

# Smooth Dynamics for Distributed Constrained Optimization with Heterogeneous Delays

Mengmou Li, Shunya Yamashita, Takeshi Hatanaka, and Graziano Chesi

**Abstract**—This work investigates the distributed constrained optimization problem under inter-agent communication delays from the perspective of passivity. First, we propose a continuous-time algorithm for distributed constrained optimization with general convex objective functions. The asymptotic stability under general convexity is guaranteed by the phase lead compensation. The inequality constraints are handled by adopting a projection-free generalized Lagrangian, whose primal-dual gradient dynamics preserves passivity and smoothness, enabling the application of the LaSalle's invariance principle in the presence of delays. Then, we incorporate the scattering transformation into the proposed algorithm to enhance the robustness against unknown and heterogeneous communication delays. Finally, a numerical example of a matching problem is provided to illustrate the results.

**Index Terms**—Distributed optimization, Passivity, Delays.

## I. INTRODUCTION

**D**ISTRIBUTED convex optimization over multi-agent systems aims to drive agents to cooperatively optimize the sum of local objective functions that are only accessible to their local agents. Ever since the pioneer work [1] that provides a control-theoretic perspective for the proportional-integral (PI) consensus-based distributed algorithms, many works have been carried out in the continuous-time scheme [2]. Recently, some problems in distributed optimization have been analyzed via passivity-based techniques [3]–[7]. Passivity-based techniques usually enjoy good scalability to large-scale networks owing to the preservation of passivity in parallel or negative feedback interconnection of passive components [8].

Distributed optimization under communication delays has been an important subject in recent years since delays can cause instability for algorithmic dynamics [3], [4], [9]. The work [9] addresses time-varying delays, but it only considers an identical delay known in advance for all communication channels and does not treat inequality constraints. The problem under unknown and heterogeneous communication delays is addressed via passivity techniques in [3], [4]. However, to ensure optimality in the presence of inequality constraints, delays are assumed to be homogeneous and an additional assumption on the graph is needed in [3], which is not always easy to verify in large scale networks. Besides, the objective function is assumed to be strictly convex in [3], [4], which does not hold for a large class of convex optimization problems.

M. Li and G. Chesi are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China (e-mail: mengmou\_li@hku.hk; chesi@eee.hku.hk).

S. Yamashita and T. Hatanaka are with School of Engineering, Tokyo Institute of Technology, 2-12-1 S5-16, Ookayama, Meguro-ku, Tokyo, Japan (e-mail: yamashita.s.ag@hfg.sc.e.titech.ac.jp; hatanaka@sc.e.titech.ac.jp). This work was in part supported by JST-Mirai Program Grant Number 18077648.

The LaSalle's invariance principle is widely used for convergence analysis of distributed algorithms. Algorithms derived from the classical Lagrange multiplier method usually adopt projected operations to guarantee the non-negativeness of the multipliers for inequality constraints. As a result, it leads to non-smooth dynamics, which is analyzed by the invariance principle for Carathéodory systems [10]. However, the discontinuous nature hinders the application of the invariance principle when delays are introduced into the systems, which results in the additional restrictive assumptions in [3], [4]. It is worth noting that a projection-free Lagrangian is adopted to solve local and couple inequalities in [11], which enables a smooth dynamics and the application of the LaSalle's invariance principle under delays. Another important issue is that the primal-dual gradient dynamics may cause oscillations when the objective function lacks strict convexity [12]. To cope with this problem, some modification methods are introduced [12], [13]. However, these methods are either restricted to affine constraints or not in a distributed structure. Recently, a phase lead compensation technique is adopted as a generalized method to ensure convergence [5].

In this work, we address unknown and heterogeneous inter-agent communication delays in distributed constrained optimization without the strictly convex assumption by combining techniques used in [5], [11] from the perspective of passivity. First, we propose a smooth continuous-time algorithm for distributed constrained optimization with general convex objective functions without delays. Then, we incorporate the scattering transformation into the proposed algorithm to enhance the robustness against unknown and heterogeneous communication delays.

## II. PRELIMINARIES

Notations: Let  $\mathbb{R}$  ( $\mathbb{R}_{\geq 0}$ ) be the set of (non-negative) real numbers.  $\text{col}(v_1, \dots, v_m) := (v_1^T, \dots, v_m^T)^T$  denotes the column vector stacked with vectors  $v_1, \dots, v_m$ .  $I_n$  denotes the  $n \times n$  identity matrix,  $\mathbf{1}_n := \text{col}(1, \dots, 1) \in \mathbb{R}^n$ , and  $\mathbf{0}$  denotes the zero matrix of proper dimension. The notation “ $\circ$ ” denotes the Hadamard product and “ $\otimes$ ” denotes the Kronecker product.  $\nabla_k f$  denotes the gradient of  $f$  along the variable  $k$ , whose subscript can be omitted if there is only one variable.

We first introduce some knowledge of convex analysis. A differentiable function  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  is *convex* over a convex set  $\mathcal{X} \subset \mathbb{R}^m$  iff  $[\nabla f(x) - \nabla f(y)]^T (x - y) \geq 0$ ,  $\forall x, y \in \mathcal{X}$ , and is *strictly convex* iff the strict inequality holds for any  $x \neq y$ . The function  $f$  is said to be *concave* if  $-f$  is convex. For a function  $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  with  $\mathcal{X} \subset \mathbb{R}^n$ ,  $\mathcal{Y} \subset \mathbb{R}^m$  being

closed and convex,  $(x^*, y^*) \in \mathcal{X} \times \mathcal{Y}$  is called a *saddle point* of  $\mathcal{L}$  if  $\mathcal{L}(x^*, y) \leq \mathcal{L}(x^*, y^*) \leq \mathcal{L}(x, y^*)$ ,  $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$ .

Next, let us present some basic concepts in graph theory. An *undirected* communication graph is represented by  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ , where  $\mathcal{N} = \{1, \dots, N\}$  is the node set of all agents,  $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$  is the edge set. The edge  $(i, j) \in \mathcal{E}$  means that agent  $i$  and  $j$  can exchange information. The adjacency matrix  $\mathcal{A} := [a_{ij}]$  satisfies  $a_{ii} = 0$ , and  $a_{ij} = a_{ji} > 0$  if  $(i, j) \in \mathcal{E}$  and  $a_{ij} = 0$ , otherwise. The graph  $\mathcal{G}$  is said to be *connected* if there exists a sequence of successive edges between any two agents. When  $\mathcal{G}$  is connected and undirected, its corresponding Laplacian matrix  $L := \text{diag}\{\mathcal{A} \cdot \mathbf{1}_N\} - \mathcal{A}$  is positive semidefinite and has zero as its simple eigenvalue associated with eigenvector  $v = \alpha \mathbf{1}_N$ ,  $\forall \alpha \in \mathbb{R}$ .

We conclude this section by giving the definition of passivity [5]. Consider a system  $\Sigma$  described by a state model with state  $x \in \mathbb{R}^m$ , input  $u \in \mathbb{R}^n$  and output  $y \in \mathbb{R}^n$ . The system  $\Sigma$  is said to be *passive* if there exists a positive semidefinite differentiable function  $S(x) : \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$  called *storage function*, such that  $\dot{S}(x) \leq y^T u$  holds for all inputs  $u(t)$ , all initial states  $x(0)$ , and all  $t \geq 0$ .

### III. PASSIVITY-BASED ALGORITHM ON CONSTRAINED DISTRIBUTED OPTIMIZATION

Let us consider a constrained distributed optimization problem in a network of  $N$  agents in the node set  $\mathcal{N} = \{1, \dots, N\}$

$$\min_z \sum_{i \in \mathcal{N}} f_i(z), \text{ s.t. } g_i(z) \leq 0, h_i(z) = 0, i \in \mathcal{N} \quad (1)$$

where  $z \in \mathbb{R}^n$  is a decision variable,  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  are local objective function, inequality constraint and affine equality constraint for the  $i$ th agent, respectively. Just for simplicity, we only consider one local inequality and equality constraint for each agent, while it is trivial to extend subsequent results to the case with multiple local constraints. Next, we adopt the following assumptions.

**Assumption 1.** *The functions  $f_i$  and  $g_i$  are convex and twice differentiable. The Slater's condition holds and there exists a finite optimal solution to problem (1).*

This assumption ensures that the problem is well-defined.  $f_i$  is only required to be convex, implying that there may exist more than one optimal solution to problem (1).

**Assumption 2.** *The communication graph  $\mathcal{G}$  is undirected and connected.*

Denote  $x = \text{col}(x_1, \dots, x_N)$ , where  $x_i \in \mathbb{R}^n$ , then problem (1) is equivalent to

$$\begin{aligned} \min_x f(x) &:= \sum_{i \in \mathcal{N}} f_i(x_i) + x^T \mathbf{L}x \\ \text{s.t. } g_i(x_i) &\leq 0, h_i(x_i) = 0, i \in \mathcal{N}, \mathbf{L}x = \mathbf{0} \end{aligned} \quad (2)$$

where  $\mathbf{L} = L \otimes I_n$ , and  $L$  is the Laplacian matrix.

#### A. Generalized Lagrange Multiplier Method

In this subsection, we briefly review the generalized Lagrange multiplier method (GLMM) in [11] for solving problem

(2). Define compact variables  $\lambda = \text{col}(\lambda_1, \dots, \lambda_N)$ ,  $\mu = \text{col}(\mu_1, \dots, \mu_N)$ ,  $\xi = \text{col}(\xi_1, \dots, \xi_N)$  with  $\lambda_i, \mu_i \in \mathbb{R}$  and  $\xi_i \in \mathbb{R}^n$ ,  $i \in \mathcal{N}$ . Adopt a Lagrangian for problem (2),

$$\begin{aligned} \mathcal{L}(x, \xi, \lambda, \mu) &:= \\ f(x) + \sum_{i \in \mathcal{N}} \lambda_i^2 g_i(x_i) + \sum_{i \in \mathcal{N}} \mu_i h_i(x_i) - \xi^T \mathbf{L}x + \frac{1}{2} x^T \mathbf{L}x \end{aligned} \quad (3)$$

where  $\mathcal{L}(x, \xi, \lambda, \mu)$  is a class of the generalized Lagrangian;  $\lambda_i^2$  is the generalized multiplier for the inequality constraint  $g_i \leq 0$ ;  $\mu_i$  is the multiplier for the equality constraint  $h_i = 0$ ;  $\xi$  is the multiplier for the consensus constraint  $\mathbf{L}x = \mathbf{0}$ . Then, by applying the primal-dual gradient flow to  $\mathcal{L}(x, \xi, \lambda, \mu)$ , we obtain the following projection-free distributed algorithm

$$\begin{aligned} \dot{x}_i &= -\nabla_{x_i} \mathcal{L} = -\nabla f_i(x_i) - \lambda_i^2 \nabla g_i(x_i) - \mu_i \nabla h_i(x_i) \\ &\quad + \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i) - \sum_{j \in \mathcal{N}_i} a_{ij}(\xi_j - \xi_i) \end{aligned} \quad (4a)$$

$$\dot{\xi}_i = \nabla_{\xi_i} \mathcal{L}, \quad \dot{\lambda}_i = \nabla_{\lambda_i} \mathcal{L}, \quad \dot{\mu}_i = \nabla_{\mu_i} \mathcal{L} \quad (4b)$$

where  $a_{ij}$  is the  $(i, j)$ -th entry of the adjacency matrix. The above algorithm is said to be projection-free since the non-negativeness of multipliers  $\lambda_i^2$  is already guaranteed without any projection operator.

Let  $(x^*, \xi^*, \lambda^*, \mu^*) \in \mathcal{H}^*$  denotes an optimal solution of the problem, where  $\mathcal{H}^*$  is the set satisfying the generalized KKT condition for problem (2) corresponding to  $\mathcal{L}(x, \xi, \lambda, \mu)$ , i.e.,

$$\mathbf{L}x^* = \mathbf{0}, \quad (5a)$$

$$h_i(x_i^*) = 0, \quad g_i(x_i^*) \leq 0, \quad \lambda_i^{*2} g_i(x_i^*) = 0, \quad (5b)$$

$$\nabla f_i(x_i^*) + \lambda_i^{*2} \nabla g_i(x_i^*) + \mu_i^* \nabla h_i(x_i^*) + \sum_{j=1}^N a_{ij}(\xi_j^* - \xi_i^*) = \mathbf{0} \quad (5c)$$

where the term  $\sum_{j \in \mathcal{N}_i} a_{ij}(x_j^* - x_i^*)$  in (5c) is omitted since (5a) implies  $x_i^* = x_j^*$ ,  $\forall i, j$ . Next, let us give the following lemma derived from [11].

**Lemma 1.** *Under Assumption 1, a fixed point  $(x^*, \xi^*, \lambda^*, \mu^*)$  solves problem (2) if and only if it satisfies condition (5).*

Then, denote  $z^* = x_i^*$ ,  $\forall i$ . Obviously,  $z^*$  is the optimal solution to problem (1).

#### B. GLMM With Phase Lead Compensation

Originally,  $f_i$  needs to be strictly convex to ensure the convergence of algorithm (4) [11]. When this restriction does not hold, i.e., when  $f_i$  lacks strict convexity, an extra modification is needed. In this subsection, we add the phase lead compensator into the dynamics (4), which serves to provide stable zeros and avoid possible oscillations for the algorithm under general convexity [5].

Define  $\nu = \text{col}(\nu_1, \dots, \nu_N)$  with  $\nu_i \in \mathbb{R}^n$  and

$$\begin{aligned} \nu_i &= -\nabla f_i(x_i) - \lambda_i^2 \nabla g_i(x_i) - \mu_i \nabla h_i(x_i) \\ &\quad + \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i) - \sum_{j \in \mathcal{N}_i} a_{ij}(\xi_j - \xi_i). \end{aligned} \quad (6)$$

We add the phase lead compensator to the integrator in the primal gradient dynamics (4a), then the dynamics for the  $i$ th agent is reformulated in the frequency domain as

$$x_i(s) = (M_i(s) \cdot I_n) \nu_i(s) \quad (7)$$



holds only if  $\rho_k^i \equiv \mathbf{0}$ ,  $\forall k = 2, \dots, m$ , which implies that  $\nu_i = \mathbf{0}$  and thus  $x_i$  is unchanged.  $-(x - x^*)^T \mathbf{L}(x - x^*) = 0$  implies that  $\mathbf{L}x = \mathbf{0}$  and  $\dot{\xi} = \mathbf{0}$ . Then  $x$  satisfies (5c).

Next, let us look at the dynamics (9e) when the states converges to  $\Omega_c$ . If  $\lambda_i^* = 0$ , then the constraint  $g_i(x_i)$  is inactive, meaning that  $g_i(x_i) \leq 0$  when the primal gradient of  $\mathcal{L}$  with respect to  $x_i$  vanishes, i.e., when (5c) holds. If  $\lambda_i^* > 0$ , recalling the definition of  $S_i^g$ , we obtain that  $\lambda_i > 0$ , for all  $t > 0$ , because if  $\lambda_i \rightarrow 0$ , then  $S_i^g \rightarrow +\infty$  due to the term  $-\frac{1}{2}\lambda_i^* \ln \lambda_i$ , which contradicts the fact that  $V$  is decreasing. Since  $\lambda_i$  is nonzero,  $g_i(x_i)$  should be zero to ensure stability and boundedness of the dynamics  $\dot{\lambda}_i = 2\lambda_i g_i$ . Similarly, it can be observed that  $h_i(x_i)$  should be zero to ensure boundedness. Therefore, the KKT condition (5) is satisfied, i.e.,  $(x, \xi, \lambda, \mu) \in \mathcal{H}^*$  and is unchanged whenever  $(x, \xi, \lambda, \mu, \rho)$  converges to  $\Omega_c$ . In conclusion, the trajectories generated by the algorithm will asymptotically converge to a constant equilibrium point that solves problem (2).  $\square$

#### IV. CONSTRAINED DISTRIBUTED OPTIMIZATION WITH HETEROGENEOUS COMMUNICATION DELAYS

Let us consider the presence of unknown and heterogeneous inter-agent communication delays in this section. For  $(i, j) \in \mathcal{E}$ , let the communication delay from agent  $i$  to  $j$  be denoted by a constant  $T_{ij}$ . In this case, each agent cannot catch the current variable of its neighboring agents. Then the algorithm under delays becomes

$$\nu_i = -\nabla f_i(x_i) - \lambda_i^2 \nabla g_i(x_i) - \mu_i \nabla h_i(x_i) + \sum_{j \in \mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) - \sum_{j \in \mathcal{N}_i} a_{ij}(r_{ij}^\xi - \xi_i) \quad (11a)$$

$$\dot{\xi}_i = \sum_{j \in \mathcal{N}_i} a_{ij}(r_{ij}^x - x_i) \quad (11b)$$

$$(9a), (9b), (9e)$$

where (11a) and (11b) are the modified update expression against (9c) and (9d) by replacing the neighbor's information with  $r_{ij}^x, r_{ij}^\xi$ , which denote the information agent  $i$  receives from agent  $j$ . The dynamics (11a), (11b) can be rewritten as

$$\begin{bmatrix} \nu_i \\ \dot{\xi}_i \end{bmatrix} = \begin{bmatrix} -\nabla f_i(x_i) - \zeta_i \\ \mathbf{0} \end{bmatrix} + p_i \quad (12)$$

where  $\zeta_i$  is defined in Lemma 3,  $p_i := \sum_{j \in \mathcal{N}_i} p_{ij}$ ,  $p_{ij} := \begin{bmatrix} p_{ij}^x \\ p_{ij}^\xi \end{bmatrix} = E_{ij} \begin{bmatrix} r_{ij}^x - x_i \\ r_{ij}^\xi - \xi_i \end{bmatrix}$  and  $E_{ij} := \begin{bmatrix} a_{ij} & -a_{ij} \\ a_{ij} & 0 \end{bmatrix} \otimes I_n$ ,  $j \in \mathcal{N}_i$ .

Define  $\bar{x}_i := x_i - z^*$ ,  $\bar{\xi}_i := \xi_i - 2\xi_i^*$ ,  $\bar{p}_i := p_i - \sum_{j \in \mathcal{N}_i} p_{ij}^* = \sum_{j \in \mathcal{N}_i} \bar{p}_{ij}$ ,  $p_{ij}^* := a_{ij} \begin{bmatrix} \xi_i^* - \xi_j^* \\ \mathbf{0} \end{bmatrix}$ ,  $r_{ij} := \begin{bmatrix} r_{ij}^x \\ r_{ij}^\xi \end{bmatrix}$ ,  $\bar{r}_{ij} := \begin{bmatrix} \bar{r}_{ij}^x \\ \bar{r}_{ij}^\xi \end{bmatrix} = r_{ij} - r_{ij}^*$ ,  $r_{ij}^* := \begin{bmatrix} z^* \\ \xi_i^* - \xi_j^* \end{bmatrix}$ . We show that the dynamics under communication delays preserves passivity-like properties.

**Lemma 4.** *Under Assumption 1, the system given by (11) has the passivity-like property  $\dot{S}_i \leq \sum_{j \in \mathcal{N}_i} \bar{r}_{ij}^T \bar{p}_{ij}$ , where  $S_i = S_i^c + S_i^g + \frac{1}{2} \|\bar{\xi}_i\|^2$ ,  $S_i^c$  and  $S_i^g$  are defined in Lemmas 2 and 3.*

*Proof.* From the former lemmas,  $S_i \geq 0$ . The time derivative of  $S_i$  satisfies

$$\begin{aligned} \dot{S}_i &\leq -\sum_{k=2}^m \frac{b_k^i}{c_k^i} \rho_k^i T \rho_k^i + \bar{x}_i^T [-\nabla f_i(x_i) - \zeta_i] \\ &\quad + \bar{x}_i^T (\zeta_i - \zeta_i^*) + [\bar{x}_i \quad \bar{\xi}_i] p_i \\ &\leq -\bar{x}_i^T (\zeta_i - \zeta_i^*) + \bar{x}_i^T (\zeta_i - \zeta_i^*) - \sum_{k=2}^m \frac{b_k^i}{c_k^i} \rho_k^i T \rho_k^i \\ &\quad + \bar{x}_i^T \sum_{j \in \mathcal{N}_i} a_{ij} (\xi_j^* - \xi_i^*) + \sum_{j \in \mathcal{N}_i} [\bar{x}_i \quad \bar{\xi}_i] p_{ij} \\ &= -\sum_{k=2}^m \frac{b_k^i}{c_k^i} \rho_k^i T \rho_k^i + \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \bar{x}_i + \bar{r}_{ij}^x - \bar{r}_{ij}^x \\ \bar{\xi}_i + \bar{r}_{ij}^\xi - \bar{r}_{ij}^\xi \end{bmatrix}^T \bar{p}_{ij} \\ &\leq \sum_{j \in \mathcal{N}_i} \bar{r}_{ij}^T \bar{p}_{ij} - \sum_{j \in \mathcal{N}_i} a_{ij} \|\bar{x} - \bar{r}_{ij}^x\|^2 \\ &\leq \sum_{j \in \mathcal{N}_i} \bar{r}_{ij}^T \bar{p}_{ij} \end{aligned} \quad (13)$$

where the first inequality follows from (10), (12) and Lemma 3.  $\square$

If state variables are exchanged, then agent  $i$  at time  $t$  receives  $r_{ij}^x = x_j(t - T_{ji})$  and  $r_{ij}^\xi = \xi_j(t - T_{ji})$  from agent  $j$  due to the existence of delays. However, this may cause instability and divergence to the dynamics [3]. Thus, we do not directly exchange original state variables here. To ensure stability under delays, a scattering transformation method is introduced [8], [14]. The scattering transformation in this work is defined as

$$s_{ij}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-p_{ij} + \eta r_{ij}), \quad s_{ij}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(p_{ij} + \eta r_{ij}) \quad (14a)$$

$$s_{ji}^{\leftarrow} = \frac{1}{\sqrt{2\eta}}(p_{ji} + \eta r_{ji}), \quad s_{ji}^{\rightarrow} = \frac{1}{\sqrt{2\eta}}(-p_{ji} + \eta r_{ji}) \quad (14b)$$

for  $(i, j) \in \mathcal{E}$ , where  $\eta > 0$ . Specifically,  $s_{ij}^{\rightarrow}$  denotes the signal that agent  $i$  sends to agent  $j$  while  $s_{ji}^{\leftarrow}$  represents the signal  $j$  receives from  $i$ . The other notations are defined similarly. Due to the delays, these signals should satisfy

$$s_{ji}^{\leftarrow}(t) = s_{ij}^{\rightarrow}(t - T_{ij}), \quad s_{ij}^{\leftarrow}(t) = s_{ji}^{\rightarrow}(t - T_{ji}). \quad (15)$$

Instead of directly exchanging  $x$  and  $\xi$ , scattering variables (14) are exchanged between agent  $i$  and  $j$  for  $(i, j) \in \mathcal{E}$ . Then the input variables  $r_{ij}^x, r_{ij}^\xi$  for each agent are computed from these scattering variables. For simplicity, we suppose that  $s_{ij}^{\leftarrow}(t) = s_{ij}^{\rightarrow}(t) = 0, \forall t < 0$ .

It has been proved that the system of the scattering transformation inherits passivity properties.

**Lemma 5** ([3]). *The system consisting of (14) and (15) is passive from  $-\begin{bmatrix} \bar{p}_{ij}^T & \bar{p}_{ji}^T \end{bmatrix}^T$  to  $\begin{bmatrix} \bar{r}_{ij}^T & \bar{r}_{ji}^T \end{bmatrix}^T$  with respect to*

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

i.e.,  $\dot{V}_{ij}(t) \leq -\bar{r}_{ij}^T \bar{p}_{ij} - \bar{r}_{ji}^T \bar{p}_{ji}$ , where  $\gamma_{ij}^* := \frac{1}{\sqrt{2\eta}} (p_{ij}^* - \eta r_{ij}^*)$ ,  $\delta_{ij}^* := \frac{1}{\sqrt{2\eta}} (p_{ij}^* + \eta r_{ij}^*)$ .

Since  $\bar{p}_{ij} = p_{ij} - p_{ij}^*$ , we can easily obtain the following lemma with  $p_{ij}^*, p_{ji}^*$  replaced by  $\mathbf{0}$  in the proof of Lemma 5.

**Lemma 6.** *The system consisting of (14) and (15) is passive from  $-\begin{bmatrix} \bar{p}_{ij}^T & \bar{p}_{ji}^T \end{bmatrix}^T$  to  $\begin{bmatrix} \bar{r}_{ij}^T & \bar{r}_{ji}^T \end{bmatrix}^T$  with respect to  $V_{ij}(t)$ , i.e.,  $\dot{V}_{ij}(t) \leq -\bar{r}_{ij}^T \bar{p}_{ij} - \bar{r}_{ji}^T \bar{p}_{ji}$ .*



denotes the matching label for Robot  $l$  and Target  $k$ . The term  $\|w_l - q_k\|$  denotes the distance from Robot  $l$  to Target  $k$ , which is regarded as a constant (given by sensing). Note that the linear programming problem (18) with continuous variables is a strictly relaxation from integer programming. The optimal solution for  $z_{lk}$  is either 1 or 0, which implies the matching status between Robot  $l$  and Target  $k$  [16]. We reformulate (18) as a consensus-based distributed optimization problem and use notation  $x_{i(lk)}$  to denote the estimation of  $z_{lk}$  from agent  $i$ .

Consider an area of  $100 \times 100[\text{m}^2]$  with  $N = 5$  Robots and  $M = 5$  Targets. The positions of Robots and Targets are shown in Figure 2(a). The communication graph is set to be a ring graph with  $a_{ij} = 4, \forall (i, j) \in \mathcal{E}$ , the other parameters are defined as  $m = 2, b_2^i = 5, c_1^i = 1, c_2^i = 10, \eta = 1$ , the initial condition is  $(x_i, \xi_i, \lambda_i, \mu_i, \rho^i)|_{t=0} = (\mathbf{0}, \mathbf{0}, 0.01, 0, \mathbf{0})$  and the stepsize is set to 0.001 in Simulink. The trajectories of  $x_i, i = 1, \dots, 5$  in algorithm (9) without communication delay converge to the optimal solution, as shown in Figure 3(a), which validates Theorem 1.

Next, we assume that there exist unknown and heterogeneous constant delays between any two neighboring agents. The communication delays can bring in instability for algorithm (9) (see, e.g., [3]). Thus, let us adopt algorithm (11) with scattering transformation (14) to enhance robustness against delays, and consider heterogeneous communication delays  $T_{ij} \in [0.2, 0.3]$  [s],  $\forall (i, j) \in \mathcal{E}$ . The trajectories of  $x_i, i = 1, \dots, 5$  converge to the optimal solution as shown in Figure 3(b), validating Theorem 2. Note that the agents' states diverge for the present delays in the absence of the scattering transformation. Therefore, the scattering transformation serves as a key technique for delay robustification. The corresponding results are shown in Figure 2(b), illustrating the matching between Robots and Targets.

## VI. CONCLUSION

We have addressed the distributed constrained optimization problem under inter-agent communication delays from the perspective of passivity. We have proposed smooth optimization dynamics incorporating scattering transformation and proved the robustness against unknown and heterogeneous delays. Future work will include the extension to coupled constraints, geometric constraints, and directed communication graphs.

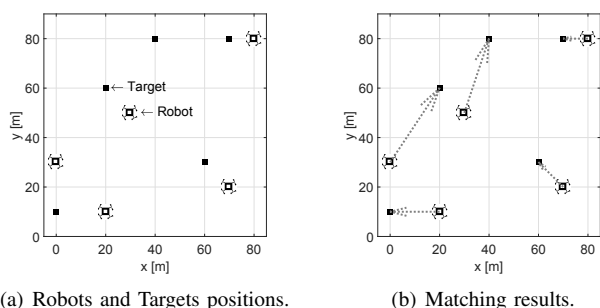
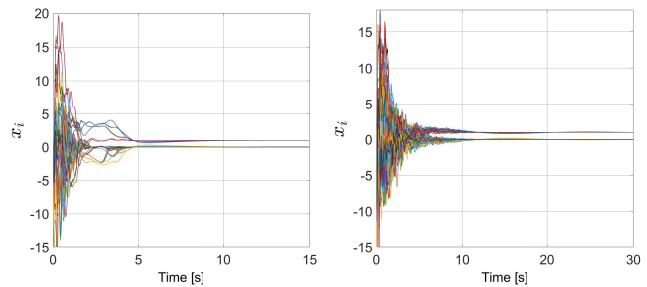


Fig. 2. Robots and Targets in the  $100 \times 100[\text{m}^2]$  area.



(a) The algorithm (9) without communication delay. (b) The algorithm (11) with scattering transformation under heterogeneous communication delays.

Fig. 3. The trajectories of  $x_i, i = 1, \dots, 5$ .

## ACKNOWLEDGMENT

The authors would like to thank the editors and reviewers for their valuable comments.

## REFERENCES

- [1] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *50th IEEE conference on decision and control and European control conference*, 2011, pp. 3800–3805.
- [2] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson, "A survey of distributed optimization," *Annual Reviews in Control*, 2019.
- [3] T. Hatanaka, N. Chopra, T. Ishizaki, and N. Li, "Passivity-based distributed optimization with communication delays using PI consensus algorithm," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4421–4428, 2018.
- [4] T. Miyano, S. Yamashita, T. Hatanaka, K. Shibata, T. Jimbo, and M. Fujita, "Continuous-time optimization dynamics mirroring ADMM architecture and passivity-based robustification against delays," *IEEE Transactions on Control of Network Systems*, 2020, early access.
- [5] S. Yamashita, T. Hatanaka, J. Yamauchi, and M. Fujita, "Passivity-based generalization of primal–dual dynamics for non-strictly convex cost functions," *Automatica*, vol. 112, p. 108712, 2020.
- [6] M. Li, G. Chesi, and Y. Hong, "Input-feedforward-passivity-based distributed optimization over jointly connected balanced digraphs," *arXiv preprint arXiv:1905.03468*, 2019.
- [7] L. Su, M. Li, V. Gupta, and G. Chesi, "Distributed resource allocation over time-varying balanced digraphs with discrete-time communication," *arXiv preprint arXiv:1907.13003*, 2019.
- [8] T. Hatanaka, N. Chopra, M. Fujita, and M. W. Spong, *Passivity-based control and estimation in networked robotics*. Springer, 2015.
- [9] S. Yang, Q. Liu, and J. Wang, "Distributed optimization based on a multiagent system in the presence of communication delays," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 5, pp. 717–728, 2016.
- [10] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of constrained primal–dual dynamics," *Systems & Control Letters*, vol. 87, pp. 10–15, 2016.
- [11] M. Li, "Generalized Lagrange multiplier method and KKT conditions with an application to distributed optimization," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 2, pp. 252–256, 2019.
- [12] T. Holding and I. Lestas, "On the emergence of oscillations in distributed resource allocation," *Automatica*, vol. 85, pp. 22–33, 2017.
- [13] 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.
- [14] R. J. Anderson and M. W. Spong, "Bilateral control of teleoperators with time delay," *IEEE Transactions on Automatic control*, vol. 34, no. 5, pp. 494–501, 1989.
- [15] H. L. Smith, *An introduction to delay differential equations with applications to the life sciences*. Springer New York, 2011, vol. 57.
- [16] T. Miyano, S. Yamashita, T. Hatanaka, K. Shibata, T. Jimbo, and M. Fujita, "Design of continuous-time ADMM and convergence analysis based on passivity," *Transactions of the Society of Instrument and Control Engineers*, vol. 55, pp. 286–293, 2019, (in Japanese).