Achieving Optimal Output Consensus for Discrete-time Linear Multi-agent Systems with Disturbances Rejection

Yutao Tang, Hao Zhu, and Xiaoyong Lv*

Abstract: In this paper, an optimal output consensus problem is studied for discrete-time linear multi-agent systems subject to external disturbances. Each agent is assigned with a local cost function which is known only to itself. Distributed protocols are to be designed to guarantee an output consensus for these high-order agents and meanwhile minimize the aggregate cost as the sum of these local costs. To overcome the difficulties brought by high-order dynamics and external disturbances, we develop an embedded design and constructively present a distributed rule to solve this problem. The proposed control includes three terms: an optimal signal generator under a directed information graph, an observer-based compensator to reject these disturbances, and a reference tracking controller for these linear agents. It is shown to solve the formulated problem with some mild assumptions. Numerical examples are also provided to illustrate the effectiveness of our proposed distributed control laws.

1 Introduction

In recent years, a lot of efforts have been made to study the distributed coordination of multi-agent systems. As one of the most important problems, distributed optimization has drawn growing attention due to its wide applications in power systems and sensor networks [1, 2]. In a typical setting of this problem, a network of interconnected nodes are associated with a group of convex functions, while each node only knows one component of these functions. The design goal is to drive all these nodes to reach an optimal consensus minimizing the sum of these functions by exchanging information with each other.

There has been a plenty of publications on this topic and many significant results have been reported in literature. For instance, the authors in [3] investigated the distributed consensus optimization problem through a novel combination of average consensus algorithms with subgradient methods. Extensions with global or local constraints on the decision variables were further studied in [4, 5]. Efforts have also been made to derive distributed algorithms with fast convergence rate in [6–8]. Paralleled with these discrete-time results, continuous-time solvers to this distributed optimization were also considered under various conditions in [9,10].

However, it is observed that most of the above results are derived only for single-integrator agents from the viewpoint of mathematical programming. In practical applications, the decision variables might be determined by or depend upon physical plants, which can not be described well by single integrators, e.g., a group of mobile robots to achieve rendezvous [12]. In [13], a numerical example was provided to show that direct use of distributed rules for single integrators might fail to achieve the optimal goal for

*This work was supported by National Natural Science Foundation of China under Grants 61973043. Yutao Tang and Hao Zhu are with the School of Automation, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China. Email: yttang@bupt.edu.cn, severuszhu@gmail.com. Xiaoyong Lv is with the College of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao, P. R. China. Email: lengfeng0506@163.com
agents with unity relative degree. Thus, we should take these high-order dynamics into account when seeking an optimal consensus in a distributed manner.

Note that the gradient-based rules are basically nonlinear, achieving optimal (output) consensus might be challenging due to the couplings between the high-order dynamics of agents with the distributed optimization requirement. To solve this problem, some interesting attempts have been made for continuous-time high-order dynamics. For example, the authors in [14, 15] extend existing rules to continuous-time second-order agents by adding some integral terms. Similar ideas have been used in [16] to achieve optimal consensus for high-order integrators by bounded controls. In a recent work [17], the authors proposed an embedded control method to solve this problem for linear agents in a modular way. Some nonlinear multi-agent systems were also investigated to achieve such an optimal consensus in [18, 19]. However, in contrast to these papers for continuous-time high-order agents, there is still no general result to our best knowledge on achieving optimal consensus for discrete-time multi-agent systems with non-integrator dynamics.

The objective of this paper is to develop distributed rules for discrete-time high-order agents to achieve an optimal output consensus. To be specific, we assume that the agents are of general linear time-invariant dynamics and each agent is assigned with a local cost function and can exchange information through a communication topology represented by a directed graph. All agents are to be designed to achieve an output consensus and meanwhile minimize the aggregate cost as the sum of local ones. Moreover, we further consider the cases when agents are subject to external disturbances, which are inevitably encountered in practical circumstances.

Based on aforementioned observations, the contribution of this work can be summarized in at least three aspects.

Firstly, an optimal output consensus problem is formulated and solved for a high-order multi-agent system in discrete time. In contrast to plenty of optimal consensus results for single-integrators, this problem is extended to multi-agent systems having linear dynamics, which can be taken as the discrete-time counterpart of existing optimal consensus results for high-order agents in continuous time.

Secondly, we develop novel distributed rules to achieve optimal (output) consensus under weight-balanced directed graphs. Compared with similar results for undirected graphs, this problem is certainly more challenging. Moreover, the proposed algorithm is free of initialization in contrast to many existing works for digraphs, which might be more favorable in large scale multi-agent systems.

Thirdly, disturbances rejection issue is discussed in achieving optimal consensus for this discrete-time multi-agent system. Note that the considered external disturbances are modeled by an autonomous linear system, which are general enough to cover many typical signals, e.g., step sequence and sinusoidal sequence. An observer based method is used to reject these harmful disturbances to facilitate our optimal consensus design.

The rest of this paper is organized as follows. We first give some preliminaries about graph notations and convex analysis in Section 2 and then formulate our problem in Section 3. The main design with proofs is presented in Section 4 with a numerical example in Section 5. Finally, some concluding remarks are given in Section 6.

2 Preliminaries

In this section, we first provide some preliminaries about graph theory [20] and convex analysis [21].
2.1 Graph theory

Let $\mathbb{R}^N$ be the $N$-dimensional Euclidean space. $\text{col}(a_1, \ldots, a_N) = [a_1^T, \ldots, a_N^T]^T$ for column vectors $a_i$ ($i = 1, \ldots, N$). For a vector $a$ (or a matrix $A$), $|a|$ (or $|A|$) denotes its Euclidian (or spectral) norm. $1_N$ (or $0_N$) denotes an $N$-dimensional all-one (or all-zero) column vector, and $I_N$ denotes the $N$-dimensional identity matrix. Denote $r_N = \frac{1}{\sqrt{N}}1_N$ and $\Pi_N = I_N - r_Nr_N^T$. Let $R_N \in \mathbb{R}^{N \times (N-1)}$ be the matrix satisfying $R_N^2 = 0_N - 1$, $R_N^3R_N = I_N - 1$ and $R_NR_N^T = \Pi_N$. We may omit the subscript when it is self-evident.

A weighted directed graph (digraph) is described by a triplet $G = (N, \mathcal{E}, \mathcal{A})$ with the node set $\mathcal{N} = \{1, \ldots, N\}$ and the edge set $\mathcal{E}$. $(i, j) \in \mathcal{E}$ denotes an edge from nodes $i$ to $j$. The weighted adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is defined by $a_{ii} = 0$ and $a_{ij} > 0$. Here $a_{ij} > 0$ iff there is an edge $(j, i)$ in this graph.

The neighbor set of node $i$ is defined as $\mathcal{N}_i = \{j \mid (j, i) \in \mathcal{E}\}$ for $i = 1, \ldots, N$. A directed path an alternating sequence $i_1e_1i_2e_2\ldots e_{k-1}i_k$ of nodes $i_l$ and edges $e_m = (i_m, i_{m+1}) \in \mathcal{E}$ for $l = 1, 2, \ldots, k$. If there is a directed path between any two vertices, then the graph is said to be strongly connected. The in-degree and out-degree of node $i$ is defined by $d_i^\text{in} = \sum_{j=1}^{N} a_{ij}$ and $d_i^\text{out} = \sum_{j=1}^{N} a_{ji}$. A directed graph is weight-balanced if $d_i^\text{in} = d_i^\text{out}$ holds for any $i = 1, \ldots, N$.

Note that $L1_N = 0_N$ for any digraph. If this digraph is weight-balanced, it also holds that $1_N^T L = 0_N^T$ and the matrix $\text{Sym}(L) \triangleq \frac{L + L^T}{2}$ is positive semidefinite. If this weight-balanced digraph is strongly connected, 0 is a simple eigenvalue of $\text{Sym}(L)$ and all other eigenvalues are positive real numbers. In this case, we order these eigenvalues as $\lambda_1 = 0 < \lambda_2 \leq \cdots \leq \lambda_N$ and have that the matrix $R^T\text{Sym}(L)R$ is positive definite with eigenvalues $\lambda_2, \ldots, \lambda_N$.

2.2 Convex analysis

A function $f: \mathbb{R}^m \to \mathbb{R}$ is said to be convex if for $0 \leq a \leq 1$,

$$f(ax_1 + (1-a)x_2) \leq af(x_1) + (1-a)f(x_2), \forall x_1, x_2 \in \mathbb{R}^m$$

When the function $f$ is differentiable, it is verified that $f$ is convex if the following inequality holds,

$$f'(\zeta_1) - f'(\zeta_2) \succeq \nabla f(\zeta_2)^T(\zeta_1 - \zeta_2), \forall \zeta_1, \zeta_2 \in \mathbb{R}^m$$

and is strictly convex if this inequality is strict whenever $\zeta_1 \neq \zeta_2$. A function $f$ is $\omega$-strongly convex ($\omega > 0$) over $\mathbb{R}^m$ if we have

$$\nabla f(\zeta_1) - \nabla f(\zeta_2)^T(\zeta_1 - \zeta_2) \succeq \omega|\zeta_1 - \zeta_2|^2, \forall \zeta_1, \zeta_2 \in \mathbb{R}^m$$

A vector-valued function $f: \mathbb{R}^m \to \mathbb{R}^m$ is Lipschitz with constant $\vartheta > 0$ (or simply $\vartheta$-Lipschitz) if we have

$$|f(\zeta_1) - f(\zeta_2)| \leq \vartheta|\zeta_1 - \zeta_2|, \forall \zeta_1, \zeta_2 \in \mathbb{R}^m$$

3 Problem Formulation

Consider a multi-agent system consisting of $N$ discrete-time linear agents of the following form:

$$x_i(t+1) = Ax_i(t) + Bu_i(t) + d_i(t), \quad i = 1, 2, \ldots, N$$

$$y_i(t) = Cx_i(t), \quad t = 0, 1, 2, \ldots$$

(1)
where \( x_i(t) \in \mathbb{R}^{n_x} \) is the state, \( u_i(t) \in \mathbb{R}^{n_u} \) is the input and \( y_i(t) \in \mathbb{R} \) is the output of agent \( i \). The system matrices \((C, A, B)\) are assumed to be minimal with compatible dimensions. Here, \( d_i(t) \in \mathbb{R}^{n_x} \) represents external disturbances modeled by

\[
d_i(t) = E w_i(t), \quad w_i(t+1) = S w_i(t), \quad t = 0, 1, 2, \ldots
\]

where \( w_i \in \mathbb{R}^{n_w} \) is the full internal state of external disturbances. As usual, we assume that \( S \) has no eigenvalue inside the unit circle on the complex plane [22]. In fact, the components of \( w_i \) corresponding to the eigenvalues inside the unit circle will exponentially converge to zero and thus in no way affect the designed goal.

Each agent is assigned with a local cost function \( f_i : \mathbb{R} \to \mathbb{R} \). We consider the following optimization problem

\[
\min_{s \in \mathbb{R}} f(s) = \sum_{i=1}^{N} f_i(s)
\]

This aggregate objective function \( f \) is called the global cost of this multi-agent system.

Here is an assumption to guarantee the solvability of \((3)\).

**Assumption 1.** For any \( 1 \leq i \leq N \), the cost function \( f_i \) is \( \mathcal{L} \)-strongly convex and \( \nabla f_i \) is \( \mathcal{L} \)-Lipschitz.

Assumption [1] implies the existence and uniqueness of optimal point to \((3)\). As usual [10, 14], we assume this point is finite and denote it as \( y^* \). We aim to develop distributed algorithms to drive the outputs of all agents to achieve a consensus on the minimizer of \( f \) without setting up a centralized working station, which might be expensive and inhibitive in some circumstances.

For this purpose, a weighted digraph \( G = (V, \mathcal{E}, \mathcal{A}) \) is used to describe the information sharing relationships among these agents with a node set \( N = \{1, \ldots, N\} \) and a weight matrix \( \mathcal{A} \in \mathbb{R}^{N \times N} \). If agent \( i \) can get the information of agent \( j \), then there is an edge \((j, i)\) in the graph, i.e., \( a_{ij} > 0 \).

The optimal output consensus problem for discrete-time multi-agent system \((1)\) can be formulated as follows. Given agent \((1)\), cost function \( f_i(\cdot) \), graph \( G \) and disturbance \((2)\), the optimal output consensus problem is to find a feedback control \( u_i \) for agent \( i \) by using its own and neighboring information such that all trajectories of agents are bounded and the associated outputs satisfy

\[ \lim_{t \to +\infty} |y_i(t) - y^*| = 0. \]

In this formulation, these agents are required to achieve an output consensus minimizing the aggregate global cost function. When agents are all single integrators without disturbances, our formulated problem is coincided with the well-studied distributed optimization problem [3]. Here, we further consider multi-agent systems having high-order dynamics subject to external disturbances.

To guarantee that any agent’s information can reach any other agents through a directed information flow, we suppose the following assumption is fulfilled as in many publications [10, 14, 23].

**Assumption 2.** \( G \) is strongly connected and weight-balanced.

Note that when this optimal output consensus problem is solved, we have \( \lim_{t \to \infty} y_i(t) = y^* \). It is natural for agent \( i \) to reach some steady state. Thus, we make another assumption to ensure this point.

**Assumption 3.** There are constant matrices \( X_1, X_2 \) and \( U_1, U_2 \) with compatible dimensions satisfying that:

\[
X_1 = AX_1 + BU_1, \quad 1 = CX_1
\]

\[
X_2 S = AX_2 + BU_2 + E, \quad 0^T = CX_2
\]
Assumption 3 is known as the solvability of regulator equations to achieve set-point regulation and disturbances rejection for discrete-time linear systems [22], which play a crucial role in resolving our optimal consensus problem. Under this assumption, one can obtain the steady-state state and input for each agent as \( X_1 y^* + X_1 w(t) \) and \( U_1 y^* + U_2 w(t) \) when the optimal output consensus is achieved at \( y^* \) with disturbances rejection.

To reject the external disturbances, we suppose the following condition holds without loss of generality.

**Assumption 4.** The pair \( \left[ \begin{bmatrix} C & 0 \\ A & E \\ 0 & S \end{bmatrix} \right] \) is observable.

This assumption implies that the external disturbances can indeed affect our regulated output \( y_i \). A sufficient condition to ensure this assumption is the observability of \( (E, S) \), which can be trivially verified by PBH-test.

### 4 Main Results

To avoid the difficulties brought by the high-order linear dynamics, we develop an embedded design in two steps as in [17] to achieve optimal output consensus for these agents over directed graphs.

#### 4.1 Optimal signal generation

To begin with, we consider an optimal consensus problem with (3) for a virtual agent

\[
z_i(t + 1) = z_i(t) + \mu_i(t)
\]

with the same cost function \( f_i \) and information sharing graph \( G \).

Note that under Assumption 1, the optimization problem (3) can be reformulated to the following equivalent form:

\[
\min \tilde{f}(y) \triangleq \sum_{i=1}^{N} f_i(y_i), \quad \text{subject to } Ly = 0
\]

with \( y \triangleq \text{col}(y_1, \ldots, y_N) \).

There have been some distributed rules to achieve optimal consensus or compute the optimal solution to this problem under directed graphs, e.g. [10] and [24]. However, most of these algorithms require some initialization process. Since there might be disturbances or round-off errors, such an initialization could fail to be fulfilled during the implementation of these rules. Thus, we are more interested to construct optimal signal generators free of such initializations.

Note that its associated Lagrangian is \( L(y, \Lambda) = \tilde{f}(y) + \Lambda^T Ly \). When the graph is undirected and connected, the Laplacian \( L \) is symmetric and the optimal point can be readily derived by a primal-dual dynamics (e.g. [25]). As for digraphs, we lose such a symmetry and the original primal-dual method fails to generate the optimal point. Here, we extend the primal-dual dynamic to weight-balanced graphs by adding a proportional terms as follows.

\[
z_i(t + 1) = z_i(t) - \gamma (\alpha \nabla f_i(z_i(t)) + \beta L z_i(t) + L \lambda_i(t))
\]

\[
\lambda_i(t + 1) = \lambda_i(t) + \gamma \alpha \beta L z_i(t)
\]

where \( \alpha, \beta \) and \( \gamma \) are positive constants to be specified later. This algorithm has been partially investigated in [25] when \( \alpha = \beta = 1 \). Here we add these two tunable parameters to ensure its efficiency with directed graphs.
Putting system (7) into a compact form gives that

\[ Z(t + 1) = Z(t) - \gamma [\alpha \nabla \tilde{f}(Z(t)) + \beta LZ(t) + L\Lambda(t)] \]

\[ \Lambda(t + 1) = \Lambda(t) + \gamma \alpha \beta LZ(t) \]

(8)

where \( Z(t) = \text{col}(z_1(t), \ldots, z_n(t)) \), \( \Lambda(t) = \text{col}(\lambda_1(t), \ldots, \lambda_n(t)) \), and \( \nabla \tilde{f}(Z(t)) \triangleq \text{col}(\nabla f_1(z_1(t)), \ldots, \nabla f_n(z_n(t))) \).

Under Assumption 1, the function \( \nabla \tilde{f} \) is \( \bar{l} \)-Lipschitz in \( Z \).

Let \( (Z^*, \Lambda^*) \) be an equilibrium point of system (8). The following lemma shows that these agents can achieve an optimal consensus at the equilibrium \( (Z^*, \Lambda^*) \).

**Lemma 1.** Under Assumptions 1–2, we have \( Z^* = y^* 1_n \).

**Proof.** At the equilibrium point of system (8), we have

\[ \alpha \nabla \tilde{f}(Z^*) + \beta LZ^* + L\Lambda^* = 0, \quad LZ^* = 0 \]

By Assumption 2, \( LZ^* = 0 \) implies that there exists a constant \( z_0^* \) such that \( Z^* = z_0^* 1_n \). Multiplying both sides by \( 1^T \), we have \( 1_n^T \alpha \nabla \tilde{f}(Z^*) + \beta LZ^* + L\Lambda^* = 1_n^T \nabla \tilde{f}(Z^*) = 0 \). That is, \( \nabla f(z_0^*) = 0 \). From the strong convexity of cost functions, the optimal solution to problem (3) is unique. This means that \( z_0^* = y^* \) and thus \( Z^* = y^* 1_n \). The proof is complete.

To develop effective optimal signal generators, we have to choose appropriate parameters \( \alpha, \beta, \gamma \) such that the equilibrium of (8) is attractive. Here is a key lemma to ensure this point. Its proof can be found in Appendix.

**Lemma 2.** Under Assumptions 1–2, the trajectory of \( z_i(t) \) along system (7) exponentially converges to the optimal solution \( y^* \) of problem (3) from any initial value if the chosen parameters satisfy

\[ \alpha \geq \max \{1, \frac{1}{I \lambda_2^2} \} \]

\[ \beta \geq \max \{1, \frac{4\alpha^2 \lambda_2^4}{\lambda_2^2} \} \]

\[ 0 < \gamma < \frac{1}{\beta^2 (\lambda_2^2 + I^2)} \]

(9)

**Remark 1.** Condition (9) is only sufficient to guarantee the efficiency of generator (7) to reproduce \( y^* \), which can be conservative. One may prefer to select these parameters from repeated simulations by first increasing \( \alpha, \beta \) and then decreasing \( \gamma \) sequentially.

**Remark 2.** Compared with existing optimal consensus design, the designed generator actually solve a distributed optimization problem under weight-balanced directed graphs. Unlike similar rules in [10, 14], the developed algorithm is initialization free and more favorable in large scale networks.

### 4.2 Solvability of optimal consensus problem

With the above optimal signal generator, we are going to solve the associated reference tracking and disturbance rejection problem for agent \( i \) with reference \( z_i(t) \) and disturbance \( d_i(t) \).

Recalling some classical output regulation results [22], a full-information control for each agent to achieve optimal output consensus is written as follows:

\[ u_i^0(t) = Kx_i(t) + K_1 y^* + K_2 w_i(t), \quad t = 0, 1, \ldots \]
where $K$ is chosen such that $A + BK$ is Schur stable and $K_1 = U_1 - KX_1$, $K_2 = U_2 - KX_2$. In fact, under this full-information control, we can obtain an error system by letting $\bar{x}_i(t) = x_i(t) - X_1y^* - X_2w_i(t)$ of the following form:

$$\bar{x}_i(t + 1) = (A + BK)\bar{x}_i(t), \quad t = 0, 1, \ldots$$

along which the regulated output $e_i(t) \triangleq y_i(t) - y^* = C\bar{x}_i(t)$ converges to zero as $t$ goes to infinity.

However, the disturbance $w_i(t)$ is not available to us and the global optimal solution $y^*$ is also unknown due to the distributedness of the global cost function. Thus, the above full-information control is not applicable to our problem.

In Lemma 2 we have shown that the optimal solution $y^*$ can be generated by (7) exponentially fast. This motivates us to replace $y^*$ by the reference signal $z_i(t)$. Considering those external disturbances, we can estimate them by observer based methods to complete the whole design.

Under Assumption 4, a full-state Luenberger observer is constructed to estimate these disturbances as follows.

$$\bar{x}_i(t + 1) = (A + L_1C)\bar{x}_i(t) + Bu_i(t) + E\tilde{w}_i(t) - L_1y_i$$

$$\tilde{w}_i(t + 1) = S\tilde{w}_i(t) + L_2(C\bar{x}_i(t) - y_i)$$

where $L_1$ and $L_2$ are chosen gain matrices such that $\begin{bmatrix} A + L_1 C & E \\ L_2 C & S \end{bmatrix}$ is Schur stable.

Here is a lemma to guarantee its efficiency and we omit its proof.

**Lemma 3.** Under Assumption 4 for any $1 \leq i \leq N$, the signals $\bar{x}_i(t)$ and $\tilde{w}_i(t)$ along the trajectory of system (10) exponentially converge to $x_i(t)$ and $w_i(t)$ as $t$ goes to infinity.

Based on this observer and signal generator (7), we propose the following control for agent $i$ to achieve the optimal consensus goal.

$$u_i(t) = K_1\bar{x}_i(t) + K_2\tilde{w}_i(t)$$

$$\bar{x}_i(t + 1) = (A + L_1C)\bar{x}_i(t) + Bu_i(t) + E\tilde{w}_i(t) - L_1y_i$$

$$\tilde{w}_i(t + 1) = S\tilde{w}_i(t) + L_2(C\bar{x}_i(t) - y_i)$$

$$z_i(t + 1) = z_i(t) - \gamma(\alpha \nabla f_i(z_i(t)) + \beta Lz_i(t) + L\lambda_i(t))$$

$$\lambda_i(t + 1) = \lambda_i(t) + \gamma \alpha \beta Lz_i(t), \quad t = 0, 1, \ldots$$

where matrices $L_1$, $L_2$ and parameters $\alpha$, $\beta$, $\gamma$ are chosen as above. This control law is distributed in sense of using only its own and neighboring information of each agent.

To show its effectiveness, we derive the new error system under control (11) by some mathematical manipulations as follows:

$$\bar{x}_i(t + 1) = (A + BK)\bar{x}_i(t) + B\Xi(t)$$

$$e_i(t) = C\bar{x}_i(t), \quad t = 0, 1, \ldots$$

where $\Xi(t) \triangleq K(\bar{x}_i(t) - x_i(t)) + K_1(z_i(t) - y^*) + K_2(\tilde{w}_i(t) - w_i(t))$. It can be found that $\Xi_i(t)$ represents the discrepancy between our actual control and full-information control.

It is ready to give our main theorem of this paper.

**Theorem 1.** Suppose Assumptions 4 hold. Then the optimal output consensus problem for discrete-time linear multi-agent system (1), (2), and (3) is solved by a distributed controller (11).
Proof. To prove this theorem, we first claim that there exists a constant $c > 0$ such that $\limsup_{t \to \infty} |e_i(t)| \leq c \limsup_{t \to \infty} |\xi_i(t)|$. Moreover, if $\lim_{t \to \infty} |\xi_i(t)| = 0$, we have $\lim_{t \to \infty} e_i(t) = 0$. This property is a variant of input-to-output stability of system (12) with input $\xi_i(t)$ and output $e_i(t)$.

To prove it, we denote $G = A + BK$ for short. By the iteration (12), one can obtain that
\[
\tau_i(t) = G^t \tau_i(0) + \sum_{j=0}^{t-1} G^{t-1-j} B \xi_i(j), \quad i = 1, \ldots, N
\]
Under Assumption 4 the regulated output $e_i$ is derived as follows.
\[
e_i(t) = CG^t \tau_i(0) + C \sum_{j=0}^{t-1} G^{t-1-j} B \xi_i(j)
\tag{13}
\]
Since $G$ is Schur stable, $\lim_{t \to \infty} CG^t \tau_i(0) = 0$. To estimate the limit superior of $\{e_i(t)\}$, we can neglect this term without affecting the conclusion. Without loss of generality, we assume $b_i \triangleq \limsup_{t \to \infty} |\xi_i(t)|$ is finite.

By its definition, for any $\varepsilon > 0$, there exists a large enough integer $M > 0$ such that $|\xi_i(t)| \leq \varepsilon$ holds for any $t > M$. Splitting the last term of (13) into two parts gives that
\[
|e_i(t)| = |C\sum_{j=0}^{M} G^{t-1-j} B \xi_i(j) + C \sum_{j=M+1}^{t-1} G^{t-1-j} B \xi_i(j)|
\leq |C||B||\sum_{j=0}^{M} |G|^{t-1-j} |\xi_i(j)| + \sum_{j=M+1}^{t-1} |G|^{t-1-j} |
\leq |C||B||G|^{t-1-M} \sum_{j=0}^{M} |G|^M |\xi_i(j)| + (b_i + \varepsilon) \sum_{j=M+1}^{t-1} |G|^{t-1-j}
\]
Note that
\[
\sum_{j=M+1}^{t-1} |G|^{t-1-j} = \frac{1 - |G|^{t-M}}{1 - |G|} < \frac{1}{1 - |G|}
\]
and $|G|^{t-1-M} \to 0$ as $t \to \infty$ by its Schurness. We have that
\[
\limsup_{t \to \infty} |e_i| \leq \frac{1}{1 - |G|}(b_i + \varepsilon)
\]
Since the constant $\varepsilon$ can be chosen arbitrarily, this further implies that $\limsup_{t \to \infty} |e_i(t)| \leq c \limsup_{t \to \infty} |\xi_i(t)|$ for $c = \frac{1}{1 - |G|}$. That is, the claim is correct.

With the above claim, we only have to show $\xi_i(t) \to 0$ as $t \to \infty$ under the control (11), which trivially holds by Lemmas 2 and 3. Thus, one can conclude that $e_i(t) = y_i(t) - y^*$ converges to 0 as $t$ goes to infinity. The proof is thus complete.

Remark 3. This theorem can be regarded as a discrete-time companion of existing optimal output consensus results for continuous-time agents derived in [17, 26]. Compared with the well-studied optimal consensus problem for discrete-time single integrators [3, 7, 8, 25], we extend them to a more general case with high-order linear multi-agent systems subject to nontrivial disturbances. Particularly, we achieve an output average consensus for these linear agents by letting $f_i(s) = (s - y_i(0))^2$ with disturbance rejection.

5 Simulations

In this section, we provide a numerical example to illustrate the effectiveness of our previous designs.
Consider a multi-agent system including four agents as follows.

\[
x_i(t+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_i(t) + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} u_i(t) + d_i(t)
\]

\[
y_i(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x_i(t), \quad t = 0, 1, 2, \ldots.
\]

where the external disturbance \(d_i(t)\) is generated by an exosystem of the form (2) with

\[
E = \begin{bmatrix} 0.5 & 0.5 \\ \sin(1) - \cos(1) & -\cos(1) - \sin(1) \end{bmatrix}
\]

and

\[
S = \begin{bmatrix} \cos(1) & \sin(1) \\ -\sin(1) & \cos(1) \end{bmatrix}
\]

Assumption 3 holds for this multi-agent system with

\[
X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad U_1 = 0, \quad X_2 = \begin{bmatrix} 0 & 0 \\ -1 & -1 \end{bmatrix}, \quad U_2 = \begin{bmatrix} 2 & 2 \end{bmatrix}
\]

The observability of \((E, S)\) is verified and implies Assumption 4.

The communication topology among these agents is represented by a directed ring graph depicted as Fig. 1 with unity edge weights, which satisfies Assumption 2 with \(\lambda_2 = 1\) and \(|L| = 2\). The local cost functions are chosen as below.

\[
f_1(y) = (y - 8)^2
\]

\[
f_2(y) = \frac{y^2}{20\sqrt{y^2 + 1}} + y^2
\]

\[
f_3(y) = \frac{y^2}{80\ln(y^2 + 2)} + (y - 5)^2
\]

\[
f_4(y) = \ln(e^{-0.005y} + e^{0.005y}) + y^2
\]

All these functions are strongly convex with Lipschitz gradients. In fact, Assumption 1 is verified with \(l = 1\) and \(\bar{l} = 3\). By minimizing \(f(y) = \sum_{i=1}^{4} f_i(y)\), the global optimal point is \(y^* = 3.24\).

According to Theorem 1, the associated optimal output consensus problem can be solved by a control (11). For simulations, we choose let \(\alpha = 1, \beta = 15, \gamma = 0.004\). The efficiency of the developed signal generator is shown in Figs. 2 and 3 where the optimal point \(y^*\) can be reproduced quickly.

Next, we choose the gain matrices in control (11) as follows.

\[
K = \begin{bmatrix} -0.4345 \\ -1.0285 \end{bmatrix}^T, \quad L_1 = \begin{bmatrix} -1.8184 \\ -0.3543 \end{bmatrix}, \quad L_2 = \begin{bmatrix} -0.1527 \\ -0.3141 \end{bmatrix}
\]

As shown in Fig. 4, the optimal output consensus for all agents is achieved on the global optimal solution \(y^*\). Also, we shut down the disturbance rejection part in the controller (i.e., set \(K_2 = 0\)) between \(t = 2000\) and \(t = 2250\) and find that these agents fail to achieve a consensus. After we restart this part, the optimal output consensus is quickly recovered, which verifies the efficiency of our control to reject these periodic disturbances.
Figure 2: Profiles of $z_i$ under the generator (7).

Figure 3: Profiles of $\lambda_i$ under the generator (7).
6 Conclusions

The paper studied an optimal output consensus for discrete-time linear multi-agent systems subject to external disturbances. Following an embedded control design, we employed a primal-dual rule with fixed step size to generate the optimal point under directed information flow and developed effective observer based distributed controllers for agents to achieve optimal output consensus. Future works will include time-varying directed graphs.

7 Appendix: Proof of Lemma 2

We first perform the coordinate transformations: \( \bar{Z}_1(t) = r^T(Z(t) - Z^*) \), \( \bar{Z}_2(t) = R^T(Z(t) - Z^*) \), \( \bar{L}_1(t) = r^T\Lambda(t) \), and \( \bar{L}_2(t) = R^T\Lambda(t) \). The translated system is given as

\[
\begin{align*}
\bar{Z}_1(t+1) &= \bar{Z}_1(t) - \gamma \alpha r^T \Delta_1 \\
\bar{Z}_2(t+1) &= \bar{Z}_2(t) - \gamma R^T [\alpha \Delta_1 + \beta L \bar{Z}(t) + LR\bar{L}_2(t)] \\
\bar{L}_2(t+1) &= \bar{L}_2(t) + \gamma \beta R^T L \bar{Z}(t)
\end{align*}
\]

where \( \bar{Z}(t) = \text{col}(\bar{Z}_1(t), \bar{Z}_2(t)) \) and \( \Delta_1 \triangleq \nabla \tilde{f}(Z(t)) - \nabla \tilde{f}(Z^*) \). It can be verified that \( \Delta_1 \) is \( L \)-Lipschitz with respect to \( \bar{Z}(t) \) under Assumption 1.

By further letting \( \xi(t) = \bar{L}_2(t) + \alpha R^T \bar{Z}_2(t) \), we have

\[
\begin{align*}
\bar{Z}_1(t+1) &= \bar{Z}_1(t) - \gamma \alpha r^T \Delta_1 \\
\bar{Z}_2(t+1) &= \bar{Z}_2(t) - \gamma \alpha R^T \Delta_1 - \gamma \Delta_2 \\
\xi(t+1) &= \xi(t) - \gamma \Delta_3 \tag{14}
\end{align*}
\]

where \( R_L = R^T L R \), \( \Delta_2 \triangleq \beta R_L \bar{Z}_2(t) + R_L \xi(t) - \alpha R_L \bar{Z}_2(t) \), and \( \Delta_3 \triangleq \alpha R_L \xi(t) + \alpha^2 R^T \Delta_1 - \alpha^2 R_L \bar{Z}_2(t) \).

Next, we present a Lyapunov analysis to prove the exponential stability of system (14). Take a quadratic Lyapunov function \( V(\bar{Z}(t), \xi(t)) = |\bar{Z}(t)|^2 + \frac{1}{\sigma^2} |\xi(t)|^2 \) with \( \sigma > 0 \) to be specified later and

\[
\text{Figure 4: Output trajectories of agents under control (11)}.
\]
denote it as \( V(t) \) for short. The difference of \( V(t) \) along the trajectory of system (14) satisfies:

\[
\Delta(t) \triangleq V(t + 1) - V(t) \\
= |Z_1(t) - \gamma \alpha R^T \Delta_1|^2 + |Z_2(t) - \gamma \alpha R^T \Delta_1 - \gamma \Delta_2|^2 - |\hat{Z}(t)|^2 + |\xi(t) - \gamma \Delta_1|^2 - |\xi(t)|^2 \\
\leq -2\gamma \alpha \bar{L} |\hat{Z}(t)|^2 + \gamma^2 \alpha^2 |\Delta_1|^2 + \gamma^2 |\Delta_2|^2 - 2\gamma \bar{Z}_2^T(t) \Delta_2 - 2\gamma^2 \alpha R^T \Delta_2 - \frac{2\gamma}{\alpha^3} \xi(t)^T \Delta_3 + \frac{\gamma^2}{\alpha^3} |\Delta_3|^2
\]

Note that \( R^T \text{Sym}(L)R = \frac{R_L + R^T L}{2} \). By Young’s inequality, we have

\[
-2\gamma \bar{Z}_2^T(t) \Delta_2 = -2\gamma \bar{Z}_2^T(t) [\beta R_L \bar{Z}_2(t) + R_L \xi(t) - \alpha R_L \bar{Z}_2(t)] \\
\leq -2\gamma \beta \lambda_2 |\bar{Z}_2(t)|^2 + 2\gamma \alpha \lambda_N |\bar{Z}_2(t)|^2 - 2\gamma \bar{Z}_2^T R_L \xi(t) \\
\leq -\gamma (2\beta \lambda_2 - 2\alpha \lambda_N - \frac{3\gamma^2 \lambda_2^3}{\lambda_2^2}) |\bar{Z}_2(t)|^2 + \frac{\gamma \lambda_2}{\lambda_2} |\xi(t)|^2
\]

and

\[
-2\gamma \xi(t)^T \Delta_3 \leq -2\gamma \xi(t)^T [\alpha R_L \xi(t) + \alpha^2 R^T \Delta_1 - \alpha^2 R_L \bar{Z}_2(t)] \\
\leq -2\gamma \alpha \lambda_2 |\xi(t)|^2 - 2\gamma \alpha^2 \xi(t)^T R^T \Delta_1 + 2\gamma \alpha^2 \xi(t)^T R_L \bar{Z}_2(t) \\
\leq -\frac{4}{3} \gamma \alpha \lambda_2 |\xi(t)|^2 + \frac{3\gamma \alpha^2}{\lambda_2} |\bar{Z}(t)|^2 + \frac{3\gamma \alpha^2}{\lambda_2} |\bar{Z}_2(t)|^2
\]

Using the above two inequalities, one can derive that

\[
\Delta(t) \leq -2\gamma \bar{L} |\hat{Z}(t)|^2 + \gamma^2 \alpha^2 |\Delta_1|^2 + \gamma^2 |\Delta_2|^2 - 2\gamma^2 \alpha R^T \Delta_2 - \gamma (2\beta \lambda_2 - 2\alpha \lambda_N - \frac{3\gamma^2 \lambda_2^3}{\lambda_2}) |\bar{Z}_2(t)|^2 \\
+ \frac{\gamma \lambda_2}{\lambda_2} |\xi(t)|^2 - \frac{4\gamma \alpha \lambda_2}{\lambda_2} |\xi(t)|^2 + \frac{3\gamma^2 \lambda_2^3}{\lambda_2} |\bar{Z}(t)|^2 + \frac{3\gamma \alpha^2}{\lambda_2} |\bar{Z}_2(t)|^2 + \frac{\gamma^2}{\alpha^3} |\Delta_3|^2
\]

\[
\leq -\gamma (2\alpha \bar{L} - \frac{3\gamma^2}{\lambda_2} |\hat{Z}(t)|^2 - \frac{\gamma \lambda_2}{\lambda_2} |\xi(t)|^2 - \gamma (2\beta \lambda_2 - 2\alpha \lambda_N - \frac{3\gamma^2 \lambda_2^3}{\lambda_2}) |\bar{Z}_2(t)|^2 \\
+ 2\gamma^2 \alpha^2 |\Delta_1|^2 + 2\gamma^2 |\Delta_2|^2 + \frac{\gamma^2}{\alpha^3} |\Delta_3|^2
\]

Letting \( \alpha \geq \max \{1, \frac{1}{2}, \frac{3\gamma^2}{\lambda_2} \} \) and \( \beta \geq \max \{1, \frac{4\gamma^2 \lambda_2^3}{\lambda_2} \} \) gives

\[
\Delta(t) \leq -\frac{\gamma}{2} V(t) + 2\gamma^2 \alpha^2 |\Delta_1|^2 + 2\gamma^2 |\Delta_2|^2 + \frac{\gamma^2}{\alpha^3} |\Delta_3|^2
\]

To deal with the last three terms, we use Young’s inequality again. It follows that

\[
|\Delta_1|^2 \leq \bar{L} |\hat{Z}(t)|^2 \\
|\Delta_2|^2 \leq 2 \max \{(\beta - \alpha^2, 1)\lambda_N^2 (|Z_2(t)|^2 + |\xi(t)|^2) \\
\leq 2\beta^2 \lambda_N^2 (|\bar{Z}_2(t)|^2 + |\xi(t)|^2) \\
\frac{|\Delta_3|^2}{\alpha^3} \leq \frac{1}{\alpha} [R_L \xi(t) + \alpha R^T \Delta_1 - \alpha R_L \bar{Z}_2(t)]^2 \\
\leq \frac{3\lambda_N^2}{\alpha} |\xi(t)|^2 + 3\alpha \bar{L} |\hat{Z}(t)|^2 + 3\alpha \lambda_N^2 |\bar{Z}_2(t)|^2
\]

Combining these inequalities gives

\[
\Delta(t) \leq -\frac{\gamma}{2} V(t) + 2\gamma^2 \alpha^2 |\hat{Z}(t)|^2 + 4\gamma^2 \beta^2 \lambda_N^2 |\bar{Z}_2(t)|^2 + 4\gamma^2 \beta^2 \lambda_N^2 |\xi(t)|^2 + \frac{3\gamma^2 \lambda_N^2}{\alpha} |\xi(t)|^2 \\
+ 3\gamma^2 \alpha \bar{L} |\hat{Z}(t)|^2 + 3\gamma^2 \alpha \lambda_N^2 |\bar{Z}_2(t)|^2 \\
\leq -\frac{\gamma}{2} V(t) + 5\gamma^2 \beta^2 (\lambda_N^2 + \bar{L}) (|\hat{Z}(t)|^2 + |\xi(t)|^2) \\
\leq -\frac{\gamma}{2} V(t) + 5\gamma^2 \beta^2 (\lambda_N^2 + \bar{L}) \alpha^3 V(t) \\
\leq -\frac{\gamma}{2} V(t) + \frac{5}{16} \gamma^2 \beta^4 (\lambda_N^2 + \bar{L})^2 V(t)
\]
By setting $0 < \gamma < \frac{1}{\beta (\lambda^2 + \bar{l}^2)}$, it follows that

$$\Delta(t) \leq -\frac{3\gamma}{16} V(t)$$

According to Theorem 2 in [27], we obtain the exponential convergence of $V(t)$ and thus $\bar{Z}(t)$ to the origin as $t \to \infty$. By Lemma [27], the proof is complete.

References

[1] B. Gharesifard and J. Cortés, “Distributed continuous-time convex optimization on weight-balanced digraphs,” IEEE Transactions on Automatic Control, vol. 59, no. 3, pp. 781–786, 2014.

[2] A. Simonetto and G. Leus, “Distributed maximum likelihood sensor network localization,” IEEE Transactions on Signal Processing, vol. 62, no. 6, pp. 1424–1437, 2014.

[3] A. Nedic and A. Ozdaglar, “Distributed subgradient methods for multi-agent optimization,” IEEE Transactions on Automatic Control, vol. 54, no. 1, pp. 48–61, 2009.

[4] M. Zhu and S. Martínez, “On distributed convex optimization under inequality and equality constraints,” IEEE Transactions on Automatic Control, vol. 57, no. 1, pp. 151–164, 2012.

[5] Q. Liu, S. Yang, and Y. Hong, “Constrained consensus algorithms with fixed step size for distributed convex optimization over multiagent networks,” IEEE Transactions on Automatic Control, vol. 62, no. 8, pp. 4259–4265, 2017.

[6] D. Jakovetić, J. Xavier, and J. M. Moura, “Fast distributed gradient methods,” IEEE Transactions on Automatic Control, vol. 59, no. 5, pp. 1131–1146, 2014.

[7] W. Shi, Q. Ling, G. Wu, and W. Yin, “Extra: An exact first-order algorithm for decentralized consensus optimization,” SIAM Journal on Optimization, vol. 25, no. 2, pp. 944–966, 2015.

[8] K. Scaman, F. Bach, S. Bubeck, L. Massoulié, and Y. T. Lee, “Optimal algorithms for non-smooth distributed optimization in networks,” in Advances in Neural Information Processing Systems, pp. 2745–2754), Curran Associates, 2018.

[9] J. Lu and C. Y. Tang, “Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case,” IEEE Transactions on Automatic Control, vol. 57, no. 9, pp. 2348–2354, 2012.

[10] S. S. Kia, J. Cortés, and S. Martínez, “Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication,” Automatica, vol. 55, pp. 254–264, 2015.

[11] X. Zeng, P. Yi, and Y. Hong, “Distributed continuous-time algorithm for constrained convex optimizations via nonsmooth analysis approach,” IEEE Transactions on Automatic Control, vol. 62, no. 10, pp. 5227–5233, 2017.

[12] W. Ren and R. W. Beard, Distributed Consensus in Multi-vehicle Cooperative Control. London: Springer, 2007.

[13] Y. Tang and P. Yi, “Distributed coordination for a class of non-linear multi-agent systems with regulation constraints,” IET Control Theory Applications, vol. 12, no. 1, pp. 1–9, 2018.
[14] Y. Zhang, Z. Deng, and Y. Hong, “Distributed optimal coordination for multiple heterogeneous Euler–Lagrangian systems,” *Automatica*, vol. 79, pp. 207–213, 2017.

[15] Z. Qiu, L. Xie, and Y. Hong, “Distributed optimal consensus of multiple double integrators under bounded velocity and acceleration,” *Control Theory and Technology*, vol. 17, no. 1, pp. 85–98, 2019.

[16] Y. Xie and Z. Lin, “Global optimal consensus for higher-order multi-agent systems with bounded controls,” *Automatica*, vol. 99, pp. 301–307, 2019.

[17] Y. Tang, Z. Deng, and Y. Hong, “Optimal output consensus of high-order multiagent systems with embedded technique,” *IEEE Transactions on Cybernetics*, vol. 49, no. 5, pp. 1768–1779, 2019.

[18] X. Wang, Y. Hong, and H. Ji, “Distributed optimization for a class of nonlinear multiagent systems with disturbance rejection,” *IEEE Transactions on Cybernetics*, vol. 46, no. 7, pp. 1655–1666, 2016.

[19] Y. Tang, “Distributed optimization for a class of high-order nonlinear multiagent systems with unknown dynamics,” *International Journal of Robust Nonlinear Control*, vol. 28, no. 17, pp. 5545–5556, 2018.

[20] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.

[21] R. T. Rockafellar, *Convex Analysis*. Princeton: Princeton University Press, 2015.

[22] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia: SIAM, 2004.

[23] Y. Tang, Y. Hong, and X. Wang, “Distributed output regulation for a class of nonlinear multi-agent systems with unknown-input leaders,” *Automatica*, vol. 62, pp. 154–160, 2015.

[24] A. Nedic, A. Olshevsky, and W. Shi, “Achieving geometric convergence for distributed optimization over time-varying graphs,” *SIAM Journal on Optimization*, vol. 27, no. 4, pp. 2597–2633, 2017.

[25] J. Lei, H.-F. Chen, and H.-T. Fang, “Primal–dual algorithm for distributed constrained optimization,” *Systems & Control Letters*, vol. 96, pp. 110–117, 2016.

[26] Y. Xie and Z. Lin, “Global optimal consensus for multi-agent systems with bounded controls,” *Systems & Control Letters*, vol. 102, pp. 104–111, 2017.

[27] Z.-P. Jiang and Y. Wang, “A converse Lyapunov theorem for discrete-time systems with disturbances,” *Systems & Control Letters*, vol. 45, no. 1, pp. 49–58, 2002.