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. This perspective allows us to handle communication delays while using scattering transformation. Moreover, we extend the results to constrained distributed optimization, where it is shown that the problem is solved by just adding one more feedback loop of a passive system to the solution of the unconstrained ones. We also show that delays can be incorporated in the same way as the unconstrained problems. 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]. In the last decade, passivity has also been actively studied in control of network systems [2], [3], where inherent passivity preservation and energy dissipative property allow one to systematically analyze the network properties and design cooperative controllers. The passivity-based approach is also employed in more advanced research fields including transportation networks [4], building control [5], [6], power grids [7], [8], cyber-security [9], cyber-physical systems [10] and human-swarm interactions [11].

Recently, passivity has been also used in optimization and related application fields [12]–[18]. Wen and Arcak [12] address Internet congestion control, and present a unifying passivity-based design framework for network utility maximization problems. The framework is applied to CDMA power control in [13]. Yamamoto and Tsumura [16] apply a distributed algorithm called Uzawa's algorithm to smart grid control based on passivity. Relations between passivity-based cooperative control and network flow optimization are also discussed in [14]. Bürger and Persis [15] present a passivity-based solution to distributed optimization with dynamic elements. Ishizaki et al. [17] prove equivalence between convex gradients and incrementally passive systems. Bravo [18] presents a fully distributed algorithm, which does not require a subject to gather information from all agents. The modularized property and constructive design procedure of the passivity-based design allows one to improve the performance and robustness [12], [13], or integrate another passive components like physical dynamics while ensuring stability and optimality for the total interconnected systems [6], [8].

All of the above papers address so-called resource allocation problems [19] or its variations. In this paper, we deal with another type of distributed optimization studied in [20]–[29]. Nedić and Ozdaglar [20] present a distributed algorithm which combines consensus algorithms and subgradient methods. The results are extended to constrained problems in [21], where a variation of the algorithm in [20] is shown to ensure exact convergence to the optimal solution using a diminishing step size. A solution to the problem with globally defined inequality and equality constraints is presented by Zhu and Martinez [22]. The results of [22] are further extended to problems with partial knowledge on the global constraints in [23]. The proposed solution, termed primal-dual perturbed subgradient method, is shown to outperform [22] in terms of the convergence speed. Acceleration of the convergence speed is also addressed in [24]. Extensions to random networks and a stochastic subgradient method are found in [25], [26].

While the above works present discrete-time recursive processes to compute the solution, Wang and Elia [27], [28] take a continuous-time algorithm and provide a control theoretic perspective for the distributed optimization algorithms, which are the most closely related to this paper. The solution in [27], [28] is further extended to dynamic topologies by Droge and Egerstedt [29].

All of the solutions in [20]–[29] rely on the inter-agent information exchanges, which is typically implemented using communication technology. However, the problems caused by the communication have not been fully explored in the literature. Among many of such problems, this paper treats communication delays inherent in communication, which may slow down the convergence to wait for the arrivals of messages or even destabilize the solution processes in the worst case.

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@mei.titech.ac.jp (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 [39], but communication delays are not addressed there.

To address the issue, we employ the notion of passivity.

We start with presenting a passivity-based perspective for the algorithms based on the consensus algorithm and so-called PI (Proportional-Integral) consensus algorithm [30], [31]. In particular, it is revealed that the PI consensus-based solution is regarded as a feedback connection of passive systems. Passivity-based formalism presented above allows one to utilize rich knowledge established in the history of passivity-based control.

We next treat communication delays using our passivity-based perspective of the distributed algorithm. Specifically, we show that the delays are successfully integrated with the above solution by using the techniques in [32] 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. In this part, we start with a delay free case, and show that the problem is solvable by just adding one more feedback loop of a passive system originating from the gradient-based update of the Lagrange multiplier. Note that the resulting architecture is similar to the one presented in [28]. However, [28] achieves only convergence to an approximate solution in the presence of the nonlinear inequality constraints, due to the use of barrier functions. In contrast, our solution ensures exact convergence to the optimal solution while avoiding the gain tuning of the barrier functions. The results are also extended to the case with communication delays by following the same procedure as the unconstrained problem.

Finally, the present algorithm is applied to a visual human localization problem using a pedestrian detection algorithm.

Distributed optimization with delays is investigated in a relatively few papers [34]–[37]. Terelius et al. [33] addresses a problem compatible with ours, and Agarwal and Duchi [34] a stochastic version without constraints, where both papers assume constant delays. The solutions in both papers presume information broadcasting through multi-hop communication, whereas the present algorithm is purely distributed. Tsianos and Rabbat [35] presents a distributed solution to the problem. However, this approach requires prior knowledge on the entire network structure together with delays at all links to correct cost functions, while our solution does not need such information. Hale et al. [36] develops a distributed solution to a resource allocation-like problem with possibly time-varying delays. The architecture is however specialized to the resource allocation and it is essentially not applicable to the problem of [20]–[29]. In addition, all of the above works need to take a diminishing step size, which is known to result in slow convergence [37]. We et al. [37] presents a distributed algorithm with a constant step size for time-varying delays, where an upper bound of the delays is assumed to be available. However, the model in [37] is not always more general than ours since just applying the buffering technique in [38] using the upper bound would reduce the delays to the constant delay model. In addition, they consider only unconstrained problems, while constrained problems are solved in this paper.

The contribution of this paper is summarized as below: The primary contribution is to present a novel purely distributed algorithm with robustness against communication delays. The secondary contribution is to reveal that the problem in [20]–[29] can be treated within the passivity paradigm. Thirdly, we handle general convex inequality constraints in this paper, while the other passivity-based approaches [12]–[18] take only linear and/or scalar constraints.

II. Preliminary

In this section, we introduce some terminologies used in this paper.

We first introduce passivity. Consider a system with a state-space representation

$$\dot{x} = \phi(x, u), \quad y = \varphi(x, u), \tag{1}$$

where $x(t) \in \mathbb{R}^N$ is the state, $u(t) \in \mathbb{R}^p$ is the input and $y(t) \in \mathbb{R}^p$ is the output. Then, passivity is defined as below.

Definition 1 *The system (1) is said to be passive if there exists a positive semi-definite function $S : \mathbb{R}^N \to \mathbb{R}_+ := [0, \infty)$, called storage function, such that*

$$S(x(t)) - S(x(0)) \leq \int_0^t y^T(\tau) u(\tau) d\tau \tag{2}$$

holds for all inputs $u : [0, t] \to \mathbb{R}^p$, all initial states $x(0) \in \mathbb{R}^N$ and all $t \in \mathbb{R}^+$.

In the case of a static system $y = \varphi(u)$, it is passive if $y^T u = \varphi^T(u) u \geq 0$ for all $u \in \mathbb{R}^p$, which can be shown by taking $S \equiv 0$ as the storage function. As widely known, if S is differentiable, (2) can be replaced by

$$\dot{S}(x(t)) \leq y^T(t) u(t). \tag{3}$$

Passivity is known to be preserved for feedback interconnections of passive systems as follows. Consider two passive systems from u_1 to y_1 and from u_2 to y_2 with storage functions S_1 and S_2 respectively. Then, the feedback interconnections $u_1 = r - y_1$ and $u_2 = y_1 + d$ of these systems with exogenous inputs r and d provides

$$\dot{S}_1 + \dot{S}_2 \leq y_1^T u_1 + y_2^T u_2 = y_1^T r + y_2^T d, \tag{4}$$

which means passivity from $[r^T \ d^T]^T$ to $[y_1^T \ y_2^T]^T$. In the absence of r and d ($r = d = 0$), the energy dissipation $\dot{S}_1 + \dot{S}_2 \leq 0$ is moreover proved, and accordingly closed-loop stability is ensured under additional assumptions on strict energy dissipation and observability. Please refer to [1]–[3] or other seminal books cited therein for more details on passivity.

We next introduce another notion closely related to passivity, namely incremental passivity. In the context of this paper, it is sufficient to define incremental passivity for a static system.

Definition 2 *A static system $y = \varphi(u)$ is said to be incrementally passive if the function φ satisfies*

$$(\varphi(u_1) - \varphi(u_2))^T (u_1 - u_2) \geq 0$$

for all $u_1 \in \mathbb{R}^p$ and $u_2 \in \mathbb{R}^p$.

We also use the following fundamental tool in convex optimization.

Definition 3 *A function $f : \mathbb{R}^N \to \mathbb{R}$ is said to be convex if the following inequality holds for any $x, y \in \mathbb{R}^N$.*

$$(\nabla f(x))^T (y - x) \leq f(y) - f(x) \tag{5}$$

The function is said to be strictly convex if the inequality (5) strictly holds whenever $x \neq y$.

The following well known result links the convex functions and passivity theory.

Lemma 1 *[19] 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 $x, y \in \mathbb{R}^N$.*

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0 \tag{6}$$

If f is strictly convex, the inequality strictly holds as long as $x \neq y$.

III. Passivity-Based Perspective for Unconstrained Distributed Optimization

In this section, we consider the following distributed optimization problem investigated in [20].

$$\min_{z \in \mathbb{R}^N} f(z) := \sum_{i=1}^{n} f_i(z) \tag{7}$$

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 unaccessible 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, i.e., . for every point $x_0 \in \mathbb{R}^N$, there exists its neighborhood $\mathcal{X}_0$ such that $\|\phi_i(x) - \phi_i(y)\| \leq L_0(x_0)\|x - y\|$ holds for all $x, y \in \mathcal{X}_0$ with some constant $L_0(x_0)$.*

Throughout this paper, we assume that the set of the optimal solutions is not empty, and the corresponding minimal value of f is finite. An optimal solution to (7) is denoted by $z^* \in \mathbb{R}^N$. Since f is also continuously differentiable and convex under Assumption 1, a vector $z^* \in \mathbb{R}^N$ is an optimal solution to the problem (7) if and only if the following equation holds [19].

$$\nabla f(z^*) = \sum_{i=1}^{n} \phi_i(z^*) = 0 \tag{8}$$

A. Consensus-Based Distributed Optimization

Suppose that each agent i has an estimate of the optimal solution z^*, denoted by $x_i \in \mathbb{R}^N$, and that x_i is updated so that it converges to the set of optimal solutions. The agents are assumed to be able 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.*

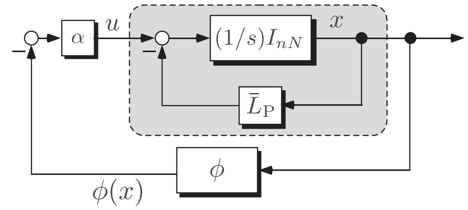


Fig. 1. Block diagram of the consensus-based distributed optimization algorithm for unconstrained optimization. The system enclosed by the dashed line is passive from u to x, and the bottom block ϕ is incrementally passive (Lemma 1).

The set of all neighbors of agent i is denoted by $\mathcal{N}_i$. This assumption is weaker than [33] and [34] if the multi-hop communication is identified with all-to-all communication, and is compatible with [35] and [37]. Assuming a fixed graph may be problematic for networks with mobility, but the extension of the subsequent results to changing topologies is left as a future work.

Nedić and Ozdaglar [20] present an update rule of x_i combining consensus algorithms and gradient descent algorithms whose continuous-time representation is given as

$$\dot{x}_i = -\alpha \phi_i(x_i) + \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i),$$

where α is a positive scalar and a_{ij} is the (i, j)-element of an adjacency matrix for G, where $a_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise, and $a_{ij} = a_{ji}$ $\forall i, j$. The graph Laplacian associated with the adjacency matrix with elements a_{ij} is denoted by $L_{\rm P}$. The matrix $L_{\rm P}$ is symmetric and positive semidefinite with a simple zero eigenvalue corresponding to the eigenvector $\mathbf{1}_N$, where $\mathbf{1}_N$ is the N dimensional real vector whose elements all equal to 1.

Collecting estimates x_i as $x := [x_1^T \ x_2^T \ \cdots \ x_n^T]^T$, the evolution of x is described by

$$\dot{x} = -\alpha \phi(x) - \bar{L}_{\rm P} x, \tag{9}$$

where $\phi(x) := [\phi_1^T(x_1) \ \cdots \ \phi_n^T(x_n)]^T$, $\bar{L}_{\rm P} := L_{\rm P} \otimes I_N$, the symbol $\otimes$ describes the Kronecker product, and I_N is the N-by-N identity matrix. The block diagram of the system is illustrated in Fig. 1.

Let us now consider the system enclosed by the dashed line in Fig. 1 whose input is denoted by u. The dynamics of the system is described by

$$\dot{x} = -\bar{L}_{\rm P} x + u. \tag{10}$$

Take a storage function $S_{\rm P} := \frac{1}{2}\|x\|^2$. The time derivative of S along the trajectories of (10) is then given as

$$\dot{S}_{\rm P} = -x^T \bar{L}_{\rm P} x + x^T u \leq x^T u. \tag{11}$$

We see from (3) that the system is passive. Lemma 1 ensures that the bottom block ϕ in Fig. 1 is incrementally passive, but this does not mean passivity of ϕ from x to $\phi(x)$. Accordingly, the energy dissipation as in (4) with $e = d = 0$ is not trivially proved due to the increments in the input and output in (6). Indeed, the trajectories of x_i do not always converge to the solution as confirmed below. The goal state is now formulated

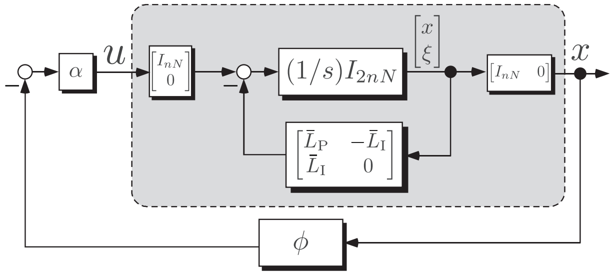


Fig. 2. Block diagram of the PI consensus-based distributed optimization algorithm for unconstrained optimization. The system enclosed by the dashed line is passive from $\tilde{u} = u - u^*$ to $\tilde{x} = x - x^*$ with $u^* := -\alpha\phi(x^*)$ as confirmed in (18). The bottom block ϕ is incrementally passive (Lemma 1), and hence passive from $x - x^*$ to $\phi(x) - \phi(x^*)$.

as $x = x^*$ for $x^* := \mathbf{1}_n \otimes z^*$. If $x = x^*$, the right-hand side of (9) is equal to

$$-\alpha\phi(x^*) - (L_{\rm P}\mathbf{1}_n) \otimes z^* = -\alpha\phi(x^*).$$

Although the sum of the elements $\phi_1(z^*), \phi_2(z^*), \ldots, \phi_n(z^*)$, of $\phi(x^*)$ is zero from (8), each element is not always zero, which implies that $x = x^*$ is not generally an equilibrium of (9). Thus, the state trajectories do not converge to the goal state x^*.

B. PI Consensus-Based Distributed Optimization

In this subsection, we focus on a distributed algorithm based on a PI consensus algorithm [30], formulated as

$$\dot{x}_i = \sum_{j\in\mathcal{N}_i} a_{ij}(x_j - x_i) - \sum_{j\in\mathcal{N}_i} b_{ij}(\xi_j - \xi_i) + u_i, \quad (12a)$$

$$\dot{\xi}_i = \sum_{j\in\mathcal{N}_i} b_{ij}(x_j - x_i), \quad (12b)$$

where $u_i \in \mathbb{R}^N$ is an external input, $\xi_i \in \mathbb{R}^N$ is an additional variable which generates the integral of the consensus input $\sum_{j\in\mathcal{N}_i} b_{ij}(x_j - x_i)$, and b_{ij} is the (i,j)-element of an adjacency matrix for G. The graph Laplacian associated with the adjacency matrix with elements b_{ij} is denoted by $L_{\rm I}$.

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

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \end{bmatrix} = -\begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix}\begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} u, \quad (13)$$

where $\bar{L}_{\rm I} := L_{\rm I} \otimes I_N$. We can prove the following lemma by making use of the skew symmetry of the non-diagonal blocks of $\begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix}$.

Lemma 2 *The system (13) is passive from u to x with respect to the storage function $S = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|\xi\|^2$.*

Proof: The time derivative of S along the system trajectories is given by

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

This completes the proof. ■

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

$$\begin{bmatrix} \dot{x} \\ \dot{\xi} \end{bmatrix} = -\begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix}\begin{bmatrix} x \\ \xi \end{bmatrix} - \alpha\begin{bmatrix} I_{nN} \\ 0 \end{bmatrix}\phi(x), \quad (14)$$

whose block diagram is illustrated in Fig. 2.

C. Equilibrium/Convergence Analysis: Unconstrained Case

Regarding the equilibrium of (14), we have the following lemma.

Lemma 3 *Under Assumptions 1 and 2, there exists ξ^* such that $[(x^*)^T \; (\xi^*)^T]^T$ is an equilibrium of (14).*

Proof: See [27]. ■

Lemma 3 means that, if we define $u^* := -\alpha\phi(x^*)$, the following equation holds.

$$(0 =)\begin{bmatrix} \dot{x}^* \\ \dot{\xi}^* \end{bmatrix} = -\begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix}\begin{bmatrix} x^* \\ \xi^* \end{bmatrix} + \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix} u^* \quad (15)$$

Define $\tilde{x} := x - x^*$, $\tilde{\xi} := \xi - \xi^*$ and $\tilde{u} := u - u^*$. Then, subtracting (15) from (13) yields

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{\xi}} \end{bmatrix} = -\begin{bmatrix} \bar{L}_{\rm P} & -\bar{L}_{\rm I} \\ \bar{L}_{\rm I} & 0 \end{bmatrix}\begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix} + \begin{bmatrix} I_{nN} \\ 0 \end{bmatrix}\tilde{u}. \quad (16)$$

Since the system matrices of (16) are the same as those of (13), passivity of the system (16) from $\tilde{u}$ to $\tilde{x}$ can be immediately proved by modifying the storage function as

$$\tilde{S} := \frac{1}{2}\|\tilde{x}\|^2 + \frac{1}{2}\|\tilde{\xi}\|^2. \quad (17)$$

More precisely, the following inequality holds.

$$\dot{\tilde{S}} = -\tilde{x}^T\bar{L}_{\rm P}\tilde{x} + \tilde{x}^T\tilde{u} \le \tilde{x}^T\tilde{u}. \quad (18)$$

Remark that, differently from (11), the input and output are defined by the increments $\tilde{u} = u - u^*$ and $\tilde{x} := x - x^*$.

We are now ready to use the passivity interpretation of the system (13) and gradient of convex functions to prove the following convergence result.

Theorem 1 *Consider the system (14). If Assumptions 1 and 2 hold, then x_i asymptotically converges to the set of optimal solutions to (7) for all $i = 1, 2, \ldots, n$.*

Proof: From $u = -\alpha\phi(x)$ and $u^* = -\alpha\phi(x^*)$, (18) is rewritten as

$$\dot{\tilde{S}} = -(x - x^*)^T\bar{L}_{\rm P}(x - x^*) - \alpha(x - x^*)^T(\phi(x) - \phi(x^*)).$$

Since $\bar{L}_{\rm P}x^* = 0$, this is further rewritten as

$$\dot{\tilde{S}} = -x^T\bar{L}_{\rm P}x - \alpha\sum_{i=1}^{n}(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*))$$

Using Lemma 1 and $\bar{L}_{\rm P} \ge 0$, we can prove $\dot{\tilde{S}} \le 0$.

Since the function $\tilde{S}$ is radially unbounded and positive definite, any sublevel set of the function is positively invariant.

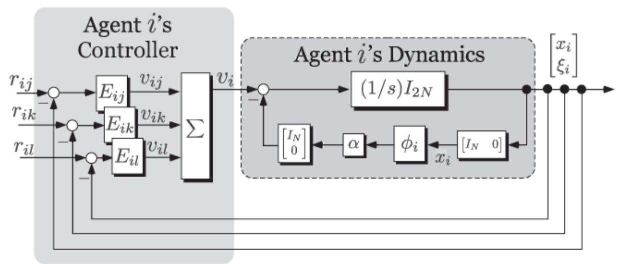


Fig. 3. Block diagram of agent i's dynamics for unconstrained optimization. The system enclosed by the dashed line is passive from $\bar{v}_i$ to $[\bar{x}_i^T \ \bar{\xi}_i^T]^T$ (Lemma 4), where $\bar{v}_i$, $\bar{x}_i$ and $\bar{\xi}_i$ are defined in (25) and (26).

Hence LaSalle's principle is applicable. Consider the state trajectories such that $\dot{\bar{S}} \equiv 0$, i.e., both of $x^T \bar{L}_{\rm P} x = 0$ and

$$\sum_{i=1}^{n} (x_i - z^*)^T (\phi_i(x_i) - \phi_i(z^*)) = 0 \tag{19}$$

identically hold. The former equation means consensus of the variable x_i, namely there exists $c(\cdot)$ such that $x_i(t) = c(t)\ \forall i, t$. In this case, (19) is rewritten as

$$\begin{aligned} 0 &= (c(t) - z^*)^T \sum_{i=1}^{n} (\phi_i(c(t)) - \phi_i(z^*)) \\ &= (c(t) - z^*)^T \sum_{i=1}^{n} \phi_i(c(t)) = (c(t) - z^*)^T \nabla f(c(t)), \end{aligned} \tag{20}$$

where the third equation holds from the optimality condition (8). Using (5), (20) is further rewritten as

$$0 = (\nabla f(c(t)))^T (c(t) - z^*) \geq f(c(t)) - f(z^*) \geq 0,$$

which implies that $f(c(t)) = f(z^*)$ holds for all t, namely the trajectories of c must be contained in the set of optimal solutions to (7). Thus, LaSalle's invariance principle proves the theorem. ■

Remark 1 *The above algorithm together with the convergence result compatible with ours was already presented in [27]. The contribution of this section is not to prove convergence itself but to provide a passivity-based perspective that (14) is regarded as feedback connection of two passive systems with incremental inputs and outputs. It will be shown in the subsequent sections that this perspective provides fruitful design concepts.*

IV. Unconstrained Distributed Optimization with Communication Delay

In this section, we suppose that the inter-agent communication suffers from delays which are assumed to be constant but heterogeneous. The delay from agent i to j is denoted by T_{ij} for any pair $(i,j) \in \mathcal{E}$. 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 [38] as long as an upper bound of the delays are available.

Let us focus on the individual agent's dynamics (12) 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

$$\dot{x}_i = \sum_{j\in\mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) - \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^\xi - \xi_i) - \alpha\phi_i(x_i), \tag{21a}$$

$$\dot{\xi}_i = \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^x - x_i) \tag{21b}$$

whose block diagram is illustrated in Fig. 3, where

$$E_{ij} := \begin{bmatrix} a_{ij}I_N & -b_{ij}I_N \\ b_{ij}I_N & 0 \end{bmatrix}, \ j \in \mathcal{N}_i. \tag{22}$$

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 VI. 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.

We start with the following restrictive assumption in order to enhance readability, and then it will be relaxed.

Assumption 3 *The functions $f_1, f_2, \ldots, f_n$ are strictly convex, continuously differentiable, and their gradients are locally Lipschitz.*

Under the assumption together with the existence of the optimal solution, the solution z^* is uniquely determined [19].

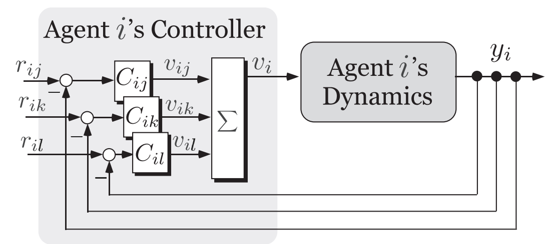


Fig. 4. Block diagram of agent i's dynamics for output synchronization with delays presented in [32].

A. Passivity-Based Output Synchronization with Delays

Let us first emphasize that our problem is regarded as a synchronization problem of the variable x_i $(i = 1, 2, \ldots, n)$ to the optimal solution z^*. Inspired by the fact, we first introduce a passivity-based architecture for output synchronization with communication delays which provide a foundation for the control architecture presented in this section. In particular, we focus on one of the most typical architectures originally presented in [32].

The authors of [32] consider a network of passive systems with a positive definite storage function where the input and output of agent i are now denoted by v_i and y_i, respectively. They first form a local feedback system for each agent i using passive controllers C_{ij} $(j \in \mathcal{N}_i)$ with a positive definite storage function which generates signal v_{ij} from $r_{ij} - y_i$ with an exogenous input r_{ij}. The input v_i is then determined by the sum of v_{ij}, i.e., $v_i = \sum_{j\in\mathcal{N}_i} v_{ij}$ as illustrated in Fig. 4.

Let the summation of the storage functions of agent i's dynamics and controllers be denoted by S_i. Using a variation of (4), they prove the passivity-like property

$$\dot{S}_i \leq \sum_{j\in\mathcal{N}_i} v_{ij}^T r_{ij}. \tag{23}$$

Accordingly, we have $\sum_{i=1}^n \dot{S}_i \leq \sum_{(i,j)\in\mathcal{E}} v_{ij}^T r_{ij}$, which means that if the communication block for any pair $(i,j)\in\mathcal{E}$ were designed so as to ensure passivity from r_{ij} to $-v_{ij}$, Lyapunov stability of the origin would be achieved. To this end, the authors of [32] let every agent i exchange the controller output v_{ij}, instead of the output y_i, with neighbors $j \in \mathcal{N}_i$ through scattering transformation which is well-known to passify the communication block including the delays [3]. Moreover, when all of the local controllers C_{ij} are memoryless, they prove output synchronization in the sense of $y_i(t) - y_j(t - T_{ji}) \to 0\ \forall j \in \mathcal{N}_i$ for all i.

We present a distributed optimization algorithm based on the above architecture. For this purpose, we first point out analogy between Fig. 3 and Fig. 4.

B. Individual Dynamics as A Feedback System with Passive Dynamics and A Collection of Passive Controllers

Consider the system in Fig. 3. The system is then regarded as a feedback system with the agent dynamics

$$\begin{bmatrix}\dot{x}_i\\ \dot{\xi}_i\end{bmatrix} = v_i - \alpha\begin{bmatrix}\phi_i(x_i)\\ 0\end{bmatrix} \tag{24}$$

and the controller

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

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

Let us cut the loop of Fig. 3 and focus on the open-loop system (24) with input v_i enclosed by the dashed line in Fig. 3. Then, we have the following lemma.

Lemma 4 *Suppose that Assumption 3 holds. Then, the system (24) 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*

$$\bar{x}_i := x_i - z^*,\ \bar{\xi}_i := \xi_i - 2\xi_i^*, \tag{25}$$

$$\bar{v}_i := v_i - \sum_{j\in\mathcal{N}_i} v_{ij}^*,\ v_{ij}^* := b_{ij}\begin{bmatrix}\xi_i^* - \xi_j^*\\ 0\end{bmatrix}. \tag{26}$$

The notation $\xi_i^* \in \mathbb{R}^N$ is a vector such that the stack vector $[(\xi_1^*)^T\ \cdots\ (\xi_n^*)^T]^T$ is equal to ξ^* in Lemma 3.

Proof: From (15), it follows that

$$\begin{aligned}\dot{x}_i^* &= \sum_{j\in\mathcal{N}_i} a_{ij}(z^* - z^*) - \sum_{j\in\mathcal{N}_i} b_{ij}(\xi_j^* - \xi_i^*) - \alpha\phi_i(z^*)\\ &= \sum_{j\in\mathcal{N}_i} b_{ij}(\xi_i^* - \xi_j^*) - \alpha\phi_i(z^*),\end{aligned} \tag{27}$$

$$\dot{\xi}_i^* = \sum_{j\in\mathcal{N}_i} b_{ij}(z^* - z^*) = 0. \tag{28}$$

Subtracting (27) and (28) from (24) yields

$$\begin{bmatrix}\dot{\bar{x}}_i\\ \dot{\bar{\xi}}_i\end{bmatrix} = \bar{v}_i - \alpha\begin{bmatrix}\phi_i(x_i) - \phi_i(z^*)\\ 0\end{bmatrix}. \tag{29}$$

because of the definition of $\bar{v}_i$ in (26). The time derivative of $\bar{S}_i$ along the trajectories of (29) is then given by

$$\dot{\bar{S}}_i = \begin{bmatrix}\bar{x}_i\\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \leq \begin{bmatrix}\bar{x}_i\\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i, \tag{30}$$

where we use Lemma 1 and Assumption 3. ■

Lemma 4 and passivity of E_{ij} mean that Fig. 3 takes the same form as Fig. 4 except that the input and output contain the biases v_i^*, z^* and $2\xi_i^*$.

We next derive an inequality corresponding to (23) in the present case. To this end, we close the loop between (24) and the controller (22). Define

$$\bar{r}_{ij} := \begin{bmatrix}\bar{r}_{ij}^x\\ \bar{r}_{ij}^\xi\end{bmatrix} = \begin{bmatrix}r_{ij}^x\\ r_{ij}^\xi\end{bmatrix} - r_{ij}^*,\ r_{ij}^* := \begin{bmatrix}z^*\\ \xi_i^* + \xi_j^*\end{bmatrix}. \tag{31}$$

and $\bar{v}_{ij} := v_{ij} - v_{ij}^*$. Then, from (22), (25) and (26), we have

$$\begin{aligned}\bar{v}_{ij} &= E_{ij}\begin{bmatrix}r_{ij}^x - x_i\\ r_{ij}^\xi - \xi_i\end{bmatrix} - b_{ij}\begin{bmatrix}\xi_i^* - \xi_j^*\\ 0\end{bmatrix}\\ &= E_{ij}\begin{bmatrix}r_{ij}^x - x_i\\ r_{ij}^\xi - \xi_i + (\xi_i^* - \xi_j^*)\end{bmatrix}\\ &\qquad \left(\because -b_{ij}\begin{bmatrix}\xi_i^* - \xi_j^*\\ 0\end{bmatrix} = E_{ij}\begin{bmatrix}0\\ \xi_i^* - \xi_j^*\end{bmatrix}\right)\\ &= E_{ij}\begin{bmatrix}(r_{ij}^x - z^*) - (x_i - z^*)\\ (r_{ij}^\xi - (\xi_i^* + \xi_j^*)) - (\xi_i - 2\xi_i^*)\end{bmatrix} = E_{ij}\begin{bmatrix}\bar{r}_{ij}^x - \bar{x}_i\\ \bar{r}_{ij}^\xi - \bar{\xi}_i\end{bmatrix}.\end{aligned}$$

Substituting this together with $\bar{v}_i = \sum_{j\in\mathcal{N}_i} \bar{v}_{ij}$ into (30) proves the following passivity-like property.

$$\begin{aligned}\dot{\bar{S}}_i &= \sum_{j\in\mathcal{N}_i}\begin{bmatrix}\bar{x}_i\\ \bar{\xi}_i\end{bmatrix}^T\begin{bmatrix}a_{ij}(\bar{r}_{ij}^x - \bar{x}_i) + b_{ij}(\bar{\xi}_i - \bar{r}_{ij}^\xi)\\ b_{ij}(\bar{r}_{ij}^x - \bar{x}_i)\end{bmatrix}\\ &\qquad - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*))\\ &= \sum_{j\in\mathcal{N}_i}\{a_{ij}(\bar{x}_i^T\bar{r}_{ij}^x - \|\bar{x}_i\|^2) + b_{ij}(-\bar{x}_i^T\bar{r}_{ij}^\xi + \bar{\xi}_i^T\bar{r}_{ij}^x)\}\\ &\qquad - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*))\\ &= \sum_{j\in\mathcal{N}_i}\bar{r}_{ij}^T\bar{v}_{ij} - \sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2\\ &\qquad - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \leq \sum_{j\in\mathcal{N}_i}\bar{r}_{ij}^T\bar{v}_{ij}.\end{aligned} \tag{32}$$

Remark that, from (26) and (31), the following equations hold, which plays an important role in deriving the subsequent theoretical results in this section.

$$v_{ji}^* = -v_{ij}^*,\ r_{ji}^* = r_{ij}^*. \tag{33}$$

C. Scattering Transformation

In view of the architecture of [32], the passivity-like property (32) 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 (32) 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

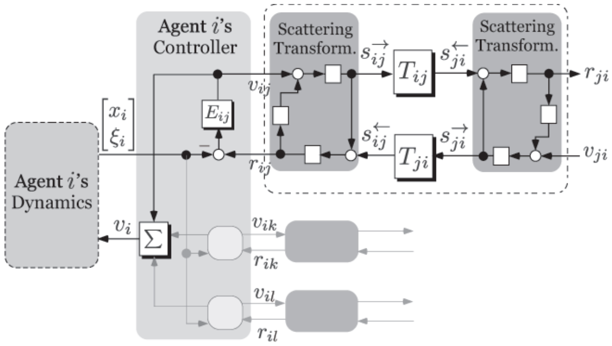


Fig. 5. Block diagram of agent i's dynamics including the scattering transformation and communication delays. The system enclosed by the dash-dotted line is passive from $-[\bar{v}_{ij}^T\ \bar{v}_{ji}]^T$ to $[\bar{r}_{ij}^T\ \bar{r}_{ji}^T]^T$. (Lemma 5), where $\bar{v}_{ij}$ and $\bar{r}_{ij}$ are defined in (26) and (31).

Algorithm 1 Algorithmic Description of in Agent i's Operation.

1: **for** $j \in \mathcal{N}_i$ **do**
2: Set the most recent message from $j \in \mathcal{N}_i$ to $s_{ij}^{\leftarrow}$.
3: $r_{ij} \leftarrow (\eta I_{2N} + E_{ij})^{-1}(\sqrt{2\eta}s_{ij}^{\leftarrow} + E_{ij}[x_i^T\ \xi_i^T]^T)$.
4: $v_{ij} \leftarrow E_{ij}(r_{ij} - [x_i^T\ \xi_i^T]^T)$ and input v_i to the dynamics (24).
5: $s_{ij}^{\rightarrow} \leftarrow \frac{1}{\sqrt{2\eta}}(-v_{ij} + \eta r_{ij})$ and send it to agent j.

let agent i send v_{ij} through the scattering transformation to eliminate such unavailable terms.

The scattering transformation in the present case is then defined as

$$s_{ij}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-v_{ij} + \eta r_{ij}),\ s_{ij}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(v_{ij}+\eta r_{ij}), \quad (34)$$

$$s_{ji}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(v_{ji} + \eta r_{ji}),\ s_{ji}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-v_{ji}+\eta r_{ji}), \quad (35)$$

where $r_{ij} := [(r_{ij}^x)^T\ (r_{ij}^\xi)^T]^T$ and $\eta > 0$. Note that the subscript ij of the notation s means that it is a signal (on neighbor j) in the agent i's side, and the superscripts $\rightarrow$ and $\leftarrow$ mean that it is a signal sent to j and a signal received from j, respectively. Then, as illustrated in Fig. 5, agent i sends $s_{ij}^{\rightarrow}$ instead of v_{ij} and it is received by agent j as $s_{ji}^{\leftarrow}$ after the delay T_{ij}. On the other hand, agent j sends $s_{ji}^{\rightarrow}$ and it is received by agent i as $s_{ij}^{\leftarrow}$ after the delay T_{ji}. Namely, these signals satisfy the following relations.

$$s_{ji}^{\leftarrow}(t) = s_{ij}^{\rightarrow}(t - T_{ij}),\ s_{ij}^{\leftarrow}(t) = s_{ji}^{\rightarrow}(t - T_{ji}). \quad (36)$$

Once agent i receives $s_{ij}^{\leftarrow}$, it computes r_{ij} from the second equation of (34) and adds the resulting r_{ij} to the controller (22). Remark that, in computation of r_{ij}, v_{ij} in (22) is also determined by r_{ij} and hence an algebraic loop occurs. In implementation, we determine r_{ij} by substituting (22) into the second equation of (34) as

$$r_{ij} = (\eta I_{2N} + E_{ij})^{-1}(\sqrt{2\eta}s_{ij}^{\leftarrow} + E_{ij}[x_i^T\ \xi_i^T]^T). \quad (37)$$

An algorithmic description of agent i's operation is shown in Algorithm 1.

The system with the scattering transformation (34) and (35) and the delay blocks (36) 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 (34) and (35) and the delay blocks (36) is passive from $-[\bar{v}_{ij}^T\ \bar{v}_{ji}]^T$ to $[\bar{r}_{ij}^T\ \bar{r}_{ji}^T]^T$.*

Proof: See Appendix. ■

D. Convergence Analysis

We are now ready to prove the convergence result.

Lemma 6 *Consider the system (21) for all i, and the scattering transformation (34) and (35) and the delays (36) 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 (7) for all $i = 1, 2, \ldots, n$.*

Proof: Define

$$V := \sum_{i=1}^{n} \bar{S}_i + \sum_{(i,j)\in\mathcal{E}} V_{ij}.$$

Then, combining (32) and (94), we obtain

$$\dot{V} = -\sum_{i=1}^{n}\sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2 - \alpha\sum_{i=1}^{n}(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \le 0. \quad (38)$$

Now, define $X := [x^T\ \xi^T]^T$ and X_t such that $X_t(\theta) = X(t + \theta)$ for $\theta \in [-\max_{i,j} T_{ij}, 0]$. Then, the extension of the LaSalle's principle for time delay systems [40] is applicable and any solution X_t to the system converges to the largest invariant set in the set of trajectories satisfying $\dot{V} \equiv 0$. Under Assumption 3, the gradient ϕ_i satisfies $(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) > 0$ whenever $x_i \neq z^*$. Thus, $\dot{V} = 0$ means $x_i \equiv z^*$ for all i, and hence we can conclude $x_i \to z^*\ \forall i = 1, 2, \ldots, n$. ■

The above proof, in particular (38), means that the communication delay blocks are successfully integrated with the distributed optimization algorithm as interconnections of passive systems. However, Assumption 3 requires that f_i is strictly convex on all elements of z, which is fairly strong and may limit applications. The assumption can be relaxed as below. Suppose now that f_i depends 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. Then, we assume the following.

Assumption 4 *The functions $f_1, f_2, \ldots, f_n$ 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.*

Remark that $\mathcal{Z}_i$ can be empty for some i.

Under Assumption 4 instead of 3, we have the following theorem.

Theorem 2 *Consider the same system as Lemma 6. If Assumptions 2 and 4 hold, then x_i asymptotically converges to the optimal solution z^* to (7) for all $i = 1, 2, \ldots, n$.*

Proof: Suppose that $l \in \mathcal{Z}_i$. It is then sufficient to prove that the l-th element of x_j converges to z_l^* for all j, where $z_l^* \in \mathbb{R}$ is the l-th element of z^*. Similarly to Lemma 6, LaSalle's principle for time delay systems [40] is applicable and hence we consider the set of solutions satisfying $\dot{V} \equiv 0$. In the set, $\|\bar{x}_i - \bar{r}_{ij}^x\| = 0\ \forall j \in \mathcal{N}_i$ holds, which means $\dot{\xi} = 0$. Thus, LaSalle's principle implies, under Assumption 4 and $l \in \mathcal{Z}_i$, that

$$\text{(i)} \quad \lim_{t\to\infty}(x_i - r_{ij}^x) = 0\ \forall j \in \mathcal{N}_i,\ \forall i = 1, 2, \ldots, n \tag{39}$$

(ii) the l-th element of x_i converges to z_l^*, and (iii) ξ_i has a limit $\lim_{t\to\infty}\xi_i$. From (22) and (34)–(36), we obtain

$$\begin{aligned} r_{ij}^x =& r_{ji}^x(t - T_{ji}) + (1/\eta)\{-a_{ij}d_{ij} + b_{ij}(r_{ij}^\xi - \xi_i) \\ & + b_{ij}(r_{ji}^\xi(t - T_{ji}) - \xi_j(t - T_{ji}))\}, \end{aligned} \tag{40}$$

$$\begin{aligned} r_{ji}^x =& r_{ij}^x(t - T_{ij}) + (1/\eta)\{-a_{ij}d_{ji} + b_{ij}(r_{ji}^\xi - \xi_j) \\ & + b_{ij}(r_{ij}^\xi(t - T_{ij}) - \xi_i(t - T_{ij}))\}, \end{aligned} \tag{41}$$

$$r_{ij}^\xi = r_{ji}^\xi(t - T_{ji}) - (b_{ij}/\eta)d_{ij}, \tag{42}$$

$$r_{ji}^\xi = r_{ij}^\xi(t - T_{ij}) - (b_{ij}/\eta)d_{ji}, \tag{43}$$

with

$$\begin{aligned} d_{ij}(t) &:= (r_{ij}^x - x_i) + (r_{ji}^x(t - T_{ji}) - x_j(t - T_{ji})), \\ d_{ji}(t) &:= (r_{ji}^x - x_j) + (r_{ij}^x(t - T_{ij}) - x_i(t - T_{ij})). \end{aligned}$$

Summing (43) at time $t - T_{ji}$ and (42) yields

$$r_{ij}^\xi - r_{ij}^\xi(t - \bar{T}_{ij}) = -(b_{ij}/\eta)(d_{ij} + d_{ji}(t - T_{ji})), \tag{44}$$

where $\bar{T}_{ij} := T_{ij} + T_{ji}$. From (39),

$$\lim_{t\to\infty} d_{ij} = 0 \text{ and } \lim_{t\to\infty} d_{ji}(t - T_{ji}) = 0 \tag{45}$$

hold. Thus, taking the limit of (44), it follows

$$\lim_{t\to\infty}(r_{ij}^\xi - r_{ij}^\xi(t - \bar{T}_{ij})) = 0. \tag{46}$$

Subtracting (41) at time $t - T_{ji}$ from (40) yields

$$\begin{aligned} r_{ij}^x + & r_{ij}^x(t - \bar{T}_{ij}) - 2r_{ji}^x(t - T_{ji}) \\ &= (1/\eta)\{-a_{ij}d_{ij} + a_{ij}d_{ji}(t - T_{ji}) + b_{ij}(r_{ij}^\xi - \xi_i) \\ &\quad - b_{ij}(r_{ij}^\xi(t - \bar{T}_{ij}) - \xi_i(t - \bar{T}_{ij}))\}. \end{aligned} \tag{47}$$

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

$$\lim_{t\to\infty}(\xi_i - \xi_i(t - \bar{T}_{ij})) = 0. \tag{48}$$

Taking the limit of (47) and using (45), (46), and (48), we have

$$\lim_{t\to\infty}(r_{ij}^x + r_{ij}^x(t - \bar{T}_{ij}) - 2r_{ji}^x(t - T_{ji})) = 0. \tag{49}$$

It is confirmed from (ii) and (39) that the l-th element of $\lim_{t\to\infty}(r_{ij}^x + r_{ij}^x(t - \bar{T}_{ij}))$ in (49) is equal to $2z_l^*$, which implies that the l-th element of r_{ji}^x converges to z_l^*. This and (39) 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 $j = 1, 2, \ldots, n$ because of Assumption 2. Convergence of the other elements is also proved in the same way. ■

V. Extension to Constrained Distributed Optimization

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

$$\min_{z\in\mathbb{R}^N} f(z) \text{ subject to } g_i(z) \le 0\ \forall i = 1, 2, \ldots, n, \tag{50}$$

where f is defined in the same way as (7). The functions f_i and $g_i : \mathbb{R}^N \to \mathbb{R}^{m_i}$ $(i = 1, 2, \ldots, n)$ are assumed to be private information of agent i and the other agents do not have access to these functions. 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 $g_{il}(l = 1, 2, \ldots, m_i) : \mathbb{R}^N \to \mathbb{R}$, we assume the following assumptions.

Assumption 5 *The functions f_i $(i = 1, 2, \ldots, n)$ are convex and twice differentiable. The functions g_{il} $(l = 1, 2, \ldots, m_i,\ i = 1, 2, \ldots, n)$ are convex and continuously differentiable and their gradients, denoted by $\Gamma_i := \nabla g_i \in \mathbb{R}^{N\times m_i}$ $(i = 1, 2, \ldots, n)$, are locally Lipschitz. In addition, there exists z such that $g_i(z) < 0\ \ \forall i = 1, 2, \ldots, n$.*

Assumption 6 *The function f is strictly convex.*

Note that, if Assumptions 5 and 6 are satisfied, the optimal solution z^* to (50) is uniquely determined [19]. It is well-known that, under these assumptions, z^* is the optimal solution to (50) if and only if there exist $\lambda_i^* \in \mathbb{R}^{m_i}$ $(i = 1, 2, \ldots, n)$ satisfying the KKT condition [19]:

$$\nabla f(z^*) + \sum_{i=1}^n \Gamma_i(z^*)\lambda_i^* = 0, \tag{51}$$

$$\lambda_i^* \ge 0,\ g_i(z^*) \le 0\ \ \forall i = 1, 2, \ldots, n, \tag{52}$$

$$\lambda_{il}^* g_{il}(z^*) = 0\ \ \forall l = 1, 2, \ldots, m_i,\ \forall i = 1, 2, \ldots, n, \tag{53}$$

where λ_{il}^* is the l-th element of λ_i^*. We finally define the Lagrangian H for the problem (50) as

$$H(z, \lambda) := \sum_{i=1}^n H_i(z, \lambda_i),\ \ H_i(z, \lambda_i) := f_i(z) + \lambda_i^T g_i(z). \tag{54}$$

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

$$[g]_\rho^+ := \begin{cases} 0, & \text{if } \rho = 0 \text{ and } g < 0, \\ g, & \text{otherwise} \end{cases}. \tag{55}$$

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

A. PI Consensus-Based Distributed Optimization for Constrained Problem

Let us now define the dual function $\min_z H(z,\lambda)$ and $g(z) := [g_1^T(z) \ \cdots \ g_n^T(z)]^T \in \mathbb{R}^m$ ($m := \sum_{i=1}^n m_i$). Then, inspired by the dual ascent method, we consider the following dynamical system formulated based on [13].

$$\dot\rho = [g(\hat z^*(\rho))]_\rho^+, \quad \rho(0) \ge 0, \tag{56}$$

where $\rho = [\rho_1^T \ \cdots \ \rho_n^T]^T \in \mathbb{R}^m$ ($\rho_i \in \mathbb{R}^{m_i}$) and

$$\hat z^*(\rho) := \arg\min_z H(z,\rho). \tag{57}$$

Extracting the dynamics of ρ_i from (56) yields

$$\dot\rho_i = [g_i(\hat z^*(\rho))]_{\rho_i}^+, \ \rho_i(0) \ge 0. \tag{58}$$

Let us next consider the problem (57). We see from (54) and (57) that, once ρ is fixed, the problem (57) to be solved here takes the same form as (7) except for the definition of the cost functions. Inspired by the fact, we let each agent i run the PI consensus algorithm (12) with

$$u_i = -\alpha\nabla_z H_i(x_i,\rho_i) = -\alpha\phi_i(x_i) - \alpha(\Gamma_i(x_i))\rho_i. \tag{59}$$

Remark that (59) relies only on the private functions f_i and g_i. Since x_i is regarded as an estimate of $\hat z^*(\rho)$, we replace $\hat z^*(\rho)$ in (58) by x_i and reformulate the dynamics as

$$\dot\rho_i = [g_i(x_i)]_{\rho_i}^+, \ \rho_i(0) \ge 0, \tag{60}$$

which depends only on the private function g_i. Thus, (12) with (59) and (60) can be locally executed by agent i. Note that a solution to (60) is known to exist and be unique despite discontinuity in the right-hand side of (60) [42]. It is then immediately confirmed that if $\rho_i(0) \ge 0$ then $\rho_i(t) \ge 0$ for all subsequent time t regardless of the trajectory of $g_i(x_i)$.

The collective dynamics of all agents is given by

$$\begin{bmatrix}\dot x\\ \dot\xi\end{bmatrix} = -\begin{bmatrix}\bar L_{\rm P} & -\bar L_{\rm I}\\ \bar L_{\rm I} & 0\end{bmatrix}\begin{bmatrix}x\\ \xi\end{bmatrix} - \alpha\begin{bmatrix}\phi(x)+\Gamma(x)\rho\\ 0\end{bmatrix}, \tag{61}$$

$$\dot\rho = [\bar g(x)]_\rho^+, \quad \rho(0) \ge 0 \tag{62}$$

where $\Gamma(x)$ is the block diagonal matrix with diagonal blocks $\Gamma_1(x_1),\ldots,\Gamma_n(x_n)$, and $\bar g(x) := [g_1^T(x_1) \ \cdots \ g_n^T(x_n)]^T$. The dynamics (62) is obtained by stacking (60) for $i = 1,2,\ldots,n$. The block diagram of the system is illustrated in Fig. 6. Notice that an outer feedback path is added to the solution of the unconstrained problems in Fig. 2.

B. Equilibrium/Convergence Analysis: Constrained Case

We consider Fig. 6 and cut the loop at the left of the block $\Gamma(x)$ and focus on the systems encircled by light and dark gray, where the input of the former is denoted by μ and output of the latter is by ν. Then, these systems are formulated as

$$\begin{bmatrix}\dot x\\ \dot\xi\end{bmatrix} = -\begin{bmatrix}\bar L_{\rm P} & -\bar L_{\rm I}\\ \bar L_{\rm I} & 0\end{bmatrix}\begin{bmatrix}x\\ \xi\end{bmatrix} - \alpha\begin{bmatrix}\phi(x)\\ 0\end{bmatrix} + \alpha\begin{bmatrix}\mu\\ 0\end{bmatrix} \tag{63}$$

with output x and

$$\dot\rho = [\bar g(x)]_\rho^+, \quad \nu = \Gamma(x)\rho. \tag{64}$$

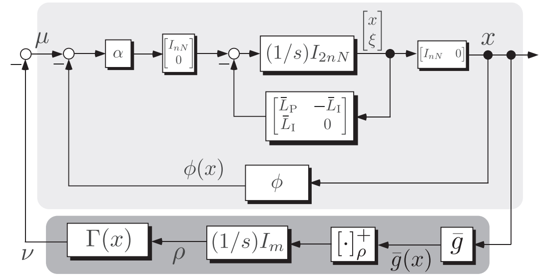


Fig. 6. Block diagram of the PI consensus-based distributed optimization algorithm for constrained optimization. The system colored by light gray is passive from $\tilde\mu = \mu - \mu^*$ to $\tilde x = x - x^*$ with $\mu^* := \Gamma(x^*)\lambda^*$ (Lemma 8). The system colored by dark gray is also passive from $\tilde x = x - x^*$ to $\tilde\nu = \nu - \nu^*$ with $\nu^* := \Gamma(x^*)\lambda^*$ (Lemma 9).

Hereafter, we prove passivity of (63) and (64). Before that, we present the following lemma on the equilibrium of (63).

Lemma 7 *Suppose that Assumptions 2, 5 and 6 hold. Denote a solution to (51)–(53) by (z^*,λ^*), where $\lambda^* := [(\lambda_1^*)^T \ \cdots \ (\lambda_n^*)^T]^T \in \mathbb{R}^m$. Then, there exists ξ^* such that the pair of $x^* = (\boldsymbol{I}_n \otimes z^*)$ and ξ^* is an equilibrium of (63) for the equilibrium input $\mu^* := \Gamma(x^*)\lambda^*$.*

Proof: Consider the right-hand side of (63). Substituting $x = x^*$ yields $\bar L_{\rm I} x^* = (L_{\rm I}\mathbf{1}_N)\otimes z^* = 0$ and hence $\dot\xi = 0$. Replacing μ by μ^*, we obtain

$$\dot x^* = \bar L_{\rm I}\xi - \alpha(\phi(x^*) + \Gamma(x^*)\lambda^*). \tag{65}$$

It is thus sufficient to prove that there exists ξ such that the right-hand side of (65) is zero. In other words, we have only to prove that $\alpha(\phi(x^*)+\Gamma(x^*)\lambda^*)$ is included in the image of the matrix $\bar L_{\rm I}$, which is equivalent to $(\mathbf{1}_n \otimes I_N)^T(\phi(x^*) + \Gamma(x^*)\lambda^*) = 0$ since L_I has a simple eigenvalue 0 for connected graphs. This equation is immediately proved from (51). ■

We are now ready to prove passivity of (63).

Lemma 8 *Suppose that Assumptions 2, 5 and 6 hold. Then, the system (63) is passive from $\tilde\mu := \mu - \mu^*$ to $\tilde x := \xi - \xi^*$ with respect to the storage function $\frac{1}{\alpha}\tilde S$, with $\tilde S$ in (17).*

Proof: Define $\tilde\xi := \xi - \xi^*$ for a fixed ξ^*. Then, it follows from (63) and Lemma 7 that

$$\begin{bmatrix}\dot{\tilde x}\\ \dot{\tilde\xi}\end{bmatrix} = -\begin{bmatrix}\bar L_{\rm P} & -\bar L_{\rm I}\\ \bar L_{\rm I} & 0\end{bmatrix}\begin{bmatrix}\tilde x\\ \tilde\xi\end{bmatrix} - \alpha\begin{bmatrix}\phi(x)-\phi(x^*)\\ 0\end{bmatrix} + \alpha\begin{bmatrix}\tilde\mu\\ 0\end{bmatrix}. \tag{66}$$

Now, following the same procedure as Lemma 2, we can immediately prove

$$\begin{aligned}\frac{1}{\alpha}\dot{\tilde S} &= -\frac{1}{\alpha}\tilde x^T\bar L_{\rm P}\tilde x - (x-x^*)^T(\phi(x)-\phi(x^*)) + \tilde\mu^T\tilde x \\ &\le \tilde\mu^T\tilde x.\end{aligned} \tag{67}$$

This completes the proof. ■

We next prove passivity of (64) as follows.

Lemma 9 *Suppose that Assumptions 5 and 6 hold. Then, the system (64) with $\rho(0) \ge 0$ is passive from $\tilde x$ to $\tilde\nu := \nu - \nu^*$*

with $\nu^* := \Gamma(x^*)\lambda^*$ for the storage function $\sum_{i=1}^n U_i$, where $U_i := \frac{1}{2}\|\rho_i - \lambda_i^*\|^2$.

Proof: Denote l-th element of ρ_i by ρ_{il} ($l = 1, 2, \ldots, m_i$). Then, the dynamics of ρ_{il} in (60) is given as

$$\dot{\rho}_{il} = [g_{il}(x_i)]^+_{\rho_{il}}, \ \rho_{il}(0) \geq 0, \tag{68}$$

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 (55) 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

$$(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}} = (\rho_{il} - \lambda_{il}^*)g_{il}(x_i) \tag{69}$$

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 (52) and $g_{il}(x_i) < 0$ from (55), the term $\lambda_{il}^* g_{il}(x_i)$ is non-positive and hence we obtain

$$(\rho_{il} - \lambda_{il}^*)[g_{il}(x_i)]^+_{\rho_{il}} \leq (\rho_{il} - \lambda_{il}^*)g_{il}(x_i). \tag{70}$$

Let us next consider the time when a mode switch happens in (60) for some i and l. In this case, $U_{il}(t) := \frac{1}{2}\|\rho_{il}(t) - \lambda_{il}^*\|^2$ can be indifferentiable in the standard sense. We thus introduce the upper Dini derivative[1] denoted by D^+U_{il}. Then, it 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

$$D^+U_i \leq \sum_{l=1}^{m_i} (\rho_{il} - \lambda_{il}^*)g_{il}(x_i). \tag{71}$$

From (69) and (70), the inequality (71) holds for all time $t \in \mathbb{R}^+$. This is rewritten as

$$D^+U_i \leq (\rho_i - \lambda_i^*)^T\{g_i(x_i) - g_i(z^*)\} + (\rho_i - \lambda_i^*)^T g_i(z^*) \tag{72}$$

Noticing that $\rho_i \geq 0$ and $g_i(z^*) \leq 0$ from (52), the inequality $\rho_i^T g_i(z^*) \leq 0$ holds. In addition, $(\lambda_i^*)^T g_i(z^*) = 0$ is true from (53). Thus, (72) is further rewritten as

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

Because of the convexity of g_{il}, the following inequalities hold.

$$\begin{aligned}g_{il}(x_i) - g_{il}(z^*) &\geq (\nabla g_{il}(z^*))(x_i - z^*), \\ g_{il}(x_i) - g_{il}(z^*) &\leq (\nabla g_{il}(x_i))(x_i - z^*).\end{aligned}$$

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

Using these together with $\rho_i \geq 0$ and $\lambda_i^* \geq 0$, we can prove

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

Thus, we have

$$\begin{aligned}D^+\left(\sum_{i=1}^n U_i\right) &\leq \{(\Gamma(x))\rho - (\Gamma(x^*))\lambda^*\}^T(x - x^*) \\ &= \tilde{\nu}^T\tilde{x}.\end{aligned} \tag{73}$$

Integrating (73) in time proves the original definition (2)[2]. ■ Namely, the closed-loop system (61) and (62) is regarded as a feedback interconnection of two passive systems.

We are now ready to state the following convergence result.

Theorem 3 *Consider the system (61) and (62). If Assumptions 2, 5 and 6 hold, then x_i asymptotically converges to the optimal solution z^* to (50) for all $i = 1, 2, \ldots, n$.*

Proof: Define $U := \frac{1}{\alpha}\tilde{S} + \sum_{i=1}^n U_i$. From (67) and (73),

$$\begin{aligned}D^+U &= \frac{1}{\alpha}\dot{\tilde{S}} + D^+\left(\sum_{i=1}^n U_i\right) \\ &\leq -\frac{1}{\alpha}x^T\bar{L}_{\rm P}x - (x - x^*)^T(\phi(x) - \phi(x^*)) \leq 0\end{aligned}$$

holds. Integrating this in time, we have

$$\int_0^\infty \Big(\frac{1}{\alpha}x^T\bar{L}_{\rm P}x + (x - x^*)^T(\phi(x) - \phi(x^*)\Big)dt < \infty. \tag{74}$$

We also see $x, \xi, \rho \in \mathcal{L}_\infty$ since U is positive definite. Thus, from (61), $\dot{x} \in \mathcal{L}_\infty$ also holds. The time derivative of $\frac{1}{\alpha}x^T\bar{L}_{\rm P}x + (x - x^*)^T(\phi(x) - \phi(x^*))$ is given as

$$\frac{2}{\alpha}x^T\bar{L}_{\rm P}\dot{x} + (\phi(x) - \phi(x^*))^T\dot{x} + (x - x^*)^T(\nabla\phi(x))\dot{x},$$

which is also bounded. Thus, invoking Barbalat's lemma, we can prove $\lim_{t\to\infty} x^T\bar{L}_{\rm P}x = 0$ and $\lim_{t\to\infty}(x - x^*)^T(\phi(x) - \phi(x^*)) = 0$. The first equation means that there exists a trajectory $c(\cdot)$ such that $\lim_{t\to\infty}(x_i - c) = 0$ for all i. The second equation is then rewritten as

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

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 equation, we can also prove $c \to z^*$ in the same way as asymptotic stability in Lyapunov theorem. It is thus concluded that $x_i \to z^*$ hold for all i. ■

[2]If a function $U(t)$ is continuous and has a finite D^+U for every t, D^+U is integrable and satisfies $U(t) - U(t_0) = \int_{t_0}^t D^+U(\tau)d\tau$ holds for any interval $[t_0, t]$ [41].

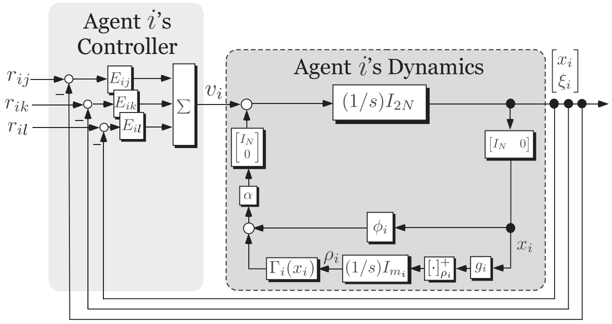


Fig. 7. Block diagram of agent i's dynamics for the constrained optimization in the presence of communication delays. The system enclosed by the dashed line is passive from $\bar{v}_i$ to $[\bar{x}_i^T \ \bar{\xi}_i^T]^T$ (Lemma 10), where $\bar{v}_i$, $\bar{x}_i$ and $\bar{\xi}_i$ are defined in (25) and (26).

The present results can be easily extended to the problem with equality constraints using techniques e.g. in [16], but we omit including the result to reduce complexity of the paper. Inequality constraints for another class of distributed optimization are also addressed in [43] using another energy functions, where the authors employ LaSalle's principle for hybrid systems [44]. The completed version of the proof of the result in [43] is presented in [42]. Besides the difference of the problem formulation, a more explicit contribution relative to [43], [42] is presented in the next subsection.

C. Constrained Distributed Optimization with Delay

In this subsection, we extend the results in Section IV to the constrained problem (50).

Let us revisit the agent i's dynamics (12), (59) and (60). Similarly to Section IV, x_j and ξ_j are not directly received from agent $j \in \mathcal{N}_i$ in the presence of delays and we again denote the received information as r_{ij}^x and r_{ij}^ξ, respectively. Then, the system is formulated as

$$\dot{x}_i = \sum_{j\in\mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) - \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^\xi - \xi_i) - \alpha(\phi_i(x_i) + (\Gamma_i(x_i))\rho_i), \quad (75)$$

$$\dot{\xi}_i = \sum_{j\in\mathcal{N}_i} b_{ij}(r_{ij}^x - x_i), \quad (76)$$

$$\dot{\rho}_i = [g_i(x_i)]^+_{\rho_i}, \ \ \rho_i(0) \geq 0, \quad (77)$$

whose block diagram is illustrated in Fig. 7. The previous subsection shows that $r_{ij}^x = x$ and $r_{ij}^\xi = \xi$ achieves asymptotic optimality in the delay free case but this is not applied to the case with delays similarly to Section IV, namely $r_{ij}^x(t) = x(t - T_{ji})$ and $r_{ij}^\xi(t) = \xi(t - T_{ji})$ can destabilize the system.

Now, we focus on the blocks encircled by the dashed line in Fig. 7 whose system formulation is given by

$$\begin{bmatrix}\dot{x}_i \\ \dot{\xi}_i\end{bmatrix} = v_i - \alpha \begin{bmatrix}\phi_i(x_i) + \Gamma_i(x_i)\rho_i \\ 0\end{bmatrix}. \quad (78)$$

Then, we have the following lemma.

Lemma 10 *Suppose that Assumptions 2, 4, and 5 hold. Then, the system (78) 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 9, respectively, and $\bar{x}_i$, $\bar{\xi}_i$ and $\bar{v}_i$ are defined in the same way as (25) and (26).*

Proof: From Lemma 7 and (78), we have

$$\begin{bmatrix}\dot{\bar{x}}_i \\ \dot{\bar{\xi}}_i\end{bmatrix} = \bar{v}_i - \alpha \begin{bmatrix}\phi_i(x_i) - \phi_i(z^*) + \Gamma_i(x_i)\rho_i - \Gamma_i(z^*)\lambda_i^* \\ 0\end{bmatrix} \quad (79)$$

Similarly to the proof of Lemma 4, it follows

$$\dot{\bar{S}}_i = \begin{bmatrix}\bar{x}_i \\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T\{\phi_i(x_i) - \phi_i(z^*) + (\Gamma_i(x_i))\rho_i - (\Gamma_i(z^*))\lambda_i^*\}. \quad (80)$$

We also see from Lemma 9 that

$$\alpha D^+\dot{U}_i \leq \alpha(x_i - z^*)^T\{(\Gamma_i(x_i))\rho_i - (\Gamma_i(z^*))\lambda_i^*\}. \quad (81)$$

Combining (80) and (81), we obtain

$$D^+W_i \leq \begin{bmatrix}\bar{x}_i \\ \bar{\xi}_i\end{bmatrix}^T \bar{v}_i - \alpha(x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)). \quad (82)$$

Integrating this in time proves the lemma. ■

The above lemma means that the agent's passive dynamics in Fig. 3 is just replaced by another passive dynamics with the same input-output pair. Thus, we take the same inter-agent communication strategy as Algorithm 1. Then, Lemma 5 again holds true and hence we have the following result.

Theorem 4 *Consider the system (75), (76) and (77) for all i with the scattering transformation (34) and (35) and delays (36) for all $j \in \mathcal{N}_i$ and all i. If Assumptions 2, 3 and 5 hold, then x_i asymptotically converges to the optimal solution z^* to (50) 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 (82) and (94), we obtain

$$D^+W \leq -\sum_{i=1}^n \sum_{j\in\mathcal{N}_i} a_{ij}\|\bar{x}_i - \bar{r}_{ij}^x\|^2 - \alpha\sum_{i=1}^n (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) \leq 0, \quad (83)$$

which implies that $x_i, \xi_i \in \mathcal{L}_\infty \ \forall i$.

Using (22), (34), (35), and (36), we can derive

$$r_{ij}(t) = \bar{E}_{ij}^2 r_{ij}(t - T_{ij} - T_{ji}) + \beta_{ij}(t) \quad (84)$$

$$r_{ji}(t) = \bar{E}_{ij}^2 r_{ji}(t - T_{ij} - T_{ji}) + \beta_{ji}(t) \quad (85)$$

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 (84) and (85) are stable difference equations with bounded inputs and hence $r_{ij}, r_{ji} \in \mathcal{L}_\infty$. Therefore, $\dot{x}_i \in \mathcal{L}_\infty \ \forall i$.

Integrating both sides of (83) proves that

$$\int_0^\infty (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) < \infty$$

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 bounded. Thus, using Barbalat's lemma, we can prove

$$\lim_{t\to\infty} (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) = 0 \ \forall i = 1, \ldots, n. \quad (86)$$

Under Assumption 3 and 5, it is equivalent to $x_i \to z^*$. ∎

Assumption 3 can be relaxed to 4 if we add the following stronger assumption on the communication. To state the assumption, we define a new graph $G' = (\{1, 2, \ldots, n\}, \mathcal{E}')$ such that $(j, k) \in \mathcal{E}'$ iff there exists i satisfying $j \in \mathcal{N}_i$ and $k \in \mathcal{N}_i$. Then, we assume the following.

Assumption 7 *The delays are homogeneous, namely $T_{ij} = T \ \forall i, j$ for some T. In addition, the graph G' is connected.*

Theorem 5 *Consider the same system as Theorem 4. If Assumptions 4, 5 and 7 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 (50) for all $i = 1, 2, \ldots, n$.*

Proof: The inequality (83) means that $(x_i - r_{ij}^x) \in \mathcal{L}_2$ for all $j \in \mathcal{N}_i$ and $i = 1, 2, \ldots, n$. As proved in Theorem 4, $\dot{x}_i \in \mathcal{L}_\infty \ \forall i$, and we can prove $\dot{\xi}_i \in \mathcal{L}_\infty \ \forall i$ in the same way. Solving (84), $r_{ij}(t)$ is given by a convolution sum of the input β_{ij} whose time derivative is bounded. Thus, we have $\dot{r}_{ij} \in \mathcal{L}_\infty \ \forall i$ for all $j \in \mathcal{N}_i$ and $i = 1, 2, \ldots, n$. Invoking Barbalat's lemma, we obtain

$$\lim_{t\to\infty} (x_i - r_{ij}^x) = 0 \quad \forall j \in \mathcal{N}_i \text{ and } \forall i = 1, 2, \ldots, n. \quad (87)$$

Under Assumption 7, there exists i such that $|\mathcal{N}_i| \geq 2$. Take two neighbors $j, k \in \mathcal{N}_i$ of such i. Then, (87) implies that

$$\lim_{t\to\infty} (r_{ij}^x - r_{ik}^x) = 0. \quad (88)$$

Following the same procedure as Theorem 2, the equations (46) and (47) hold. From (87), we can also prove (45). Remark that (48) is not proved here. Taking the limit of (47) with $T_{ij} = T_{ji} = T$ and $b_{ij} = b$ and using (45) and (46) yield

$$\lim_{t\to\infty} \{r_{ij}^x + r_{ij}^x(t - 2T) - 2r_{ji}^x(t - T) + (b/\eta)(\xi_i - \xi_i(t - 2T))\} = 0 \quad (89)$$

under Assumption 7. The same equation holds for k as

$$\lim_{t\to\infty} \{r_{ik}^x + r_{ik}^x(t - 2T) - 2r_{ki}^x(t - T) + (b/\eta)(\xi_i - \xi_i(t - 2T))\} = 0. \quad (90)$$

Subtracting (90) from (89) and using (88), we have

$$\lim_{t\to\infty} (r_{ji}^x - r_{ki}^x) = 0. \quad (91)$$

From (87), (91) 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 7, we can conclude that there exists a trajectory $c(\cdot)$ such that $\lim_{t\to\infty}(x_i - c) = 0$ for all i. The remaining of the proof is the same as Theorem 3. ∎

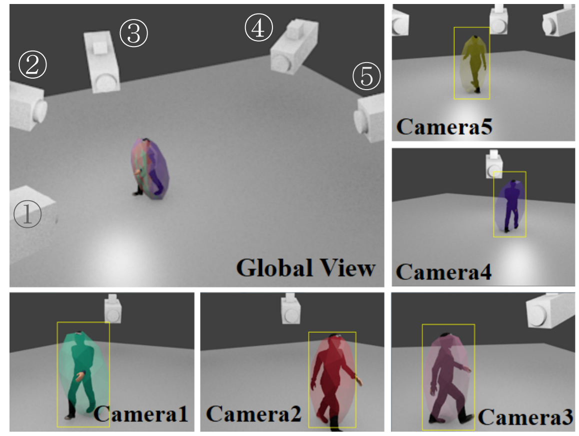


Fig. 8. Intended scenario of 3-D visual human localization. Each camera acquires a rectangle enclosing the human (yellow boxes in the small windows) using a pedestrian detection algorithm. The transparent colored ellipsoids in the small windows are final estimates generated by the algorithm presented in Subsection V-C, and all of them are also shown in the large window.

In Theorem 4, convergence to the optimal solution is directly proved from the second term of (83), since $x_i = z^*$ holds if $\sum_{i=1}^n (x_i - z^*)^T(\phi_i(x_i) - \phi_i(z^*)) = 0$ under Assumption 3. However, it is not true under Assumption 4, and hence further investigations to prove synchronization of x_i are needed, where the additional Assumption 7 is required.

VI. Application to Human Localization Using Pedestrian Detection Algorithm

We finally apply the proposed algorithms to the visual human localization problem investigated in [39]. Here, multiple networked cameras are assumed to be distributed over the 3-D Euclidean space to monitor a human as shown in Fig. 8. Each camera acquires 2-D rectangles on its own image plane in which the human lives, as shown in the small windows of Fig. 8, by executing a pedestrian detection algorithm e.g. in [45]. 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. Please refer to [39] for the case where some cameras do not detect.

If the human is modeled as an ellipsoid $\Omega(q, Q) = \{p \in \mathbb{R}^3 | \ (p - q)^T Q^{-2}(p - q) \leq 1\}$, the decision variable z consists of the elements of $q \in \mathbb{R}^3$ and $Q \in \mathbb{S}^{3\times 3}$. Note that the symmetric matrices are parametrized by 6 variables. Then, we formulate the local cost function

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

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 camera 1–3, and 1st and 3rd elements to camera 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 scalar $w > 0$ is a weighting coefficient, and it is set to $w = 1$ in this

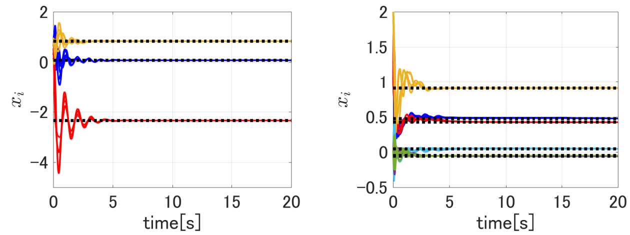


Fig. 9. Trajectories of $x_1, \ldots, x_5$ without communication delays (left: vector variable q, right: matrix variable Q).

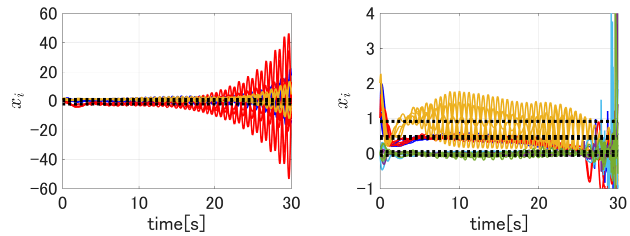


Fig. 10. Trajectories of $x_1, \ldots, x_5$ with communication delays and without the scattering transformation (left: vector variable q, right: matrix variable Q).

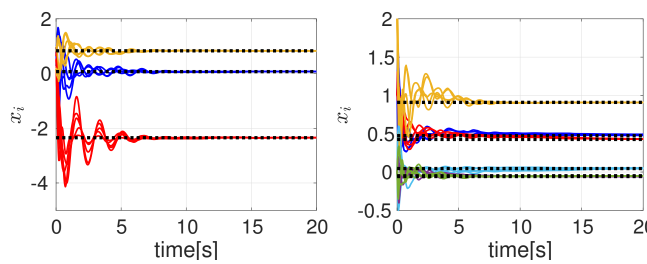


Fig. 11. Trajectories of $x_1, \ldots, x_5$ with communication delays and scattering transformation (left: vector variable q, right: matrix variable Q).

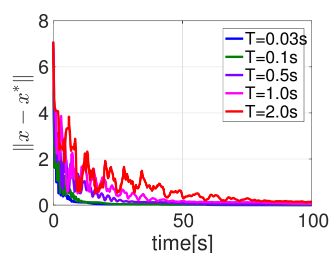


Fig. 12. Trajectories of $\|x^* - x_i\|$ for a variety of delays T.

simulation. The local constraints are given by $\Omega(q, Q) \subseteq \mathcal{H}_i$ and $Q > 0$, which are reduced to the form of $g_i(z) \leq 0$. See [39] for more details on the formulation. Note that the problem satisfies Assumptions 4–6 in a realistic situation [3].

We first run the algorithm with $\eta = 1$ presented in Subsection V-A without adding communication delays, 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 $a_{ij} = 1$ 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 the estimates $x_1, x_2, \ldots, x_5$ are illustrated in Fig. 9, where the dashed line describes the actual optimal solution. We see that the estimates converge to the solution.

[3]The cost function can be neither differentiable nor convex in the region that Q is not positive definite. However, this does not matter since it is proved in Lemma 5 of [39] that if the initial estimate of Q is positive definite, they remain positive definite for all subsequent time.

Let us next add communication delays, where every T_{ij} is set to $T = 0.03$s. Then, we run the above algorithm under the same setting. In the presence of delays, the trajectories of $x_1, x_2, \ldots, x_5$ are changed to Fig. 10, namely they diverge and the simulation stops with errors.

We then implement the algorithm presented in Subsection V-C. The resulting trajectories $x_1, x_2, \ldots, x_5$ are shown in Fig. 11. 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. 8, where we see that every camera successfully computes an ellipsoid tightly enclosing the human.

In order to demonstrate the effect of the delay length on the system performance, we draw trajectories of $\|x - x_i\|$ a variety of delays in Fig. 12. It is seen from the figure that the error converges to zero for all of the delays, but the convergence is deaccelerated as T increases. Acceleration of the convergence is left as a future work but potential solutions are listed below. The option considered to be the most promising is the loop shaping strategy in [12], [13]. The optimal edge weight design technique in [46] is also expected to be useful for this purpose. Taking the Newton-Raphson approach as in [47], instead of the gradient descent algorithm, would accelerate the algorithm if it would successfully fit the present framework.

VII. 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 the inter-agent communication delays can be integrated while ensuring the convergence property using scattering transformation. Moreover, we have extended the results to distributed optimization with local inequality constraints, and presented a passivity-based solution both in the absence and presence of delays. 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] S. Knorn, A. Donaire, J.C. Aguero and R.H. Middleton, "Passivity-based control for multi-vehicle systems subject to string constraints," Automatica vol. 50, no. 12, pp. 3224–3230, 2014.

[5] S. Mukherjee, S. Mishra and J.T. Wen, "Building temperature control: A passivity-based approach," Proc. IEEE Conf. Decision and Control, pp. 6902-6907, 2012.

[6] 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, to be presented, 2017.

[7] M. Ilic, K. Bachovchin, M. Cvetkovic and M. Xia, "Physics-based foundations for cyber and market design in complex electric energy systems," Proc. IEEE Conf. Decision and Control, pp. 4635–4654, 2014.

[8] 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.

[9] P. Lee, A. Clark, L. Bushnell and R. Poovendran, "A passivity framework for modeling and mitigating wormhole attacks on networked control systems," IEEE Trans. Autom. Control, vol. 59, no. 12, pp. 3224–3237, 2014.

[10] P.J. Antsaklis, B. Goodwine, V. Gupta, M.J. McCourt, Y. Wang, P. Wu, M. Xia, H. Yu and F. Zhu, "Control of cyberphysical systems using passivity and dissipativity based methods," Eur. J. Control, vol. 19, no. 5, 379–388, 2013.

[11] T. Hatanaka, N. Chopra and M. Fujita, "Passivity-based bilateral human-swarm-interactions for cooperative robotic networks and human passivity analysis," Proc. of 54th IEEE Conf. Decision and Control, pp.1033–1039, 2015.

[12] 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.

[13] 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.

[14] M. Bürger, D. Zelazo and F. Allgöwer, "Duality and network theory in passivity-based cooperative control," Automatica, vol. 50, no. 8, pp. 2051–2061, 2014.

[15] M. Bürger and C.D. Persis, "Dynamic coupling design for nonlinear output agreement and time-varying flow control," Automatica, vol. 51, no. 1, pp. 210–222, 2015.

[16] H. Yamamoto and K. Tsumura, "Control of smart grids based on price mechanism and network structure," The University of Tokyo, Mathematical Engineering Technical Reports, 2012.

[17] T. Ishizaki, A. Ueda and J. Imura, "Convex gradient controller design for incrementally passive systems with quadratic storage functions ," Proc. 55th IEEE Conf. Decision and Control, pp. 1721–1726, 2016.

[18] G.D.O. Bravo, *Distributed Methods for Resource Allocation : A Passivity Based Approach*, PhD thesis, Universidad de Los Andes, Ecole des Mines de Nantes, 2015.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] G. Qu and N. Li, "Harnessing smoothness to accelerate distributed optimization," arXiv:1605.07112v1, 2016.

[25] I. Lobel and A. Ozdaglar, "Distributed subgradient methods for convex optimization over random networks," IEEE Trans. Autom. Control, vol. 56, no. 6, pp. 1291–1306, 2011.

[26] S.S. Ram, A. Nedić and V.V. Veeravalli, "Distributed stochastic subgradient projection algorithms for convex optimization," J. Optim. Theory Appl., vol. 147, pp. 516–545, 2010.

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

[28] 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.

[29] G. Droge and M. Egerstedt, "Proportional integral distributed optimization for dynamic network topologies," Proc. 2014 American Control Conference, pp. 3621–3626, 2014.

[30] 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.

[31] H. Bai, R.A. Freeman, K.M. Lynch, "Robust dynamic average consensus of time-varying inputs," Proc. 49th IEEE Conf. Decision and Control, pp. 3104–3109, 2010.

[32] 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.

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

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

[35] 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.

[36] M. T Hale, A. Nedić and M. Egerstedt, "Cloud-based centralized/decentralized multi-agent optimization with communication delays," Proc. 53rd IEEE Conference on Decision and Control, pp. 6635–6640, 2015.

[37] 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.

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

[39] 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, to be uploaded, 2016.

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

[41] J.W. Hagood and B.S. Thomson, "Recovering a function from a Dini derivative," Amer. Math. Monthly, vol. 113, pp. 34–46, 2006.

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

[43] D. Feijer and F. Paganini, "Stability of primal–dual gradient dynamics and applications to network optimization," Automatica, vol. 46, no. 12, pp. 1974–1981, 2010.

[44] J. Lygeros, K.H. Johansson, S.N. Simic, J. Zhang and S.S. Sastry, "Dynamical properties of hybrid automata," IEEE Trans. Autom. Control, vol. 48, pp. 2–17, 2003.

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

[46] X. Wu, F. Dörfler, and M.R. Jovanović, "Topology identification and design of distributed integral action in power networks," Proc. 2016 American Control Conference, pp. 5921–5926, 2016.

[47] R. Carli, G. Notarstefano, L. Schenato and D. Varagnolo, "Analysis of Newton-Raphson Consensus for multi-agent convex optimization under asynchronous and lossy communi- cation," Proc. IEEE Conference on Decision and Control 2015, pp. 418-424 2015.

APPENDIX A
PROOF OF LEMMA 5

Define

$$V_{ij}(t) = \frac{1}{2}\int_0^t \Big(\|s_{ij}^{\rightarrow}(\tau)+\gamma_{ij}^*\|^2 - \|s_{ji}^{\leftarrow}(\tau)+\gamma_{ij}^*\|^2 + \|s_{ji}^{\rightarrow}(\tau)-\delta_{ij}^*\|^2 - \|s_{ij}^{\leftarrow}(\tau)-\delta_{ij}^*\|^2\Big)d\tau + \frac{T_{ij}}{2}(\gamma_{ij}^*)^2 + \frac{T_{ji}}{2}(\delta_{ij}^*)^2, \quad (92)$$

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 (36) and $s_{ij}(t) = 0\ \forall t < 0$, the second term of (92) is equal to

$$-\frac{1}{2}\int_0^t \|s_{ji}^{\leftarrow}(\tau)+\gamma_{ij}^*\|^2 d\tau = -\frac{1}{2}\int_{-T_{ij}}^{t-T_{ij}} \|s_{ij}^{\rightarrow}(\tau)+\gamma_{ij}^*\|^2 d\tau$$
$$= -\frac{1}{2}\int_0^{t-T_{ij}} \|s_{ij}^{\rightarrow}(\tau)+\gamma_{ij}^*\|^2 d\tau - \frac{T_{ij}}{2}(\gamma_{ij}^*)^2$$

Similarly, the forth term of (92) is rewritten as

$$-\frac{1}{2}\int_{0}^{t-T_{ji}} \|s_{ji}^{\rightarrow}(\tau)-\delta_{ij}^{*}\|^2 d\tau - \frac{T_{ji}}{2}(\delta_{ij}^{*})^2$$

Substituting them into (92) yields

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

which means $V_{ij}(t) \geq 0\ \forall t$.

Let us get back to the description (92). The time derivative of V_{ij} is then given by

$$\begin{aligned} \dot{V}_{ij}(t) =& \frac{1}{2}\Big(\|s_{ij}^{\rightarrow}(t)+\gamma_{ij}^{*}\|^2 - \|s_{ji}^{\leftarrow}(t)+\gamma_{ij}^{*}\|^2 \\ &+ \|s_{ji}^{\rightarrow}(t)-\delta_{ij}^{*}\|^2 - \|s_{ij}^{\leftarrow}(t)-\delta_{ij}^{*}\|^2\Big). \end{aligned} \tag{93}$$

Substituting (34), (35), $\gamma_{ij}^{*} := \frac{1}{\sqrt{2\eta}}(v_{ij}^{*}-\eta r_{ij}^{*})$ and $\delta_{ij}^{*} := \frac{1}{\sqrt{2\eta}}(v_{ij}^{*}+\eta r_{ij}^{*})$ into (93), we obtain

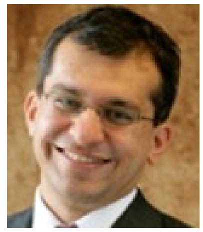
$$\begin{aligned} \dot{V}_{ij}(t) =& \frac{1}{4\eta}\Big(\big\|(-v_{ij}(t)+v_{ij}^{*})+\eta(r_{ij}(t)-r_{ij}^{*})\big\|^2 \\ &- \big\|(v_{ji}(t)+v_{ij}^{*})+\eta(r_{ji}(t)-r_{ij}^{*})\big\|^2 \\ &+ \big\|(-v_{ji}(t)-v_{ij}^{*})+\eta(r_{ji}(t)-r_{ij}^{*})\big\|^2 \\ &- \big\|(v_{ij}(t)-v_{ij}^{*})+\eta(r_{ij}(t)-r_{ij}^{*})\big\|^2\Big). \\ =& \frac{1}{4\eta}\Big(\big\|(-v_{ij}(t)+v_{ij}^{*})+\eta(r_{ij}(t)-r_{ij}^{*})\big\|^2 \\ &- \big\|(v_{ji}(t)-v_{ji}^{*})+\eta(r_{ji}(t)-r_{ji}^{*})\big\|^2 \\ &+ \big\|(-v_{ji}(t)+v_{ji}^{*})+\eta(r_{ji}(t)-r_{ji}^{*})\big\|^2 \\ &- \big\|(v_{ij}(t)-v_{ij}^{*})+\eta(r_{ij}(t)-r_{ij}^{*})\big\|^2\Big), \end{aligned}$$

where the latter equation holds from (33). Since $\bar{v}_{ij} = v_{ij}-v_{ij}^{*}$ and $\bar{r}_{ij} = r_{ij}-r_{ij}^{*}$ for all i, j, we also have

$$\begin{aligned} \dot{V}_{ij}(t) =& \frac{1}{4\eta}\Big(\|-\bar{v}_{ij}(t)+\eta\bar{r}_{ij}(t)\|^2 - \|\bar{v}_{ji}(t)+\eta\bar{r}_{ji}(t)\|^2 \\ &+ \|-\bar{v}_{ji}(t)+\eta\bar{r}_{ji}(t)\|^2 - \|\bar{v}_{ij}(t)+\eta\bar{r}_{ij}(t)\|^2\Big) \\ =& -\bar{v}_{ij}^{T}(t)\bar{r}_{ij}(t) - \bar{v}_{ji}^{T}(t)\bar{r}_{ji}(t). \end{aligned} \tag{94}$$

This completes the proof.

Takeshi Hatanaka received the B.Eng. degree in informatics and mathematical science, the M.Inf. and Ph.D. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 2002, 2004, and 2007, respectively. Since 2007, he has been with Tokyo Institute of Technology, where he is currently an Associate Professor. His research interests include networked robotics and energy management systems. He is the coauthor of "Passivity-Based Control and Estimation in Networked Robotics" (Springer, 2015). He received 2017 Kimura Award, 2014 Pioneer Award and 2009, 2015 Outstanding Paper Award from SICE and 10th Asian Control Conference Best Paper Prize Award. He is a member of the Conference Editorial Board of IEEE CSS and an Associate Editor of IEEE Transactions on Control Systems Technology.



Nikhil Chopra received the bachelor's (Hons.) degree in mechanical engineering from the Indian Institute of Technology, Kharagpur, Kharagpur, India, and the Ph.D. degree in systems and entrepreneurial engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2001 and 2006, respectively. He was a Post-Doctoral Research Associate at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign from 2006 to 2007. He is currently an Associate Professor with the Department of Mechanical Engineering, University of Maryland, College Park, MD, USA. His current research interests include networked control systems, cooperative control of networked robots, and bilateral teleoperation. Dr. Chopra was the Co-Chair of the IEEE RAS Technical Committee on Telerobotics from 2006 to 2009. He is an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Takayuki Ishizaki was born in Aichi, Japan, in 1985. He received the B.Sc., M.Sc., and Ph.D. degrees in Engineering from Tokyo Institute of Technology, Tokyo, Japan, in 2008, 2009, and 2012, respectively. He served as a Research Fellow of the Japan Society for the Promotion of Science from April 2011 to October 2012. From October to November 2011, he was a Visiting Student at Laboratoire Jean Kuntzmann, Universite Joseph Fourier, Grenoble, France. From June to October 2012, he was a Visiting Researcher at School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. Since November 2012, he has been with the Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, where he is currently an Assistant Professor. His research interests include the development of model reduction and its applications. Dr. Ishizaki is a member of IEEE, SICE, and ISCIE. He was a finalist of the 51st IEEE Conference on Decision and Control Best Student-Paper Award.

Na Li received the B.S. degree in mathematics and applied mathematics from Zhejiang University, China and the Ph.D. degree in control and dynamical systems from California Institute of Technology, Pasadena, CA, USA, in 2013. She is an Assistant Professor with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. She was a Postdoctoral Associate with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. Her research interests include design, analysis, optimization and control of distributed network systems, with particular applications to power networks and systems biology/physiology. She was the Best Student Paper Award finalist in the 2011 IEEE Conference on Decision and Control.