear time-invariant system is controllable. The result is the same as the rank test of controllable matrix, verifying the correctness of the proposed algorithm.

In Fig. 3, the trajectories of row elements of one agent's estimate X are displayed, which shows that the element values converge to one solution. Fig. 4 shows logarithmic trajectories of $\|AXA' - X + Q\|_F$ over time-varying topologies and Fig. 5 shows trajectories of $D(\bar{X}) = \sum_{i=1}^m \left\| \sum_{j=1}^m (X_i - X_j) \right\|_F, \forall i, j \in \{1, 2, 3, 4, 5\}$. The lines in Fig. 4 and Fig. 5 are demonstrated in Table II. In Fig. 4, all curves over time-varying topologies converge to 0, implying that the convergent matrix is a solution to the discrete-time Lyapunov matrix equation (1). In addition, the curve with half step sizes $sp/2$ converges slower than the curve over uniformly connected

Lines	Step sizes	Topological graphs
line1	sp	uniformly connected
line2	sp	connected
line3	$sp/2$	uniformly connected

TABLE II
PARAMETER TABLE

graphs, demonstrating the influence of step sizes on the convergence rate of proposed algorithm. All curves in Fig. 5 converge to 0, which implies that the solution states of all agents converge to the same matrix. Since every graph in connected topologies is connected, the curve converges more smoothly by the detailed diagram in Fig. 5.

VI. CONCLUSION

The paper has studied the problem of solving large scale discrete-time Lyapunov matrix equations in a distributed way over dynamic multi-agent networks. A distributed algorithm with constant step sizes has been proposed and proofs for convergence properties over time-varying topologies have been presented. Constant step sizes of the proposed algorithm are uncoordinated, depending on every agent's local information. Variable estimates generated by the proposed algorithm converge to an optimal solution of the solution set at a linear convergence rate.

REFERENCES

- [1] B. Anderson, S. Mou, A S. Morse, and U Helmke. Decentralized gradient algorithm for solution of a linear equation. *Numerical Algebra, Control and Optimization*, 6(3):319–328, 2016.
- [2] P. Benner, M. Heinkenschloss, J. Saak, and H. K. Weichelt. An inexact low-rank Newton-ADI method for large-scale algebraic Riccati equations. *Applied Numerical Mathematics*, 108:125–142, 2016.
- [3] D. P. Bertsekas. *Nonlinear Programming: 3rd Edition*. Athena Scientific, Belmont, Massachusetts, USA, 2016.
- [4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
- [5] C. Chang, J. Cortes, and S. Martinez. Scheduled-asynchronous distributed optimization for optimal power flow. *IEEE Transactions on Control of Network Systems*, pages 1–1, 2018.
- [6] A. Falsone, K. Margellos, S. Garatti, and M. Prandini. Dual decomposition for multi-agent distributed optimization with coupling constraints. *Automatica*, 84:149 – 158, 2017.
- [7] P. Giesl and S. Hafstein. Review on computational methods for Lyapunov functions. *Discrete and Continuous Dynamical Systems - Series B*, 20(8):2291–2331, 2015.
- [8] A. Haber and M. Verhaegen. Sparse solution of the Lyapunov equation for large-scale interconnected systems. *Automatica*, 73:256–268, 2016.

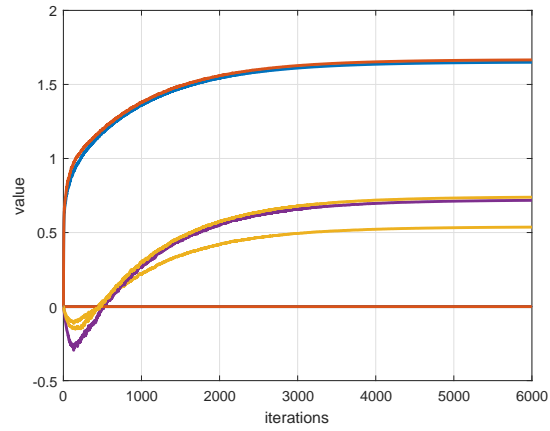


Fig. 3. The trajectories of row elements of X

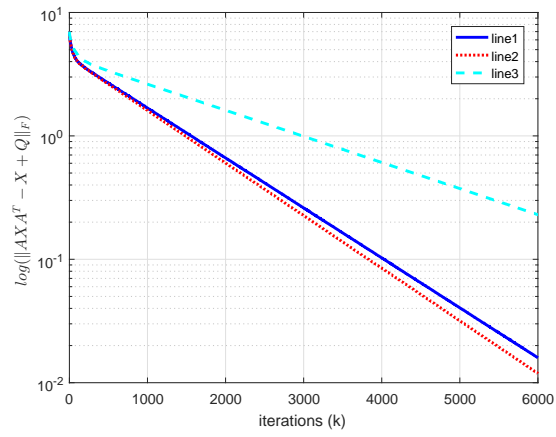


Fig. 4. The trajectories of $\|AXA^T - X + Q\|_F$ versus iterations

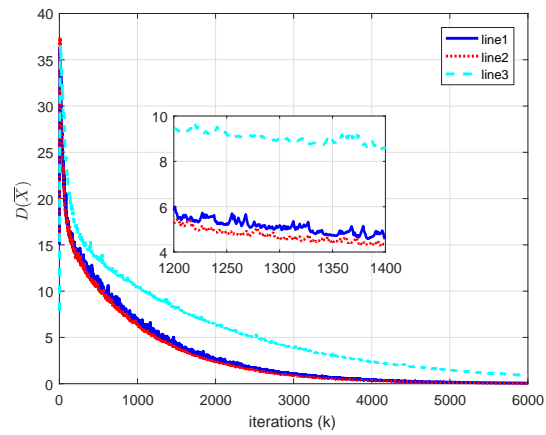


Fig. 5. The trajectories of $D(\bar{X})$ versus iterations

- [9] S. J. Hammarling. Numerical solution of the stable, non-negative definite Lyapunov equation. *IMA Journal of Numerical Analysis*, 2(3):303–323, 1982.
- [10] J. P. Hespanha. *Linear Systems Theory*. Princeton University Press, 2009.
- [11] Q. Hui. Semistability and robustness analysis for switched systems. *European Journal of Control*, 17(1):73–88, 2011.
- [12] Q. Hui, W. M. Haddad, and S. P. Bhat. Semistability, finite-time stability, differential inclusions, and discontinuous dynamical systems having a continuum of equilibria. *IEEE Transactions on Automatic Control*, 54(10):2465–2470, Oct 2009.
- [13] X. Jiang, X. Zeng, J. Sun, and J. Chen. Distributed algorithm for discrete-time Lyapunov equations. In *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 768–773, Singapore, Nov 2018.
- [14] S. Liang, L. Wang, and G. Yin. Distributed quasi-monotone subgradient algorithm for nonsmooth convex optimization over directed graphs. *Automatica*, 101:175–181, 2019.
- [15] J. Liu, A. S. Morse, A. Nedic, and T. Basar. Exponential convergence of a distributed algorithm for solving linear algebraic equations. *Automatica*, 83:37–46, 2017.
- [16] Y. Liu, C. Lageman, B. D. O. Anderson, and G. Shi. An Arrow-Hurwicz-Uzawa type flow as least squares solver for network linear equations. *Automatica*, 100:187–193, 2019.
- [17] I. Lobel and A. Ozdaglar. Distributed subgradient methods for convex optimization over random networks. *IEEE Transactions on Automatic Control*, 56(6):1291–1306, 2011.
- [18] V. S. Mai and E. H. Abed. Distributed optimization over directed graphs with row stochasticity and constraint regularity. *Automatica*, 102:94 – 104, 2019.
- [19] S. Mou, J. Liu, and A. S. Morse. A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions on Automatic Control*, 60(11):2863–2878, 2015.
- [20] I. Necoara and V. Nedelcu. On linear convergence of a distributed dual gradient algorithm for linearly constrained separable convex problems. *Automatica*, 55:209–216, 2015.
- [21] A. Nedic, A. Ozdaglar, and P. A. Parrilo. Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 55(4):922–938, April 2010.
- [22] J. Shen, J. Hu, and Q. Hui. Semistability of switched linear systems with application to PageRank algorithms. *European Journal of Control*, 20(3):132–140, 2014.
- [23] G. Shi, B. D. O. Anderson, and U. Helmke. Network flows that solve linear equations. *IEEE Transactions on Automatic Control*, 62(6):2659–2674, June 2017.
- [24] W. Shi, Q. Ling, G. Wu, and W. Yin. Extra: An exact first-order algorithm for decentralized consensus optimization. *SIAM Journal on Optimization*, 25(2):944–966, 2015.
- [25] A. Wu, H. Sun, and Y. Zhang. An SOR implicit iterative algorithm for coupled Lyapunov equations. *Automatica*, 97:38–47, 2018.
- [26] Z. Wu, M. Fu, Y. Xu, and R. Lu. A distributed Kalman filtering algorithm with fast finite-time convergence for sensor networks. *Automatica*, 95:63–72, 2018.
- [27] C. Xi, V. S. Mai, R. Xin, E. H. Abed, and U. A. Khan. Linear convergence in optimization over directed graphs with row-stochastic matrices. *IEEE Transactions on Automatic Control*, 63(10):3558–3565, Oct 2018.
- [28] W. Xia, Z. Meng, G. Shi, and K. H. Johansson. Optimal tradeoff between instantaneous and delayed neighbor information in consensus algorithms. *Automatica*, 83:116–123, 2017.
- [29] J. Xu, S. Zhu, Y. C. Soh, and L. Xie. Augmented distributed gradient methods for multi-agent optimization under uncoordinated constant stepsizes. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 2055–2060, Japan, 2015.

- [30] X. Zeng, S. Liang, Y. Hong, and J. Chen. Distributed computation of linear matrix equations: An optimization perspective. *IEEE Transactions on Automatic Control*, pages 1–1, 2018.