Optimal Finite-Time Distributed Linear Averaging

Qing Hui

***Abstract*— A new optimal finite-time distributed linear averaging (OFTDLA) problem is presented in this paper. This problem is motivated from the distributed averaging problem which arises in the context of distributed algorithms in computer science and coordination of groups of autonomous agents in engineering. The aim of the OFTDLA problem is to compute the average of the initial values in finite-time steps at nodes of a graph through an optimal distributed algorithm in which the nodes in the graph can only communicate with their neighbors. Optimality is given by a minimization problem of a quadratic cost functional under finite-time horizon. We show that this problem has a very close relationship with the notion of semistability. By developing new necessary and sufficient conditions for semistability of discrete-time linear systems, we convert the original OFTDLA problem into two equivalent optimization problems. One of them is a convex optimization problem and can be solved by using semidefinite programming methods.**

I. Introduction

The distributed averaging problem arises in the context of coordination of networks of autonomous agents, and in particular, the consensus or agreement problem among the agents. Distributed consensus problems have been studied extensively in the computer science literature [1]–[4]. Recently it has found a wide range of applications, in areas such as formation control of underwater autonomous vehicles [5], coordination of mobile robots [6], [7], and sensor networks [8], [9].

Even though many consensus protocol algorithms have been developed over the last several years in the literature (see [10]–[19] and the numerous references therein), cost-functional optimality properties of these algorithms have been largely ignored. Cost-functional optimality here refers to minimization of a quadratic cost functional for coordination algorithms subject to stability and network connectivity constraints. Furthermore, optimality over finite horizon for consensus algorithms has never been considered in the literature. It is important to note that this optimality is different from the standard LQR problem and the one in [20]–[22] since the standard LQR does not consider semistability, and [20]–[22] neither consider optimality of quadratic cost functionals nor consider optimality under finite horizon. In this paper, we address the equivalent formulation of optimal finite-time semistable linear iteration algorithms with a quadratic cost functional, which is named for the *optimal finite-time distributed linear averaging* (OFTDLA) problem. To this end, we develop necessary and sufficient conditions for semistability of linear discrete-time systems characterized by *recursive* Lyapunov equations and the notion of *weak semiobservability*. Built on these relevant results, we present two equivalent optimization problems regarding the OFTDLA problem and one of them can be solved by using semidefinite programming.

The organization of this paper is as follows. Section II gives the formulation of the OFTDLA problem motivated from the distributed averaging problem. Section III explores necessary conditions for optimality and semistability of the OFTDLA problem. Based on these results, necessary and sufficient conditions for semistability of the OFTDLA problem are developed by introducing a new recursive Lyapunov equation and the notion of weak semiobservability in Section IV. These necessary and sufficient conditions turn out to be the bridge establishing the equivalent OFTDLA problems. The new convex optimization equivalent formulation is shown in Section V in which two equivalent OFTDLA problems are proposed. Finally, some concluding remarks are provided in Section VI.

This work was supported in part by a research grant from Texas Tech University.

Q. Hui is with the Department of Mechanical Engineering, Texas Tech University, Lubbock, TX 79409-1021, USA (qing.hui@ttu.edu).

II. Problem Formulation: Optimal Finite-Time Distributed Linear Averaging

The notion we use in this paper is fairly standard. Specifically, $\mathbb{R}$ (resp., $\mathbb{C}$) denotes the set of real (resp., complex) numbers, $\mathbb{Z}_+$ denotes the set of positive integers, $\overline{\mathbb{Z}}_+$ denotes the set of nonnegative integers, $\mathbb{R}^n$ (resp., $\mathbb{C}^n$) denotes the set of $n \times 1$ real (resp., complex) column vectors, $\mathbb{R}^{n\times m}$ (resp., $\mathbb{C}^{n\times m}$) denotes the set of $n \times m$ real (resp., complex) matrices, $(\cdot)^{\mathrm{T}}$ denotes transpose, $(\cdot)^*$ denotes complex conjugate transpose, $(\cdot)^{\#}$ denotes the group generalized inverse, and I_n or I denotes the $n \times n$ identity matrix. Furthermore, we write $\|\cdot\|$ for the Euclidean vector norm, $\mathcal{R}(A)$ and $\mathcal{N}(A)$ for the range space and the null space of a matrix A, $\mathrm{spec}(A)$ for the spectrum of the square matrix A, $\mathrm{tr}(\cdot)$ for the trace operator, $\mathbb{E}$ for the expectation operator, and $A \geq 0$ (resp., $A > 0$) to denote the fact that the Hermitian matrix A is nonnegative (resp., positive) definite. Finally, we write $\mathcal{B}_\varepsilon(x)$, $x \in \mathbb{R}^n$, $\varepsilon > 0$, for the open ball with *radius* ε and *center* x, $\otimes$ for the Kronecker product, and $\mathrm{vec}(\cdot)$ for the column stacking operator.

We consider a network characterized by a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of the set of nodes $\mathcal{V} = \{1, \ldots, q\}$ and the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, where each edge $\{i, j\} \in \mathcal{E}$ is an unordered pair of distinct nodes. The set of neighbors of node i is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} : \{i, j\} \in \mathcal{E}\}$. Finally, we denote the value of the node $i \in \{1, \ldots, q\}$ at time t by $x_i(t) \in \mathbb{R}$.

ISBN 978-963-311-370-7

Each node i holds an initial value on the network $x_i(0) \in \mathbb{R}$. The network gives the allowed communication between two nodes if and only of they are neighbors. We are interested in computing the average of the initial values, $(1/q)\sum_{i=1}^{q} x_i(0)$, via an *optimal* distributed algorithm in which the nodes only communicate with their neighbors.

In this paper, we consider distributed linear iterations given by the form

$$x_i(t+1) = W_{(i,i)}x_i(t) + \sum_{j\in\mathcal{N}_i} W_{(i,j)}x_j(t), \quad i = 1,\dots,q, \quad t \in \overline{\mathbb{Z}}_+, \tag{1}$$

where $W_{(i,j)}$ denotes the weight on x_j at node i. Letting $W_{(i,j)} = 0$ for $j \neq \mathcal{N}_i$, this iteration can be rewritten as a compact form

$$x(t+1) = Wx(t), \quad t \in \overline{\mathbb{Z}}_+, \tag{2}$$

where $x(t) = [x_1(t),\dots,x_q(t)]^{\mathrm{T}} \in \mathbb{R}^q$. The constraint on the matrix W can be expressed as $W \in \mathcal{W}$, where

$$\mathcal{W} = \{W \in \mathbb{R}^{q\times q} : W_{(i,j)} = 0 \text{ if } \{i,j\} \notin \mathcal{E} \text{ and } i \neq j\}. \tag{3}$$

Recall from [23] that a matrix $A \in \mathbb{R}^{q\times q}$ is (discrete-time) *semistable* if $\mathrm{spec}(A) \subset \{s \in \mathbb{C} : |s| < 1\} \cup \{1\}$ and, if $1 \in \mathrm{spec}(A)$, then 1 is semisimple.

Lemma 2.1: Consider the linear iteration (2). Then $\lim_{t\to\infty} x(t)$ exists if and only if W is semistable. If W is semistable, then $\lim_{t\to\infty} W^t = I_q - (W - I_q)(W - I_q)^{\#}$.

Proof: The result is Proposition 11.9.2 of [23]. ■

The linear iteration (2) implies that $x(t) = W^t x(0)$ for $t \in \overline{\mathbb{Z}}_+$. We want to choose the weight matrix W so that for any initial value $x(0)$, $\lim_{t\to\infty} x(t) = \alpha\mathbf{1}$ and the cost functional

$$\begin{aligned} J(W, x(0)) &= \sum_{t=0}^{N} [(x(t) - \alpha\mathbf{1})^{\mathrm{T}} Q(x(t) - \alpha\mathbf{1}) \\ &\quad + (Wx(t) - \alpha W\mathbf{1})^{\mathrm{T}} R(Wx(t) - \alpha W\mathbf{1})] \\ &= \sum_{t=0}^{N} \|Q^{1/2}(x(t) - \alpha\mathbf{1})\|^2 \\ &\quad + \|R^{1/2}(Wx(t) - \alpha W\mathbf{1})\|^2 \end{aligned} \tag{4}$$

is minimized, where N is a positive integer, $\alpha \in \mathbb{R}$, $\mathbf{1} = [1,\dots,1]^{\mathrm{T}} \in \mathbb{R}^q$, $Q = Q^{\mathrm{T}} \geq 0$, and $R = R^{\mathrm{T}} > 0$. The cost functional (4) is motivated by the quadratic cost functional in the LQR control theory whereas the control input $u(t)$ can be viewed as $Wx(t)$.

Lemma 2.2: For any $x(0) \in \mathbb{R}^q$, $\lim_{t\to\infty} x(t) = \alpha\mathbf{1}$ if and only if

$$\mathbf{1}^{\mathrm{T}} W = \mathbf{1}^{\mathrm{T}}, \tag{5}$$

$$W\mathbf{1} = \mathbf{1}, \tag{6}$$

$$W \text{ is semistable.} \tag{7}$$

Proof: The assertion follows from Lemma 2.1 and Theorem 1 of [20]. ■

Lemma 2.3: If for any $x(0) \in \mathbb{R}^q$, $\lim_{t\to\infty} x(t) = \alpha\mathbf{1}$, then

$$\alpha = \frac{1}{q}\mathbf{1}^{\mathrm{T}} x(0) = \frac{1}{q}\sum_{i=1}^{q} x_i(0). \tag{8}$$

Proof: The result follows from Lemma 2.2. ■

Hence, the *optimal distributed linear averaging* (ODLA) problem [24] over finite horizon can be formulated as

$$\begin{aligned} &\text{minimize } J(W, x(0)) \\ &\text{subject to } W \in \mathcal{W}, \quad \mathbf{1}^{\mathrm{T}} W = \mathbf{1}^{\mathrm{T}}, \quad W\mathbf{1} = \mathbf{1}, \\ &\qquad W \text{ is semistable}, \quad x(t+1) = Wx(t). \end{aligned}$$

In this paper, we further consider *finite-time distributed averaging* [25], [26] via optimal linear iterations, that is, distributed linear averaging in finite-time steps.

Lemma 2.4: Consider (2). If W satisfies the conditions in Lemma 2.2, then for every $i = 1,\dots,q$,

$$\lim_{t\to\infty} x_i(t) = \alpha = \frac{[x_i(q-1), x_i(q-2), \dots, x_i(0)]S}{\mathbf{1}^{\mathrm{T}} S}, \tag{9}$$

where $S \triangleq [1, 1+\alpha_{q-1}, 1+\alpha_{q-2}+\alpha_{q-1}, \dots, 1+\sum_{j=1}^{q-1}\alpha_j]^{\mathrm{T}}$ and $|\lambda I_q - W| = \lambda^q + \alpha_{q-1}\lambda^{q-1} + \cdots + \alpha_1\lambda + \alpha_0$. In other words, if it is possible for node i to obtain the necessary information to calculate the linear averaging $(1/q)\sum_{i=1}^{q} x_i(0)$ via the linear interation (2), then it will require at most $q+1$ time-steps.

Proof: It follows from the Cayley-Hamilton theorem that

$$W^q + \alpha_{q-1}W^{q-1} + \cdots + \alpha_1 W + \alpha_0 I_q = 0. \tag{10}$$

Note that from (2) we have $x(t) = W^t x(0)$ for all $t \in \overline{\mathbb{Z}}_+$. However, from (10), we can see that the value $x(q)$ can equivalently be obtained as

$$\begin{aligned} x(q) &= W^q x(0) \\ &= -(\alpha_{q-1}W^{q-1} + \cdots + \alpha_1 W + \alpha_0 I_q)x(0) \\ &= -\alpha_{q-1}x(q-1) - \cdots - \alpha_1 x(1) - \alpha_0 x(0). \end{aligned} \tag{11}$$

Thus, for all $t \in \overline{\mathbb{Z}}_+$ and for all $i = 1,\dots,q$, $x_i(t)$ satisfies a linear difference equation of the form

$$x_i(t+q) + \alpha_{q-1}x_i(t+q-1) + \cdots + \alpha_1 x_i(t+1) + \alpha_0 x_i(t) = 0. \tag{12}$$

In particular, after time-steps q, each node only needs to know its own q previous values in order to compute the linear iteration (2), and does not need to receive further information from its neighbors. Taking the Z-transform of the above expression, we obtain

$$(z^q + \alpha_{q-1}z^{q-1} + \cdots + \alpha_1 z + \alpha_0)X_i(z) = \sum_{j=0}^{q-1} x_i(j)z^{q-j} + \alpha_{q-1}\sum_{j=0}^{q-2} x_i(j)z^{q-1-j} + \cdots + \alpha_1 z x_i(0). \tag{13}$$

The expression on the right-hand side of the above equation represents the initial conditions of the linear difference equation, and arises from the Z-transform. The term multiplying $X_i(z)$ on the left-hand side of the equation is the characteristic polynomial $|\lambda I_q - W| = \lambda^q + \alpha_{q-1}\lambda^{q-1} + \cdots + \alpha_1\lambda + \alpha_0$. In particular, if W satisfies the conditions in Lemma 2.2, the characteristic polynomial will have a single root at $z = 1$, and all other roots will have magnitude less than 1 (since the roots of the characteristic polynomial are the eigenvalues of W). We can therefore factor $\lambda^q + \alpha_{q-1}\lambda^{q-1} + \cdots + \alpha_1\lambda + \alpha_0$ as

$$z^q + \alpha_{q-1}z^{q-1} + \cdots + \alpha_1 z + \alpha_0 = (z-1)p(z), \quad (14)$$

where

$$\begin{aligned} p(z) &= z^{q-1} + (1+\alpha_{q-1})z^{q-2} \\ &\quad + (1+\alpha_{q-1}+\alpha_{q-2})z^{q-3} \\ &\quad + \cdots + \left(1 + \sum_{j=2}^{q-1}\alpha_j\right)z + \left(1 + \sum_{j=1}^{q-1}\alpha_j\right). \end{aligned} \quad (15)$$

Since the roots of $p(z)$ have magnitude strictly less than 1, we can now use (13) and (15) in the final value theorem to obtain

$$\begin{aligned} \lim_{t\to\infty} x_i(t) &= \lim_{z\to 1}(z-1)X_i(z) \\ &= \frac{[x_i(q-1), x_i(q-2), \ldots, x_i(0)]S}{\mathbf{1}^{\mathrm{T}}S}, \end{aligned} \quad (16)$$

where $S = [1, 1+\alpha_{q-1}, 1+\alpha_{q-2}+\alpha_{q-1}, \ldots, 1+\sum_{j=1}^{q-1}\alpha_j]^{\mathrm{T}}$. ■

Remark 2.1: It was shown in [26] that if W is given, then the conclusion of Lemma 2.4 can be strengthened as at most $D+1$ time-steps, where $D+1$ denotes the degree of the *minimal polynomial* of W (i.e., the unique monic polynomial $\mathbb{P}(t)$ of smallest degree such that $\mathbb{P}(W) = 0$). However, since W is not given *a priori* in our problem, the best upper bound for time-steps is $q+1$.

Lemma 2.4 states that finite-time linear averaging can be achieved by iterating (2) $q-1$ times and then using

$$x_i(q) = \frac{[x_i(q-1), x_i(q-2), \ldots, x_i(0)]S}{\mathbf{1}^{\mathrm{T}}S} \quad (17)$$

to obtain the average value α at time $t = q$.

We assume that W satisfies the conditions in Lemma 2.2. Since in this case, W is semistable, it follows that $\lim_{t\to\infty} x(t) = [I_q - (I_q - W)(I_q - W)^{\#}]x(0)$. Moreover, since by Lemma 2.4, $\lim_{t\to\infty} x(t) = \alpha\mathbf{1}$, it follows that $\alpha\mathbf{1} = [I_q - (I_q - W)(I_q - W)^{\#}]x(0)$.

Next, consider the cost functional given by (4) where $N = q-1$, namely,

$$\begin{aligned} \hat{J}(W, x(0)) &= \sum_{t=0}^{q-1}[(x(t)-\alpha\mathbf{1})^{\mathrm{T}}Q(x(t)-\alpha\mathbf{1}) \\ &\quad + (Wx(t) - \alpha W\mathbf{1})^{\mathrm{T}}R(Wx(t) - \alpha W\mathbf{1})] \\ &= \sum_{t=0}^{q-1}\|Q^{1/2}(x(t)-\alpha\mathbf{1})\|^2 \\ &\quad + \|R^{1/2}(Wx(t) - \alpha W\mathbf{1})\|^2. \end{aligned} \quad (18)$$

Hence, the *optimal finite-time distributed linear averaging* (OFTDLA) problem can be formulated as

$$\begin{aligned} &\text{minimize } \hat{J}(W, x(0)) \\ &\text{subject to } W \in \mathcal{W}, \quad \mathbf{1}^{\mathrm{T}}W = \mathbf{1}^{\mathrm{T}}, \quad W\mathbf{1} = \mathbf{1}, \\ &\qquad W \text{ is semistable}, \quad x(t+1) = Wx(t). \end{aligned}$$

III. Necessary Conditions for Optimality

In this section, we derive necessary conditions for optimality of the linear iteration algorithm (2) provided that all the conditions in Lemma 2.2 are guaranteed. Suppose W in the linear iteration (2) satisfies all the conditions in Lemma 2.2. Then it follows from semistability that $\lim_{t\to\infty} x(t) = x_{\mathrm{e}}$, where $x_{\mathrm{e}} = [I_q - (I_q - W)(I_q - W)^{\#}]x(0)$. Now, it follows that

$$\begin{aligned} \hat{J}(W, x_0) &= \sum_{t=0}^{q-1}[x(t) - x_{\mathrm{e}}]^{\mathrm{T}}(Q + W^{\mathrm{T}}RW)[x(t) - x_{\mathrm{e}}] \\ &= \sum_{t=0}^{q-1}[x(0) - x_{\mathrm{e}}]^{\mathrm{T}}(W^t)^{\mathrm{T}}\tilde{R}W^t[x(0) - x_{\mathrm{e}}] \\ &= \sum_{t=0}^{q-1}x^{\mathrm{T}}(0)[(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}(W^t)^{\mathrm{T}}\tilde{R} \\ &\quad W^t(I_q - W)(I_q - W)^{\#}x(0) \\ &= \mathrm{tr}\, x^{\mathrm{T}}(0)\left[\sum_{t=0}^{q-1}[(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}\right. \\ &\quad \left.(W^t)^{\mathrm{T}}\tilde{R}W^t(I_q - W)(I_q - W)^{\#}\right]x(0) \\ &= \mathrm{tr}\left[\sum_{t=0}^{q-1}[(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}(W^t)^{\mathrm{T}}\tilde{R}\right. \\ &\quad \left.W^t(I_q - W)(I_q - W)^{\#}\right]x(0)x^{\mathrm{T}}(0) \\ &= \mathrm{tr}\, PV, \end{aligned} \quad (19)$$

where we assume that the initial state x_0 is a random variable such that $\mathbb{E}[x_0] = 0$ and $\mathbb{E}[x_0x_0^{\mathrm{T}}] = V$, we used the fact that $x(t) - x_{\mathrm{e}} = W^t[x(0) - x_{\mathrm{e}}]$, $\tilde{R} \triangleq Q + W^{\mathrm{T}}RW$, and P is defined by

$$\begin{aligned} P &\triangleq \sum_{t=0}^{q-1}[(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}(W^t)^{\mathrm{T}} \\ &\quad \tilde{R}W^t(I_q - W)(I_q - W)^{\#}. \end{aligned} \quad (20)$$

Clearly, P is well defined.

Next, we define a sequence of matrices by

$$\begin{aligned} P_N &\triangleq \sum_{t=0}^{N}[(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}(W^t)^{\mathrm{T}} \\ &\quad \tilde{R}W^t(I_q - W)(I_q - W)^{\#}. \end{aligned} \quad (21)$$

Clearly, $P_{q-1} = P$, $P_{k+1} \geq P_k$ for all $k \geq 0$, and $P_0 = [(I_q - W)(I_q - W)^{\#}]^{\mathrm{T}}\tilde{R}(I_q - W)(I_q - W)^{\#}$. Furthermore, it follows from [24] that $\lim_{k\to\infty} P_k$ exists.

Lemma 3.1: Let $P_{-1} = 0$. If W in the linear iteration (2) is semistable, then P_k, $k \geq -1$, $k \in \mathbb{Z}$, given by (21) satisfies

$$(I_q-W)^{\mathrm{T}}P_{k+1}(I_q-W) \\ =(I_q-W)^{\mathrm{T}}W^{\mathrm{T}}P_kW(I_q-W) \\ +(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W). \quad (22)$$

Proof: Note that $(I_q-W)(I_q-W)^{\#}(I_q-W)=I_q-W$ and $W(I_q-W)=(I_q-W)W$. Then it follows that

$$(I_q-W)^{\mathrm{T}}P_{k+1}(I_q-W) \\ =\sum_{t=0}^{k+1}(I_q-W)^{\mathrm{T}}(W^t)^{\mathrm{T}}\tilde{R}W^t(I_q-W) \\ =\sum_{t=0}^{k+1}(W^t)^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)W^t \quad (23)$$

and

$$(I_q-W)^{\mathrm{T}}W^{\mathrm{T}}P_kW(I_q-W) \\ =W^{\mathrm{T}}(I_q-W)^{\mathrm{T}}P_k(I_q-W)W \\ =\sum_{t=0}^{k}(W^{t+1})^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)W^{t+1} \\ =\sum_{t=1}^{k+1}(W^t)^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)W^t \\ =\sum_{t=0}^{k+1}(W^t)^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)W^t \\ -(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W). \quad (24)$$

Hence,

$$(I_q-W)^{\mathrm{T}}P_{k+1}(I_q-W)-(I_q-W)^{\mathrm{T}}W^{\mathrm{T}}P_kW(I_q-W) \\ =(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W). \quad (25)$$

■

The following result is immediate.

Lemma 3.2: If W in the linear iteration (2) is semistable, then for every positive semidefinite matrix $\tilde{R}=\tilde{R}^{\mathrm{T}}\in\mathbb{R}^{q\times q}$, there exists a bounded sequence consisting of positive semidefinite matrices $\tilde{P}_k=\tilde{P}_k^{\mathrm{T}}\in\mathbb{R}^{q\times q}$ for all $k\geq -1$, where $\tilde{P}_{-1}=0$, such that $\tilde{P}_{k+1}\geq\tilde{P}_k$ and

$$\tilde{P}_{k+1}=W^{\mathrm{T}}\tilde{P}_kW+(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W). \quad (26)$$

Proof: The existence of $\tilde{P}_k$ is a direct consequence of Lemma 3.1 by setting $\tilde{P}_k=(I_q-W)^{\mathrm{T}}P_k(I_q-W)$. ■

IV. Necessary and Sufficient Conditions for Semistability

Based on the necessary conditions developed in Section III, we come up with a linear-matrix-inequality (LMI) characterization for semistability. Before we state this necessary and sufficient condition, we need the following notion.

Definition 4.1: Let $A\in\mathbb{R}^{n\times n}$ and $C\in\mathbb{R}^{m\times n}$. The pair (A,C) is *weakly semiobservable* if

$$\bigcap_{k=1}^{n}\mathcal{N}(C(A-I_n)^k)=\mathcal{N}(A-I_n). \quad (27)$$

Recall from Definition 2.3 of [27] that the pair (A,C) is *semiobservable* if and only if

$$\bigcap_{k=1}^{n}\mathcal{N}(C(A-I_n)^{k-1})=\mathcal{N}(A-I_n). \quad (28)$$

It is easy to see from the definitions of semiobservability and weak semiobservability that (A,C) is weakly semiobservable if and only if (A,CA) is semiobservable. Motivated by Lemma 2.1 of [27], we have the following lemma.

Lemma 4.1: Let $\tilde{P}_{-1}=0$. Assume that there exist matrices $\tilde{P}_r\geq 0$ and $\tilde{R}\geq 0$ in $\mathbb{R}^{q\times q}$ such that (26) holds and the pair $(W,\tilde{R})$ is weakly semiobservable for every $r\geq -1$, $r\in\mathbb{Z}$.

i) Then $\mathcal{N}(\tilde{P}_k)\subseteq\mathcal{N}(W-I_q)\subseteq\mathcal{N}(\tilde{R}(W-I_q))$.

ii) If, in addition, $\mathcal{N}((W-I_q)^2)\subseteq\mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k\geq 0$, then $\mathcal{N}(W-I_q)\cap\mathcal{R}(W-I_q)=\{0\}$ for every $k\geq 0$.

Proof: *i*) Note that $\mathcal{N}(W-I_q)\subseteq\mathcal{N}(\tilde{R}(W-I_q))$. If $\tilde{P}_kx=0$, then

$$\begin{aligned} 0 &\leq x^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)x \\ &= x^{\mathrm{T}}(\tilde{P}_k-W^{\mathrm{T}}\tilde{P}_{k-1}W)x \\ &= -x^{\mathrm{T}}W^{\mathrm{T}}\tilde{P}_{k-1}Wx\leq 0, \end{aligned} \quad (29)$$

and hence, $x^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_g-W)x=0$ or, equivalently, $\tilde{R}(I_q-W)x=0$. Thus, $\mathcal{N}(\tilde{P}_k)\subseteq\mathcal{N}(\tilde{R}(I_q-W))$.

Next, let $x\in\mathcal{N}(\tilde{P}_{k+1})\subseteq\mathcal{N}(\tilde{R}(W-I_q))$. In this case, it follows from (26) that $0=x^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)x=x^{\mathrm{T}}(\tilde{P}_{k+1}-W^{\mathrm{T}}\tilde{P}_kW)x=-x^{\mathrm{T}}W^{\mathrm{T}}\tilde{P}_kWx=-x^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{P}_k(W-I_q)x$, which implies that $(W-I_q)x\in\mathcal{N}(\tilde{P}_k)\subseteq\mathcal{N}(\tilde{R}(W-I_q))$.

Moreover, if $(W-I_q)^nx\in\mathcal{N}(\tilde{P}_{k+1})\subseteq\mathcal{N}(\tilde{R}(W-I_q))$ for some $n\geq 0$, then

$$\begin{aligned} 0 &= x^{\mathrm{T}}(W^{\mathrm{T}}-I_q)^n\tilde{R}(W-I_q)^nx \\ &= x^{\mathrm{T}}(W^{\mathrm{T}}-I_q)^n(\tilde{P}_{k+1}-W^{\mathrm{T}}\tilde{P}_kW)(W-I_q)^nx \\ &= -x^{\mathrm{T}}(W^{\mathrm{T}}-I_q)^nW^{\mathrm{T}}\tilde{P}_kW(W-I_q)^nx \\ &= -x^{\mathrm{T}}(W^{\mathrm{T}}-I_q)^{n+1}\tilde{P}_k(W-I_q)^{n+1}x, \end{aligned} \quad (30)$$

and hence, $\tilde{P}_k(W-I_q)^{n+1}x=0$, which implies that $(W-I_q)^{n+1}x\in\mathcal{N}(\tilde{P}_k)\subseteq\mathcal{N}(\tilde{R}(W-I_q))$. By induction on n, it follows that $x\in\bigcap_{n=1}^{q}\mathcal{N}(\tilde{R}(W-I_q)^n)$. Now weak semiobservability of $(W,\tilde{R})$ implies that $x\in\mathcal{N}(W-I_q)$. Thus, $\mathcal{N}(\tilde{P}_k)\subseteq\mathcal{N}(W-I_q)\subseteq\mathcal{N}(\tilde{R}(W-I_q))$.

ii) Consider $x\in\mathcal{N}(W-I_q)\cap\mathcal{R}(W-I_q)$. Then $(W-I_q)x=0$ and there exists $z\in\mathbb{R}^q$ such that $x=(W-I_q)z$. It is clear that $\tilde{R}(W-I_q)x=0$ and $z\in\mathcal{N}((W-I_q)^2)$. If $\mathcal{N}((W-I_q)^2)\subseteq\mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k\geq 0$, then

$$\begin{aligned} 0 &= z^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)x \\ &= z^{\mathrm{T}}(\tilde{P}_{k+1}-W^{\mathrm{T}}\tilde{P}_kW)x \\ &= -z^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{P}_kx \\ &= -x^{\mathrm{T}}\tilde{P}_kx, \end{aligned} \quad (31)$$

and hence, $\tilde{P}_k x = 0$. Finally,

$$\begin{aligned} z^{\mathrm{T}}(W-I_q)\tilde{R}(W-I_q)z &= z^{\mathrm{T}}(\tilde{P}_{k+1}-W^{\mathrm{T}}\tilde{P}_kW)z \\ &= z^{\mathrm{T}}\tilde{P}_{k+1}z-(x+z)^{\mathrm{T}}\tilde{P}_k(x+z) \\ &= -x^{\mathrm{T}}\tilde{P}_kx-x^{\mathrm{T}}\tilde{P}_kz-z^{\mathrm{T}}\tilde{P}_kx=0, \end{aligned} \quad (32)$$

and hence, $\tilde{R}(W-I_q)z = 0$. This implies that $z \in \bigcap_{n=1}^{q}\mathcal{N}(\tilde{R}(W-I_q)^n)$. Hence, by weak semiobservability, $(W-I_q)z = x = 0$ as required. ■

The part of the converse result for Lemma 3.2 can be stated as follows.

Lemma 4.2: Assume that for every positive semidefinite matrix $\tilde{R} = \tilde{R}^{\mathrm{T}} \in \mathbb{R}^{q\times q}$ satisfying $(W,\tilde{R})$ is weakly semiobservable, there exists a bounded sequence consisting of positive semidefinite matrices $\tilde{P}_k = \tilde{P}_k^{\mathrm{T}} \in \mathbb{R}^{q\times q}$ such that (26) holds. Furthermore, assume that $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ and $\tilde{P}_{k+1} \geq \tilde{P}_k$ for all $k \geq 0$. Then (2) is semistable.

Proof: Since, by Lemma 4.1, $\mathcal{N}(W-I_q)\cap\mathcal{R}(W-I_q) = \{0\}$, it follows from [28, p. 119] that $W-I_q$ is group invertible. Let $L \triangleq I_q-(W-I_q)(W-I_q)^{\#}$ and note that $L^2 = L$. Hence, L is the unique $q\times q$ matrix satisfying $\mathcal{N}(L) = \mathcal{R}(W-I_q)$, $\mathcal{R}(L) = \mathcal{N}(W-I_q)$, and $Lx = x$ for all $x \in \mathcal{N}(W-I_q)$.

Consider the multiple nonnegative functions

$$V_k(x) = x^{\mathrm{T}}\tilde{P}_kx + x^{\mathrm{T}}L^{\mathrm{T}}Lx. \quad (33)$$

If $V_k(x) = 0$ for some $x \in \mathbb{R}^q$, then $\tilde{P}_kx = 0$ and $Lx = 0$. It follows from *i*) of Lemma 4.1 that $x \in \mathcal{N}(W-I_q)$, while $Lx = 0$ implies $x \in \mathcal{N}(W-I_q)$. Now, it follows from *ii*) of Lemma 4.1 that $x = 0$. Hence, $V_k(\cdot)$ is positive definite. Since $\{\tilde{P}_k\}_{k=0}^{\infty}$ is bounded and monotonic, it follows that $\lim_{k\to\infty}\tilde{P}_k = P_\infty$ exists. Furthermore, $P_\infty \geq \tilde{P}_k$ for all $k \geq 0$ and P_∞ satisfies

$$P_\infty = W^{\mathrm{T}}P_\infty W + (W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q). \quad (34)$$

Define

$$V(x) = x^{\mathrm{T}}P_\infty x + x^{\mathrm{T}}L^{\mathrm{T}}Lx. \quad (35)$$

Clearly, $V(x) \geq V_k(x)$ for all $k \geq 0$, and hence, $V(\cdot)$ is positive definite.

Next, since $L(W-I_q) = W-I_q-(W-I_q)(W-I_q)^{\#}(W-I_q) = 0$, it follows that

$$\begin{aligned} V(x(t+1)) - V(x(t)) &= -x^{\mathrm{T}}(t)(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)x(t) \\ &\quad +x^{\mathrm{T}}(t)(W-I_q)^{\mathrm{T}}L^{\mathrm{T}}L(W-I_q)x(t) \\ &\quad +x^{\mathrm{T}}(t)(W-I_q)^{\mathrm{T}}L^{\mathrm{T}}Lx(t) \\ &\quad +x^{\mathrm{T}}(t)L^{\mathrm{T}}L(W-I_q)x(t) \\ &= -x^{\mathrm{T}}(t)(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)x(t) \\ &\leq 0. \end{aligned} \quad (36)$$

To find the largest invariant subset $\mathcal{M}$ of $\mathcal{N}(\tilde{R}(W-I_q))$, consider a solution x of (2) such that $\tilde{R}(W-I_q)x(t) = 0$ for all $t \in \overline{\mathbb{Z}}_+$. Then, $\tilde{R}(W-I_q)x(t+1)-\tilde{R}(W-I_q)x(t) = 0$, that is, $\tilde{R}(W-I_q)^2x(t) = 0$. Similarly, $\tilde{R}(W-I_q)^2x(t+1)-\tilde{R}(W-I_q)^2x(t) = \tilde{R}(W-I_q)^3x(t) = 0$, and so on. This implies $\tilde{R}(W-I_q)^ix(t) = 0$ for all $t \in \overline{\mathbb{Z}}_+$ and $i = 1,2,\ldots$. Weak semiobservability of $(W,\tilde{R})$ now implies that $x(t) \in \mathcal{N}(W-I_q)$ for all $t \in \overline{\mathbb{Z}}_+$. Thus, $\mathcal{M} \subseteq \mathcal{N}(W-I_q)$. However, $\mathcal{N}(W-I_q)$ consists of only equilibrium points and, hence, is invariant. Hence, $\mathcal{M} = \mathcal{N}(W-I_q)$.

Now, let $x_{\mathrm{e}} \in \mathcal{N}(W-I_q)$ be an equilibrium point of (2) and consider the function $U(x) = V(x-x_{\mathrm{e}})$, which is positive definite with respect to x_{e}. Then it follows that $U(x(t+1))-U(x(t)) = -(x(t)-x_{\mathrm{e}})^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)(x(t)-x_{\mathrm{e}}) \leq 0$, $x(t) \in \mathbb{R}^q$, $t \in \overline{\mathbb{Z}}_+$. Thus, it follows that x_{e} is Lyapunov stable, and hence, by Proposition 2.1 of [27], (2) is semistable. ■

Lemma 4.3: Consider the linear iteration (2). Then (2) is semistable if and only if for every weakly semiobservable pair $(W,\tilde{R})$ with positive semidefinite $\tilde{R}$, there exists a bounded sequence consisting of $q\times q$ positive semidefinite matrices $\tilde{P}_k \geq 0$ such that (26) holds, $\tilde{P}_{k+1} \geq \tilde{P}_k$ for all $k \geq 0$, and $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k \geq 0$.

Proof: The existence of $\tilde{P}_k$ satisfying (26) and monotonicity is a combined result of Lemmas 3.2 and 4.2. To show that $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k \geq 0$, let $x \in \mathcal{N}((W-I_q)^2)$ and $z = (W-I_q)x$. Then it follows that $(W-I_q)z = 0$. Note that $\tilde{P}_k = (W-I_q)^{\mathrm{T}}P_k(W-I_q)$, then it follows that $\tilde{P}_{k+1}-\tilde{P}_k = (W-I_q)^{\mathrm{T}}(P_{k+1}-P_k)(W-I_q) = (W^{k+1})^{\mathrm{T}}(W-I_q)^{\mathrm{T}}\tilde{R}(W-I_q)W^{k+1} = (W-I_q)^{\mathrm{T}}(W^{k+1})^{\mathrm{T}}\tilde{R}W^{k+1}(W-I_q)$. Hence, $z^{\mathrm{T}}(\tilde{P}_{k+1}-\tilde{P}_k)z = 0$, which implies that $z \in \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$. Thus, $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k \geq 0$. ■

Theorem 4.1: Consider the linear iteration (2). Then (2) is semistable if and only if for every weakly semiobservable pair $(W,\tilde{R})$ with positive semidefinite $\tilde{R}$, there exists a bounded sequence consisting of $q\times q$ matrices $\hat{P}_k > 0$ such that $\hat{P}_{k+1} \geq \hat{P}_k$ for all $k \geq 0$, $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\hat{P}_{k+1}-\hat{P}_k)$ for all $k \geq 0$, and

$$\hat{P}_{k+1} = W^{\mathrm{T}}\hat{P}_kW + (I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W). \quad (37)$$

for all $k \geq 0$. Such a sequence is not unique. Furthermore, if $(W,\tilde{R})$ is weakly semiobservable and $\hat{P}_k$ satisfies (37), then

$$\begin{aligned} \hat{P}_k &= \sum_{t=0}^{k}(W^t)^{\mathrm{T}}(I_q-W)^{\mathrm{T}}\tilde{R}(I_q-W)W^t \\ &\quad +(W^{\mathrm{T}})^{k+1}\hat{P}_{-1}W^{k+1}, \quad k \geq 0, \end{aligned} \quad (38)$$

where $\hat{P}_{-1}$ satisfies $\hat{P}_{-1} \geq 0$ and $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}}\hat{P}_{-1}W-\hat{P}_{-1})$.

Proof: Suppose $(W,\tilde{R})$ is weakly semiobservable. Then it follows from Lemma 4.3 that there exists a bounded sequence of $q\times q$ matrices $\tilde{P}_k \geq 0$ such that (26) holds, $\tilde{P}_{k+1} \geq \tilde{P}_k$ for all $k \geq 0$, and $\mathcal{N}((W-I_q)^2) \subseteq \mathcal{N}(\tilde{P}_{k+1}-\tilde{P}_k)$ for all $k \geq 0$. Since, by Lemma 4.1, $\mathcal{N}(W-I_q)\cap\mathcal{R}(W-I_q) = \{0\}$, it follows from [28, p. 119] that $W-I_q$ is group invertible. Thus, let $L \triangleq I_q-(W-I_q)(W-I_q)^{\#}$ and note that $L^2 = L$. Hence, L is the unique $q\times q$ matrix satisfying $\mathcal{N}(L) = \mathcal{R}(W-I_q)$, $\mathcal{R}(L) = \mathcal{N}(W-I_q)$, and $Lx = x$ for all $x \in \mathcal{N}(W - I_q)$. Now, define

$$\hat{P}_k \triangleq \tilde{P}_k + L^{\mathrm{T}}L. \qquad (39)$$

Clearly, $\hat{P}_{k+1} \geq \hat{P}_k$ for all $k \geq 0$ and $\mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(\hat{P}_{k+1} - \hat{P}_k)$ for all $k \geq 0$. Next, we show that $\hat{P}_k$ is positive definite. Consider the multiple functions $V_k(x) = x^{\mathrm{T}}\hat{P}_k x$, $x \in \mathbb{R}^q$. If $V_k(x) = 0$ for some $x \in \mathbb{R}^q$, then $\tilde{P}_k x = 0$ and $Lx = 0$. It follows from i) of Lemma 4.1 that $x \in \mathcal{N}(W - I_q)$, and $Lx = 0$ implies that $x \in \mathcal{R}(W - I_q)$. Now, it follows from ii) of Lemma 4.1 that $x = 0$. Hence, $\hat{P}_k$ is positive definite. Next, since $L(W - I_q) = W - I_q - (W - I_q)(W - I_q)^{\#}(W - I_q) = 0$, it follows that

$$\begin{aligned}\hat{P}_{k+1} - W^{\mathrm{T}}\hat{P}_k W &= \tilde{P}_{k+1} + L^{\mathrm{T}}L - W^{\mathrm{T}}\tilde{P}_k W \\ &\quad -W^{\mathrm{T}}L^{\mathrm{T}}LW \\ &= (W - I_q)^{\mathrm{T}}\tilde{R}(W - I_q) + L^{\mathrm{T}}L \\ &\quad -L^{\mathrm{T}}L \\ &= (I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W). \qquad (40)\end{aligned}$$

Conversely, if there exists a bounded sequence of $\hat{P}_k > 0$ such that (37) holds and $\hat{P}_{k+1} \geq \hat{P}_k$ for all $k \geq 0$, consider the function $U(x) = x^{\mathrm{T}}\hat{P}x$, where $\hat{P} = \lim_{k\to\infty}\hat{P}_k$ and $x \in \mathbb{R}^q$. Then $U(x(t+1)) - U(x(t)) = -x^{\mathrm{T}}(t)(I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W)x(t) \leq 0$, $t \in \overline{\mathbb{Z}}_+$, and $\{x \in \mathbb{R}^q : x^{\mathrm{T}}(I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W)x = 0\} = \mathcal{N}(\tilde{R}(I_q - W))$. To obtain the largest invariant set $\mathcal{M}$ contained in $\mathcal{N}(\tilde{R}(I_q - W))$, consider a solution $x(t)$ of (2) such that $\tilde{R}(I_q - W)x(t) = 0$ for all $t \in \overline{\mathbb{Z}}_+$. Then, $\tilde{R}(I_q - W)x(t+1) - \tilde{R}(I_q - W)x(t) = 0$, that is, $\tilde{R}(I_q - W)^2 x(t) = 0$. This implies $\tilde{R}(I_q - W)^i x(t) = 0$ for all $t \in \overline{\mathbb{Z}}_+$ and $i = 1, 2, \ldots$. Now, it follows from weak semiobservability of $(W, \tilde{R})$ that $x(t) \in \mathcal{N}(I_q - W)$ for all $t \in \overline{\mathbb{Z}}_+$. Thus, $\mathcal{M} \subseteq \mathcal{N}(I_q - W)$. Since $\mathcal{N}(I_q - W)$ consists of only equilibrium points, it follows that $\mathcal{M} = \mathcal{N}(I_q - W)$. For $x_{\mathrm{e}} \in \mathcal{N}(I_q - W)$, Lyapunov stability of x_{e} now follows by considering the Lyapunov function $U(x - x_{\mathrm{e}})$.

For any $\hat{P}_k > 0$ satisfying (37) and $M \geq 0$, let $\overline{P}_k \triangleq \hat{P}_k + L^{\mathrm{T}}ML$. Clearly, $\overline{P}_k \geq \hat{P}_k > 0$. It is easy to verify that $\overline{P}_k$ is a solution to (37), and hence, such a $\hat{P}_k$ is not unique.

Finally, since W is semistable, it follows from the above result that there exist $q \times q$ positive-definite matrices $\hat{P}_k$ such that (37) holds. Note that (37) is a recursive equation. By induction, we obtain the expression (38) for $\hat{P}_k$. ■

V. A Convex Formulation for the OFTDLA Problem

In this section, we formulate the OFTDLA problem into a convex optimization problem. To do this, first we note that the OFTDLA problem can be recast as the following optimization problem.

Theorem 5.1: Consider the linear iteration (2). Assume $(W, \tilde{R})$ is weakly semiobservable for positive semidefinite $\tilde{R}$. Let $S_{\min}$ be a solution to the minimization problem

$$\begin{aligned}\min\Big\{&\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}S_q(I_q - W)^{\#}V : S_k > 0, \\ &S_{-1} \geq 0,\ \mathcal{N}((W - I_q)^2) = \mathcal{N}(W^{\mathrm{T}}S_{-1}W - S_{-1}), \\ &S_{k+1} = W^{\mathrm{T}}S_k W + (I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W), \\ &k \in \overline{\mathbb{Z}}_+\Big\}. \qquad (41)\end{aligned}$$

Then for P given by (20),

$$\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}S_{\min}(I_q - W)^{\#}V = \mathrm{tr}\,PV. \qquad (42)$$

Furthermore, such a $S_{\min}$ is not unique.

Proof: Since $(W, \tilde{R})$ is weakly semiobservable, it follows from Theorem 4.1 that all the solutions S_k of $S_{k+1} = W^{\mathrm{T}}S_k W + (I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W) = 0$ have the form

$$\begin{aligned}S_k &= \sum_{t=0}^{k}(W^t)^{\mathrm{T}}(I_q - W)^{\mathrm{T}}\tilde{R}(I_q - W)W^t \\ &\quad +(W^{k+1})^{\mathrm{T}}S_{-1}W^{k+1}, \qquad (43)\end{aligned}$$

where S_{-1} satisfies $S_{-1} \geq 0$ and $\mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}}S_{-1}W - S_{-1})$. Hence,

$$\begin{aligned}&\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}S_q(I_q - W)^{\#}V \\ &= \mathrm{tr}\sum_{t=0}^{q}((I_q - W)^{\#})^{\mathrm{T}}(W^t)^{\mathrm{T}}(I_q - W)^{\mathrm{T}} \\ &\quad \tilde{R}(I_q - W)W^t(I_q - W)^{\#}V \\ &\quad +\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}(W^{q+1})^{\mathrm{T}}S_{-1}W^{q+1}(I_q - W)^{\#}V \\ &= \mathrm{tr}\sum_{t=0}^{q}((I_q - W)(I_q - W)^{\#})^{\mathrm{T}}(W^t)^{\mathrm{T}} \\ &\quad \tilde{R}W^t(I_q - W)(I_q - W)^{\#}V \\ &\quad +\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}(W^{q+1})^{\mathrm{T}}S_{-1}W^{q+1}(I_q - W)^{\#}V \\ &= \mathrm{tr}\,PV + \mathrm{tr}\,V^{1/2}((I_q - W)^{\#})^{\mathrm{T}}(W^{q+1})^{\mathrm{T}}S_{-1}W^{q+1} \\ &\quad (I_q - W)^{\#}V^{1/2} \\ &\geq \mathrm{tr}\,PV. \qquad (44)\end{aligned}$$

Now, let $S_{\min} = (W - I_q)^{\mathrm{T}}P(W - I_q) + L^{\mathrm{T}}L$. It follows from Theorem 4.1 that $S_{\min} > 0$. Next,

$$\begin{aligned}S_{\min} &= \sum_{t=0}^{q}(I_q - W)^{\mathrm{T}}((I_q - W)(I_q - W)^{\#})^{\mathrm{T}} \\ &\quad (W^t)^{\mathrm{T}}\tilde{R}W^t(I_q - W)(I_q - W)^{\#}(I_q - W) \\ &\quad +L^{\mathrm{T}}L \\ &= \sum_{t=0}^{q}(I_q - W)^{\mathrm{T}}(W^t)^{\mathrm{T}}\tilde{R}W^t(I_q - W) \\ &\quad +L^{\mathrm{T}}L. \qquad (45)\end{aligned}$$

Hence,

$$\begin{aligned}&\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}S_{\min}(I_q - W)^{\#}V \\ &= \mathrm{tr}\sum_{t=0}^{q}((I_q - W)(I_q - W)^{\#})^{\mathrm{T}}(W^t)^{\mathrm{T}} \\ &\quad \tilde{R}W^t(I_q - W)(I_q - W)^{\#}V \\ &\quad +\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}L^{\mathrm{T}}L(I_q - W)^{\#}V \\ &= \mathrm{tr}\,PV + \mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}}L^{\mathrm{T}}L(I_q - W)^{\#}V. \qquad (46)\end{aligned}$$

Note that $L(I_q - W)^{\#} = (I_q - (I_q - W)(I_q - W)^{\#})(I_q - W)^{\#} = (I_q - W)^{\#} - (I_q - W)^{\#}(I_q - W)(I_q - W)^{\#} = 0$. Hence, $\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} L^{\mathrm{T}} L (I_q - W)^{\#} V = 0$. Thus,

$$\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} S_{\min} (I_q - W)^{\#} V = \mathrm{tr}\, PV. \tag{47}$$

To show nonuniqueness of $S_{\min}$, let $\hat{S} \triangleq S_{\min} + L^{\mathrm{T}} M L$ for any $M \geq 0$. Clearly, $\hat{S} \geq S_{\min} > 0$. It is easy to verify that $\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} \hat{S} (I_q - W)^{\#} V = \mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} S_{\min} (I_q - W)^{\#} V = \mathrm{tr}\, PV$, and hence, such a $S_{\min}$ is not unique. ■

It is important to note that the optimization problem presented in Theorem 5.1 is not a convex optimization problem due to the involvement of $(I_q - W)^{\#}$ in the cost functional. Hence, many convex optimization methods cannot be applied to solving this problem. Next, we show that this obstacle can be removed by the following result.

Theorem 5.2: Consider the linear iteration (2). Assume $(W, \tilde{R})$ is weakly semiobservable for positive semidefinite $\tilde{R}$. Let $S_{\min}$ be a solution to the minimization problem (41) and $Z_{\min}$ be a solution to the minimization problem

$$\min \Big\{ \mathrm{tr}\, Z_q V : Z_k > 0,\ Z_{-1} \geq 0, \\ \mathcal{N}((W - I_q)^2) = \mathcal{N}(W^{\mathrm{T}} Z_{-1} W - Z_{-1}), \\ Z_{k+1} = W^{\mathrm{T}} Z_k W + (I_q - W)^{\mathrm{T}} \tilde{R} (I_q - W), \\ k \in \overline{\mathbb{Z}}_+ \Big\}. \tag{48}$$

Then

$$\mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} S_{\min} (I_q - W)^{\#} V \\ = \mathrm{tr}\,((I_q - W)^{\#})^{\mathrm{T}} Z_{\min} (I_q - W)^{\#} V. \tag{49}$$

Proof: It follows from Theorem 4.1 that $S_{\min}$ can be characterized by

$$S_{\min} = \sum_{t=0}^{q} (I_q - W)^{\mathrm{T}} (W^t)^{\mathrm{T}} \tilde{R} W^t (I_q - W) \\ + (W^{\mathrm{T}})^{q+1} S_{-1} W^{q+1}, \tag{50}$$

where $S_{-1} \geq 0$ and $\mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}} S_{-1} W - S_{-1})$. Since W is semistable, it follows that the dimension of the null space of $I_q - W^{\mathrm{T}}$ is one. Hence, it follows from Theorem 4.1 that $Z_{\min}$ can be characterized by

$$Z_{\min} = \sum_{t=0}^{q} (I_q - W)^{\mathrm{T}} (W^t)^{\mathrm{T}} \tilde{R} W^t (I_q - W) \\ + (W^{\mathrm{T}})^{q+1} Z_{-1} W^{q+1}, \tag{51}$$

where $Z_{-1} \geq 0$ and $\mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}} Z_{-1} W - Z_{-1})$. Now, by (50) and (51), we have

$$Z_{\min} = S_{\min} + (W^{\mathrm{T}})^{q+1} (Z_{-1} - S_{-1}) W^{q+1}, \tag{52}$$

which implies (49). ■

In summary, a new formulation for the OFTLDA problem based on Theorem 5.2 is given by

$$\text{minimize } x^{\mathrm{T}}(0) Z_q x(0) \\ \text{subject to } W \in \mathcal{W} \cap \mathcal{S}, \quad \mathbf{1}^{\mathrm{T}} W = \mathbf{1}^{\mathrm{T}}, \quad W\mathbf{1} = \mathbf{1}, \\ Z_k > 0, \quad Z_{-1} \geq 0, \\ \mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}} Z_{-1} W - Z_{-1}), \\ Z_{k+1} = W^{\mathrm{T}} Z_k W + (I_q - W)^{\mathrm{T}} \\ (Q + W^{\mathrm{T}} R W)(I_q - W), \quad k \geq 0,$$

where

$$\mathcal{S} = \{ W \in \mathbb{R}^{q \times q} : (W, Q + W^{\mathrm{T}} R W) \\ \text{is weakly semiobservable} \}.$$

It follows from the above description that this new formulation of the OFTLDA problem needs to consider the fourth-order recursive matrix equation and the set $\mathcal{S}$, which are not easy to compute in practice. Next, we further boil down this formulation by imposing some simple assumptions.

Lemma 5.1: Let $Q > 0$. Then $(W, Q + W^{\mathrm{T}} R W)$ is weakly semiobservable.

Proof: Since $Q > 0$, it follows that $Q + W^{\mathrm{T}} R W > 0$. Hence, $\mathcal{N}((Q + W^{\mathrm{T}} R W)(W - I_q)^k) = \mathcal{N}((W - I_q)^k)$ for every $k = 1, \ldots, q$. Note that $\mathcal{N}(W - I_q) \subseteq \mathcal{N}((W - I_q)^k)$ for every $k = 1, \ldots, q$. Then it follows that $\bigcap_{k=1}^{q} \mathcal{N}((Q + W^{\mathrm{T}} R W)(W - I_q)^k) = \bigcap_{k=1}^{q} \mathcal{N}((W - I_q)^k) = \mathcal{N}(W - I_q)$, which means that $(W, Q + W^{\mathrm{T}} R W)$ is weakly semiobservable. ■

Proposition 5.1: Let $Q > 0$. The solution to the following optimization problem gives a solution to the OFTDLA problem:

$$\text{minimize } x^{\mathrm{T}}(0) Z_q x(0) \\ \text{subject to } W \in \mathcal{W}, \quad \mathbf{1}^{\mathrm{T}} W = \mathbf{1}^{\mathrm{T}}, \quad W\mathbf{1} = \mathbf{1}, \\ Z_k > 0, \quad Z_{-1} \geq 0, \\ \mathcal{N}((W - I_q)^2) \subseteq \mathcal{N}(W^{\mathrm{T}} Z_{-1} W - Z_{-1}), \\ Z_{k+1} = W^{\mathrm{T}} Z_k W + (I_q - W)^{\mathrm{T}} \\ (Q + W^{\mathrm{T}} R W)(I_q - W), \quad k \geq 0.$$

Proof: The assertion is a direct consequence of Lemma 5.1. ■

Proposition 5.1 can be solved via semidefinite programming (SDP) methods by converting the recursive matrix algebraic equation into a finite set of LMIs.

VI. Concluding Remarks

By stating that solving the optimization problem (48) is equivalent to the OFTDLA problem according to Theorem 5.2, one could possibly use some computational methods such as the interior-point method or the subgradient method to solve the optimization problem (48) on a large-scale graph, but problems with more than a few thousand edges are probably beyond the capabilities of current interior-point semidefinite programming solvers and the subgradient method has no simple stopping criterion that guarantees a certain level of suboptimality. Hence, we need to consider alternatives such as stochastic optimization methods [29]. The stochastic optimization methods have attracted much attention in recent years since their algorithms do not require properties such as linearity, differentiability, convexity, separability or nonexistence of constraints. Developing an efficient, distributed stochastic algorithm to solve the equivalent OFTDLA problem will be an interesting topic for the future research.

REFERENCES

[1] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Dept. Elect. Engr. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, 1984.

[2] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, pp. 803–812, 1986.

[3] N. A. Lynch, *Distributed Algorithms*. San Francisco, CA: Morgan Kaufmann, 1996.

[4] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, MA: Athena Scientific, 1997.

[5] T. R. Smith, H. Hanssmann, and N. E. Leonard, "Orientation control of multiple underwater vehicles with symmetry-breaking potentials," in *Proc. IEEE Conf. Decision Control*, Orlando, FL, 2001, pp. 4598–4603.

[6] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Trans. Robot. Automat.*, vol. 17, pp. 905–908, 2001.

[7] Z. Lin, M. E. Broucke, and B. A. Francis, "Local control strategies for groups of mobile autonomous agents," *IEEE Trans. Autom. Control*, vol. 49, pp. 622–629, 2004.

[8] P. Ogren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Trans. Autom. Control*, vol. 49, pp. 1292–1302, 2004.

[9] J. Cortés, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. Robot. Automat.*, vol. 20, pp. 243–255, 2004.

[10] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, pp. 988–1001, 2003.

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

[12] J. A. Marshall, M. E. Broucke, and B. A. Francis, "Formations of vehicles in cyclic pursuit," *IEEE Trans. Autom. Control*, vol. 49, pp. 1963–1974, 2004.

[13] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, pp. 169–182, 2005.

[14] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, pp. 401–420, 2006.

[15] Q. Hui, W. M. Haddad, and S. P. Bhat, "Finite-time semistability and consensus for nonlinear dynamical networks," *IEEE Trans. Autom. Control*, vol. 53, pp. 1887–1900, 2008.

[16] Q. Hui and W. M. Haddad, "Distributed nonlinear control algorithms for network consensus," *Automatica*, vol. 44, pp. 2375–2381, 2008.

[17] Q. Hui, W. M. Haddad, and S. P. Bhat, "Semistability theory for differential inclusions with applications to consensus problems in dynamical networks with switching topology," in *Proc. Amer. Control Conf.*, Seattle, WA, 2008, pp. 3981–3986.

[18] ——, "Semistability for time-varying discontinuous dynamical systems with application to agreement problems in switching networks," in *Proc. IEEE Conf. Decision Control*, Cancun, Mexico, 2008, pp. 2985–2990.

[19] Q. Hui and W. M. Haddad, "Uniform semistability for time-varying dynamical systems and network consensus with time-dependent communication links," in *Proc. Amer. Control Conf.*, St. Louis, MO, 2009, pp. 4050–4055.

[20] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, pp. 65–78, 2004.

[21] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inform. Theory*, vol. 52, pp. 2508–2530, 2006.

[22] L. Xiao, S. Boyd, and S. J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distributed Computing*, vol. 67, pp. 33–46, 2007.

[23] D. S. Bernstein, *Matrix Mathematics*. Princeton, NJ: Princeton Univ. Press, 2005.

[24] Q. Hui and H. Zhang, "Optimal linear iterations for distributed agreement," in *Proc. Amer. Control Conf.*, Baltimore, MD, 2010.

[25] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation and consensus using linear iterative strategies," *IEEE J. Selected Areas Commun.*, vol. 26, pp. 650–660, 2008.

[26] ——, "Finite-time distributed consensus in graphs with time-invariant topologies," in *Proc. Amer. Control Conf.*, New York, NY, 2007, pp. 711–716.

[27] Q. Hui and W. M. Haddad, "$\mathcal{H}_2$ optimal semistable stabilization for linear discrete-time dynamical systems with applications to network consensus," *Int. J. Control*, vol. 82, pp. 456–469, 2009.

[28] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic, 1979.

[29] J. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*. New York: Wiley, 2003.