Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Distributed Optimization of Nonlinear Uncertain Systems: An Adaptive Backstepping Design ⋆

Zhengyan Qin * **Tengfei Liu** * **Zhong-Ping Jiang** **

* *State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, 110004, China (e-mail: qinzhengyan163@163.com, tfliu@mail.neu.edu.cn)*
** *Department of Electrical and Computer Engineering, New York University, 370 Jay Street, Brooklyn, NY 11201, USA (e-mail: zjiang@nyu.edu)*

Abstract: This paper proposes a Lyapunov-based adaptive backstepping approach to distributed optimization of nonlinear uncertain multi-agent systems. The model of each agent is in the strict-feedback form with parametric uncertainties. By only using local objective functions, this paper aims to solve the distributed optimization problem for the multi-agent system such that the outputs of the agents converge to the optimizer of the total objective function. Based on the idea of adaptive backstepping, the distributed optimization problem for the high-order multi-agent system is decomposed into solving the optimization or control problem for multiple first-order subsystems. The technical contributions lie in a Lyapunov-based design for distributed optimization, and a refined nonlinear damping design to deal with the newly appearing nonlinear uncertain terms caused by optimization. Based on the new designs, a Lyapunov function is constructed for the entire system, and the LaSalle-Yoshizawa Theorem is employed for convergence analysis. It is shown that the objective of distributed optimization is achievable if the local objective functions are convex with at least one of them being strongly convex. Computer-based numerical simulation is employed to show the effectiveness of the proposed design.

Keywords: Distributed optimization, adaptive control, multi-agent systems, nonlinear strict-feedback systems, parametric uncertainty.

1. INTRODUCTION

In the past decades, distributed optimization has attracted increasing attention due to the demand for solving optimization problems through parallel and coordinated computation **???**. Promising applications of distributed optimization include sensor networks **?**, signal processing **?**, power systems **?** and robotic networks **?**.

Aiming at developing a systematic theory for distributed optimization, **??** have exploited discrete-time algorithms and **?????** have constructed continuous-time control laws. Several existing methods have been refined for distributed optimization, such as gradient-based control **????**, intersection computation **?**, alternating direction method of multipliers (ADMM) **?** and extremum-seeking control (ESC) **?**. The information exchange topology has been generalized from the static case **??** to the switching case **??**.

Although most of the existing results focus on multi-agent systems modeled by first-order or second-order integrators, some recent results have shown the interest in systems involving more complex dynamics. The recent works **??** consider linear multi-agent systems, and the study in **??** focus on nonlinear multi-agent systems. **??** study the distributed adaptive optimization problem for nonlinear multi-agent systems in the normal form with parametric uncertainties.

This paper studies the distributed adaptive optimization for nonlinear uncertain systems in the popular strict-feedback form. References **??** also studied the observer-based adaptive fuzzy backstepping control of uncertain nonlinear pure-feedback systems.

The challenge lies in unified Lyapunov-based designs for distributed optimization and adaptive control, as well as the convergence analysis of the resulted closed-loop system. In this paper, the problem is solved through a refined backstepping design **?**. The initial step of the recursive design fixes the distributed optimization problem with the outputs of the agents considered as the virtual control input. Then, for each agent, a constructive design is developed to deal with the subsystems one-by-one, until the true control input appears. New nonlinear damping terms are employed to the constructive design to cope with the nonlinear terms introduced by the distributed optimiza-

⋆ This work was supported in part by NSFC grants 61522305, 61633007 and 61533007, in part by NSF grant EPCN-1903781, and in part by State Key Laboratory of Intelligent Control and Decision of Complex Systems at BIT.

tion algorithm. Along with the constructive design, a Lyapunov function can be constructed for the resulted closed-loop system at the same time, and the boundedness and convergence of the states are guaranteed by the LaSalle-Yoshizawa Theorem **?**.

2. NOTATIONS AND TERMINOLOGY

$|\cdot|$ denotes the Euclidean norm for vectors and the induced norm for matrices. For $A \in \mathbb{R}^{m\times n}$, $|A|$ represents its induced 2-norm, and $\operatorname{rank} A$ represents the rank of A. For real matrices A and B, $A \otimes B$ represents the Kronecker product. We use $\mathbf{1}_N$ to represent the N-dimensional $[1,\dots,1]^T$, and use Id to represent the identity function. By default, $\sum_{i=k}^{s} a_i = 0$ for any $a_i \in \mathbb{R}$ and any $k > s$.

For a function $\varphi : \mathbb{R}^p \to \mathbb{R}^q$, $\nabla\varphi$ represents the gradient wherever it exists. A function $\varphi : \mathbb{R}^p \to \mathbb{R}$ is convex if, for any $0 < a < 1$, $\varphi(a\zeta_1 + (1-a)\zeta_2) \le a\varphi(\zeta_1) + (1-a)\varphi(\zeta_2)$ holds for all $\zeta_1, \zeta_2 \in \mathbb{R}^p$. If φ is differentiable, then its convexity can be equivalently defined by $\nabla\varphi^T(\zeta_2)(\zeta_1 - \zeta_2) \le \varphi(\zeta_1) - \varphi(\zeta_2)$ for all $\zeta_1, \zeta_2 \in \mathbb{R}^p$. A differentiable function φ is strictly convex if the inequality above is strict whenever $\zeta_1 \ne \zeta_2$, and it is called ω-strongly convex with $\omega > 0$ if $(\nabla^T\varphi(\zeta_1) - \nabla^T\varphi(\zeta_2))(\zeta_1 - \zeta_2) \ge \omega|\zeta_1 - \zeta_2|^2$ for all $\zeta_1, \zeta_2 \in \mathbb{R}^p$. A vector-valued function $\psi : \mathbb{R}^p \to \mathbb{R}^p$ is called Lipschitz with constant ϑ, or simply called ϑ-Lipschitz, if $|\psi(\zeta_1) - \psi(\zeta_2)| \le \vartheta|\zeta_1 - \zeta_2|$ for all $\zeta_1, \zeta_2 \in \mathbb{R}^p$.

A weighted digraph $\mathcal{G}$ is a triple $(\mathcal{N}, \mathcal{E}, \mathcal{A})$ where $\mathcal{N} = \{1,\dots,N\}$ is a nonempty, finite set, $\mathcal{E}$ is a subset of $\mathcal{N}\times\mathcal{N}$ with $(i,i) \notin \mathcal{E}$ for all $i \in \mathcal{N}$, and $\mathcal{A} = [a_{ij}]_{N\times N}$ is the weighted adjacency matrix with $a_{ij} = 0$ for all i,j satisfying $(j,i) \notin \mathcal{E}$ and $a_{ij} > 0$ for all i,j satisfying $(j,i) \in \mathcal{E}$. Elements of $\mathcal{N}$ are referred to as nodes, and an element (i,j) of $\mathcal{E}$ is referred to as the edge from i to j. The Laplacian of a weighted digraph $\mathcal{G}$, denoted by L, is defined as $L = [l_{ij}]_{N\times N}$ with $l_{ii} = \sum_{j\ne i} a_{ij}$ and $l_{ij} = -a_{ij}$ for $j \ne i$. Clearly, $L\mathbf{1}_N = 0$. The digraph $\mathcal{G}$ is quasi-strongly connected (QSC) if there exists some $c \in \mathcal{N}$ such that there is a directed path from c to i for each $i \in \mathcal{N}\backslash\{c\}$; the node c is called the center of $\mathcal{G}$. The weighted digraph $\mathcal{G}$ is weight-balanced if $\sum_{j\in\mathcal{N}} a_{ij} = \sum_{j\in\mathcal{N}} a_{ji}$ holds for all $i \in \mathcal{N}$. For a weight-balanced digraph, $L^T\mathbf{1}_N = L\mathbf{1}_N = 0$. See Figure 1 in Section 6 for an example of a weighted digraph.

3. PROBLEM FORMULATION

Given a multi-agent system composed of N agents with each agent described by

$$\dot{x}_{ik} = x_{i(k+1)} + g_{ik}^T(\bar{x}_{ik})\theta_i, \quad k = 1,\cdots,n_i - 1 \tag{1}$$

$$\dot{x}_{in_i} = u_i + g_{in_i}^T(\bar{x}_{in_i})\theta_i, \tag{2}$$

$$y_i = x_{i1} \tag{3}$$

for $i \in \mathcal{N} = \{1,\dots,N\}$, where $x_{ik} \in \mathbb{R}^{n_y}$ for $k = 1,\cdots,n_i$ is the state, $u_i \in \mathbb{R}^{n_y}$ is the control input, $\bar{x}_{ik} = [x_{i1}^T,\cdots,x_{ik}^T]^T$, $y_i \in \mathbb{R}^{n_y}$ is the output, $g_{ik} : \mathbb{R}^{kn_y} \to \mathbb{R}^{n_y}$ is a vector valued locally Lipschitz function, and $\theta_i \in \mathbb{R}^{n_{\theta i}}$ are unknown parameters. Notice that the outputs of all the agents have the same dimension while their states could have different dimensions.

Consider the optimization problem

$$\min_{r\in\mathbb{R}^{n_y}} c(r), \quad c(r) = \sum_{i\in\mathcal{N}} c_i(r) \tag{4}$$

where each $c_i : \mathbb{R}^{n_y} \to \mathbb{R}$, referred to as local objective function, is differentiable. $c(r)$ is known as the (total) objective function.

We make an assumption on the local objective functions.

Assumption 1. The objective function c satisfies that

(1) each c_i with $i \in \mathcal{N}$ is convex, and at least one of c_i is ω-strongly convex with constant $\omega > 0$;
(2) for each $i \in \mathcal{N}$, the gradient ∇c_i is ϑ-Lipschitz with constant $\vartheta > 0$.

Remark 1. The first condition of Assumption 1 guarantees the existence of the unique solution to (4). The second condition is used for the design of a distributed optimization algorithm. Note that the second-order differentiability is not required for c_i and c, and at least one of the local objective functions is required to be strongly convex.

Denote $y_* \in \mathbb{R}^{n_y}$ as the global minimizer of c in (4). For the multi-agent system (1)–(2), *distributed optimal output agreement* aims to design a control law for each agent by using the local measurements x_i and $\nabla c_i(y_i)$ and the information exchange between the agents such that all the signals in the closed-loop system are bounded, and the outputs of the agents converge to y^*, i.e.,

$$\lim_{t\to\infty} y_i(t) = y_*. \tag{5}$$

The information exchange topology of the multi-agent system is described by a digraph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{A})$. Specifically, each agent is represented by a node in $\mathcal{G}$, and $(i,j) \in \mathcal{E}$ if the information of agent i is available to agent j. The element a_{ij} in the adjacency matrix $\mathcal{A}$ represents the weight of the edge (j,i). The following assumption is made on the digraph.

Assumption 2. The digraph $\mathcal{G}$ is QSC and weight-balanced.

4. BACKSTEPPING DESIGN FOR DISTRIBUTED ADAPTIVE OPTIMIZATION

4.1 Initial Step for Distributed Optimization

This subsection presents the design for the initial step, mainly focusing on distributed optimization. A Lyapunov function is constructed to characterize the proposed distributed optimization algorithm, and will be used later to construct a Lyapunov function for the entire system.

Consider the distributed optimization algorithm

$$\dot{y}_i^r = -\nabla c_i(y_i) - \sum_{j\in\mathcal{N}} a_{ij}(y_i^r - y_j^r) - \sum_{j\in\mathcal{N}} a_{ij}(q_i - q_j) \tag{6}$$

$$\dot{q}_i = \mu \sum_{j\in\mathcal{N}} a_{ji}(y_i^r - y_j^r) \tag{7}$$

where $y_i^r \in \mathbb{R}^{n_y}$ and $q_i \in \mathbb{R}^{n_y}$ are the states, y_i is the output of agent i, a_{ij} are elements of $\mathcal{A}$, and $\mu > 0$ is a design parameter.

Here, y_i^r can be considered as the virtual control input for distributed optimization. If $y_i^r = y_i$ and $a_{ij} = a_{ji}$, then the existing results, e.g., **?**, can readily guarantee the convergence of the algorithm.

For convenience of discussions, rewrite the system (6)–(7) in the compact form

$$\dot{y}^r = -\Delta_c(y) - L_\otimes y^r - L_\otimes q \tag{8}$$

$$\dot{q} = \mu L_\otimes^T y^r \tag{9}$$

where $y^r = [y_1^{rT}, \ldots, y_N^{rT}]^T$, $q = [q_1^T, \ldots, q_N^T]^T$, $y = [y_1^T, \ldots, y_N^T]^T$, $\Delta_c(y) = [\nabla^T c_1(y_1), \ldots, \nabla^T c_N(y_N)]^T$, and $L_\otimes = L \otimes I_{n_y}$.

The existing result (?, Theorem 4.1) gives the equilibrium of the system (8)–(9) with $y \equiv y^r$. The result is given here to make the paper self-contained.

Proposition 1. Under Assumptions 1 and 2, $[y_0^{rT}, q_0^T]^T$ is an equilibrium of the system (8)–(9) with $y \equiv y^r$ if

$$L_\otimes q_0 = -\Delta_c(\mathbf{1}_N \otimes y_*), \tag{10}$$

$$y_0^r = \mathbf{1}_N \otimes y_*. \tag{11}$$

Taking any $[y_0^{rT}, q_0^T]^T$ satisfying (10)–(11), define

$$\bar{y} = y^r - y_0^r, \quad \bar{q} = q - q_0. \tag{12}$$

Then, the system (8)–(9) is transformed into

$$\dot{\bar{y}} = -(\Delta_c(y^r) - \Delta_c(y_0^r)) - L_\otimes \bar{y} - L_\otimes \bar{q} + (\Delta_c(y^r) - \Delta_c(y)), \tag{13}$$

$$\dot{\bar{q}} = \mu L_\otimes^T \bar{y} \tag{14}$$

where $\Delta_c(y^r) = [\nabla^T c_1(y_1^r), \ldots, \nabla^T c_N(y_N^r)]^T$ and the property $L\mathbf{1}_N = 0$ is used. Denote $Z = [\bar{y}^T, \bar{q}^T]^T$ as the state of the system.

Define $\bar{y}_i = y_i^r - y_*$ for $i \in \mathcal{N}$. By using the definitions of y_0^r in (11) and $\bar{y}_i$ above, we have $\bar{y} = [\bar{y}_1^T, \ldots, \bar{y}_N^T]^T$. Also, define $\tilde{y} = [\tilde{y}_1^T, \ldots, \tilde{y}_N^T]^T$ with

$$\tilde{y}_i = y_i - y_i^r. \tag{15}$$

The following result gives a method to construct a Lyapunov function for the system (13)–(14).

Proposition 2. Consider the system (13)–(14). Under Assumptions 1 and 2, there exist constants $k_1, k_2, \chi_2, \chi_3 > 0$ and $0 < \chi_1 < 2/3$ such that

$$P = \begin{bmatrix} \mu k_1 I_N & k_2 L \\ k_2 L^T & k_1 I_N \end{bmatrix},$$

$$Q_1 = \mu k_1 (H^T + H - \chi_2 I_N) - k_2 \left(\frac{\vartheta^2}{\chi_1} I_N + 2\mu L L^T \right)$$

are positive definite, where $H = B + L$ and $B = \mathrm{diag}(b_1, \cdots, b_N)$ with

$$b_i = \begin{cases} \omega & \text{if } c_i \text{ is } \omega\text{-strongly convex}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Moreover, with

$$V_0(Z) = Z^T (P \otimes I_{n_y}) Z, \tag{17}$$

it holds that

$$\lambda_{\min}(P)|Z|^2 \le V_0(Z) \le \lambda_{\max}(P)|Z|^2 \tag{18}$$

$$\dot{V}_0(Z) \le -Z^T (Q \otimes I_{n_y}) Z + d|\tilde{y}|^2 \tag{19}$$

where $d = \vartheta^2(\mu k_1/\chi_2 + k_2/\chi_3)$ and

$$Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} \tag{20}$$

is positive semidefinite with $Q_2 = k_2 L^T L$ and $Q_3 = (2 - \chi_1 - \chi_3) k_2 L^T L$.

4.2 Recursive Steps: Handling High Order Uncertain Dynamics

This subsection presents the constructive design procedure for desired nonlinear adaptive control laws. It is shown that the distributed optimization algorithm leads to a nonlinear uncertain term appearing at each step of the recursive design, and the nonlinear damping technique is refined to handle it. To simplify the discussions, this subsection only gives the design for the case of $n_y = 1$.

For $i = 1, \cdots, N$, introduce the state transformation

$$e_{ik} = x_{ik} - \alpha_{i(k-1)}, \quad k = 1, \ldots, n_i \tag{21}$$

$$\tilde{\theta}_i = \theta_i - \hat{\theta}_i \tag{22}$$

where $\alpha_{i0} = y_i^r$, α_{ik} are virtual control laws of e_{ik}-subsystem for $k = 1, \cdots, n_i$, and $\hat{\theta}_i$ represents an estimate of the unknown parameter θ_i. Furthermore, for $k = 1, \cdots, n_i$ define $\bar{e}_{ik} = [e_{i1}, \cdots, e_{ik}]^T$ and

$$V_{ik}(\bar{e}_{ik}, \tilde{\theta}_i) = \frac{1}{2}|\bar{e}_{ik}|^2 + \frac{1}{2}\tilde{\theta}_i^T \Gamma_i^{-1} \tilde{\theta}_i. \tag{23}$$

Then, we recursively design α_{ik} as well as the update law for $\hat{\theta}_i$ until the appearance of the true control input u_i.

Step 1. From (1) and (21), the dynamics of e_{i1}-subsystem is

$$\dot{e}_{i1} = e_{i2} + \alpha_{i1} + w_{i1}^T \theta_i - \dot{y}_i^r, \tag{24}$$

where $w_{i1} = g_{i1}$. Taking the time derivative of V_{i1}, one has

$$\dot{V}_{i1} = e_{i1}(e_{i2} + \alpha_{i1} + w_{i1}^T \hat{\theta}_i - \dot{y}_i^r) - \tilde{\theta}_i^T \Gamma_i^{-1}(\dot{\hat{\theta}}_i - \Gamma_i w_{i1} e_{i1}). \tag{25}$$

Consider the virtual control law

$$\alpha_{i1} = -p_{i1} e_{i1} - w_{i1}^T \hat{\theta}_i - m_{i1} e_{i1}. \tag{26}$$

Then, (25) becomes

$$\dot{V}_{i1} = -p_{i1} e_{i1}^2 - m_{i1} e_{i1}^2 + e_{i1} e_{i2} - e_{i1} \dot{y}_i^r - \tilde{\theta}_i^T \Gamma_i^{-1}(\dot{\hat{\theta}}_i - \Gamma_i w_{i1} e_{i1}). \tag{27}$$

Step $k+1$ for $1 \le k \le n_i - 1$. Suppose that for the $\bar{e}_{ik}$-subsystem we have designed virtual control laws

$$\alpha_{ij} = -e_{i(j-1)} - p_{ij} e_{ij} + \sum_{s=1}^{j-1} \frac{\partial \alpha_{i(j-1)}}{\partial x_{is}} x_{i(s+1)} - w_{ij}^T \hat{\theta}_i + \frac{\partial \alpha_{i(j-1)}}{\partial \hat{\theta}_i} \Gamma_i \tau_{ij} - m_{ij} e_{ij} \left(\frac{\partial \alpha_{i(j-1)}}{\partial y_i^r} \right)^2 + v_{ij}, \tag{28}$$

where $w_{ij} = g_{ij} - \sum_{s=0}^{j-1} \frac{\partial \alpha_{i(j-1)}}{\partial x_{is}} g_{is}$ and

$$v_{ij} = \sum_{s=1}^{j} e_{i(s-1)} \frac{\partial \alpha_{i(s-2)}}{\partial \hat{\theta}_i} \Gamma_i w_{ij} \tag{29}$$

with $\alpha_{i(-1)} = e_{i0} = 0$, $\frac{\partial \alpha_{i0}}{\partial x_{i0}} = g_{i0} = 0$, $\tau_{ij} = \bar{w}_{ij}^T \bar{e}_{ij}$, $\bar{w}_{ij} = [w_{i1}^T, \cdots, w_{ij}^T]^T$ and $1 \le j \le k$ such that

$$\dot{V}_{ik} = -\sum_{s=1}^{k} p_{is} e_{is}^2 + e_{ik} e_{i(k+1)} + \tilde{\theta}_i^T (\tau_{ik} - \Gamma_i^{-1} \dot{\hat{\theta}}_i) + \left(\sum_{s=1}^{k} \frac{\partial \alpha_{i(s-1)}}{\partial \hat{\theta}_i} e_{is} \right) (\Gamma_i \tau_{ik} - \dot{\hat{\theta}}_i) - \sum_{s=1}^{k} m_{is} e_{is}^2 \left(\frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} \right)^2 - \sum_{s=1}^{k} \frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} e_{is} \dot{y}_i^r \tag{30}$$

Property (27) ensures the validity of the hypothesis for $k = 1$.

Consider the $e_{i(k+1)}$-subsystem. From (1) and (21), we have

$$\dot{e}_{i(k+1)} = e_{i(k+2)} + \alpha_{i(k+1)} - \sum_{s=1}^{k} \frac{\partial \alpha_{ik}}{\partial x_{is}} x_{i(s+1)} + w_{i(k+1)}^T \theta_i - \frac{\partial \alpha_{ik}}{\partial \hat{\theta}_i} \dot{\hat{\theta}}_i - \frac{\partial \alpha_{ik}}{\partial y_i^r} \dot{y}_i^r. \quad (31)$$

Then, along the trajectories of the system (31), the derivative of $V_{i(k+1)}$ satisfies

$$\begin{aligned} \dot{V}_{i(k+1)} =& -\sum_{s=1}^{k} p_{is} e_{is}^2 + \left(\sum_{s=1}^{k-1} \frac{\partial \alpha_{is}}{\partial \hat{\theta}_i} e_{i(s+1)} \right) (\Gamma_i \tau_{ik} - \dot{\hat{\theta}}_i) \\ &+ e_{i(k+1)} \Bigg[e_{ik} + e_{i(k+2)} + \alpha_{i(k+1)} \\ &- \sum_{s=1}^{k} \frac{\partial \alpha_{ik}}{\partial x_{is}} x_{i(s+1)} + w_{i(k+1)}^T \hat{\theta}_i - \frac{\partial \alpha_{ik}}{\partial \hat{\theta}_i} \dot{\hat{\theta}}_i \Bigg] \\ &+ \tilde{\theta}_i^T (\tau_{i(k+1)} - \Gamma_i^{-1} \dot{\hat{\theta}}_i) - \sum_{s=1}^{k+1} \frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} e_{is} \dot{y}_i^r. \end{aligned} \quad (32)$$

Consider the virtual control law

$$\begin{aligned} \alpha_{i(k+1)} =& -e_{ik} - p_{i(k+1)} e_{i(k+1)} + \sum_{s=1}^{k} \frac{\partial \alpha_{ik}}{\partial x_{is}} x_{i(s+1)} \\ &- w_{i(k+1)}^T \hat{\theta}_i + \frac{\partial \alpha_{ik}}{\partial \hat{\theta}_i} \Gamma_i \tau_{i(k+1)} \\ &- m_{i(k+1)} e_{i(k+1)} \left(\frac{\partial \alpha_{ik}}{\partial y_i^r} \right)^2 + v_{i(k+1)}. \end{aligned} \quad (33)$$

Then, the equation (32) becomes

$$\begin{aligned} \dot{V}_{i(k+1)} =& -\sum_{s=1}^{k+1} p_{is} e_{is}^2 + e_{i(k+1)} e_{i(k+2)} + \tilde{\theta}_i^T (\tau_{i(k+1)} \\ &- \Gamma_i^{-1} \dot{\hat{\theta}}_i) + \left(\sum_{s=1}^{k} \frac{\partial \alpha_{is}}{\partial \hat{\theta}_i} e_{i(s+1)} \right) (\Gamma_i \tau_{i(k+1)} - \dot{\hat{\theta}}_i) \\ &- \sum_{s=1}^{k+1} m_{is} e_{is}^2 \left(\frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} \right)^2 - \sum_{s=1}^{k+1} \frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} e_{is} \dot{y}_i^r, \end{aligned} \quad (34)$$

where we used

$$\begin{aligned} \dot{\hat{\theta}}_i - \Gamma_i \tau_{i(k-1)} &= \dot{\hat{\theta}}_i - \Gamma_i \tau_{ik} + \Gamma_i \tau_{ik} - \Gamma_i \tau_{i(k-1)} \\ &= \dot{\hat{\theta}}_i - \Gamma_i \tau_{ik} + \Gamma_i w_{ik} e_{ik}. \end{aligned}$$

Clearly, the $\dot{V}_{i(k+1)}$ defined by (34) is in the form of the $\dot{V}_{ik}$ defined by (30).

Step n_i. When $k = n_i$, set $e_{i(n_i+1)} = 0$ and $\alpha_{in_i} = u_i$. Then, equation (30) is rewritten as

$$\begin{aligned} \dot{V}_{in_i} =& -\sum_{s=1}^{n_i} p_{is} e_{is}^2 + e_{in_i} e_{i(n_i+1)} + \tilde{\theta}_i^T (\tau_{in_i} - \Gamma_i^{-1} \dot{\hat{\theta}}_i) \\ &+ \left(\sum_{s=1}^{n_i-1} \frac{\partial \alpha_{is}}{\partial \hat{\theta}_i} e_{i(s+1)} \right) (\Gamma_i \tau_{in_i} - \dot{\hat{\theta}}_i) \\ &- \sum_{s=1}^{n_i} m_{is} e_{is}^2 \left(\frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} \right)^2 - \sum_{s=1}^{n_i} \frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} e_{is} \dot{y}_i^r. \end{aligned} \quad (35)$$

Consider the control law for the true control input

$$\begin{aligned} u_i = \alpha_{in_i} =& -e_{i(n_i-1)} - p_{in_i} e_{in_i} + \sum_{s=1}^{n_i-1} \frac{\partial \alpha_{i(n_i-1)}}{\partial x_{is}} x_{i(s+1)} \\ &- w_{in_i}^T \hat{\theta} + \frac{\partial \alpha_{i(n_i-1)}}{\partial \hat{\theta}_i} \Gamma_i \tau_{in_i} \\ &- m_{in_i} e_{in_i} \left(\frac{\partial \alpha_{i(n_i-1)}}{\partial y_i^r} \right)^2 + v_{in_i} \end{aligned} \quad (36)$$

where $w_{in_i} = g_{in_i} - \sum_{s=1}^{n_i-1} \frac{\partial \alpha_{i(n_i-1)}}{\partial x_{is}} g_{is}$ and $v_{in_i} = \sum_{s=1}^{n_i} e_{i(s-1)} \frac{\partial \alpha_{i(s-2)}}{\partial \hat{\theta}_i} \Gamma_i w_{in_i}$.

To eliminate the terms of $\tilde{\theta}$ from the right-hand side of (35), we design the update law for $\hat{\theta}_i$ as

$$\dot{\hat{\theta}}_i = \Gamma_i \tau_{in_i}(\bar{e}_{in_i}, \hat{\theta}_i) = \Gamma_i \bar{w}_{in_i}^T \bar{e}_{in_i}. \quad (37)$$

Then, equation (35) becomes

$$\begin{aligned} \dot{V}_{in_i} =& -\sum_{s=1}^{n_i} p_{is} e_{is}^2 - \sum_{s=1}^{n_i} m_{is} e_{is}^2 \left(\frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} \right)^2 \\ &- \sum_{s=1}^{n_i} \frac{\partial \alpha_{i(s-1)}}{\partial y_i^r} e_{is} \dot{y}_i^r \end{aligned} \quad (38)$$

Remark 2. The proposed recursive design introduces nonlinear damping terms **?** to the virtual control laws to address the terms of $\dot{y}_i^r$, to avoid the usage of high-order derivatives of y_i^r.

5. STABILITY ANALYSIS

This section gives the main result of the paper.

Theorem 1. Consider the multi-agent system (1)–(3) with the controller defined by (6)–(7) and (21)–(37). Under Assumptions 1 and 2, the objective of distributed optimal output agreement is achievable by appropriately choosing p_{ij} and m_{ij} for $i = 1, \ldots, N$ and $j = 1, \ldots, n_i$.

Proof. Denote $X = [Z^T, \bar{e}_{1n_1}^T, \tilde{\theta}_1^T, \cdots, \bar{e}_{Nn_N}^T, \tilde{\theta}_i^T]^T$. Define

$$V(X) = bV_0(Z) + \sum_{i=1}^{N} V_{in_i}(\bar{e}_{in_i}, \tilde{\theta}_i) \quad (39)$$

with $b > 0$ to be determined later. Then, along the trajectories of the closed-loop system considered by Theorem 1, it holds that

$$\dot{V}(X) = b\dot{V}_0(Z) + \sum_{i=1}^{N} \dot{V}_{in_i}(\bar{e}_{in_i}, \tilde{\theta}_i^T)$$
$$\leq -bZ^T(Q \otimes I_{n_y})Z + bd|\tilde{y}|^2 - \sum_{i=1}^{N}\sum_{s=1}^{n_i} p_{is}e_{is}^2$$
$$+ \sum_{i=1}^{N}\sum_{s=1}^{n_i} \frac{1}{4m_{is}}|\dot{y}_i^r|^2. \quad (40)$$

By using $\dot{\bar{y}}_i = \dot{y}_i^r - \dot{y}_* = \dot{y}_i^r$, (13) and $L\bar{q} = \bar{L}\hat{q}$, we have
$$|\dot{y}_i^r| \leq \vartheta(|\bar{y}_i| + |\tilde{y}_i|) + |L_i||\bar{y}| + |\bar{L}_i||\hat{q}|$$
which used the ϑ-Lipschitz property of ∇c_i. Then, it holds that
$$|\dot{y}_i^r|^2 \leq 4\vartheta^2(|\bar{y}_i|^2 + |\tilde{y}_i|^2) + 4|L_i|^2|\bar{y}|^2 + 4|\bar{L}_i|^2|\hat{q}|^2. \quad (41)$$
Substituting (41) into (40) yields
$$\dot{V}(X) \leq -bZ^T(Q \otimes I_{n_y})Z + bd|\tilde{y}|^2 - \sum_{i=1}^{N}\sum_{s=1}^{n_i} p_{is}e_{is}^2$$
$$+ \sum_{i=1}^{N}\sum_{s=1}^{n_i} \frac{1}{4m_{is}}\Big(4\vartheta^2(|\bar{y}_i|^2 + |\tilde{y}_i|^2) + 4|L_i|^2|\bar{y}|^2$$
$$+ 4|\bar{L}_i|^2|\hat{q}|^2\Big). \quad (42)$$
By using (19), $\bar{Q}_3 = (2 - \chi_1 - \chi_3)k_2\bar{L}^T\bar{L}$ and $L\bar{q} = \bar{L}\hat{q}$, inequality (42) becomes
$$\dot{V}(X) \leq -bZ'^T(S \otimes I_{n_y})Z' - b\tilde{y}^T\bar{d}\tilde{y} - \sum_{i=1}^{N}\sum_{s=1}^{n_i} p_{is}e_{is}^2 \quad (43)$$
with $Z' = [\bar{y}^T, \hat{q}^T]^T$, $S = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{bmatrix}$, and
$$S_1 = bQ_1 - \sum_{i=1}^{N} \bar{m}_i|L_i|^2 I_N - \vartheta^2 \mathrm{diag}(\bar{m}_1, \cdots, \bar{m}_N),$$
$$S_2 = k_2\bar{L}^T\bar{L},\ S_3 = b\bar{Q}_3 - \sum_{i=1}^{N} \bar{m}_i|\bar{L}_i|^2 I_N,\ \bar{m}_i = \sum_{s=1}^{n_i}\frac{1}{m_{is}},$$
$$\bar{d} = diag(p_{11} - \bar{m}_1\vartheta^2, \cdots, p_{N1} - \bar{m}_N\vartheta^2) - bdI_N.$$

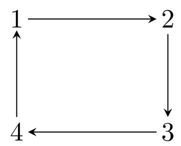
Proposition 2 guarantees the positive definiteness of Q_1 and $\bar{Q}_3$, which means that given specific m_{is}, L_i and $\bar{L}_i$, we can choose b large enough such that S is positive definite. With such b, we can choose p_{i1} large enough such that $\bar{d}$ is positive definite. Then, (43) implies
$$\dot{V}(X) \leq -bZ'^T(S \otimes I_{n_y})Z' - b\bar{d}|\tilde{y}|^2 - \sum_{i=1}^{N}\sum_{s=1}^{n_i} p_{is}e_{is}^2 \leq 0. \quad (44)$$

This guarantees the boundedness of X (**?**, Theorem 4.1). Then, the boundedness of x_{ik} for $i = 1, \ldots, N$ and $k = 1, \ldots, n_i$ can be proved following the standard analysis of the backstepping approach **?**.

With (44), by using the LaSalle-Yoshizawa Theorem (**?**, Theorem 2.1), we have
$$\lim_{t\to\infty}\Big(-bZ'^T(S \otimes I_{n_y})Z' - b\bar{d}|\tilde{y}|^2 - \sum_{i=1}^{N}\sum_{s=1}^{n_i} p_{is}e_{is}^2\Big) = 0$$
Recall that S and $\bar{d}$ are positive definite, and $p_{ik} > 0$ for $k = 2, \cdots, n_i$. Then, it holds that
$$\lim_{t\to\infty}|\bar{y}(t)| = \lim_{t\to\infty}|\tilde{y}(t)| = \lim_{t\to\infty}|\hat{q}(t)| = \lim_{t\to\infty}|e_{ik}| = 0, \quad (45)$$
for $k = 2, \cdots, n_i$, which, together with the fact $y_i - y_* = y_i - y_i^r + y_i^r - y_* = \tilde{y}_i + \bar{y}_i$, leads to $\lim_{t\to\infty}(y_i(t) - y_*) = 0$, and equivalently the satisfaction of (5). This ends the proof. ◇

6. AN EXAMPLE

In this section, we employ a multi-agent system composed of four pendulums to show the effectiveness of the proposed design. For $i = 1, \ldots, 4$, each pendulum is described by (**?**, p. 5)
$$\dot{x}_{i1} = x_{i2} \quad (46)$$
$$\dot{x}_{i2} = -\frac{g}{l_i}\sin x_{i1} - \frac{\theta_i}{m_i}x_{i2} + \frac{1}{m_i l_i^2}T_i \quad (47)$$
$$y_i = x_{i1} \quad (48)$$
where x_{i1} and x_{i2} are the states, g is the acceleration due to gravity, l_i and m_i are the length and the mass of the bob separately and T_i is the control input representing the torque applied to the bob. In our example, we take $m_1 = 1, m_2 = 1.2, m_3 = 1, m_4 = 0.5, l_1 = 0.1, l_2 = 0.2, l_3 = 0.15, l_4 = 0.1, \theta_1 = 0.5, \theta_2 = 0.4, \theta_3 = 0.2, \theta_4 = 0.3$ and $g = 9.8$. The objective is to render all of the four pendulums to a desired position, i.e., the outputs y_i for $i = 1, 2, 3$ go to a common value determined by the optimization problem (4) with $c_i(r) = 0.5(r - i)^2$ for $i = 1, 2, 4$ and $c_3(r) = r - 1$. The information exchange topology and its adjacency matrix are shown in Figure 1. One can check that Assumptions 1 and 2 are satisfied and the optimal solution of (4) is $y_* = 2$.



$$\mathcal{A} = \begin{bmatrix} 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{bmatrix}$$

Fig. 1. The information exchange digraph $\mathcal{G}$ and the adjacency matrix $\mathcal{A}$.

By taking $g_{i1} = 0, g_{i2} = -x_{i2}/m_i$ and introducing a new control input $\tilde{u}_i$ such that $T_i = m_i l_i^2(g/l_i \sin x_{i1} + u_i)$, the system (46)–(48) is transformed into the form of (1)–(3).

We use the distributed coordinator (6)–(7) with $\mu = 1$ and the adaptive backstepping control laws in the form of (21)–(37) with the estimation $\hat{\theta}_i$ in the form of (37). In particular, the control law is $u_i = \alpha_{i2} = -e_{i1} - p_{i2}e_{i2} - (p_{i1} + m_{i1})x_{i2} - w_{i2}\hat{\theta}_i + w_{i1}\Gamma_i x_{i2}e_{i2}/m_i - m_{i2}e_{i2}(p_{i1} + m_{i1})^2$, $\dot{\hat{\theta}}_i = -x_{i2}e_{i2}/m_i$ with the parameters $\Gamma_i = 1, m_{11} = m_{22} = m_{32} = m_{41} = m_{42} = 0.6, m_{12} = m_{21} = 0.8, m_{31} = 0.7, p_{12} = p_{32} = 0.002, p_{22} = p_{42} = 0.001, p_{11} = 79.98, p_{21} = 79.99, p_{31} = 80.1$ and $p_{41} = 80$. The trajectories of y_i are shown in Figure 2. It is shown that all the y_i converge to $y_* = 2$.

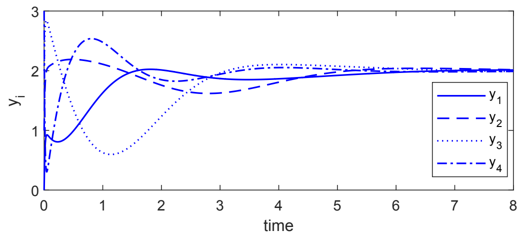


Fig. 2. Trajectories of the outputs of four pendulums with initial states $y^r(0) = [1, 2, 3, 0]$, $q(0) = [1, -1, 1, 0]$, $x_1(0) = [0, 1, 2, 3]$ and $x_2(0) = [1, 1, 0.1, 0]$.

7. CONCLUSIONS

This paper studies the distributed optimal output agreement problem of nonlinear multi-agent systems taking the parametric strict-feedback form. A class of distributed optimal coordinators is proposed and shown to have robustness with respect to tracking errors. This kind of coordinators can estimate the optimizer exponentially if there exists no tracking error under a mild assumption on the optimal objective function and the interconnection topologies. By means of a refined adaptive backstepping method, state-feedback control laws are constructed that render the closed-loop multi-agent system stable. In addition, it is shown that the distributed optimal output agreement problem can be solved.