Robustification of Continuous-Time ADMM against Communication Delays under Non-Strict Convexity: A Passivity-Based Approach

Shunya YAMASHITA *, Mengmou Li **, and Takeshi HATANAKA *

Abstract: In this paper, we address a class of distributed optimization problems with non-strictly convex cost functions in the presence of communication delays between an agent and a coordinator. To this end, we focus on a continuous-time optimization algorithm that mirrors the alternating direction method of multipliers. We first redesign the algorithm so that the dynamics ensures smoothness and a sub-block for primal optimization includes stable zeros. It is then revealed that the algorithm is composed of feedback interconnection of passive systems. We next robustify the algorithm against communication delays by applying the so-called scattering transformation. The smoothness of the dynamics allows one to use the invariance principle for delay systems, and accordingly, the state trajectories are shown to converge to an optimal solution even without the strict convexity assumption. Finally, the presented method is demonstrated via simulation of an environmental-monitoring problem.

Key Words: convex optimization, alternating direction method of multipliers, communication delays, passivity.

1. Introduction

Due to demands for real-time solutions to large-scale optimization problems, continuous-time algorithms have gained increasing interests because of their light online computational cost [1]. One of the most typical examples of such algorithms is so-called primal-dual dynamics [2], which is also known to be suitable for distributed implementation for network optimization problems [1],[3]. A variation of the primal-dual dynamics is also studied in [4], wherein the authors employ the augmented Lagrangian instead of the Lagrangian. The authors of [5],[6] present continuous-time dynamics that mirrors another (distributed) optimization algorithm, namely the alternating direction method of multiplier (ADMM).

Recent publications have pointed out that these continuous-time algorithms are strongly related to passivity [7], and some of them have studied the passivity-based perspective of the optimization dynamics. The perspective allows one to study interconnected systems [8]–[10], to improve the transient responses [11],[12], and to relax the strict convexity assumption on the cost function [12].

Distributed implementation of the above optimization algorithms relies on communication technology, which poses additional issues to be addressed like delays [13]–[15], packet losses [16], and noises [17]. In this paper, we address delays in distributed optimization algorithms. To this end, we revisit the continuous-time ADMM presented in [6], where the algorithm is shown to take the same architecture as bilateral teleoperation [7] from the viewpoint of passivity. Based on this perspective, the authors also robustify the algorithm against delays by using the so-called scattering transformation that has been extensively studied in teleoperation [7]. However, differently from the delay-free case, strict convexity is assumed for the cost function in the presence of delays in order to complete the convergence proof. The same is applied to [13]. Recently, the authors succeeded in eliminating the strict convexity assumption in the presence of delays, but the communication structure assumed therein is different from that in [6].

In this paper, we focus on constrained convex optimization investigated in [6]. We start with formulating the optimization problem in the ADMM form. We then redesign the continuous-time ADMM in [6] so that stable zeros are supplied to a sub-process in the primal optimization and the overall dynamics ensures the local Lipschitz condition. It is then proved that the algorithm is composed of feedback interconnection of passive systems. Based on this passivity paradigm, we next apply the scattering transformation [7] known to passify the communication channel including delays. The Lipschitz condition of all subprocesses renders the invariance principle for delay systems [18] available for convergence analysis. Accordingly, the algorithm is shown to ensure asymptotic optimality even without assuming a strictly convex cost function. Finally, we demonstrate the effectiveness of the proposed method via simulation of an environmental-monitoring problem.

2. Problem Formulation

2.1 Preliminaries

Let us first introduce the notion of passivity [7].

Definition 1 Consider a system $\Sigma$ described by a state space model with state $x \in \mathbb{R}^n$, 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 semi-definite function $S : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} := [0, \infty)$, called storage function, such that $S(x) \leq y^Tu$ holds for all states $x \in \mathbb{R}^n$ and all inputs $u \in \mathbb{R}^N$.

Next, we introduce convex functions as follows [19].
Definition 2 A differentiable function \( f : \mathbb{R}^n \to \mathbb{R} \) is said to be convex if \( f(x) - f(y) \leq (x - y)^\top \nabla f(x) \) holds for all \( x, y \in \mathbb{R}^n \).

From the above inequality, we immediately have the so-called monotone condition:

\[
(x - y)^\top (\nabla f(x) - \nabla f(y)) \geq 0 \quad \forall x, y \in \mathbb{R}^n.
\] (1)

Let us next introduce a projection operator. For a vector \( v \in \mathbb{R}^n \), define a projection operator on a closed convex set \( \Omega \subset \mathbb{R}^n \) as \( \mathcal{P}_\Omega(v) := \arg\min_{y \in \Omega} \|v - y\|^2 \). Regarding the operator, the following projection theorem is known to hold [20, §3.12].

Fact 1 ([20]) Let \( \Omega \subset \mathbb{R}^n \) be a closed convex set. For a given point \( y \in \mathbb{R}^n \), \( x = \mathcal{P}_\Omega(y) \) is satisfied if and only if \( x \in \Omega \) and \((y - x)^\top (x - z) \geq 0 \) holds for any \( z \in \Omega \).

In particular, if we consider the projection onto \( \mathbb{R}^n_{\geq 0} \), Fact 1 immediately yields the following inequality for any \( y \in \mathbb{R}^n \) and any \( z \in \mathbb{R}^n_{\geq 0} \):

\[
(y - \max(0,y))^\top (\max(0,y) - z) \geq 0.
\] (2)

2.2 Convex Optimization Problem and ADMM Form

In this paper, we consider the following constrained optimization problem:

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g(x) \leq 0, \quad Ax = b,
\end{align*}
\] (3)

where \( x \in \mathbb{R}^n \) is the decision variable, \( f : \mathbb{R}^n \to \mathbb{R} \) is the cost function, \( g : \mathbb{R}^n \to \mathbb{R}^m \) is the constraint function, \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \). The above inequality constraint implies the element-wise condition. The \( l \)-th element of the function \( g \) is denoted by \( g_l (l = 1, 2, \ldots, m) : \mathbb{R}^n \to \mathbb{R} \). For convenience, we define the set \( C_E := \{ x \in \mathbb{R}^n \mid Ax = b \} \). Throughout this paper, we adopt the following assumption.

Assumption 1 The functions \( f, g_l \) \((l = 1, 2, \ldots, m)\) are convex, continuously differentiable, and their gradients \( \nabla f, \nabla g_l \) are locally Lipschitz. The matrix \( A \) is of full row rank. The set \( \{ x \in \mathbb{R}^n \mid g_l(x) < 0, \quad Ax = b \} \) is nonempty. The function \( f \) has a minimum value over the feasible set.

Under Assumption 1, (3) is a convex optimization problem. Then, the problem (3) can be equivalently rewritten as the following ADMM form [21]:

\[
\begin{align*}
\text{minimize} & \quad f(x) + h(z) \\
\text{subject to} & \quad g(x) \leq 0, \quad x = z,
\end{align*}
\] (4)

where \( z \in \mathbb{R}^n \) is the auxiliary decision variable, and \( h : \mathbb{R}^n \to [0, \infty) \) is the indicator function of the set \( C_E \), defined as

\[
h(z) := \begin{cases} 
0 & \text{if } Az = b, \\
\infty & \text{otherwise},
\end{cases}
\]

Notice that the problem (4) is also a convex optimization problem under Assumption 1. Hence, \((x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}^n \) is an optimal solution of (4) if and only if there exists \((\lambda^*, \mu^*) \in \mathbb{R}_{\geq 0}^n \times \mathbb{R}^n \) satisfying the following optimality conditions [19],[21]:

\[
\nabla f(x^*) + \nabla g(x^*)\lambda^* + \mu^* = 0,
\] (5a)

\[
\lambda^* \geq 0, \quad g(x^*) \leq 0, \quad \lambda^* \circ g(x^*) = 0, \quad (5b)
\]

\[
x^* = z^*, \quad A^*b = b, \quad (I_n - A^*A)\lambda^* = 0, \quad (5c)
\]

where \( \nabla g(x) := [\nabla g_1(x) \nabla g_2(x) \cdots \nabla g_m(x)] \), the symbol \( \circ \) denotes the Hadamard product, \( I_n \in \mathbb{R}^{n \times n} \) is the identity matrix, and \( A^* := (A^\top A)^{-1} \) is the pseudo-inverse matrix of \( A \). The set of the optimal solutions \( \lambda^* \) is now defined as \( \lambda^* := \{ (x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}^n \times \mathbb{R}^n \mid (5) \text{ holds} \} \).

3. Continuous-Time ADMM with Passivity-Based Generalization

In this section, we present a continuous-time optimization algorithm inspired by ADMM. Throughout this paper, we suppose that there are two types of decision makers: one is a coordinator which manages the equality constraint with \( A \) and \( b \), and the other is an agent which has the information of \( f \) and \( g \). This formulation is widely adopted. See [6] for more details.

Similar to [6], we take the following augmented Lagrangian:

\[
L(x, \lambda, \mu, z) := f(x) + h(z) + \sum_{l=1}^{m} G_l g_l(x), \lambda_l
\]

\[
+ \mu^\top (z - x) + \frac{\rho}{2} \|z - x\|^2,
\] (6)

where \( \lambda \in \mathbb{R}_{\geq 0}^m \) and \( \mu \in \mathbb{R}^n \) are the Lagrange multipliers for constraints \( g(x) \leq 0 \) and \( x = z \), respectively. The constant \( \rho > 0 \) implies the penalty parameter for the constraint \( x = z \). The function \( G_l : \mathbb{R} \times \mathbb{R}_{\geq 0}^n \to \mathbb{R} \) is a penalty function for \( g_l(x) \leq 0 \), employed in [22],[23], given as

\[
G_l(g_l(x), \lambda_l) := \begin{cases} 
\lambda_l g_l(x) + \frac{\rho}{2} g_l(x)^2 & \text{if } g_l(x) \geq -\lambda_l, \\
-\frac{1}{2\rho} \lambda_l^2 & \text{if } g_l(x) < -\lambda_l,
\end{cases}
\]

where \( \beta > 0 \) is a free parameter. We employ \( G_l \) in (6) to make the gradient of \( L \) continuous.

To solve the optimization problem (4), we propose the optimization dynamics based on the augmented Lagrangian (6). In the same way as [6], the agent is assumed to update \( x, \lambda, \) and \( \mu \), while the coordinator is assumed to update \( z \). In particular, the agent dynamics consists of the three subprocesses:

- Primal dynamics: updating \( x \).
- Dual dynamics: updating \( \lambda \).
- Proportional-integral (PI) controller: updating \( \mu \).

Let us first design the agent dynamics. As the update of the primal variable \( x \), we employ the generalized gradient dynamics proposed in [12] formulated as

\[
x(s) = M(s)u(s),
\] (7a)

\[
u = -\nabla f(x) - \psi - y_a,
\] (7b)

where \( M(s) \) is the diagonal transfer function matrix \( M(s) = \text{diag} (M_1(s), M_2(s), \ldots, M_n(s)) \). The signals \( \psi \in \mathbb{R}^n \) and \( y_a \in \mathbb{R}^n \) will be designed below. The \( i \)-th diagonal element of \( M(s) \), denoted by \( M_i(s) \), is formulated as

\[
M_i(s) = \sum_{k=1}^{n_i} \frac{c_{ik}}{s + a_{ik}} + d_i,
\]

where \( n_i \geq 1, a_{mi} > \cdots > a_{i2} > a_{i1} = 0, \quad c_{i1} > 0 \quad (k = 1, 2, \ldots, n_i) \) and \( d_i \geq 0 \) are the design parameters. Remark that
$M_i(s)$ inevitably includes one integrator due to $a_{ii} = 0$. The $i$-th transfer function $M_i(s)$ is expressed as the following state space representation:

$$\dot{\xi}_{ik} = -a_{ii} \xi_{ik} + c_{ik} \mu_i, \quad k = 1, 2, \ldots, n_i, \quad (8a)$$

$$x_i = \mathbf{I}_{n_i} \xi_i + d_i \mu_i, \quad (8b)$$

where $\xi_{ik}$ is the $k$-th element of the state $\xi_i \in \mathbb{R}^{n_i}$ and $\mathbf{I}_{n_i}$ is the $n_i$ dimensional all-ones vector.

We next design the dual dynamics as follows:

$$\lambda = \alpha \lambda_{\text{max}}(-\lambda, \beta g(x)), \quad \lambda(0) \in \mathbb{R}^{m_n}, \quad (9a)$$

$$\psi = \nabla g(x) (\lambda + \alpha \beta g(x)), \quad (9b)$$

where $\alpha > 0$ is a scalar parameter. The right hand side of (9a) guarantees $\lambda \in \mathbb{R}^{m_n}$ for all time. Although (9) includes a logic $\max(-\lambda, \beta g(x))$, the right hand sides of (9) satisfy locally Lipschitz continuity under Assumption 1 [23].

In the same way as [6], the PI controller is designed as

$$\mu = \alpha \mu_{\text{max}}(x - v_a), \quad (10a)$$

$$y_a = \mu + \rho(x - v_a), \quad (10b)$$

where $\alpha > 0$ is a scalar parameter. The right hand side of (10a) coincides with $\nabla_x L$. The signals $v_a \in \mathbb{R}^n$ and $y_a \in \mathbb{R}^n$ are the input and output of the agent to be exchanged with the coordinator.

By summarizing the above, the agent dynamics is composed of (7), (9), and (10). In the sequel, we fix $(x^*, \lambda, \mu, z^*) \in \mathcal{X}$, and take the notations $\bar{x} := x - x^*, \bar{\psi} := \psi - \nabla g(x^*) \lambda$, $\bar{y}_a := y_a - \mu$, and $\bar{y} := y_a - x^*$.

**Lemma 1** ([12]) Under Assumption 1, the system (7) is passive from $-\bar{y} = \bar{y}_a$ to $\bar{x}$ for the storage function $S := \sum_{i=1}^{n_i} S_i$, $S_i := \frac{1}{2\alpha_i} (\xi_{i1}^2 + \sum_{k=2}^{n_i} \frac{1}{\gamma_{ik}} \xi_{ik}^2)$, where $\xi_{ik}^2$ is the $i$-th element of $x^*$.

**Lemma 2** Under Assumption 1, the system (9) is passive from $\bar{x}$ to $\bar{\psi}$ for the storage function $H := \frac{1}{2\alpha} \| \lambda - \lambda^* \|^2$.

**Lemma 3** Under Assumption 1, the system (10) is passive from $\bar{x} - \bar{v}_a$ to $\bar{y}_a$ for the storage function $U := \frac{1}{2\alpha} \| \mu - \mu^* \|^2$.

**Proof.** See Appendix A. \hfill $\square$

Lemmas 1–3 imply that the agent’s dynamics given by (7), (9), and (10) is regarded as a feedback interconnection of passive systems. Accordingly, we can immediately confirm that the agent’s dynamics is passive from $-\bar{v}_a$ to $\bar{y}_a$.

The coordinator takes the role of updating $z$ so that it is projected onto $\mathcal{C}_E$. To this end, we employ the dynamics [6]:

$$\dot{z} = \alpha_c (-A^* A z + A^* b + (I_n - A^* A) v_c), \quad (11a)$$

$$y_c = (I_n - A^* A) z + A^* b, \quad (11b)$$

where $v_c \in \mathbb{R}^n$ is the input signal, $y_c \in \mathbb{R}^n$ is the output signal, and $\alpha_c > 0$ is a scalar parameter. As shown in [6], $y_c \in \mathcal{C}_E$ is always satisfied.

Let us now take the notations $\bar{v}_c := v_c - z^*$ and $\bar{y}_c := y_c - \mu^*$. Similar to [6], we obtain the lemma below.

![Block diagram of the whole optimization dynamics without communication delay](image)

**Fig. 1** Block diagram of the whole optimization dynamics without communication delay. The operator $\psi^\dagger$ implies $\psi^\dagger(\cdot) = \max(-\lambda, \beta g(\cdot))$.

**Lemma 4** ([6]) Under Assumption 1, the system (11) is passive from $\bar{v}_c$ to $\bar{y}_c$ for the storage function $W := \frac{1}{2\alpha} \| z - z^* \|^2$.

**Proof.** See Appendix A.

If the communication between the agent and the coordinator obeys $v_c = y_a$ and $v_d = y_d$, the dynamics composed of (7) and (9)–(11) can be regarded as a feedback interconnection of passive systems, as illustrated in Fig. 1: The light gray and dark gray part imply the agent and the coordinator, respectively. Then, we obtain the following convergence result.

**Theorem 1** Consider the system (7) and (9)–(11) with $v_c = y_a$ and $v_d = y_d$. If Assumption 1 holds, then $(x, \lambda, \mu, z)$ approaches to one of the constants included in $\mathcal{X}$ as $t \to \infty$.

Since the dynamics composed of (7) and (9)–(11) satisfies a locally Lipschitz condition, this theorem can be proved by using LaSalle’s invariance principle [24]. Except for the slight difference in the invariance principle, the theorem is proved in the same way as [6, Theorem 1].

**Remark 1** In view of the similarity to [6, Section III], the result of the asymptotic convergence in Theorem 1 cannot be a main contribution of this paper. However, the present algorithm itself contains two novel contributions. First, the primal dynamics is generalized to (7), which provides additional design flexibility that may be useful for improving the transient responses as exemplified in [12]. The second one is the smoothness of the dual dynamics (9), which will play a key role in the convergence proof in the next section.

### 4. Continuous-Time ADMM under Delays

In this section, we consider the situation that the communication between the agent and coordinator suffers from unknown constant delays. We assume that they exchange $n$ real values as communication messages. The delay that the agent’s $i$-th message to the coordinator suffers from is denoted by $\tau_{a_{ii}}$, and that from the coordinator to the agent is denoted by $\tau_{c_{ii}}$. We now denote $\tau_{ac} := [\tau_{ac_1}, \tau_{ac_2}, \ldots, \tau_{ac_n}] \in \mathbb{R}^{n_n}$, and $\tau_{ca} := [\tau_{ca_1}, \tau_{ca_2}, \ldots, \tau_{ca_n}] \in \mathbb{R}^{n_n}$. In the sequel, for a signal $\zeta \in \mathbb{R}^n$, time $t \geq 0$, and a delay vector $\tau_D = [\tau_{D1}, \tau_{D2}, \ldots, \tau_{Dn}] \in \mathbb{R}^{n_n}$, we denote $\zeta(t - \tau_D) := [\zeta(t - \tau_{D1}), \zeta(t - \tau_{D2}), \ldots, \zeta(t - \tau_{Dn})]^T$.

Inspired by the architectural analogy between the continuous-time ADMM and bilateral teleoperation, the
Lemma 5 ([6]) The system given by (12) and (13) is passive from $[\tilde{y}_a \ - \ \tilde{y}_c]$ to $[\tilde{v}_a \ \ \tilde{v}_c]$ for the following storage function:

$$T := \frac{1}{2} \int_{t-\tau_a}^{t} \left( s_{\text{ca}}(t) - \frac{1}{\sqrt{2}} (x^* + \mu^*) \right)^2 dt + \frac{1}{2} \int_{t-\tau_c}^{t} \left( s_{\text{va}}(t) - \frac{1}{\sqrt{2}} (x^* - \mu^*) \right)^2 dt.$$  

Proof. The function $T$ is clearly positive semi-definite. From (12) and (13),

$$\dot{T} = \frac{1}{4} \left( \left\| \dot{\tilde{y}}_a + \tilde{y}_c \right\|^2 - \left\| \tilde{v}_c + \tilde{y}_c \right\|^2 + \left\| - \tilde{v}_c + \tilde{y}_c \right\|^2 - \left\| \tilde{v}_a - \tilde{y}_a \right\|^2 \right)$$

$$= \tilde{y}_a^\top \dot{\tilde{v}}_a - \tilde{y}_a^\top \tilde{v}_a = \tilde{v}_a^\top \tilde{v}_a^\top - \tilde{y}_a^\top \tilde{y}_a^\top$$  

holds. This completes the proof. $\square$

Lemmas 1–5 imply that the dynamics (7), (9)–(11) with (12) and (13) can be regarded as a passivity-preserving interconnection of passive systems, as illustrated in Fig. 3. As a result, we have the following lemma.

Lemma 6 Define the function $V := S + H + U + W + T$, where $S$, $H$, $U$, $W$, $T$ are the storage functions defined in Lemmas 1–5, respectively. If Assumption 1 is satisfied, $\dot{V} \leq 0$ holds under (7), (9)–(11) with (12) and (13).

Proof. From (A.1), (A.3)–(A.5), and (14), we obtain

$$\dot{V} = \dot{S} + \dot{H} + \dot{U} + \dot{W} + \dot{T} \leq -\sum_{i=1}^{n} \left( d_i u_i^2 + \sum_{k=2}^{n} \frac{a_{ik}}{\xi_k} \tilde{y}_k \right) - \frac{1}{\beta} \left\| \text{max}(-A, \beta g(x)) \right\|^2$$

$$- \rho \left\| x - y_a \right\|^2 - (AZ - b)^\top (A^\top Y)^{-1} (AZ - b) \leq 0.$$  

This completes the proof. $\square$

We are now ready for convergence analysis of the proposed dynamics, composed of (7) and (9)–(13). Due to the communication delays, the proposed method is categorized as a time delay system. Then, by using the invariance principle for delay systems [18, Theorem 4], we can prove the main result of this paper shown below.

Theorem 2 Consider the system (7), (9)–(11), and the communication block (12) and (13). Assume that for all $i$, $M_i(x)$ has one or more stable zeros. If Assumption 1 holds, then $x, \lambda, \mu, \varsigma$ approaches to one of the constants included in $X^*$ as $t \to \infty$.

Proof. Since $S$, $H$, $U$ and $W$ are radially unbounded, the result in (15) implies that $\xi_a, \lambda, \mu,$ and $\varsigma$ belong to class $\mathcal{L}_\infty$. From (8b) and (11b), $x$ and $y_a$ also belong to class $\mathcal{L}_\infty$. Combining (10b), (12), and (13) by eliminating $y_a$ and $v_c$, we obtain

$$v_a(t) = \frac{\rho - 1}{\rho + 1} v_a(t - \tau_a) + \frac{\rho}{\rho + 1} (x(t) - x(t - \tau))$$

$$+ \frac{1}{\rho + 1} (\mu(t) - \mu(t - \tau)) + \frac{2}{\rho + 1} y_a(t - \tau_a),$$  

where $\tau := \tau_a + \tau_c$. Since $\frac{\rho - 1}{\rho + 1} < 1$ is satisfied from $\rho > 0$, this difference equation is stable with bounded inputs. Accordingly, we obtain $v_a \in \mathcal{L}_\infty$, and $y_a \in \mathcal{L}_\infty$ from (10b). Besides, from (12) and (13), we have

$$v_c(t) = -y_c(t) + v_a(t - \tau_a) + y_a(t - \tau_a).$$
which implies \( v_c \in \mathcal{L}_c \).

From the above discussion, all of the signals are bounded and then the invariance principle for delay systems [18] can be applied. From (15), \( V \equiv 0 \) implies that

\[
du_i^2 + \sum_{k=2}^{n_i} \frac{\alpha_k}{c_k} \xi_{ik}^2 = 0 \quad \forall i, 1, 2, \ldots, n, \tag{18}
\]

and

\[
\max[-\lambda \beta g(x)] \equiv 0, \tag{19}
\]

\[
x - v_a \equiv 0, \tag{20}
\]

\[
(Az - b)^2(AA^T)^{-1}(Az - b) \equiv 0. \tag{21}
\]

First, we consider the system trajectories satisfying (18) and the dynamics (8). Since \( M_i(s) \) has at least one stable zero, \( d_i > 0 \) or \( n_i \geq 2 \) must hold. If \( d_i > 0 \), then (18) implies \( u_i \equiv 0 \). If \( n_i \geq 2 \) holds, we have \( \xi_{ik} \equiv 0 \) and \( \xi_{ik} \equiv 0 \) for all \( k = 2, 3, \ldots, n_i \), which also implies \( u_i \equiv 0 \). Hence, \( u_i \equiv 0 \) is satisfied for all \( i \), and then the signals must satisfy

\[
\nabla f(x) + \psi + y_a \equiv 0. \tag{22}
\]

As a result, we also see from (8) that \( \xi_{ik} \) and \( x_i \) must be constant for all \( i \).

Next, we focus on (19) and the system (9). We have immediately from (19) that \( \lambda \equiv 0 \) and \( \lambda \) is constant. From (9b),

\[
\psi \equiv \nabla g(x) \lambda \tag{23}
\]

holds. In addition, (19) implies that

\[
\lambda \geq 0, \quad g(x) \leq 0, \quad \lambda \circ g(x) = 0 \tag{24}
\]

must be identically satisfied.

Let us next consider (20). From the dynamics (10a), we immediately see that \( \mu \) must be constant. Hence,

\[
y_a \equiv \mu \tag{25}
\]

must hold and this signal is constant. Consequently, \( v_c \) must be constant when \( x \) is constant. Thus, by invoking the invariance principle [18], \( \lim_{t \rightarrow -\infty} x(t) - x(t - \tau) = 0 \), \( \lim_{t \rightarrow -\infty} \mu(t) - \mu(t - \tau) = 0 \) and \( \lim_{t \rightarrow -\infty} (v_c(t) - v_c(t - \tau)) = 0 \) hold. Then, using (16), we obtain \( \lim_{t \rightarrow -\infty} (v_c(t) - y_c(t - \tau)) = 0 \). This implies that, when \( \dot{V} \equiv 0 \),

\[
y_c \equiv x \tag{26}
\]

holds, and \( y_c \) must be constant. Thus, \( \lim_{t \rightarrow -\infty} (v_c(t) - \tau_{ac} - y_c(t)) = 0 \) is also satisfied. Accordingly, \( \lim_{t \rightarrow -\infty} (v_c(t) - y_c(t - \tau_{ac})) = 0 \) holds due to (16), which implies

\[
v_c \equiv \mu, \tag{27}
\]

and \( v_c \) is constant under \( \dot{V} \equiv 0 \).

Now we consider the system trajectories satisfying (21) and focus on the system (11). Since \( AA^T \) is a regular matrix,

\[
Az \equiv b \tag{28}
\]

holds from (21). Substituting this into (11b), we have

\[
y_c \equiv z, \tag{29}
\]

and hence \( z \) must be constant. Thus, \( \dot{z} \equiv 0 \) holds and we obtain

\[
(\mathbf{A}_s - \mathbf{A}) v_c \equiv 0. \tag{30}
\]

In summary, the invariance principle [18] proves that all of the state signals asymptotically converge to the constants satisfying \( (x, A, \mu, z) \in X^* \) since (22)–(30) imply the optimality condition (5). This completes the proof. \(\square\)

It should be emphasized that the proof is completed by three key factors in the present algorithm. The first one is the smoothness of the dual dynamics (9). It allows one to apply the invariance principle for delay systems, differently from [6], which results in a stronger result than [6]. The second key is the ADMM-like architecture including the penalty parameter \( \rho \), which stabilizes the delayed signals as shown in (16) and (17). The last one is the stable zero that can be added owing to the generalization (7). This provides the negative energies in (A.1) even without the strict convexity assumption.

5. Application: Environmental-Monitoring Problem

Finally, we apply the proposed optimization algorithm to an environmental-monitoring problem by multiple robots.

Suppose that five robots need to monitor five static targets on a 2-D plane. The issue to be considered here is how to determine the matching between the robots and targets. Although this matching problem is originally given as an integer programming, we can formulate this problem as the following linear programming problem\(^3\) [15], [26]:

\[
\begin{align*}
\text{minimize} & \quad y \sum_{i=1}^{5} \sum_{j=1}^{5} \|p_i - q_j\| & \theta_{ij} \\
\text{subject to} & \quad \sum_{j=1}^{5} \theta_{ij} = 1 & \forall j = 1, 2, \ldots, 5, \\
& \quad \sum_{i=1}^{5} \theta_{ij} = 1 & \forall i = 1, 2, \ldots, 4,
\end{align*}
\]

where \( p_i \in \mathbb{R}^2 (i = 1, 2, \ldots, 5) \) is the current position of robot \( i \), \( q_j \in \mathbb{R}^2 (1, 2, \ldots, 5) \) is the position of target \( j \), \( \theta_{ij} \) is the decision variable whose optimal value is set as \( \theta_{ij} = 1 \) when the robot \( i \) monitors the target \( j \) and otherwise \( \theta_{ij} = 0 \), and \( y \) is a positive scalar. The objective of (31) is to minimize the sum of the distance to accomplish the matching. Setting \( x \in \mathbb{R}^{25} \) as \( x_{(5i-1)+j} = \theta_{ij} \), the problem (31) is reduced to the problem (3) satisfying Assumption 1. Note that the cost function is non-strictly convex. In this simulation, we consider an area of 600 m \( \times \) 600 m. The positions of the robots and the targets are illustrated in Fig. 4 (a). Then, the optimal solution to (31) is given as Fig. 4 (b).

Let us test the following four cases:

(C1) Delay: 0, Algorithm: Section 3.

(C2) Delay: \( \tau_{ac} \) and \( \tau_{ac} \), Algorithm: Section 3.

(C3) Delay: \( \tau_{ac} \) and \( \tau_{ac} \), Algorithm: [6, Section IV].

(C4) Delay: \( \tau_{ac} \) and \( \tau_{ac} \), Algorithm: Section 4.

For (C1), (C2), and (C4), we set \( M_i(s) = \frac{100}{s + 300} \forall i, \alpha_1 = 10, \beta = 10, \alpha_2 = 10, \rho = 2, \alpha_3 = 10 \) and \( \gamma = 0.01 \). The design parameters for (C3) are tuned so that the performance in the

\(^3\) The reformulation is available thanks to the total unimodularity of the equality constraint matrix [26, Theorem 19.3].

\(^4\) The constraint (31c) does not include the case of \( i = 5 \) to satisfy the full row rank condition in Assumption 1. However, the feasible set given by (31c) is the same as the one with the case of \( i = 5 \) due to the property of this matching problem.
Theorem 2 is validated. the dynamics and see that our proposed method presented in Section 4 stabilizes absence of strict convexity. On the other hand, from Fig. 5 (d), we communication channel through the scattering transformation, trajectories are illustrated in Fig. 5 (c). Despite passivation of the conditioner delays are set as $\tau_{aci} = 0.03 \text{ s}$ $\forall i$ and $\tau_{aci} = 0.065 \text{ s}$ $\forall i$, respectively.

The trajectories of all elements of $x$ for (C1)–(C4) are respectively illustrated in Fig. 5 (a)–(d). From Fig. 5 (a), we see that the trajectories generated in (C1) converge to the optimal solution of (31) under non-delay communication as shown in Theorem 1. We see from Fig. 5 (b) that communication delays destabilize the dynamics. We next apply (C3), and the trajectories are illustrated in Fig. 5 (c). Despite passivation of the communication channel through the scattering transformation, the trajectories are oscillatory and do not converge due to the lack of strict convexity. On the other hand, from Fig. 5 (d), we see that our proposed method presented in Section 4 stabilizes the dynamics and $x$ converges to the optimal solution. Hence, Theorem 2 is validated.

6. Conclusion

In this paper, we addressed constrained convex optimization under communication delays. We redesigned the continuous-time ADMM presented in [6] and analyzed the passivity of the redesigned subsystems. We also made the algorithm robust against communication delays by applying the scattering transformation. Then, we proved asymptotic optimality without strict convex assumption by adopting the invariance principle for delay systems. Finally, the proposed algorithm was applied to an environmental-monitoring problem.

Acknowledgments

This work was supported in part by JST-Mirai Program Grant Number 18077648.

References

[1] 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, Vol. 47, pp. 278–305, 2019.
[2] A. Cherukuri, E. Mallada, and J. Cortés: Asymptotic convergence of constrained primal-dual dynamics, Systems & Control Letters, Vol. 87, pp. 10–15, 2016.
[3] J. Wang and N. Elia: A control perspective for centralized and distributed convex optimization, Proceedings of the 50th IEEE Conference on Decision and Control, pp. 3800–3805, 2011.
[4] T. Stegink, C.D. Persis, and A. van der Schaft: A unifying energy-based approach to stability of power grids with market dynamics, IEEE Transactions on Automatic Control, Vol. 62, No. 6, pp. 2612–2622, 2017.
[5] G. França, D.P. Robinson, and R. Vidal: ADMM and accelerated ADMM as continuous dynamical systems, Proceedings of the 35th International Conference on Machine Learning, pp. 1559–1567, 2018.
[6] 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, Vol. 7, No. 3, pp. 1296–1307, 2020.
[7] T. Hatanaka, N. Chopra, M. Fujita, and M.W. Spong: Passivity-Based Control and Estimation in Networked Robotics, Springer-Verlag, 2015.
[8] T. Hatanaka, X. Zhang, W. Shi, M. Zhu, and N. Li: Physics-integrated hierarchical/distributed HVAC optimization for multiple buildings with robustness against time delays, Proceedings of 56th IEEE Conference on Decision and Control, pp. 6573–6579, 2017.
[9] K. Tsumura, S. Baros, K. Okano, and A.M. Annaswamy: Design and stability of optimal frequency control in power networks: A passivity-based approach, Proceedings of 2018 European Control Conference, pp. 2581–2586, 2018.
[10] K. Yoshida, M. Inoue, and T. Hatanaka: Instant MPC for linear systems and dissipativity-based stability analysis, IEEE Control Systems Letters, Vol. 3, No. 4, pp. 811–816, 2019.
[11] J.T. Wen and M. Arcak: A unifying passivity framework for network flow control, IEEE Transactions on Automatic Control, Vol. 49, No. 2, pp. 162–174, 2004.
[12] 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, 108712, 2020.
[13] K. Muto and T. Namierikawa: Passivity-short-based stability analysis on electricity market trading with communication delay, Proceedings of 2019 12th Asian Control Conference, pp. 25–30, 2019.
[14] 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.
[15] M. Li, S. Yamashita, T. Hatanaka, and G. Chesi: Smooth dynamics for distributed constrained optimization with heterogeneous delays, IEEE Control Systems Letters, Vol. 4, No. 3, pp. 626–631, 2020.
[16] O. Ratanapinunchai, H. Ishii, and S. Hara: Feedback control through networks with packet loss: Mixed H2/H∞ approach and application to a teleoperating system, Proceedings of 2007 IEEE International Conference on Control Applications, pp. 36–41, 2007.
[17] L.T.H. Nguyen, T. Wada, I. Masubuchi, T. Asai, and Y. Fujisaki: Probabilistic consensus over directed two-layered networks with communication noise, Transactions of the Institute of Systems, Control and Information Engineers, Vol. 28,
Appendix Proof of Lemmas

We firstly prove Lemma 1. Using (8), we obtain $\dot{S} = \bar{x}^T \dot{u} - \sum_{i=1}^{n} d_{ai}u_{ai}^2 + \sum_{k=1}^{m} a_{ik}^2 \xi_{ik}^2$. From (7b) and (5a), $\dot{u} = -\nabla f(x) + \nabla f(x') - \bar{y} - \bar{y}_x, \quad \nabla f(x) + \nabla f(x') \geq 0$ from (1). Thus, 

$$\dot{S} \leq \bar{x}^T \left( -\bar{y} - \bar{y}_x \right) - \sum_{k=1}^{n} d_{ai}u_{ai}^2 + \sum_{k=1}^{m} a_{ik}^2 \xi_{ik}^2 \leq \bar{x}^T \left( -\bar{y} - \bar{y}_x \right)$$

(A.1)

is satisfied. This completes the proof of Lemma 1.

Next, we prove Lemma 2. Time derivative of $H$ along (9) satisfies $\dot{H} = \frac{1}{2} (\lambda - \lambda^*)^T \max(\nabla \beta(x))$. Noticing $\lambda^* \in \mathbb{R}_{\geq 0}$ and (2), we have 

$$0 \leq (\lambda + \beta(x) - \max(0, \lambda + \beta(x)) - \lambda^*)^T \max(\nabla \beta(x)) = (\lambda + \beta(x) - \max(\nabla \beta(x)) - \lambda^*)^T (\lambda + \beta(x) + \lambda^*) + \beta(x)^T \max(\nabla \beta(x)) + \lambda - \lambda^* \leq \max(\nabla \beta(x))$$

Hence, 

$$\dot{H} \leq g(x)^T \max(\nabla \beta(x)) + \lambda - \lambda^* = \frac{1}{2} \max(\nabla \beta(x))$$

(A.2)

holds. Due to the convexity of $g$, $g(x) - g(x') \leq (\nabla g(x))^T \bar{x}$ and $g(x') - g(x) \leq - (\nabla g(x'))^T \bar{x}$ are satisfied. Now, because $\max(\nabla \beta(x)) + \lambda \geq 0$ and $\lambda^* \geq 0$, the first term of the right hand side of (A.2) satisfies 

$$g(x)^T \max(\nabla \beta(x)) + \lambda - g(x)^T \lambda^* \leq \bar{x}^T \nabla g(x) \max(\nabla \beta(x)) + \lambda - \bar{x}^T \nabla g(x') \lambda^* + g(x')^T \max(\nabla \beta(x)) + \lambda - g(x')^T \lambda^* = \bar{x}^T \bar{y} + g(x')^T \max(\nabla \beta(x)) + \lambda - g(x')^T \lambda^*.$$ 

Applying this inequality with $g(x') \leq 0$ and $g(x')^T \lambda^* = 0$ to (A.2), we have 

$$\dot{H} \leq \bar{x}^T \bar{y} - \frac{1}{\beta} \max(\lambda - \beta g(x))^2 \leq \bar{x}^T \bar{y}.$$ 

(A.3)

This completes the proof of Lemma 2.

We now prove Lemma 3. From (10), we have 

$$\dot{U} = (y_a - \mu^* - \rho(x - v_a))^T (x - v_a) = \tilde{y}_a^T (\bar{x} - \bar{y}_a) - \rho \| x - v_a \|^2 \leq \tilde{y}_a^T (\bar{x} - \bar{y}_a).$$

(A.4)

This completes the proof of Lemma 3.

Finally, we prove Lemma 4. From (11), we obtain 

$$\dot{W} = (z - \bar{z})^T (\bar{A} - A^T) \bar{v} + (z - \bar{z})^T (A - b)(A - b) \bar{v}. \quad \leq \tilde{y}_a^T \bar{v} - (A - b)^T (A - b) \bar{v} \leq \tilde{y}_a^T \bar{v}. \quad (A.5)$$

This completes the proof of Lemma 4.

Shunya Yamashita (Student Member)

He received the B.Eng. and M.Eng. degrees from Tokyo Institute of Technology in 2017 and 2019, respectively. He is currently a Ph.D. student at the Department of Systems and Control Engineering, Tokyo Institute of Technology. His research interests include distributed optimization and passivity-based control. He is a student member of IEEE CSS and JSME.

Mengmou Li

He received his B.Sc. degree in physics from Zhejiang University, Zhejiang, China, in 2016. He is currently a Ph.D. candidate in Electrical and Electronic Engineering at the University of Hong Kong. His research interests include distributed optimization, consensus, and synchronization.

Takeshi Hatanaka (Member)

He received the Ph.D. degree in applied mathematics and physics from Kyoto University in 2007. He then joined Tokyo Institute of Technology, where he held positions as an assistant and associate professor. He spent an eight-month sabbatical at Harvard University in 2016 and two-years faculty exchange at Osaka University in 2018-2019. Since April 2020, he is an associate professor at Tokyo Institute of Technology. He is the coauthor of “Passivity-Based Control and Estimation in Networked Robotics” (Springer, 2015). His research interests include cyber-physical & human systems and networked robotics. He received the Kimura Award (2017), Pioneer Award (2014), Outstanding Book Award (2016), Control Division Conference Award (2018), and Outstanding Paper Awards (2009, 2015) all from SICE. He is serving/served as an AE for IEEE TSCT and SICE JCMSI, and is a member of the Conference Editorial Board of IEEE CSS. He is a senior member of IEEE.