Passivity-Based Distributed Optimization with Communication Delays Using PI Consensus Algorithm

Takeshi Hatanaka, *Member, IEEE*, Nikhil Chopra, *Member, IEEE*, Takayuki Ishizaki, *Member, IEEE*, and Na Li, *Member, IEEE*

***Abstract*—In this paper, we address a class of distributed optimization problems in the presence of inter-agent communication delays based on passivity. We first focus on unconstrained distributed optimization and provide a passivity-based perspective for distributed optimization algorithms. With the help of the scattering transformation, this perspective allows us to handle arbitrary and unknown constant communication delays in the distributed optimization problem. Moreover, we extend the results to constrained distributed optimization. Finally, the algorithm is applied to a visual human localization problem using a pedestrian detection algorithm.**

***Index Terms*—Distributed optimization, Passivity, PI consensus, Communication delays, Scattering transformation**

I. Introduction

Passivity has been extensively studied for analysis and design of both linear and nonlinear systems with broad applications in physical systems [1] and network systems [2], [3]. Recently, passivity is also used in optimization and related application fields including Internet congestion control [4], CDMA power control [5], electrical grid control [6] and building control [7]. The modularized property and composability of the passivity-based design allows one to improve the performance and robustness [4], [5], or integrate another passive components like physical dynamics while ensuring stability and optimality for the overall system [6], [7].

In this paper, we deal with the class of distributed optimization studied in [9]–[14]. Nedić and Ozdaglar [9] present a distributed algorithm which combines consensus algorithms and subgradient methods. The results are extended to constrained problems in [10], where a variation of the algorithm in [9] is shown to ensure exact convergence to the optimal solution using a diminishing step size. Solution to the problem with globally defined inequality and equality constraints is presented in [11], [12]. While the above works present discrete-time recursive processes to compute the solution, Wang and Elia [13], [14] take a continuous-time algorithm and provide a control theoretic perspective for the distributed optimization algorithms.

T. Hatanaka (corresponding author) and T. Ishizaki are with School of Engineering, Tokyo Institute of Technology, 2-12-1 S5-16 (Hatanaka)/W8-1 (Ishizaki), Ookayama, Meguro-ku, Tokyo 152-8550, JAPAN. Tel: +81-3-5734-3316 (Hatanaka), +81-3-5734-2646 (Ishizaki). Email: hatanaka@ctrl.titech.ac.jp (Hatanaka), ishizaki@mei.titech.ac.jp (Ishizaki) N. Chopra is with Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA. Tel: +1-301-405-7011. Email: nchopra@umd.edu. N. Li is with Electrical Engineering and Applied Mathematics of the School of Engineering and Applied Sciences, Harvard University, 33 Oxford St, Cambridge, MA 02138, USA. Tel: +1-617-496-1441. Email: nali@seas.harvard.edu.

The results of this paper are partially presented in the authors' antecessor [24], but communication delays are not addressed there.

All of the solutions in [9]–[14] rely on the inter-agent information exchanges and are typically implemented using communication technology, which poses several challenges both in theory and practice. Among many of such problems, this paper treats delays inherent in the communication. Distributed optimization with delays is investigated e.g. in [16]–[21]. An early work [16] addresses unconstrained distributed optimization with delays and present a distributed solution combining consensus algorithms and subgradient methods. Tsianos and Rabbat [18] and Terelius et al. [17] present distributed solutions to a problem with constraints, but information broadcasting over the network is assumed in [17] and prior knowledge on the entire network structure together with delays at all links is required in [18]. While all of the above works need to take a diminishing step size to ensure asymptotic optimality, Wu et al. [19] presents a distributed algorithm with a constant step size for time-varying delays to avoid slow convergence for the diminishing step size. However, they consider only unconstrained problems. Delays in the gradient descent loop are investigated in [20], [21].

In this paper, we approach the above problem based on passivity. To this end, we first present a passivity-based perspective for the algorithm based on so-called PI (Proportional-Integral) consensus algorithm [15]. In particular, it is revealed that the PI consensus-based solution is regarded as a feedback connection of passive systems. We next treat communication delays using our passivity-based perspective. Specifically, we show that the delays are successfully integrated with the above solution by using the techniques in [22] together with the scattering transformation [3]. Exact convergence to the optimal solution is then proved based on passivity. The above results are then extended to a constrained optimization problem. Finally, the present algorithm is applied to a visual human localization problem using a pedestrian detection algorithm.

II. Passivity-Based Perspective for PI Consensus-Based Distributed Optimization

In this section, we consider a network of n agents to solve the following distributed optimization problem in [9].

$$\min_{z\in\mathbb{R}^N} f(z) := \sum_{i=1}^{n} f_i(z) \tag{1}$$

The function $f_i : \mathbb{R}^N \to \mathbb{R}$ $(i = 1, 2, \ldots, n)$ is the private cost function of agent i, which is assumed to be inaccessible from agents other than i. The subsequent discussions rely on the following assumption.

Assumption 1 *The functions $f_1, f_2, \ldots, f_n$ are convex, continuously differentiable, and their gradients denoted by $\phi_i := \nabla f_i$ $(i = 1, 2, \ldots, n)$ are locally Lipschitz.*

Throughout this paper, we assume that the set of the optimal solutions is not empty, and the corresponding minimal value of f is finite. Since f is also continuously differentiable and convex under Assumption 1, a vector $z^* \in \mathbb{R}^N$ is an optimal solution to (1) iff $\nabla f(z^*) = \sum_{i=1}^n \phi_i(z^*) = 0$ holds [8].

The agents are assumed to exchange information with neighboring agents through a network modeled by a graph $G := (\{1, 2, \ldots, n\}, \mathcal{E})$ satisfying the following assumption.

Assumption 2 *The graph G is undirected and connected.*

The set of all neighbors of agent i is denoted by $\mathcal{N}_i$.This assumption is weaker than [17] and [20] if the multi-hop communication is identified with all-to-all communication, and is compatible with [18] and [19].

In the sequel, we use the following well-known result.

Lemma 1 *Consider a convex function $f : \mathbb{R}^N \to \mathbb{R}$. Then, its gradient $\nabla f : \mathbb{R}^N \to \mathbb{R}^N$ is incrementally passive, i.e., the following inequality holds for any $x, y \in \mathbb{R}^N$.*

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0 \tag{2}$$

If f is strictly convex, (2) strictly holds as long as $x \neq y$.

In this section, we focus on a distributed algorithm based on a PI consensus algorithm [15], formulated as

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i) - \sum_{j \in \mathcal{N}_i} b_{ij}(\xi_j - \xi_i) + u_i, \tag{3a}$$

$$\dot{\xi}_i = \sum_{j \in \mathcal{N}_i} b_{ij}(x_j - x_i), \tag{3b}$$

where $x_i \in \mathbb{R}^N$ is an estimate of the optimal solution z^*, $u_i \in \mathbb{R}^N$ is an external input, $\xi_i \in \mathbb{R}^N$ is an additional variable which generates the integral of the consensus input $\sum_{j \in \mathcal{N}_i} b_{ij}(x_j - x_i)$. We assume that $a_{ij} > 0$ and $b_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ and $b_{ij} = 0$ otherwise, and $a_{ij} = a_{ji},\ b_{ij} = b_{ji}\ \forall i, j$.

Defining $\xi := [\xi_1^T\ \xi_2^T\ \cdots\ \xi_n^T]^T$ and $u := [u_1^T\ u_2^T\ \cdots\ u_n^T]^T$, the total system is described as

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \end{bmatrix} = - \begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} u, \tag{4}$$

where $\bar{L}_{\rm P} := L_{\rm P} \otimes I_N$, $\bar{L}_{\rm I} := L_{\rm I} \otimes I_N$, $L_{\rm P}$ and $L_{\rm I}$ are the graph Laplacians associated with the adjacency matrix with elements a_{ij} and b_{ij}, respectively. The matrices $L_{\rm P}$ and $L_{\rm I}$ are symmetric and positive semidefinite with a simple zero eigenvalue. We then have the following lemma.

Lemma 2 *[13] Under Assumptions 1 and 2, there exists ξ^* such that $[(x^*)^T\ (\xi^*)^T]^T$ is an equilibrium of (4) with $u \equiv u^* := -\alpha\phi(x^*)$, $\phi(x) := [\phi_1^T(x_1)\ \cdots\ \phi_n^T(x_n)]^T$ and $x^* := \mathbf{1}_n \otimes z^*$, where $\mathbf{1}_n$ denotes the n dimensional vector whose elements all equal to 1.*

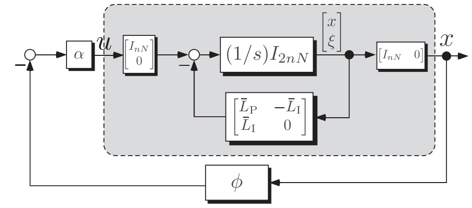

Fig. 1. Block diagram of the PI consensus-based distributed optimization algorithm for unconstrained optimization. The system enclosed by the dashed line is passive from $\tilde{u} = u - u^*$ to $\tilde{x} = x - x^*$ with $x^* := \mathbf{1}_n \otimes z^*$ and $u^* := -\alpha\phi(x^*)$ (Lemma 3). The bottom block ϕ is incrementally passive (Lemma 1), and hence passive from $x - x^*$ to $\phi(x) - \phi(x^*)$.

From Lemma 2 and linearity of (4), we immediately have

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{\xi}} \end{bmatrix} = - \begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix} + \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} \tilde{u}, \tag{5}$$

where $\tilde{x} := x - x^*$, $\tilde{\xi} := \xi - \xi^*$ and $\tilde{u} := u - u^*$. The above system (5) is proved to be passive as below.

Lemma 3 *The system (5) is passive from $\tilde{u}$ to $\tilde{x}$ with respect to the storage function $\tilde{S} := \frac{1}{2}\|\tilde{x}\|^2 + \frac{1}{2}\|\tilde{\xi}\|^2$.*

Proof: The time derivative of $\tilde{S}$ along the system trajectories is given by

$$\begin{aligned} \dot{\tilde{S}} &= - \begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix}^T \begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix} + \begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix}^T \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} \tilde{u} \\ &= -\tilde{x}^T \bar{L}_{\rm P} \tilde{x} + \tilde{x}^T \tilde{u} \leq \tilde{x}^T \tilde{u}. \end{aligned} \tag{6}$$

This completes the proof. ∎

Let us now close the loop of u by the gradient-based feedback law $u = -\alpha\phi(x)$. Then, the closed-loop system is formulated as

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \end{bmatrix} = - \begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} - \alpha \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} \phi(x), \tag{7}$$

whose block diagram is illustrated in Fig. 1. Lemma 1 means that the block ϕ in Fig. 1 is passive from $\tilde{x}$ to $\phi(x) - \phi(x^*)$. Noticing $u = -\alpha\phi(x)$ and $u^* = -\alpha\phi(x^*)$, it is concluded that the system in Fig. 1 is a feedback interconnection of passive systems, which is known to ensure energy dissipation [1]–[3]. Based on this perspective, we can prove the following convergence result.

Theorem 1 *Consider the system (7). If Assumptions 1 and 2 hold, then x_i asymptotically converges to the set of optimal solutions to (1) for all $i = 1, 2, \ldots, n$.*

Remark 1 *The above algorithm together with the convergence result compatible with ours was already presented in [13]. The contribution of this section is not to prove convergence itself but to provide a passivity-based perspective that (7) is regarded as feedback connection of two passive systems with incremental inputs and outputs. It will be shown below that this perspective provides fruitful design concepts.*

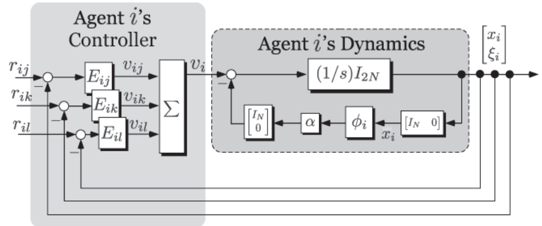

Fig. 2. Block diagram of agent i's dynamics for unconstrained optimization. The system enclosed by the dashed line is passive from $\bar{v}_i$ to $[\bar{x}_i^T\ \ \bar{\xi}_i^T]^T$ (Lemma 4), where $\bar{v}_i$, $\bar{x}_i$ and $\bar{\xi}_i$ are defined in (11) and (12).

III. Unconstrained Distributed Optimization with Communication Delay

In this section, we suppose that the inter-agent communication suffers from heterogeneous albeit constant [1] delays The delay from agent i to j is denoted by T_{ij} for any pair $(i,j) \in \mathcal{E}$.

Let us focus on the individual agent's dynamics (3) with $u_i = -\alpha\phi_i(x_i)$, while viewing the messages from neighbors $j \in \mathcal{N}_i$ as external inputs, denoted by r_{ij}^x and r_{ij}^ξ, as

$$\dot{x}_i = \sum_{j\in\mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) - \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^\xi - \xi_i) - \alpha\phi_i(x_i) \quad \text{(8a)}$$

$$\dot{\xi}_i = \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^x - x_i) \quad \text{(8b)}$$

whose block diagram is illustrated in Fig. 2, where

$$E_{ij} := \begin{bmatrix} a_{ij}I_N & -b_{ij}I_N \\ b_{ij}I_N & 0 \end{bmatrix},\ j \in \mathcal{N}_i.$$

It is easy to confirm that E_{ij} is a passive map.

In the delay free case addressed in the previous section, r_{ij}^x and r_{ij}^ξ are simply set to $r_{ij}^x = x_j$ and $r_{ij}^\xi = \xi_j$. However, this strategy with the delays, namely $r_{ij}^x(t) = x_j(t - T_{ji})$ and $r_{ij}^\xi(t) = \xi_j(t - T_{ji})$, may destabilize the system as will be confirmed in Section V. We thus need to redesign r_{ij}^x and r_{ij}^ξ to ensure convergence to the optimal solution even in the presence of the delays.

In this section, we make the following assumption.

Assumption 3 *The functions $f_i (i = 1, 2, \ldots, n)$ depend only on $(z_l)_{l\in\mathcal{Z}_i}$ for a subset $\mathcal{Z}_i \subseteq \{1, 2, \ldots, N\}$, where z_l is the l-th element of z and are continuously differentiable, and their gradients are locally Lipschitz. Every f_i $(i = 1, 2, \ldots, n)$ is strictly convex in $(z_l)_{l\in\mathcal{Z}_i}$, where $\mathcal{Z}_i \subseteq \{1, 2, \ldots, N\}$. Also, $\cup_{i=1}^N \mathcal{Z}_i = \{1, 2, \ldots, N\}$ holds.*

Under the above assumption together with the existence of the optimal solution, the solution z^* is uniquely determined [8]. Remark that $\mathcal{Z}_i$ can be empty for some i.

[1] Remark that actual delays may be time-varying but they are known to be reduced to the constant delay model using the buffering technique presented in [23] as long as an upper bound of the delays are available.

A. Passivity-like Property in Individual Dynamics

Consider the system in Fig. 2. The system is then regarded as a feedback system with the agent dynamics

$$\begin{bmatrix} \dot{x}_i \\ \dot{\xi}_i \end{bmatrix} = v_i - \alpha \begin{bmatrix} \phi_i(x_i) \\ 0 \end{bmatrix} \quad (9)$$

and the controller

$$v_i = \sum_{j\in\mathcal{N}_i} v_{ij},\ v_{ij} := \begin{bmatrix} v_{ij}^x \\ v_{ij}^\xi \end{bmatrix} = E_{ij} \begin{bmatrix} r_{ij}^x - x_i \\ r_{ij}^\xi - \xi_i \end{bmatrix}. \quad (10)$$

Let us focus on the open-loop system (9) with input v_i enclosed by the dashed line in Fig. 2. Then, we have the following lemma.

Lemma 4 *Suppose that Assumption 3 holds. Then, the system (9) is passive from $\bar{v}_i$ to $[\bar{x}_i^T\ \bar{\xi}_i^T]^T$ with respect to the storage function $\bar{S}_i := \frac{1}{2}\|\bar{x}_i\|^2 + \frac{1}{2}\|\bar{\xi}_i\|^2$, where*

$$\bar{x}_i := x_i - z^*,\ \bar{\xi}_i := \xi_i - 2\xi_i^*, \quad (11)$$

$$\bar{v}_i := v_i - \sum_{j\in\mathcal{N}_i} v_{ij}^*,\ v_{ij}^* := b_{ij} \begin{bmatrix} \xi_i^* - \xi_j^* \\ 0 \end{bmatrix}. \quad (12)$$

The symbol $\xi_i^* \in \mathbb{R}^N$ is a vector such that the stack vector $[(\xi_1^*)^T\ \cdots\ (\xi_n^*)^T]^T$ is equal to ξ^* in Lemma 2.

Proof: Subtracting the stationary equation of (7) for the equilibrium in Lemma 2 from (9) yields

$$\begin{bmatrix} \dot{\bar{x}}_i \\ \dot{\bar{\xi}}_i \end{bmatrix} = \bar{v}_i - \alpha \begin{bmatrix} \phi_i(x_i) - \phi_i(z^*) \\ 0 \end{bmatrix} \quad (13)$$

because of the definition of $\bar{v}_i$ in (12). The time derivative of $\bar{S}_i$ along the trajectories of (13) is then given by

$$\dot{\bar{S}}_i = \begin{bmatrix} \bar{x}_i \\ \bar{\xi}_i \end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \le \begin{bmatrix} \bar{x}_i \\ \bar{\xi}_i \end{bmatrix}^T \bar{v}_i, \quad (14)$$

where we use Lemma 1 and Assumption 3. ■

We next close the loop between (9) and (10). Define

$$\bar{r}_{ij} := \begin{bmatrix} \bar{r}_{ij}^x \\ \bar{r}_{ij}^\xi \end{bmatrix} = \begin{bmatrix} r_{ij}^x \\ r_{ij}^\xi \end{bmatrix} - r_{ij}^*,\ r_{ij}^* := \begin{bmatrix} z^* \\ \xi_i^* + \xi_j^* \end{bmatrix} \quad (15)$$

and $\bar{v}_{ij} := v_{ij} - v_{ij}^*$. Then, from (10), (11) and (12), we have

$$\bar{v}_{ij} = E_{ij} \begin{bmatrix} r_{ij}^x - x_i \\ r_{ij}^\xi - \xi_i \end{bmatrix} - b_{ij} \begin{bmatrix} \xi_i^* - \xi_j^* \\ 0 \end{bmatrix} = E_{ij} \begin{bmatrix} \bar{r}_{ij}^x - \bar{x}_i \\ \bar{r}_{ij}^\xi - \bar{\xi}_i \end{bmatrix}.$$

Substituting this together with $\bar{v}_i = \sum_{j\in\mathcal{N}_i} \bar{v}_{ij}$ into (14) proves the following passivity-like property.

$$\begin{aligned} \dot{\bar{S}}_i = &\sum_{j\in\mathcal{N}_i} \begin{bmatrix} \bar{x}_i \\ \bar{\xi}_i \end{bmatrix}^T \begin{bmatrix} a_{ij}(\bar{r}_{ij}^x - \bar{x}_i) + b_{ij}(\bar{\xi}_i - \bar{r}_{ij}^\xi) \\ b_{ij}(\bar{r}_{ij}^x - \bar{x}_i) \end{bmatrix} \\ &- \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \\ = &\sum_{j\in\mathcal{N}_i} \bar{r}_{ij}^T \bar{v}_{ij} - \sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2 \\ &- \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \le \sum_{j\in\mathcal{N}_i} \bar{r}_{ij}^T \bar{v}_{ij}. \quad (16) \end{aligned}$$

Remark that, from (12) and (15), the following equations hold, which plays an important role in deriving the subsequent theoretical results in this section.

$$v_{ji}^* = -v_{ij}^*,\ r_{ji}^* = r_{ij}^*. \quad (17)$$

B. Scattering Transformation

In view of the architecture of [22], the passivity-like property (16) inspires us to exchange the controller outputs $\bar{v}_{ij}$ instead of x_i and ξ_i. However, in the case of the present problem, $\bar{v}_{ij}$ ensuring (16) includes ξ_i^* and ξ_j^* in the definitions of $\bar{\xi}_i$ and $\bar{r}_{ij}^{\xi}$ which are not available for agent i. Thus, we instead let agent i send v_{ij} through the scattering transformation to eliminate such unavailable terms.

The scattering transformation is then defined as

$$s_{ij}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-v_{ij} + \eta r_{ij}),\ s_{ij}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(v_{ij}+\eta r_{ij}), \quad (18)$$

$$s_{ji}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(v_{ji} + \eta r_{ji}),\ s_{ji}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-v_{ji}+\eta r_{ji}), \quad (19)$$

where $r_{ij} := [(r_{ij}^x)^T\ (r_{ij}^{\xi})^T]^T$ and $\eta > 0$. Agent i sends $s_{ij}^{\rightarrow}$ to agent j and $s_{ji}^{\leftarrow}$ is the signal which j receives from i, while j sends $s_{ji}^{\rightarrow}$ to i and $s_{ij}^{\leftarrow}$ is the signal which i receives from i. Due to the delays, these signals satisfy

$$s_{ji}^{\leftarrow}(t) = s_{ij}^{\rightarrow}(t - T_{ij}),\ s_{ij}^{\leftarrow}(t) = s_{ji}^{\rightarrow}(t - T_{ji}). \quad (20)$$

Once agent i receives $s_{ij}^{\leftarrow}$, it computes r_{ij} from the second equation of (18) and adds the resulting r_{ij} to the controller (10). Remark however that (10) and (18) suffer from an algebraic loop. In implementation, we thus determine r_{ij} by substituting (10) into the second equation of (18) as $r_{ij} = (\eta I_{2N} + E_{ij})^{-1}(\sqrt{2\eta}s_{ij}^{\leftarrow} + E_{ij}[x_i^T\ \xi_i^T]^T)$.

The system with the scattering transformation (18) and (19) and the delay blocks (20) is known to be passive from $-[v_{ij}^T\ v_{ji}]^T$ to $[r_{ij}^T\ r_{ji}^T]^T$ [3]. The following lemma proves that the system is also passive from the input to output with the biases v_{ij}^*, v_{ji}^*, r_{ij}^* and r_{ji}^*. The subsequent results follow for any signal $s_{ij}^{\rightarrow}, s_{ji}^{\rightarrow}$ over the negative time $t < 0$, but just for simplicity, we suppose that $s_{ij}^{\rightarrow}(t) = s_{ji}^{\rightarrow}(t) = 0\ \forall t < 0$ throughout this paper.

Lemma 5 *The system consisting of the scattering transformation (18) and (19) and the delay blocks (20) is passive from* $-[\bar{v}_{ij}^T\ \bar{v}_{ji}]^T$ *to* $[\bar{r}_{ij}^T\ \bar{r}_{ji}^T]^T$.

Proof: Define

$$\begin{aligned} V_{ij}(t) = \frac{1}{2}\int_0^t &\Big(\|s_{ij}^{\rightarrow}(\tau) + \gamma_{ij}^*\|^2 - \|s_{ji}^{\leftarrow}(\tau) + \gamma_{ij}^*\|^2 \\ &+ \|s_{ji}^{\rightarrow}(\tau)-\delta_{ij}^*\|^2 - \|s_{ij}^{\leftarrow}(\tau)-\delta_{ij}^*\|^2\Big)d\tau \\ &+ \frac{T_{ij}}{2}(\gamma_{ij}^*)^2 + \frac{T_{ji}}{2}(\delta_{ij}^*)^2, \end{aligned} \quad (21)$$

where $\gamma_{ij}^* := \frac{1}{\sqrt{2\eta}}(v_{ij}^* - \eta r_{ij}^*)$, $\delta_{ij}^* := \frac{1}{\sqrt{2\eta}}(v_{ij}^* + \eta r_{ij}^*)$. From (20) and $s_{ij}(t) = 0\ \forall t < 0$, we have

$$\begin{aligned} V_{ij}(t) =& \frac{1}{2}\int_{t-T_{ij}}^t \|s_{ij}^{\rightarrow}(\tau) + \gamma_{ij}^*\|^2 d\tau \\ &+ \frac{1}{2}\int_{t-T_{ji}}^t \|s_{ji}^{\rightarrow}(\tau)-\delta_{ij}^*\|^2 d\tau \geq 0\ \forall t, \end{aligned}$$

Let us get back to the description (21). The time derivative of V_{ij} is then given by

$$\begin{aligned} \dot{V}_{ij}(t) =& \frac{1}{2}\Big(\|s_{ij}^{\rightarrow}(t) + \gamma_{ij}^*\|^2 - \|s_{ji}^{\leftarrow}(t) + \gamma_{ij}^*\|^2 \\ &+ \|s_{ji}^{\rightarrow}(t) - \delta_{ij}^*\|^2 - \|s_{ij}^{\leftarrow}(t) - \delta_{ij}^*\|^2\Big). \end{aligned} \quad (22)$$

Substituting (18), (19), $\gamma_{ij}^* := \frac{1}{\sqrt{2\eta}}(v_{ij}^* - \eta r_{ij}^*)$ and $\delta_{ij}^* := \frac{1}{\sqrt{2\eta}}(v_{ij}^* + \eta r_{ij}^*)$ into (22), we obtain

$$\begin{aligned} \dot{V}_{ij}(t) =& \frac{1}{4\eta}\Big(\ \|-\bar{v}_{ij}(t) + \eta\bar{r}_{ij}(t)\|^2 - \|\bar{v}_{ji}(t) + \eta\bar{r}_{ji}(t)\|^2 \\ &+ \|-\bar{v}_{ji}(t) + \eta\bar{r}_{ji}(t)\|^2 - \|\bar{v}_{ij}(t) + \eta\bar{r}_{ij}(t)\|^2\Big) \\ =& -\bar{v}_{ij}^T(t)\bar{r}_{ij}(t) - \bar{v}_{ji}^T(t)\bar{r}_{ji}(t), \end{aligned} \quad (23)$$

where the first equation holds from (17), $\bar{v}_{ij} = v_{ij} - v_{ij}^*$ and $\bar{r}_{ij} = r_{ij} - r_{ij}^*$ for all i, j. This completes the proof. ■

C. Convergence Analysis

We are now ready to prove the main result of this section.

Theorem 2 *Consider the system (8) for all i, the scattering transformation (18), (19), and the delays (20) for all $j \in \mathcal{N}_i$ and all i. If Assumptions 2 and 3 hold, then x_i asymptotically converges to the optimal solution z^* to (1) for all $i = 1, 2, \ldots, n$.*

Proof: Using V_{ij} in (21), define $V := \sum_{i=1}^n \bar{S}_i + \sum_{(i,j)\in\mathcal{E}} V_{ij}$. Then, combining (16) and (23), we obtain

$$\begin{aligned} \dot{V} =& -\sum_{i=1}^n \sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2 \\ &- \alpha\sum_{i=1}^n (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \leq 0, \end{aligned} \quad (24)$$

which implies $x_i, \xi_i \in \mathcal{L}_\infty\ \forall i$. Using (10), (18), (19), and (20), we can derive

$$r_{ij}(t) = \bar{E}_{ij}^2 r_{ij}(t - T_{ij} - T_{ji}) + \beta_{ij}(t) \quad (25)$$

$$r_{ji}(t) = \bar{E}_{ij}^2 r_{ji}(t - T_{ij} - T_{ji}) + \beta_{ji}(t) \quad (26)$$

by calculation, where $\bar{E}_{ij} := (E_{ij} + \eta I_{2N})^{-1}(E_{ij} - \eta I_{2N})$, and β_{ij} and β_{ji} are linear functions of the states $[x_i^T\ \xi_i]^T$ and $[x_j^T\ \xi_j]^T$ at times t, $t - T_{ij}$, $t - T_{ji}$ and $t - T_{ij} - T_{ji}$. Remark that β_{ij} and β_{ji} are both bounded since $x_i, \xi_i \in \mathcal{L}_\infty\ \forall i$. It is also confirmed by calculation that all the eigenvalues of $\bar{E}_{ij}^2$ lie within the unit circle for any $\eta > 0, a_{ij}$ and b_{ij}. Thus, both of (25) and (26) are stable difference equations with bounded inputs and hence $r_{ij}, r_{ji} \in \mathcal{L}_\infty$. Namely, all of the signals are bounded and the extension of the LaSalle's principle for time delay systems [25] is applicable. In the set of solutions satisfying $\dot{V} \equiv 0$, for every i, $\|\bar{x}_i - \bar{r}_{ij}^x\| = 0\ \forall j \in \mathcal{N}_i$ holds, which means $\dot{\xi} = 0$, and the l-th element of x_i coincides with $z_l^*\ \forall l \in \mathcal{Z}_i$ under Assumption 3, where $z_l^* \in \mathbb{R}$ is the l-th element of z^*. Thus, LaSalle's principle implies that (i) $\lim_{t\to\infty}(x_i - r_{ij}^x) = 0\ \forall j \in \mathcal{N}_i,\ \forall i = 1, 2, \ldots, n$, (ii) the l-th element of x_i converges to z_l^* for all $l \in \mathcal{Z}_i$, and (iii) ξ_i has a limit $\lim_{t\to\infty} \xi_i$.

Consider an $l \in \mathcal{Z}_i$. It is then sufficient to prove that the l-th element of x_j converges to z_l^* for all $j \neq i$. From (10) and (18)–(20), we obtain

$$r_{ij}^x = r_{ji}^x(t - T_{ji}) + (1/\eta)\{-a_{ij}d_{ij} + b_{ij}(r_{ij}^\xi - \xi_i) + b_{ij}(r_{ji}^\xi(t - T_{ji}) - \xi_j(t - T_{ji}))\}, \quad (27)$$

$$r_{ji}^x = r_{ij}^x(t - T_{ij}) + (1/\eta)\{-a_{ij}d_{ji} + b_{ij}(r_{ji}^\xi - \xi_j) + b_{ij}(r_{ij}^\xi(t - T_{ij}) - \xi_i(t - T_{ij}))\}, \quad (28)$$

$$r_{ij}^\xi = r_{ji}^\xi(t - T_{ji}) - (b_{ij}/\eta)d_{ij}, \quad (29)$$

$$r_{ji}^\xi = r_{ij}^\xi(t - T_{ij}) - (b_{ij}/\eta)d_{ji}, \quad (30)$$

where $d_{ij}(t) := (r_{ij}^x - x_i) + (r_{ji}^x(t - T_{ji}) - x_j(t - T_{ji}))$ and $d_{ji}(t) := (r_{ji}^x - x_j) + (r_{ij}^x(t - T_{ij}) - x_i(t - T_{ij}))$. From (i),

$$\lim_{t\to\infty} d_{ij} = 0 \text{ and } \lim_{t\to\infty} d_{ji}(t - T_{ji}) = 0 \quad (31)$$

hold. Since ξ_i converges to a constant from (iii), we also obtain

$$\lim_{t\to\infty}(\xi_i - \xi_i(t - \bar{T}_{ij})) = 0. \quad (32)$$

Now, summing (30) at time $t - T_{ji}$ and (29) and taking its limit, it follows from (31) that

$$\lim_{t\to\infty}(r_{ij}^\xi - r_{ij}^\xi(t - \bar{T}_{ij})) = 0. \quad (33)$$

Subtracting (28) at time $t - T_{ji}$ from (27) and taking its limit, then we have

$$\lim_{t\to\infty}(r_{ij}^x + r_{ij}^x(t - \bar{T}_{ij}) - 2r_{ji}^x(t - T_{ji})) = 0 \quad (34)$$

from (31), (32) and (33). It is confirmed from (i) and (ii) that the l-th element of $\lim_{t\to\infty}(r_{ij}^x + r_{ij}^x(t - \bar{T}_{ij}))$ in (34) is equal to $2z_l^*$, which implies that the l-th element of r_{ji}^x converges to z_l^*. This and (i) also mean that the l-th element of x_j converges to z_l^*. Following the same procedure for a neighbor k of i or j, the l-th element of x_k is proved to converge to z_l^*. Repeating the same process, we can prove that the l-th element of x_j converges to z_l^* for all $j = 1, 2, \dots, n$ because of Assumption 2. The above discussions are applied to every l. ■

IV. Extension to Constrained Distributed Optimization

In this section, we consider the following constrained optimization problem.

$$\min_{z\in\mathbb{R}^N} f(z) \text{ subject to } g_i(z) \le 0 \ \forall i = 1, 2, \dots, n, \quad (35)$$

where f is defined in the same way as (1). Not only f_i but also $g_i : \mathbb{R}^N \to \mathbb{R}^{m_i}$ $(i = 1, 2, \dots, n)$ are assumed to be private information of agent i. In this section, we also assume that the optimal solution exists and the minimal value of f is finite. Denoting the l-th element of g_i by $g_{il}(l = 1, 2, \dots, m_i) : \mathbb{R}^N \to \mathbb{R}$, we make the following assumptions in addition to Assumptions 2 and 3.

Assumption 4 *The functions f_i $(i = 1, 2, \dots, n)$ are twice differentiable. The functions g_{il} $(l = 1, 2, \dots, m_i,\ i = 1, 2, \dots, n)$ are convex and continuously differentiable and their gradients, denoted by $\Gamma_i := \nabla g_i \in \mathbb{R}^{N\times m_i}$ $(i = 1, 2, \dots, n)$, are locally Lipschitz. In addition, there exists z such that $g_i(z) < 0$ $\forall i = 1, 2, \dots, n$.*

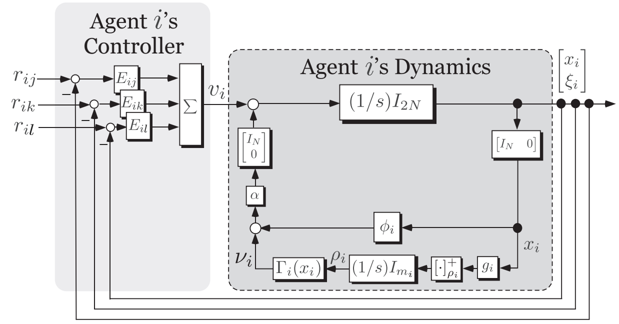

Fig. 3. Block diagram of agent i's dynamics for the constrained optimization in the presence of communication delays. The system enclosed by the dashed line is passive from $\bar{v}_i$ to $[\bar{x}_i^T\ \bar{\xi}_i^T]^T$ (Lemma 8), where $\bar{v}_i$, $\bar{x}_i$ and $\bar{\xi}_i$ are defined in (11) and (12).

Assumption 5 *The delays are homogeneous, namely $T_{ij} = T$ $\forall i, j$ for some T. Define $G' = (\{1, 2, \dots, n\}, \mathcal{E}')$ so that $(j, k) \in \mathcal{E}'$ iff there exists i satisfying $j \in \mathcal{N}_i$ and $k \in \mathcal{N}_i$. The graph G' is then connected.*

Note that, if Assumptions 3 and 4 are satisfied, the optimal solution z^* to (35) is uniquely determined [8]. It is well-known that, under these assumptions, z^* is the optimal solution to (35) if and only if there exist $\lambda_i^* \in \mathbb{R}^{m_i}$ $(i = 1, 2, \dots, n)$ satisfying the KKT condition [8]:

$$\nabla f(z^*) + \sum_{i=1}^n \Gamma_i(z^*)\lambda_i^* = 0, \quad (36)$$

$$\lambda_i^* \ge 0,\ g_i(z^*) \le 0 \ \ \forall i = 1, 2, \dots, n, \quad (37)$$

$$\lambda_{il}^* g_{il}(z^*) = 0 \ \ \forall l = 1, 2, \dots, m_i,\ \forall i = 1, 2, \dots, n, \quad (38)$$

where λ_{il}^* is the l-th element of λ_i^*.

In the sequel, we use the following notation. Given $g \in \mathbb{R}$ and $\rho > 0$, the notation $[g]_\rho^+$ provides

$$[g]_\rho^+ := \begin{cases} 0, & \text{if } \rho = 0 \text{ and } g < 0, \\ g, & \text{otherwise} \end{cases}. \quad (39)$$

If g and ρ are vectors, then $[g]_\rho^+$ is interpreted in the component-wise sense.

Based on the primal-dual gradient algorithm [26] and the distributed optimization algorithm presented in Section II, we present the following algorithm.

$$\dot{x}_i = \sum_{j\in\mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) - \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^\xi - \xi_i) - \alpha(\phi_i(x_i) + \nu_i), \quad (40)$$

$$\dot{\xi}_i = \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^x - x_i), \quad (41)$$

$$\dot{\rho}_i = [g_i(x_i)]_{\rho_i}^+,\ \nu_i = \Gamma_i(x_i)\rho_i \ \ (\rho_i(0) \ge 0), \quad (42)$$

whose block diagram is illustrated in Fig. 3. Remark that (40)–(42) depend only on the private functions f_i and g_i and local variables together with the neighbors' information x_j and ξ_j, $j \in \mathcal{N}_i$. Thus, the algorithm can be locally executed by agent i. A solution to (42) is known to exist and be unique despite discontinuity in the right-hand side of (42) [26]. It is then immediately confirmed that if $\rho_i(0) \geq 0$ then $\rho_i(t) \geq 0$ for all subsequent time t regardless of the trajectory of $g_i(x_i)$.

Let us consider the collection of (40) and (41) for all i. We then have the following lemma.

Lemma 6 *Consider the collection of (40) and (41) for all i whose states are x and ξ. If Assumptions 2, 3 and 4 hold, there exists ξ_i^* $(i = 1, 2, \ldots, n)$ such that the pair of $x^* = (\boldsymbol{1}_n \otimes z^*)$ and $\xi^* = [\xi_1^* \cdots \xi_n^*]^T$ is an equilibrium of the system with the equilibrium inputs $\nu_i \equiv \nu_i^* := \Gamma_i(z^*)\lambda_i^*$, $r_{ij}^x \equiv z^*$ and $r_{ij}^\xi \equiv \xi_j^*$ $\forall j \in \mathcal{N}_i$ for all $i = 1, 2, \ldots, n$.*

This lemma is proved by substituting these equilibrium inputs into (40) and (41) and using (36) and the fact that $L\mathbf{1}_n = 0$ under Assumption 2.

We next consider the open-loop system (42) from x_i to ν_i. We then have the following lemma.

Lemma 7 *Suppose that Assumptions 3 and 4 hold. Then, the system (42) is passive from $\tilde{x}_i$ to $\bar{\nu}_i$ with respect to the storage function $U_i := \frac{1}{2}\|\rho_i - \lambda_i^*\|^2$, where $\bar{\nu}_i := \nu_i - \nu_i^*$ and $\bar{x}_i := x_i - z^*$.*

Proof: Denote l-th element of ρ_i by ρ_{il} $(l = 1, 2, \ldots, m_i)$. Then, the dynamics of ρ_{il} in (42) is given as

$$\dot{\rho}_{il} = [g_{il}(x_i)]^+_{\rho_{il}}, \ \rho_{il}(0) \geq 0, \tag{43}$$

whose right-hand side can be discontinuous at $\rho_{il} = 0$ and $g_{il}(x_i) < 0$. For convenience, the mode satisfying the upper condition in (39) is called mode 1 and the other is mode 2.

We first consider the time when no mode switch occurs. The time derivative of U_i along the system trajectories is then given by $\dot{U}_i = \sum_{l=1}^{m_i}(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}}$. If mode 2 is active, $[g_{il}(x_i)]^+_{\rho_{il}} = g_{il}(x_i)$ and hence $(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}} = (\rho_{il} - \lambda_{il}^*)g_{il}(x_i)$ holds. If mode 1 is active, $\rho_{il} = 0$ and $[g_{il}(x_i)]^+_{\rho_{il}} = 0$ hold, and hence we have

$$\begin{aligned}(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}} = 0 &= \rho_{il} g_{il}(x_i) \\ &= (\rho_{il} - \lambda_{il}^*)g_{il}(x_i) + \lambda_{il}^* g_{il}(x_i).\end{aligned}$$

Since $\lambda_{il}^* \geq 0$ from (37) and $g_{il}(x_i) < 0$ from (39), the term $\lambda_{il}^* g_{il}(x_i)$ is non-positive and hence we obtain

$$(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}} \leq (\rho_{il} - \lambda_{il}^*)g_{il}(x_i). \tag{44}$$

Let us next consider the time when a mode switch happens in (42) for some l. In this case, U_i can be indifferentiable in the standard sense. We thus introduce the upper Dini derivative[2] denoted by D^+U_i. Then, $D^+\|\rho_{il} - \lambda_{il}^*\|^2$ is given by either of $(\rho_{il} - \lambda_{il}^*)0 = 0$ or $(\rho_{il} - \lambda_{il}^*)g_{il}(x_i)$ depending on the sign of $(\rho_{il} - \lambda_{il}^*)$. Thus, following the same procedure as above, we can confirm that the following inequality holds for all time $t \in \mathbb{R}^+$.

$$D^+U_i \leq \sum_{l=1}^{m_i}(\rho_{il} - \lambda_{il}^*)g_{il}(x_i). \tag{45}$$

[2]The upper Dini derivative $D^+U(t)$ of a scalar function U is defined as $D^+U(t) = \limsup_{h\to 0+} \frac{U(t+h)-U(t)}{h}$

This is rewritten as

$$D^+U_i \leq (\rho_i - \lambda_i^*)^T\{g_i(x_i) - g_i(z^*)\} + (\rho_i - \lambda_i^*)^T g_i(z^*) \tag{46}$$

Noticing that $\rho_i \geq 0$ and $g_i(z^*) \leq 0$ from (37), the inequality $\rho_i^T g_i(z^*) \leq 0$ holds. In addition, $(\lambda_i^*)^T g_i(z^*) = 0$ is true from (38). Thus, (46) is further rewritten as

$$\begin{aligned}D^+U_i &\leq (\rho_i - \lambda_i^*)^T\{g_i(x_i) - g_i(z^*)\} \\ &= \sum_{l=1}^{m_i}\left[\rho_{il}\{g_{il}(x_i) - g_{il}(z^*)\} - \lambda_{il}^*\{g_{il}(x_i) - g_{il}(z^*)\}\right].\end{aligned}$$

From convexity of g_{il}, we have $g_{il}(x_i) - g_{il}(z^*) \geq \nabla g_{il}(z^*)(x_i - z^*)$ and $g_{il}(x_i) - g_{il}(z^*) \leq \nabla g_{il}(x_i)(x_i - z^*)$. Using these together with $\rho_i \geq 0$ and $\lambda_i^* \geq 0$, we can prove

$$\begin{aligned}D^+U_i &\leq \left\{\sum_{l=1}^{m_i}(\nabla g_{il}(x_i)\rho_{il} - \nabla g_{il}(z^*)\lambda_{il}^*)\right\}^T (x_i - z^*) \\ &= \{\Gamma_i(x_i)\rho_i - \Gamma_i(z^*)\lambda_i^*\}^T (x_i - z^*) = \bar{\nu}_i^T \bar{x}_i.\end{aligned} \tag{47}$$

Integrating (47) in time completes the proof. ■

Now, we focus on the blocks encircled by the dashed line in Fig. 3 whose system formulation is given by

$$\begin{bmatrix}\dot{x}_i \\ \dot{\xi}_i\end{bmatrix} = v_i - \alpha\begin{bmatrix}\phi_i(x_i) \\ 0\end{bmatrix} - \alpha\begin{bmatrix}\nu_i \\ 0\end{bmatrix} \tag{48}$$

with (42). Then, we have the following lemma.

Lemma 8 *Suppose that Assumptions 2, 3, and 4 hold. Then, the system (48) and (42) is passive from $\bar{v}_i$ to $[\bar{x}_i^T \ \bar{\xi}_i^T]^T$ with respect to the storage function $W_i := \bar{S}_i + \alpha U_i$, where $\bar{S}_i$ and U_i are defined in Lemmas 4 and 7, respectively, and $\bar{x}_i$, $\bar{\xi}_i$ and $\bar{v}_i$ are the same as those in (11) and (12).*

Proof: Following the same procedure as Lemma 4 yields

$$\dot{\bar{S}}_i = \begin{bmatrix}\bar{x}_i \\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) - \alpha\bar{\nu}_i^T\bar{x}_i,$$

Combining this and (47), we have

$$D^+W_i \leq \begin{bmatrix}\bar{x}_i \\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)). \tag{49}$$

Integrating this in time completes the proof. ■

The above lemma means that the agent's passive dynamics in Fig. 2 is just replaced by another passive dynamics with the same input-output pair, and hence the passivity-like property (16) again holds. Let us now take the same inter-agent communication strategy as in Section III-B so that Lemma 5 is ensured. We then obtain the following result.

Theorem 3 *Consider the system (40), (41) and (42) for all i with the scattering transformation (18) and (19) and delays (20) for all $j \in \mathcal{N}_i$ and all i. If Assumptions 3, 4 and 5 hold and b_{ij} is common, i.e., $b_{ij} = b$ $\forall j \in \mathcal{N}_i$, $\forall i$ holds for some b, then x_i asymptotically converges to the optimal solution z^* to (35) for all $i = 1, 2, \ldots, n$.*

Proof: Define the energy function $W := \sum_{i=1}^n W_i + \sum_{(i,j)\in\mathcal{E}} V_{ij}$. Then, combining (23) and (49), we obtain

$$D^+W \le -\sum_{i=1}^n \sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2 - \alpha\sum_{i=1}^n (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \le 0, \quad (50)$$

which implies $x_i, \xi_i \in \mathcal{L}_\infty\ \forall i$, which implies $\dot{\xi}_i \in \mathcal{L}_\infty\ \forall i$. The inclusions $r_{ij}, r_{ji} \in \mathcal{L}_\infty$ are also proved in the same way as (25) and (26), and hence $\dot{x}_i \in \mathcal{L}_\infty\ \forall i$. Solving (25), $r_{ij}(t)$ is given by a convolution sum of the input β_{ij} whose time derivative is bounded. We thus have $\dot{r}_{ij} \in \mathcal{L}_\infty\ \forall i$ for all $j \in \mathcal{N}_i$ and $i = 1, 2, \ldots, n$.

The inequality (50) means that $(x_i - r_{ij}^x) \in \mathcal{L}_2$ for all $j \in \mathcal{N}_i$ and $i = 1, 2, \ldots, n$, and

$$\int_0^\infty (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) < \infty\ \forall i = 1, 2, \ldots, n.$$

We see from the above discussions that the time derivative of $x_i - r_{ij}^x$ is bounded. The derivative of $(x_i - z^*)(\phi_i(x_i) - \phi_i(z^*))$ is given as

$$(\phi_i(x_i) - \phi_i(z^*))^T\dot{x}_i + (x_i - z^*)^T\nabla^2 f_i(x_i)\dot{x}_i,$$

which is also bounded. Thus, invoking Barbalat's lemma [3], we can prove

$$\lim_{t\to\infty}(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) = 0\ \forall i = 1, \ldots, n. \quad (51)$$

$$\lim_{t\to\infty}(x_i - r_{ij}^x) = 0\quad \forall j \in \mathcal{N}_i \text{ and } \forall i = 1, 2, \ldots, n. \quad (52)$$

Under Assumption 5, there exists i such that $|\mathcal{N}_i| \ge 2$. Take two neighbors $j, k \in \mathcal{N}_i$ of such i. Then, (52) implies that

$$\lim_{t\to\infty}(r_{ij}^x - r_{ik}^x) = 0. \quad (53)$$

Following the same procedure as Theorem 2, the equations (33) hold. From (52), we can also prove (31). Subtracting (28) at time $t - T_{ji}$ from (27) with $T_{ij} = T_{ji} = T$ and $b_{ij} = b$ and taking its limit yield

$$\lim_{t\to\infty}\{r_{ij}^x + r_{ij}^x(t - 2T) - 2r_{ji}^x(t - T) + (b/\eta)(\xi_i - \xi_i(t - 2T))\} = 0 \quad (54)$$

from (31) and (33). The same equation holds for k. Subtracting (54) from that for k, and using (53), we have

$$\lim_{t\to\infty}(r_{ji}^x - r_{ki}^x) = 0. \quad (55)$$

From (52), (55) means that $\lim_{t\to\infty}(x_j - x_k) = 0$, which holds for any pair such that $(j, k) \in \mathcal{E}'$. Under Assumption 5, we can conclude that there exists a trajectory $c(\cdot)$ such that $\lim_{t\to\infty}(x_i - c) = 0$ for all i. (51) is then rewritten as

$$\sum_{i=1}^n \lim_{t\to\infty}((c - z^*)^T(\phi_i(c) - \phi_i(z^*)) + \sigma(x_i) - \sigma(c)) = 0,$$

where $\sigma(x_i) := x_i^T\phi_i(x_i) - x_i^T\phi_i(z^*) - (z^*)^T\phi_i(x_i)$. Since σ is continuous and $\lim_{t\to\infty}(x_i - c) = 0$, we have $\lim_{t\to\infty}(\sigma(x_i) - \sigma(c)) = 0$ and hence $\sum_{i=1}^n \lim_{t\to\infty}(c - z^*)^T(\phi_i(c) - \phi_i(z^*)) = 0$. From this, we can also prove $c \to z^*$ in the same way as asymptotic stability in Lyapunov theorem. It is thus concluded that $x_i \to z^*\ \forall i$. ■

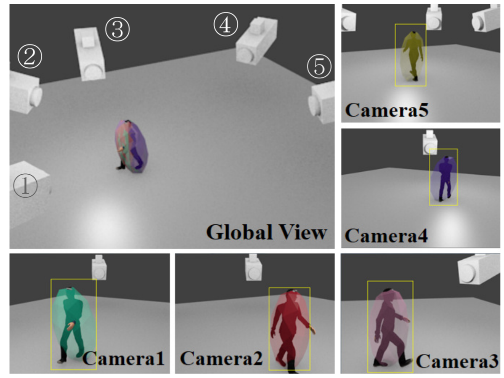

Fig. 4. Intended scenario of 3-D visual human localization. Each camera acquires a rectangle enclosing the human (yellow boxes in the small windows) using a pedestrian detection algorithm. The transparent colored ellipsoids in the small windows are final estimates generated by the algorithm presented in Section IV, and all of them are also shown in the large window.

It is immediate to see from (51) that if $\mathcal{Z}_i = \{1, 2, \ldots, N\}$ for all i, then $x_i \to z^*$ holds in the absence of Assumption 5 and the restriction that b_{ij} are common for all $j \in \mathcal{N}_i$ and i.

V. Application to Human Localization Using Pedestrian Detection Algorithm

We finally apply the proposed algorithms to the visual human localization problem investigated in [24]. Here, multiple networked cameras are assumed to be distributed over the 3-D Euclidean space to monitor a human as shown in Fig. 4. Each camera acquires 2-D rectangles on its own image plane in which the human lives, as shown in the small windows of Fig. 4, by executing a pedestrian detection algorithm e.g. in [27]. Then, if camera i detects the human, then it knows that the human must be inside of a cone $\mathcal{H}_i$ defined by connecting the focal center and the vertices of the rectangle. In this paper, we suppose that all of the five cameras detect the human.

Suppose that the human is modeled as an ellipsoid $\Omega(q, Q) = \{p \in \mathbb{R}^3 |\ (p - q)^T Q^{-2}(p - q) \le 1\}$, the decision variable z consists of the elements of $q \in \mathbb{R}^3$ and $Q \in \mathbb{S}^{3\times 3}$. We then formulate the local cost function

$$f_i(z) = -\log\det(Q) + \min_{\bar{q}\in\mathcal{C}_i}\|\bar{q} - \mathbb{E}_i q\|^2,$$

where $\mathbb{E}_i \in \mathbb{R}^{2\times 3}$ extracts two of three elements of q. In this simulation, we assign 1st and 2nd elements to cameras 1–3, and 1st and 3rd elements to cameras 4 and 5. The set $\mathcal{C}_i$ is the line segment connecting the focal center and the center of the rectangle on the image projected onto the 2-D plane such that the element not extracted by $\mathbb{E}_i$ is 0. The local constraints are given by $\Omega(q, Q) \subseteq \mathcal{H}_i$ and $Q > 0$, which are reduced to the form of $g_i(z) \le 0$. See [24] for more details on the formulation. Note that the problem satisfies Assumptions 3 and 4 in a realistic situation.

We first run the algorithm without the scattering transformation for communication delays $T_{ij} = 0.03$s for all $(i, j) \in \mathcal{E}$, where the time step is set to 4ms. The communication network is set to a ring graph, where a_{ij} and b_{ij} are selected as $a_{ij} = 1$

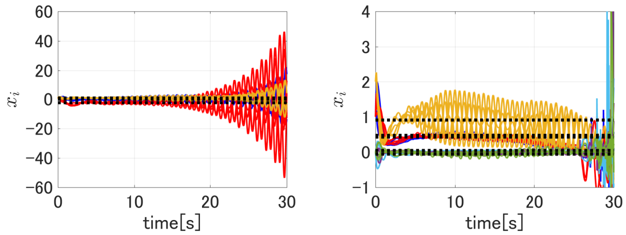

Fig. 5. Trajectories of $x_1, \ldots, x_5$ with communication delays and without the scattering transformation, where $x_i^q \in \mathbb{R}^3$ in the left describe the elements in x_i corresponding to the vector variable q, and $x_i^Q in \mathbb{R}^6$ in the right those corresponding to the matrix variable Q.

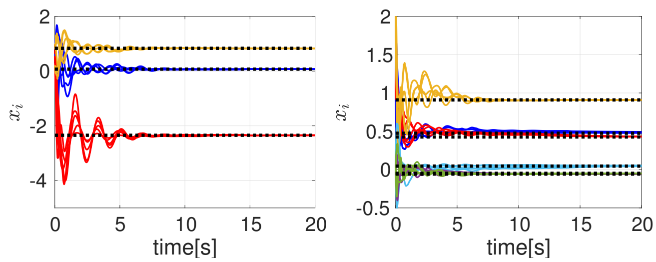

Fig. 6. Trajectories of $x_1, \ldots, x_5$ with communication delays and scattering transformation.

and $b_{ij} = 3$ for all $(i, j) \in \mathcal{E}$. The initial values of the estimates of q are randomly selected within $[0\ 1]$ and those for Q are set to a diagonal matrix with elements 1, 1, and 2 for all i. The initial values of ρ_i are also randomly selected within $[0\ 1]$, and $\xi_i(0) = [1\ 2\ 3\ 1\ 1\ 2\ 0\ 0\ 0]^T$ for all i. The gain α is set to $\alpha = 2$. Then, the trajectories of $x_1, x_2, \ldots, x_5$ are illustrated in Fig. 5, namely they diverge and the simulation stops.

We then implement the algorithm presented in Section IV with $\eta = 1$. The resulting trajectories $x_1, x_2, \ldots, x_5$ are shown in Fig. 6. We see that the system is stabilized by the scattering transformation, and they successfully converge to the optimal solution. The final estimates of the ellipsoid are illustrated in Fig. 4, where we see that every camera successfully computes an ellipsoid tightly enclosing the human.

VI. Conclusion

In this paper, we have addressed a class of distributed optimization problems in the presence of the inter-agent communication delays. To this end, we first have focused on unconstrained distributed optimization problem, and presented a passivity-based perspective for the PI consensus-based distributed optimization algorithm. We then have proved that by suitably interconnecting the PI consensus-based distributed optimization algorithm with the scattering transformation, convergence of the optimization process can be demonstrated despite arbitrary, constant, and unknown inter-agent communication delays. Moreover, we have extended the results to distributed optimization with local inequality constraints. Finally, the present algorithm has been applied to a visual human localization problem.

References

[1] R. Ortega, A. Loria, P.J. Nicklasson and H. Sira-Ramirez, *Passivity-Based Control of Euler- Lagrange Systems: Mechanical, Electrical and Electromechanical Applications*, 2nd ed. Communications and Control Engineering Series. Springer, 2010.

[2] H. Bai, M. Arcak and J.T. Wen, *Cooperative Control Design: A Systematic, Passivity-based Approach*, Communications and Control Engineering Series. Springer, New York, 2011.

[3] T. Hatanaka, N. Chopra, M. Fujita and M.W. Spong, *Passivity-Based Control and Estimation in Networked Robotics*, Communications and Control Engineering Series, Springer-Verlag, 2015.

[4] J.T. Wen and M. Arcak, "A unifying passivity framework for network flow control," IEEE Trans. Autom. Control, vol. 49, no. 2, pp. 162–174, 2004.

[5] X. Fan, T. Alpcan, M. Arcak, T.J. Wen and T. Basar, "A passivity approach to game-theoretic CDMA power control," Automatica, vol. 42, no. 11, pp. 1837–1847, 2006.

[6] T. Stegink, C.D. Persis and A. van der Schaft, "A unifying energy-based approach to stability of power grids with market dynamics," IEEE Trans. Autom. Control, vol. 62, no. 6, pp. 2612–2622, 2017.

[7] T. Hatanaka, X. Zhang, W. Shi, M. Zhu and N. Li, "An Integrated Design of Optimization and Physical Dynamics for Energy Efficient Buildings: A Passivity Approach," Proc. 1st IEEE Conf. Control Technology and Applications, pp. 1050–1057, 2017.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[9] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," IEEE Trans. Autom. Control, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[10] A. Nedić, A. Ozdaglar and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," IEEE Trans. Autom. Control, vol. 55, no. 4, pp. 922–938, 2010.

[11] M. Zhu and S. Martinez, "On distributed convex optimization under inequality and equality constraints," IEEE Trans. Autom. Control, vol. 57, no. 1, pp. 151–164, 2012.

[12] T-H. Chang, A. Nedić and A. Scaglione, "Distributed constrained optimization by consensus-based primal-dual perturbation method," IEEE Trans. Autom. Control, vol. 59, no. 6, pp. 1524–1538, 2014.

[13] J. Wang and N. Elia, "Control approach to distributed optimization," Proc. Forty-Eighth Annual Allerton Conference, pp. 557–561, 2010.

[14] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," Proc. 50th IEEE Conf. Decision and Control, pp. 557–561, 2011.

[15] R.A. Freeman, P. Yang and K.M. Lynch, "Stability and convergence properties of dynamic average consensus estimators," Proc. 45th IEEE Conf. Decision and Control, pp. 398–403, 2006.

[16] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," IEEE Trans. Automatic Control, vol. 31, no. 9, pp. 803–812, 1986.

[17] H. Terelius, U. Topcu and R.M. Murray, "Decentralized multi-agent optimization via dual decomposition," Proc. 18th IFAC World Congress, pp. 11245–11251, 2011.

[18] K.I. Tsianos and M.G. Rabbat, "Distributed dual averaging for convex optimization under communication delays," Proc. 2012 American Control Conference, pp. 1067–1072, 2012.

[19] T. Wu, K. Yuan, Q. Ling, W. Yin, and A.H. Sayed, "Decentralized consensus optimization with asynchrony and delays," arXiv preprint arXiv:1612.00150, 2016.

[20] A. Agarwal and J.C. Duchi "Distributed delayed stochastic optimization," Proc. Advances in Neural Information Processing Systems, pp. 873–881, 2011.

[21] H. Wang, X. Liao, T. Huang, and C. Li, "Cooperative distributed optimization in multiagent networks with delays," IEEE Trans. Systems, Man, and Cybernetics: Systems, vol. 45, no. 2, pp. 363–369, 2015.

[22] N. Chopra and M.W. Spong, "Output synchronization of nonlinear systems with time delay in communication," Proc. 45th IEEE Conf. Decision and Control, pp. 4986–4992, 2006.

[23] R. Luck and A. Ray, "An observer-based compensator for distributed delays," Automatica, vol. 26, no. 5, pp. 903–908, 1990.

[24] T. Hatanaka, R. Funada, G. Gezer and M. Fujita, "Distributed visual 3-D localization of a human using pedestrian detection algorithm: A passivity-based approach," Proc. 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems, pp. 210–215, 2016.

[25] J.K. Hale and S.M.V. Lunel, *Introduction to Functional Differential Equations*, Applied Mathematical Sciences Series, vol. 99, Springer, New York, 1993.

[26] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of constrained primal-dual dynamics," System and Control Letters, vol. 87, pp. 10–15, 2016.

[27] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Conference on Computer Vision and Pattern Recognition, pp. 886–893, 2005.