Distributed Resource Allocation over Time-varying Balanced Digraphs with Discrete-time Communication

Lanlan Su, Mengmou Li, Vijay Gupta, and Graziano Chesi

Abstract—**This work is concerned with the problem of distributed resource allocation in continuous-time setting but with discrete-time communication over infinitely jointly connected and balanced digraphs. We provide a passivity-based perspective for the continuous-time algorithm, based on which an intermittent communication scheme is developed. Particularly, a periodic communication scheme is first derived through analyzing the passivity degradation over output sampling of the distributed dynamics at each node. Then, an asynchronous distributed event-triggered scheme is further developed. The sampled-based event-triggered communication scheme is exempt from Zeno behavior as the minimum inter-event time is lower bounded by the sampling period. The parameters in the proposed algorithm rely only on local information of each individual node, which can be designed in a truly distributed fashion.**

Index Terms—**Resource Allocation, Input Feed-forward Passive, Time-varying Balanced Graphs, Sampling, Event Triggering**

I. INTRODUCTION

An important distributed optimization problem is one in which each node has access to a convex local cost function, and all the nodes collectively seek to minimize the sum of all the local cost functions [1]–[4]. Most optimization algorithms reported in the literature are implemented in discrete time. However, as pointed out by [5], [6], discrete-time algorithms might be insufficient for applications where the optimization algorithm is not run digitally, but rather via the dynamics of a physical system, such as collectively optimizing social, biological and natural systems, robotic systems [7]. The resource allocation, as an important class of distributed optimization problems, has been studied in continuous-time setting [8]–[13] and discrete-time setting [14]–[16]. The existing works concerned with the distributed resource allocation problem assume topology graphs to be fixed over time and/or do not take the communication cost into account. In this work, we aim at providing a passivity-based perspective for a continuous-time algorithm of distributed resource allocation over time-varying digraphs, based on which an intermittent communication scheme is developed.

L. Su is with School of Engineering, University of Leicester, Leicester, LE1 7RH, UK (ls499@leicester.ac.uk); M. Li is with Department of Engineering, University of Cambridge, CB2 1TN, UK (mmli.research@gmail.com); V. Gupta is with Department of Electrical Engineering, University of Notre Dame, IN 46556 USA (vgupta2@nd.edu); G. Chesi is with Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong (chesi@eee.hku.hk).

Partial support from the National Science Foundation under Grant No. ECCS-2020246, ARO under Grant No. W911NF1910483, and AFOSR under Grant No.FA9550-21-1-0231 are gratefully acknowledged.

Passivity serves as a useful tool to analyze multi-agent systems (MASs). This has been well illustrated by [17] which shows that a MAS of possibly heterogeneous agents can reach consensus over a time-varying balanced strongly connected graph as long as all individual agents are input-output passive. In [18], we generalize the results of [17] to MASs with all agents that can be characterized by a passivity index. The current work is rooted in the same idea, but we would like to note that it is not trivial to apply the idea of consensus for MASs to distributed optimization problems. One of the main challenges is to verify that the individual algorithmic dynamics of a particular distributed algorithm can be characterized by an input feed-forward passivity (IFP) index. On the other hand, the passivity and dissipativity (including IFP as a special case) have been recently exploited in networked control systems coping with different communication imperfections. For instance, [19] addresses the problem of output synchronization of passive systems with event-triggered communication wherein network delay and quantization are considered as well; in [20], passivity index has been used to control two-dimensional systems over digital communication network wherein output sampling and event-triggered scheme are designed; [21] uses a passivity framework to model and mitigate attacks on networked control system; [22] considers the packet drops of the communication channel. See [23] for more recent works on cyber-physical systems using passivity indices. These motivate us to provide a passivity-based perspective for the algorithm as the IFP framework opens up the new possibilities of implementing the algorithm over an imperfect digital communication network and reducing the channel usage.

Whereas there exist considerable works on designing algorithms for distributed optimization problems, there are relatively fewer works taking the communication cost into account. In this work, we are interested in intermittent communication including periodic (also called sampled-based) and event-triggered communication schemes. The existing distributed algorithms proposed in this direction is either discrete-time [14], [15], [24]–[27] (see Tables 1 and 2 in [6] for a comprehensive list) or continuous-time [12], [28]–[35] (see Table 3 in [6] for a comprehensive list). Among the works focusing on continuous-time algorithms with discrete-time communication scheme, the results in [12], [28], [30], [31], [33], [35] are limited to undirected and fixed topological graphs while [29], [32], [36] assume the graph to be strongly connected and fixed over time. [34] studies the problem of event-triggered distributed optimization over a uniformly jointly connected graph but restricts it to be undirected. To the best of our knowledge, distributed optimization problem over time-varying jointly connected digraphs has never been addressed in the continuous-time setting due to the difficulties of stability analysis for time-varying systems and lack of connectedness of topology. The passivity-based method has been shown to be powerful in handling communication imperfections and in distributed control, thus being a promising approach to treat both the time-varying graph topology and communication.

In this work, we consider the problem of distributed resource allocation over a time-varying digraph under intermittent communication. Specifically, each node has access to its own local cost function and local network resource, and the goal is to minimize the sum of the local cost functions subject to a global network resource constraint. The communication topology is described by a weight-balanced and infinitely jointly strongly connected digraph. Closest papers which have also exploited the notion of passivity to address the distributed optimization problem are [37], [38]. The results in these mentioned works are limited to fixed and undirected graphs. Our work features a novel passivity-based perspective for continuous-time algorithms, which enables us to design an intermittent communication scheme over infinitely jointly strongly connected digraphs. Starting from a continuous-time algorithm, a periodic communication scheme is first derived through analyzing the passivity degradation over output sampling of the distributed dynamics at each node. Then, an asynchronous distributed event-triggered scheme is further developed. The sampled-based event-triggered communication scheme is exempt from Zeno behavior as the minimum inter-event time is lower bounded by the sampling period. The parameters in the proposed algorithm rely only on local information of each individual node, which can be designed in a truly distributed fashion.

The rest of this paper is organized as follows. Section II introduces some preliminaries and states the problem formulation. Section III presents the main results. Specifically, Section III-A reformulates the problem into its dual distributed convex optimization problem. Section III-B proposes a continuous-time algorithm, and by providing a novel passivity-based perspective of the proposed algorithm, a distributed condition is provided for convergence over time-varying digraphs. In III-C, a periodic communication scheme based on the passivity-based notion is presented, and an event-triggered communication scheme is developed in Section III-D. The main results are illustrated by an example in Section IV. Some final remarks and future works are described in Section V.

II. Preliminaries and Problem Formulation

In this section, we first introduce our notation, some concepts of convex functions and graph theory followed by a passivity-related definition. Then, the problem to be addressed in this work is formulated.

Notation Let $\mathbb{R}$, $\mathbb{R}^+$ and $\mathbb{N}$ denote the sets of of real numbers, nonnegative real numbers, and nonnegative integers, respectively. The $m \times m$ identity matrix is denoted by I_m. For symmetric matrices A and B, the notation $A \geq B$ ($A > B$) denotes $A - B$ is positive semidefinite (positive definite). $\text{diag}(a_i)$ is the diagonal matrix with a_i being the i-th diagonal entry. $\mathbf{0}_m$ and $\mathbf{1}_m$ denote all zero and one vectors with size $m \times 1$. For column vectors $v_1, \dots, v_m$, $\text{col}(v_1, \dots, v_m) = (v_1^T, \dots, v_m^T)^T$. $\|\lambda\|$ denotes the Euclidean norm of vector λ. Given a positive semidefinite matrix $A \in \mathbb{R}^{N \times N}$, $\sigma_{min}^+(A)$ and $\sigma_N(A)$ denote the smallest positive and the largest eigenvalue of A, respectively. For a twice differentiable function $f(x)$, its gradient and Hessian are denoted by $\nabla f(x)$ and $\nabla^2 f(x)$, respectively. $\text{range}(\nabla f(x))$ denotes the range of the function $\nabla f(x)$. Given a linear mapping L, $\text{null}(L)$ denotes the null space of L. The Kronecker product is denoted by $\otimes$.

Convex function A differentiable function $f : \mathbb{R}^m \to \mathbb{R}$ over a convex set $\mathcal{X} \subset \mathbb{R}^m$ is strictly convex if and only if $(\nabla f(x) - \nabla f(y))^T (x - y) > 0, \forall x \neq y \in \mathcal{X}$, and it is μ-strongly convex if and only if $(\nabla f(x) - \nabla f(y))^T (x - y) \geq \mu \|x - y\|^2, \forall x, y \in \mathcal{X}$, if and only if $f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\mu}{2} \|y - x\|^2, \forall x, y \in \mathcal{X}$. A function $g : \mathbb{R}^m \to \mathbb{R}^m$ over a set $\mathcal{X}$ is l-Lipschitz if and only if $\|g(x) - g(y)\| \leq l \|x - y\|, \forall x, y \in \mathcal{X}$.

Algebraic graph theory A digraph is a pair $\mathcal{G} = (\mathcal{I}, \mathcal{E})$ where $\mathcal{I} = \{1, \dots, N\}$ is the node set and $\mathcal{E} \subseteq \mathcal{I} \times \mathcal{I}$ is the edge set. An edge $(i, j) \in \mathcal{E}$ means that node j can send information to node i, and i is called the out-neighbor of j while j is called the in-neighbor of i. A digraph is strongly connected if for every pair of nodes there exists a directed path connecting them. A time-varying graph $\mathcal{G}(t)$ is uniformly jointly strongly connected if there exists a constant $T \in \mathbb{R}^+$ such that for any $t_k \in \mathbb{R}^+$, the union $\cup_{t \in [t_k, t_k + T]} \mathcal{G}(t)$ is strongly connected. A time-varying graph $\mathcal{G}(t)$ is infinitely jointly strongly connected if the union $\cup_{t \in [\tau, \infty)} \mathcal{G}(t)$ is strongly connected for all $\tau \in \mathbb{R}^+$. Evidently, infinitely jointly connected graphs are less restrictive than uniformly jointly strongly connected graphs. A weighted digraph is a triple $\mathcal{G} = (\mathcal{I}, \mathcal{E}, A)$ where $A \in \mathbb{R}^{N \times N}$ is a weighted adjacency matrix defined as $A = [a_{ij}]$ with $a_{ii} = 0$, $a_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The weighted in-degree and out-degree of node i are $d_{in}^i = \sum_{j=1}^N a_{ij}$ and $d_{out}^i = \sum_{j=1}^N a_{ji}$, respectively. A digraph is said to be weight-balanced if $d_{in}^i = d_{out}^i, \forall i \in \mathcal{I}$. The Laplacian matrix of $\mathcal{G}$ is defined as $L = D_{in} - A$ where $D_{in} = \text{diag}(d_{in}^i)$.

Input Feed-forward passive Consider the following nonlinear system:

$$H : \begin{cases} \dot{s} & = & F(s, u) \\ y & = & Y(s, u), \end{cases}$$

where $s \in S \subset \mathbb{R}^n$, $u \in U \subset \mathbb{R}^m$ and $y \in \mathbb{R}^m$ are the state, input and output variables, respectively, and S, U are the state and input spaces, respectively. F and Y are state function and output function.

Definition 1: ([39]) System H is *Input Feed-forward Passive* (IFP) if there exists a nonnegative real function $V(s) : S \to \mathbb{R}^+$, called the storage function, such that for all $t_1 \geq t_0 \geq 0$, initial condition $s_0 \in S$ and $u \in U$,

$$V(s(t_1)) - V(s(t_0)) \leq \int_{t_0}^{t_1} u^T y - \nu u^T u dt \tag{1}$$

for some $\nu \in \mathbb{R}$, denoted as IFP(ν).

If the storage function $V(s)$ is differentiable, the inequality (1) is equivalent to

$$\dot{V}(s) \leq u^T y - \nu u^T u. \tag{2}$$

As it can be seen from the above definition, a positive value of ν means that the system has an excess of passivity while a negative value of ν means the system lacks passivity. The index ν can be taken as a measurement to quantify how passive a dynamic system is. This concept will play a crucial role in the subsequent results.

Problem formulation Each node i has a local cost function $f_i(x_i) : \mathbb{R}^m \to \mathbb{R}$ where $x_i \in \mathbb{R}^m$ is the local decision variable. The sum of $f_i(x_i)$ is considered as the global cost function. We make the following assumptions.

Assumption 1: Each $f_i, i \in \mathcal{I}$ is twice differentiable with $\nabla^2 f_i(x_i) > 0$ and its gradient $\nabla f_i(x_i)$ is l_i-Lipschitz.

Under Assumption 1, f_i is strictly convex and

$$\|\nabla f_i(x_i) - \nabla f_i(y_i)\| \leq l_i \|x_i - y_i\|. \tag{3}$$

Thus, its Hessian satisfies

$$0 < \nabla^2 f_i(x_i) \leq l_i I_m, \forall i \in \mathcal{I}. \tag{4}$$

Assumption 2: The time-varying communication graph $\mathcal{G}(t)$ is weight-balanced and infinitely jointly strongly connected.

The objective is to design a continuous-time distributed algorithm such that the following problem

$$\begin{array}{cl} \min\limits_{x_1,\ldots,x_N} & \sum\limits_{i=1}^N f_i(x_i) \\ \text{s.t.} & \sum\limits_{i=1}^N x_i = \sum\limits_{i=1}^N d_i \end{array} \tag{5}$$

is solved by each node using its own information only and information received from its neighbors under discrete-time communication. In fact, this problem can be used to formulate many practical applications such as network utility maximization and economic dispatch in power systems.

Let us denote $x = \text{col}(x_1, \ldots, x_N)$. It can be observed that problem (5) is feasible and has a unique optimal point x^*.

III. Main Results

A. The Lagrange dual problem

In this subsection, we show that the resource allocation problem (5) can be equivalently converted into a general distributed convex optimization problem.

Let us define a set of new variables $\lambda_i \in \mathbb{R}^m, i \in \mathcal{I}$, and denote the set of range(∇f_i) as Λ_i. It can be derived from [40] that Λ_i is convex. Under Assumption 1, we have that the inverse function of $\nabla f_i(\cdot)$ exists and is differentiable, denoted as $h_i(\cdot)$ [1], and further define

$$g_i(\lambda_i) \triangleq f_i(h_i(\lambda_i)) + \lambda_i^T (d_i - h_i(\lambda_i)) \tag{6}$$

[1] If the analytic form of the inverse function $h_i(\cdot)$ can not be obtained, one can replace $h_i(\cdot)$ by $\text{argmin}_{x_i}\{f_i(x_i) - \lambda_i^T x_i\}$ in the algorithm that will be proposed later.

when $\lambda_i \in \Lambda_i$.

Lemma 1: Problem (5) can be equivalently solved by the following convex optimization

$$\begin{array}{cl} \min\limits_{\lambda_i \in \Lambda_i, \forall i \in \mathcal{I}} & \sum\limits_{i=1}^N J_i(\lambda_i) \\ \text{s.t.} & \lambda_i = \lambda_j, \forall i, j \in \mathcal{I} \end{array} \tag{7}$$

with $J_i(\lambda_i) = -g_i(\lambda_i)$ and $\nabla J_i(\lambda_i) = h_i(\lambda_i) - d_i$. Moreover, $J_i(\lambda_i)$ is twice differentiable and $\frac{1}{l_i}$-strongly convex in the domain Λ_i, i.e., $\frac{1}{l_i} \leq \nabla^2 J_i(\lambda_i), \forall \lambda_i \in \Lambda_i$.

Proof. This result can be obtained via the duality [41]. □

Due to the strong duality, the primal optimal solution x^* is a minimizer of $\mathcal{L}(x, \lambda^*)$ which is defined as

$$\mathcal{L}(x, \lambda^*) = \sum_{i=1}^N f_i(x_i) + \lambda^{*T} \left(\sum_{i=1}^N d_i - \sum_{i=1}^N x_i \right). \tag{8}$$

This fact enables us to recover the primal solution x^* from the dual optimal solution λ^*. Specifically, since f_i is strictly convex, the function $\mathcal{L}(x, \lambda^*)$ is strictly convex in x, and therefore has a unique minimizer which is identical to x^*. Moreover, since $\mathcal{L}(x, \lambda^*)$ is separable according to (8), we can recover x_i^* from $x_i^* = h_i(\lambda^*)$.

Based on Lemma 1, we then aim to design a continuous-time algorithm with discrete-time communication for problem (7). For simplicity, we will abuse the notation by using $\lambda = \text{col}(\lambda_1, \ldots, \lambda_N)$ hereafter.

B. IFP-based Distributed Algorithm Design

For $i \in \mathcal{I}$ and with constant scalars $\alpha, \beta > 0$, let us consider the following continuous-time algorithm

$$\begin{array}{rcl} \dot{\lambda}_i &=& -\alpha(h_i(\lambda_i) - d_i) - \gamma_i \\ \dot{\gamma}_i &=& -u_i \\ u_i &=& \beta \sum_{j=1}^N a_{ij}(t)(\lambda_j - \lambda_i) \end{array} \tag{9}$$

where $\lambda_i, \gamma_i \in \mathbb{R}^m$ are the local states variables and $u_i \in \mathbb{R}^m$ is the local input. $\alpha > 0$ is a predefined constant and $\beta > 0$ is the coupling gain to be designed. $A(t) = [a_{ij}(t)]_{N \times N}$ is the adjacency matrix of the graph $\mathcal{G}(t)$.

Let $\gamma = \text{col}(\gamma_1, \ldots \gamma_N)$, $d = \text{col}(d_1, \ldots, d_N)$ and $h(\lambda) = \text{col}(h_1(\lambda_1), \ldots, h_N(\lambda_N))$. The algorithm in (9) can be rewritten in a compact form as

$$\begin{array}{rcl} \dot{\lambda} &=& -\alpha \left(h(\lambda) - d \right) - \gamma \\ \dot{\gamma} &=& \beta \mathbf{L}(t) \lambda \end{array} \tag{10}$$

where $\mathbf{L}(t) = L(t) \otimes I_m$ with $L(t)$ being the Laplacian matrix of the graph $\mathcal{G}(t)$.

The above continuous-time algorithm is a simplification of the one proposed in [28] which is motivated by the feedback control consideration. Specifically, each node evolves in the direction of gradient descent while trying to reach an agreement with its neighbors. To correct the error between the local gradient and the consensus with neighbors, the integral feedback of u_i representing the node disagreements is exploited.

In the rest of this work, we assume that $\lambda_i(0) \in \Lambda_i$ for all $i \in \mathcal{I}$. This can be trivially satisfied by letting $\lambda_i(0) = \nabla f_i(x_i(0))$.

In the following, we will first show in Lemma 2 that the optimal solution of (7) coincides with the equilibrium point of algorithm (9). Then we provide a passivity-based perspective for the error dynamics of each individual node in Theorem 1, based on which the convergence of algorithm (9) is shown in Theorem 2.

Lemma 2: Under Assumptions 1 and 2, the equilibrium point (λ^*, γ^*) of the system in (9) with the initial condition $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ is unique and λ^* is the optimal solution of problem (7).

Proof. Suppose (λ^*, γ^*) is the equilibrium of system (9) and $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$. It follows that

$$\begin{array}{rcl} \dot{\lambda^*} & = & -\alpha\left(h(\lambda^*) - d\right) - \gamma^* = \mathbf{0} \\ \dot{\gamma^*} & = & \beta \mathbf{L}(t)\lambda^* = \mathbf{0}. \end{array} \tag{11}$$

Since $\left(1_N \otimes I_m\right)^T \mathbf{L}(t) = \mathbf{0}_{Nm}^T$, we have $\left(1_N \otimes I_m\right)^T \dot{\gamma} = \beta\left(1_N \otimes I_m\right)^T \mathbf{L}(t)\lambda = \mathbf{0}$, which gives $\sum_{i=1}^N \dot{\gamma}_i = \mathbf{0}$. Hence, it can be observed that $\sum_{i=1}^N \gamma_i(t) = \sum_{i=1}^N \gamma_i(0) = \mathbf{0}_m$ for all $t \geq 0$. Next, let us multiply $\left(1_N \otimes I_m\right)^T$ from the left of $\dot{\lambda}^*$, and obtain that

$$\begin{array}{rl} & \left(1_N \otimes I_m\right)^T \dot{\lambda^*} \\ = & -\alpha\left(1_N \otimes I_m\right)^T \left(h(\lambda^*) - d\right) - \sum_{i=1}^N \gamma_i^* = \mathbf{0}, \end{array}$$

which indicates that

$$\nabla J(\lambda^*) = \sum_{i=1}^N \nabla J_i(\lambda_i^*) = \sum_{i=1}^N \left(h_i(\lambda_i^*) - d_i\right) = \mathbf{0}.$$

Moreover, since the graph $\mathcal{G}(t)$ is infinitely jointly strongly connected, $\dot{\gamma^*} = \beta \mathbf{L}(t)\lambda^* \equiv \mathbf{0}$ implies that $\lambda_1^* = \ldots = \lambda_N^*$. Under Assumption 1, problem (7) has a unique solution, which coincides with λ^* based on the optimality condition [42]. $\square$

Before proceeding to show in Theorem 2 that the algorithm converges, let us investigate the IFP property of the error dynamics of each individual node. Denote $\Delta\lambda_i = \lambda_i - \lambda_i^*$ and $\Delta\gamma_i = \gamma_i - \gamma_i^*$. Comparing (9) and (11) yields the individual error system shown as

$$\Psi_i : \left\{ \begin{array}{rcl} \Delta\dot{\lambda}_i & = & -\alpha\left(h_i(\lambda_i) - h_i(\lambda_i^*)\right) - \Delta\gamma_i \\ \Delta\dot{\gamma}_i & = & -u_i \\ u_i & = & \beta \sum_{j=1}^N a_{ij}(t)(\Delta\lambda_j - \Delta\lambda_i). \end{array} \right. \tag{12}$$

By taking u_i and $\Delta\lambda_i$ as the input and output of the error system Ψ_i, the following theorem shows that each error system Ψ_i is IFP with its proof provided in Appendix.

Theorem 1: Suppose Assumption 1 holds. Then, the system Ψ_i is IFP(ν_i) from u_i to $\Delta\lambda_i$ with $\nu_i \geq -\frac{l_i^2}{\alpha^2}$.

Remark 1: It is shown in the above theorem that for the nonlinear system (12) resulting from general strongly convex objective function $J_i(\lambda_i)$ is IFP from u_i to $\Delta\lambda_i$. Moreover, the IFP index is lower bounded by $-\frac{l_i^2}{\alpha^2}$, which means that the system (12) can have the IFP index arbitrarily close to 0 (i.e, passivity) if the coefficient α can be chosen to be arbitrarily large. However, it might be impractical to choose a large α due to the potential numerical error or higher computing cost when solving the ordinary differential equation (10) numerically. In view of this, in order to achieve larger IFP index, we can choose α as the largest positive number allowed by the error tolerance level of the available computing platform. It is worth mentioning that a similar algorithm has been shown in [28]. The contribution of Theorem 1 is to provide a novel passivity-based perspective of the proposed algorithm, and this perspective will lead to fruitful results in the remainder of this section.

The next theorem provides a condition on the coupling gain β for the convergence of algorithm (9).

Theorem 2: Under Assumptions 1 and 2, suppose the coupling gain β satisfies

$$0 < \beta < \frac{\alpha^2 \sigma_{min}^+ \left(L(t) + L(t)^T\right)}{2\sigma_N\left(L(t)^T \mathrm{diag}\left(l_i^2\right) L(t)\right)}, \tag{13}$$

where σ_{min}^+ and σ_N are the smallest positive and the largest eigenvalue respectively. Then under algorithm (9), for all $i \in \mathcal{I}$, Λ_i is a positively invariant set of λ_i, and algorithm (9) with any initial condition with $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ will converge to the optimal solution of (7).

Proof. The proof is given in Appendix. $\square$

Remark 2: Lemma 2 states that the equilibrium point of the continuous-time algorithm (9) under the initial constraint $\sum_{i=1}^N \gamma(0) = \mathbf{0}$ is identical to the optimal solution of the distributed optimization problem (7) while Theorem 2 states that algorithm (9) will converge to such an equilibrium point if the coefficients α and β are chosen to satisfy (13). As discussed in Section III-A, the optimal solution x_i^* of the original resource allocation problem (5) can be recovered from $x_i^* = h_i(\lambda^*)$. In this view, the distributed algorithm in (9) involves only local interaction by exchanging λ_i instead of the real decision variable x_i to achieve the optimal collective goal.

It should be mentioned that the condition proposed in Theorem 2 might be difficult to be examined in a time-varying graph. Nevertheless, the following distributed condition can be obtained based on Theorem 2.

Corollary 1: Under Assumptions 1 and 2, algorithm (9) with any initial condition with $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ will converge to the optimal solution of (7) if the coupling gain $\beta > 0$ satisfies

$$\beta \frac{l_i^2}{\alpha^2} d_{in}^i(t) < \frac{1}{2}, \forall i \in \mathcal{I}, \forall t > 0 \tag{14}$$

where $d_{in}^i(t)$ denotes the in-degree of the i-th node.

Proof. The proof is given in Appendix. $\square$

Remark 3: (Design of parameter β) In order to implement algorithm (9), the parameter β needs to be designed. The condition proposed in the above corollary provides a distributed strategy to design β. A heuristic solution is to let each node compute the maximum β according to (14) and search the minimum of β among them by communicating among its neighboring nodes. Repeat this procedure when a smaller β is updated (a larger $d_{in}^i(t)$ is detected) at any node due to the graph variation. However, this has to be done in an off-line manner. An easier way when $a_{ij} \in \{0,1\}, \forall i,j$ is to simply let the upper bound of β be $\frac{\alpha^2}{2\max_i\{l_i^2\}(N-1)}$.

C. Periodic Discrete-time Communication

Continuous-time communication among the nodes is required in the distributed algorithm proposed in Section III-B whereas a digital network with limited channel capacity generally allows communication only at discrete instants. Moreover, the communication cost is far larger than the computation cost in real applications such as sensor networks [43]. To separate the communication and the computation, we will investigate in this subsection the distributed algorithm design under periodic discrete-time communication by exploiting the IFP property stated in Theorem 1.

By considering a sampled-based scheme, we proceed to investigate the convergence of algorithm (9) with periodic communication.

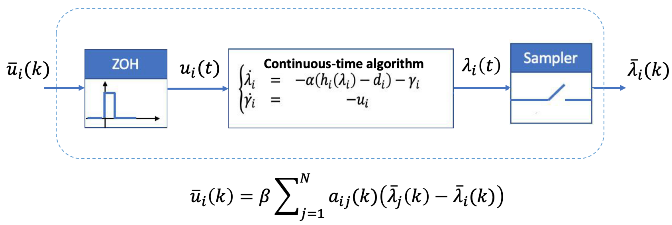


Fig. 1. Sampled continuous-time distributed algorithm.

As depicted in Figure 1, let us consider the algorithm with sampling at each output of individual node,

$$\begin{aligned} \dot{\lambda}_i &= -\alpha(h_i(\lambda_i)-d_i)-\gamma_i \\ \dot{\gamma}_i &= -u_i \\ \bar{u}_i(k) &= \beta\textstyle\sum_{j=1}^{N} a_{ij}(k)(\bar{\lambda}_j(k)-\bar{\lambda}_i(k)) \end{aligned} \tag{15}$$

where $a_{ij}(k)$ denotes $a_{ij}(t)$ at the k-th sampling instant, the output $\bar{\lambda}_i$ is obtained by sampling the continuous-time output λ_i, while the input u_i depending on the sampled $\bar{\lambda}_i, \forall i \in \mathcal{I}_i$ is applied to the continuous-time system through a zero order holder. In particular, let the sampling period be denoted as T_s, and then for all $k \in \mathbb{N}$,

$$\begin{aligned} &\bar{\lambda}_i(k) = \lambda_i(kT_s), \\ &u_i(t) = \bar{u}_i(k), \forall t \in [kT_s,(k+1)T_s). \end{aligned} \tag{16}$$

Denote the time sequence $k = \{0, T_s, 2T_s, \dots\}$. Since the communication is carried out in periodic discrete-time instants, we need to make the following additional assumption for the graph.

Assumption 3: The time-varying graph $\mathcal{G}(k)$ is balanced and infinitely jointly strongly connected, i.e., $\mathcal{G}(k)\cup\mathcal{G}(k+1)\cup\mathcal{G}(k+2)\cup\cdots$ is strongly connected for any $k \in \mathbb{N}$.

With $\Delta\bar{\lambda}_i = \bar{\lambda}_i - \lambda_i^*$ where λ_i^* is defined in (11), the error dynamics of subsystem i is

$$\bar{\Psi}_i : \begin{cases} \Delta\dot{\lambda}_i = -\alpha\left(h_i(\lambda_i)-h_i(\lambda^*)\right)-\Delta\gamma_i \\ \Delta\dot{\gamma}_i = -u_i \\ \bar{u}_i = \beta\sum_{j=1}^{N} a_{ij}(\Delta\bar{\lambda}_j-\Delta\bar{\lambda}_i). \end{cases} \tag{17}$$

In the following, we first analyze and approximate the bound of the sampling error $\Delta\lambda_i - \Delta\bar{\lambda}_i$ with respect to the input $\bar{u}_i$ in Lemmas 3 and 4. Based on these results, Theorem 3 characterizes the passivity degradation over sampling of the error dynamics at each node, and the convergence of algorithm (15) is stated in Corollary 2.

For notational simplicity, let us denote $z_i = \Delta\dot{\lambda}_i$.

Lemma 3: Suppose Assumption 1 holds. Then, under the dynamics $\bar{\Psi}_i$, it holds that for all $u_i \in \mathbb{R}^m$,

$$\frac{l_i}{\alpha}\cdot\frac{d\|z_i\|^2}{dt} \le \frac{l_i^2}{\alpha^2}\|u_i\|^2 - \|z_i\|^2. \tag{18}$$

Proof. The derivative of z_i yields that

$$\dot{z}_i = -\alpha\frac{\partial h_i(\lambda_i)}{\partial\lambda_i}z_i - \Delta\dot{\gamma}_i = -\alpha\frac{\partial h_i(\lambda_i)}{\partial\lambda_i}z_i + u_i$$

and it leads to

$$\frac{l_i}{\alpha}\cdot\frac{d\|z_i\|^2}{dt} = 2\frac{l_i}{\alpha}z_i^T\left(-\alpha\frac{\partial h_i(\lambda_i)}{\partial\lambda_i}z_i + u_i\right).$$

Observe that

$$\begin{pmatrix} \frac{2\alpha}{l_i}\frac{l_i}{\alpha}-1 & -\frac{l_i}{\alpha} \\ -\frac{l_i}{\alpha} & \frac{l_i^2}{\alpha^2} \end{pmatrix} \ge 0,$$

which follows that for all $z_i, u_i \in \mathbb{R}^m$

$$\begin{pmatrix} z_i \\ u_i \end{pmatrix}^T \left(\begin{pmatrix} \frac{2\alpha}{l_i}\frac{l_i}{\alpha}-1 & -\frac{l_i}{\alpha} \\ -\frac{l_i}{\alpha} & \frac{l_i^2}{\alpha^2} \end{pmatrix}\otimes I_m\right)\begin{pmatrix} z_i \\ u_i \end{pmatrix} \ge 0.$$

Since $\frac{1}{l_i}I_m \le \frac{\partial h_i(\lambda_i)}{\partial\lambda_i}$ under Assumption 1, we further obtain that for all $z_i, u_i \in \mathbb{R}^m$

$$\begin{pmatrix} z_i \\ u_i \end{pmatrix}^T \begin{pmatrix} 2\frac{l_i}{\alpha}\left(\alpha\frac{\partial h_i(\lambda_i)}{\partial\lambda_i}\right)-I_m & -\frac{l_i}{\alpha}I_m \\ -\frac{l_i}{\alpha}I_m & \frac{l_i^2}{\alpha^2}I_m \end{pmatrix}\begin{pmatrix} z_i \\ u_i \end{pmatrix} \ge 0,$$

which is equivalent to $\frac{l_i}{\alpha}\frac{d\|z_i\|^2}{dt} \le \frac{l_i^2}{\alpha^2}\|u_i\|^2 - \|z_i\|^2$. □

From the above lemma, it can be seen by the integration of (18) over $t \in [kT_s,(k+1)T_s]$ that

$$\begin{aligned} &\tfrac{l_i}{\alpha}\|z_i((k+1)T_s)\|^2 - \tfrac{l_i}{\alpha}\|z_i(kT_s)\|^2 \\ &\le \tfrac{l_i^2}{\alpha^2}\int_{kT_s}^{(k+1)T_s}\|u_i(t)\|^2dt - \int_{kT_s}^{(k+1)T_s}\|z_i(t)\|^2dt. \end{aligned} \tag{19}$$

Note that (18) or (19) implies that $\frac{l_i^2}{\alpha^2}$ provides an upper bound of the $\mathcal{L}_2$ gain for the mapping $u_i \to z_i$ considering the specific storage function $\frac{l_i}{\alpha}\|z_i\|^2$.

Lemma 4: Under Assumption 1, for all $k \in \mathbb{N}$, the following inequality holds

$$\begin{aligned} &\int_{kT_s}^{(k+1)T_s}\|\Delta\lambda_i(t)-\Delta\bar{\lambda}_i(k)\|^2dt \le T_s^2\cdot \\ &\left(T_s\tfrac{l_i^2}{\alpha^2}\|\bar{u}_i(k)\|^2 + \tfrac{l_i}{\alpha}\left(\|z_i(kT_s)\|^2 - \|z_i((k+1)T_s)\|^2\right)\right). \end{aligned} \tag{20}$$

Proof. First, let us observe that for all $t \in [kT_s,(k+1)T_s), \forall k \in \mathbb{N}$,

$$\begin{aligned} \left\|\int_{kT_s}^{t}\Delta\dot{\lambda}_i(s)ds\right\|^2 &\le \left\|\int_{kT_s}^{(k+1)T_s}\left\|\Delta\dot{\lambda}_i(s)\right\|ds\right\|^2 \\ &\le T_s\int_{kT_s}^{(k+1)T_s}\left\|\Delta\dot{\lambda}_i(s)\right\|^2ds \end{aligned} \tag{21}$$

where the second inequality holds by Cauchy-Schwarz inequality.

Next, it follows from (19) and (21) that

$$\begin{aligned} & \int_{kT_s}^{(k+1)T_s} \|\Delta\lambda_i(t) - \Delta\bar{\lambda}_i(k)\|^2 dt \\ = & \int_{kT_s}^{(k+1)T_s} \| \int_{kT_s}^{t} \Delta\dot{\lambda}_i(s) ds\|^2 dt \\ \le & \int_{kT_s}^{(k+1)T_s} \left(T_s \int_{kT_s}^{(k+1)T_s} \left|\left|\Delta\dot{\lambda}_i(s)\right|\right|^2 ds \right) dt \\ = & T_s^2 \int_{kT_s}^{(k+1)T_s} \left|\left|\Delta\dot{\lambda}_i(s)\right|\right|^2 ds \\ \le & T_s^2 \frac{l_i^2}{\alpha^2} \int_{kT_s}^{(k+1)T_s} \|u_i(s)\|^2 ds + T_s^2 \frac{l_i}{\alpha} \cdot \\ & \left(\|z_i(kT_s)\|^2 - \|z_i((k+1)T_s)\|^2 \right). \end{aligned}$$

Based on the relationship between $u_i(t)$ and $\bar{u}_i(k)$ shown in (16), the inequality (20) can be therefore obtained. □

Theorem 3: Under Assumption 1, the sampled system $\bar{\Psi}_i$ is IFP($\bar{\nu}_i$) from $\bar{u}_i$ to $\Delta\bar{\lambda}_i$ with $\bar{\nu}_i \ge -\left(\frac{l_i^2}{\alpha^2} + T_s \frac{l_i}{\alpha}\right)$ where T_s is the sampling period.

Proof. The proof is stated in Appendix. □

Theorem 3 shows that the lower bound of the IFP index, ν, decreases from $-\frac{l_i^2}{\alpha^2}$ to $-\frac{l_i^2}{\alpha^2} - T_s \frac{l_i}{\alpha}$ over the sampling. This passivity "degradation" is caused by sampling error, which depends on the sampling period T_s. Based on this new IFP index bound, a revised distributed condition for convergence of algorithm (15) is provided as follows.

Corollary 2: Under Assumptions 1 and 3, algorithm (15) under periodic communication with any initial condition with $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ will converge to the optimal solution of (7) if the following condition is satisfied for all $t \ge 0$:

$$\beta \left(\frac{l_i^2}{\alpha^2} + T_s \frac{l_i}{\alpha} \right) d_{in}^i(t) < \frac{1}{2}, \forall i \in \mathcal{I}. \tag{22}$$

Proof. This condition can be derived based on similar argument in the proofs of Theorem 2 and Corollary 1, and the discrete-time LaSalle's invariance principle [44]. □

As shown in the above corollary, when α and β are fixed and satisfy the condition in (14), there always exists a constant $T_s > 0$ satisfying (22). Indeed, with fixed α and β, the sampling period T_s can also be determined in a distributed way by a similar heuristic solution described in Remark 3.

D. Distributed Event-triggered Communication

Based on the sampled-based framework in the preceding subsection, we further consider an event-triggered communication strategy. Reconsider the algorithm as shown in (15) by incorporating an event-triggered communication mechanism depicted by Figure 2, i.e.,

$$\begin{aligned} \dot{\lambda}_i &= -\alpha(h_i(\lambda_i) - d_i) - \gamma_i \\ \dot{\gamma}_i &= -u_i \\ \hat{u}_i(k) &= \beta \textstyle\sum_{j=1}^N a_{ij}(k)(\hat{\lambda}_j(k) - \hat{\lambda}_i(k)) \end{aligned} \tag{23}$$

where $\hat{\lambda}_i(k),\ i \in \mathcal{I}$ denotes the last known state of node i that has been transmitted to its neighbors at the time kT_s. Similar to (16), we set

$$\begin{aligned} & \bar{\lambda}_i(k) = \lambda_i(kT_s), \\ & u_i(t) = \hat{u}_i(k), \forall t \in [kT_s, (k+1)T_s). \end{aligned} \tag{24}$$

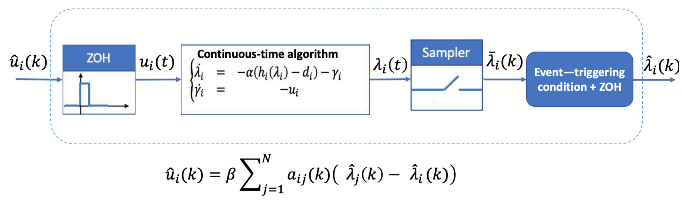


Fig. 2. Continuous-time distributed algorithm with sampled-based event-triggered communication.

The following theorem presents a triggering condition for each node to update its output while the convergence to the global optimal solution is ensured.

Theorem 4: Under Assumptions 1 and 3, consider algorithm (23). If α, β are designed such that (22) is satisfied, and the triggering instant for node $i, \forall i \in \mathcal{I}$ to transmit its current information of λ_i is chosen whenever the following condition is satisfied

$$\begin{aligned} \|e_i(k)\|^2 \ge & \frac{c_i}{d_{in}^i(k)} \left(\frac{1}{2} - \beta d_{in}^i(k) \left(\frac{l_i^2}{\alpha^2} + T_s \frac{l_i}{\alpha} \right) \right)^2 \cdot \\ & \sum_{j=1}^N a_{ij}(k) \|\hat{\lambda}_j(k) - \hat{\lambda}_i(k)\|^2 \end{aligned} \tag{25}$$

where $e_i(k) = \bar{\lambda}_i(k) - \hat{\lambda}_i(k)$ and $c_i \in (0, 1)$, then algorithm (23) with any initial condition with $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ will converge to the optimal solution of (7).

Proof. The proof is stated in Appendix. □

Under the event triggering condition in (25), each node broadcasts its current state (after sampling) $\bar{\lambda}_i(k)$ to its out-neighbors when a local "error" signal exceeds a threshold depending on its own cost function and the last received state of $\hat{\lambda}_j(k)$ from its in-neighbors. Such a triggering condition requires each node being aware of the existence of its in-neighbors. Whenever an edge between two nodes is established, the sender sends its last triggered state to the receiver, which is not considered as a "triggering". Whenever an edge is canceled or established, the receiver updates its input $\hat{u}_i(k)$ by removing or adding the corresponding entry of $\hat{\lambda}$.

Remark 4: Given fixed α, β and T_s, condition (25) is a simple and distributed one to be verified by each node over a balanced graph with very weak connectivity (Assumption 2). It is worth mentioning that this sampled-based event-triggered communication scheme is exempt from Zeno behavior as the minimum inter-event time is lower bounded by the sampling period T_s.

IV. Simulation

In this section, a numerical example is provided to illustrate the previous results.

Consider the resource allocation problem (5) with $N = 10, m = 2$,

$$\begin{aligned}
&f_1(x_1) = x_{11}^2 + \tfrac{1}{2}x_{11}x_{12} + \tfrac{1}{2}x_{12}^2 + 1; \quad f_2(\cdot) = f_1(\cdot);\\
&f_3(x_3) = \tfrac{1}{4}(x_{31}+2)^2 + x_{32}^2; \quad f_4(\cdot) = f_3(\cdot);\\
&f_5(x_5) = \tfrac{1}{2}x_{51}^2 - \tfrac{1}{2}x_{51}x_{52} + x_{52}^2; \quad f_6(\cdot) = f_5(\cdot);\\
&f_7(x_7) = \ln(e^{2x_{71}} + 1) + x_{72}^2; \quad f_8(\cdot) = f_7(\cdot);\\
&f_9(x_9) = \ln(e^{2x_{91}} + e^{-0.2x_{91}}) + \ln(e^{x_{92}} + 1); \;\; f_{10}(\cdot) = f_9(\cdot),
\end{aligned}$$

and $d_1 = d_2 = d_3 = d_4 = d_5 = [1\ 1]^T$, $d_6 = d_7 = d_8 = d_9 = d_{10} = [2\ 2]^T$. Suppose the communication graph $\mathcal{G}(t)$ is time varying, which alternates every 1s between $\mathcal{G}_1$ and $\mathcal{G}_2$ shown in Fig. 3. It can be observed that the switching graph $\mathcal{G}(t)$ is weight-balanced and infinitely jointly strongly connected, and Assumption 1 holds with $l_1 = l_2 = l_5 = l_6 = 2.21, l_3 = l_4 = 1_7 = l_8 = 2, l_9 = l_{10} = 1.21$.

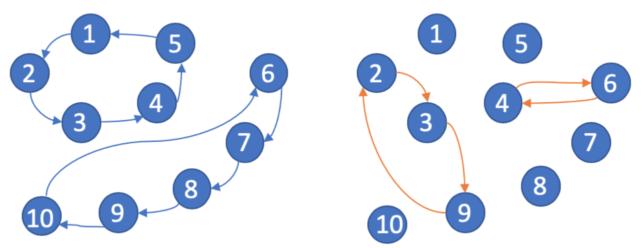


Fig. 3. The switching communication graph $\mathcal{G}(t)$.

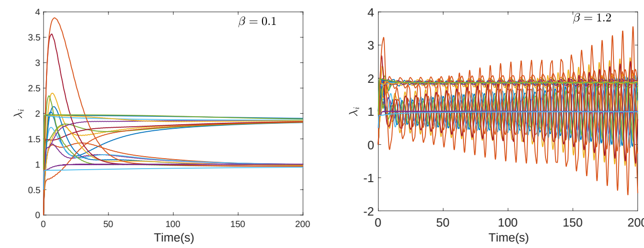
We solve the centralized convex problem (5) using Yalmip, and obtain the optimal solution $x_i^*, i = 1, \ldots, 10$. According to Lemma 1, $\lambda_1^* = \ldots = \lambda_{10}^* = \nabla f_i(x_i^*) = [1.87\ 0.992]^T$. The goal is to design a continuous-time distributed algorithm to equivalently solve the optimization problem (5) under discrete-time communication.

To start with, we recast the above problem into (7) based on Section III-A. It can be obtained that $\Delta J_i(\lambda_i) = h_i(\lambda_i) - d_i$ with

$$\begin{aligned}
h_1(\lambda_1) &= \begin{pmatrix} \frac{4}{7}\lambda_{11} - \frac{2}{7}\lambda_{12} \\ \frac{8}{7}\lambda_{12} - \frac{2}{7}\lambda_{11} \end{pmatrix}; \quad h_2(\cdot) = h_1(\cdot);\\
h_3(\lambda_3) &= \begin{pmatrix} 2\lambda_{31} - 2 \\ \frac{1}{2}\lambda_{32} \end{pmatrix}; \quad h_4(\cdot) = h_3(\cdot);\\
h_5(\lambda_5) &= \begin{pmatrix} \frac{8}{7}\lambda_{51} + \frac{2}{7}\lambda_{52} \\ \frac{2}{7}\lambda_{51} + \frac{4}{7}\lambda_{52} \end{pmatrix}; \quad h_6(\cdot) = h_5(\cdot);\\
h_7(\lambda_7) &= \begin{pmatrix} \frac{1}{2}\ln\frac{\lambda_{71}}{2-\lambda_{71}} \\ \frac{1}{2}\lambda_{72} \end{pmatrix}; \quad h_8(\cdot) = h_7(\cdot);\\
h_9(\lambda_9) &= \begin{pmatrix} \frac{5}{11}\ln\frac{5\lambda_{91}+1}{10-5\lambda_{91}} \\ \ln\frac{\lambda_{91}}{1-\lambda_{91}} \end{pmatrix}; \quad h_{10}(\cdot) = h_9(\cdot).
\end{aligned}$$

In the following simulations, we fix $\alpha = 1$, and fix $\gamma_i(0) = \mathbf{0}, \forall i \in \mathcal{I}$ to satisfy the initial condition $\sum_{i=1}^{N} \gamma_i(0) = \mathbf{0}$. To examine the effectiveness of the distributed algorithms amounts to checking whether the trajectories of $\lambda_i(t), i \in \mathcal{I}$ converge to the value $\lambda^* = [1.87\ 0.992]^T$.

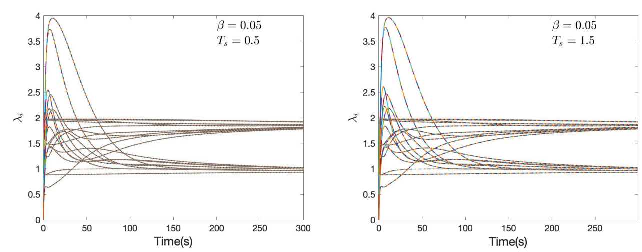
Let us first implement the distributed algorithm (9) under continuous-time communication. By condition (14) in Corollary 1, one has that algorithm (10) will converge with $0 < \beta < 0.103$. Under randomly generated initial value of $x_i(0)$, the trajectories of $\lambda_i(t), i \in \mathcal{I}$ are shown in Figure 4 with different value of β. Although condition (14) is only sufficient, it is shown in Figure 4 that the convergence is no longer ensured when β takes some larger value.



(a) (b)

Fig. 4. Trajectories of $\lambda_i(t)$ under continuous-time communication.

Next, we explore the distributed algorithm (15) under periodic communication. By exploiting condition (22), we have that algorithm (15) will converge with $0 < \beta < \frac{1}{9.74+4.41T_s}$. If we let $\beta = 0.05$, then the condition yields that $T_s < 2.3$. In this example, we let $T_s = 0.5, 1.5$ and it is obvious that Assumption 3 holds. The trajectories of $\lambda_i(t)$ are shown in Figure 5. Note that T_s here is relatively large such that communication is greatly reduced.



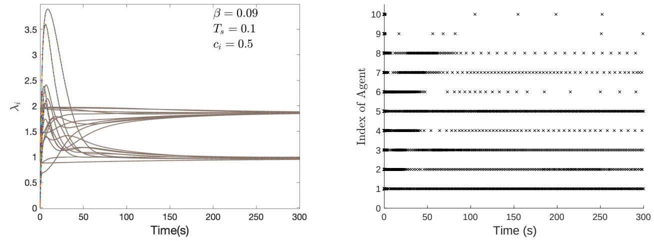
(a) (b)

Fig. 5. Trajectories of $\lambda_i(t)$ under periodic communication.

In the end, let us illustrate algorithm (23) with sample-based event-triggered communication. We select $\beta = 0.09$, $T_s = 0.1$ and $c_i = 0.5$ in (25). The trajectories and the triggering instants of $\lambda_i(t)$ are shown in Figure 6. In Figure 6(b), the largest number of triggering times is 337 for node 5 while the smallest one is only 13 for node 9 and node 10, both of which are a lot smaller than the number of periodic sampling number $300/T_s = 3000$. These show that the sample-based event-triggered control effectively reduces communication costs. Moreover, a better convergence performance is observed in Figure 6(a) than the one in Figure 5(a) with less triggering times, due to the larger coupling gain β.

V. Conclusion

We have introduced the passivity-based perspective for the continuous-time algorithm addressing the distributed resource allocation problem over weight-balanced and infinitely jointly connected digraphs. By showing that the individual algorithmic dynamics is IFP, it is shown how to redesign the algorithm with intermittent communication protocol.



(a) (b)

Fig. 6. Trajectories and triggering instants of $\lambda_i(t)$ under sample-based event-triggered communication.

The passivity-based analysis in this work is based on an existing algorithm that considers the distributed optimization without set constraints and with strong assumptions on cost functions. An interesting future direction is to explore the passivity property for more advanced algorithms and investigate the performance of the algorithms affected by uncertainty. Another promising direction is to explore the compatibility of the passivity-based approach with other network-induced imperfections such as time delay and packet drops.

Appendix

Proof of Theorem 1

Since the Jacobian of $h_i(\lambda_i)$ satisfies $\frac{1}{l_i}I \leq \frac{\partial h_i(\lambda_i)}{\partial \lambda_i}$, it follows from Mean Value Theorem that $h_i(\lambda_i) - h_i(\lambda_i^*) = B_{\lambda_i}(\lambda_i - \lambda_i^*)$ where B_{λ_i} is a symmetric λ_i-dependent matrix defined as $B_{\lambda_i} = \int_0^1 \frac{\partial h_i}{\partial \lambda_i}(\lambda_i + t(\lambda_i - \lambda_i^*))dt$ and $\frac{1}{l_i}I \leq B(\lambda_i)$. Therefore, system (12) can be rewritten as

$$\begin{cases} \Delta\dot{\lambda}_i &= -\alpha B_{\lambda_i}\Delta\lambda_i - \Delta\gamma_i \\ \Delta\dot{\gamma}_i &= -u_i \\ u_i &= \beta\sum_{j=1}^N a_{ij}(t)(\Delta\lambda_j - \Delta\lambda_i). \end{cases}$$

Consider the storage function

$$\begin{aligned} V_i \;=\; & \tfrac{\eta_i}{2}\|\Delta\dot{\lambda}_i\|^2 - \Delta\lambda_i^T\Delta\gamma_i + \alpha(J_i(\lambda_i^*) - J_i(\lambda_i) \\ & + (h_i(\lambda^*) - d_i)^T\Delta\lambda_i) \end{aligned} \tag{26}$$

where η_i is chosen to satisfy $\eta_i > \frac{l_i}{\alpha}$.

First, let us verify the positive definiteness of V_i.

It can be observed that $\frac{\eta_i}{2}\|\Delta\dot{\lambda}_i\|^2 = \frac{\eta_i}{2}\|\alpha B_{\lambda_i}\Delta\lambda_i + \Delta\gamma_i\|^2$, and the strong convexity of $J_i(\lambda_i)$ provides that

$$J_i(\lambda_i^*) - J_i(\lambda_i) \geq -(h_i(\lambda_i) - d_i)^T\Delta\lambda_i + \frac{1}{2l_i}\|\Delta\lambda_i\|^2,$$

which follows that the last term in the storage function V_i satisfies

$$\begin{aligned} & \alpha\left(J_i(\lambda_i^*) - J_i(\lambda_i) + (h_i(\lambda_i^*) - d_i)^T\Delta\lambda_i\right) \\ \geq\; & \alpha\left(-(h_i(\lambda_i) - h_i(\lambda_i^*))^T\Delta\lambda_i + \tfrac{1}{2l_i}\|\Delta\lambda_i\|^2\right) \\ =\; & \Delta\lambda_i^T\left(-\alpha B_{\lambda_i} + \tfrac{\alpha}{2l_i}I\right)\Delta\lambda_i. \end{aligned}$$

It can be derived that

$$\begin{aligned} V_i \geq\; & \tfrac{\eta_i}{2}\|\alpha B_{\lambda_i}\Delta\lambda_i + \Delta\gamma_i\|^2 - \Delta\lambda_i^T\Delta\gamma_i \\ & + (\tfrac{\alpha}{2l_i}I - \alpha B_{\lambda_i})\|\Delta\lambda_i\|^2 \\ =\; & \begin{pmatrix}\Delta\lambda_i \\ \Delta\gamma_i\end{pmatrix}^T \underbrace{\begin{pmatrix} \frac{\alpha^2\eta_i}{2}B_{\lambda_i}^2 - \alpha B_{\lambda_i} + \frac{\alpha}{2l_i}I & * \\ \frac{\alpha\eta_i}{2}B_{\lambda_i} - \frac{1}{2}I & \frac{\eta_i}{2}I \end{pmatrix}}_{W} \begin{pmatrix}\Delta\lambda_i \\ \Delta\gamma_i\end{pmatrix}. \end{aligned} \tag{27}$$

Since $\frac{\eta_i}{2}I > 0$, $\eta_i > \frac{l_i}{\alpha}$ and $\frac{\alpha^2\eta_i}{2}B_{\lambda_i}^2 - \alpha B_{\lambda_i} + \frac{\alpha}{2l_i}I - \left(\frac{\alpha\eta_i}{2}B_{\lambda_i} - \frac{1}{2}I\right)\left(\frac{\eta_i}{2}I\right)^{-1}\left(\frac{\alpha\eta_i}{2}B_{\lambda_i} - \frac{1}{2}I\right) = -\frac{1}{2\eta_i}I + \frac{\alpha}{2l_i}I > 0$, it can be concluded based on Schur Complement Lemma that $W > 0$. Therefore, it can be claimed that $V_i \geq 0$ and $V_i = 0$ if and only if $(\lambda_i, \gamma_i) = (\lambda_i^*, \gamma_i^*)$.

The next step is to show that with the defined storage function V_i, the system Ψ_i is IFP(ν_i) from u_i to $\Delta\lambda_i$.

Let us observe that

$$\begin{aligned} \frac{\eta_i}{2}\cdot\frac{d\|\Delta\dot{\lambda}_i\|^2}{dt} &= \eta_i\Delta\dot{\lambda}_i^T\left(-\alpha\frac{dh_i(\lambda_i)}{dt} - \Delta\dot{\gamma}_i\right) \\ &= \eta_i\Delta\dot{\lambda}_i^T\left(-\alpha\frac{\partial h_i(\lambda_i)}{\partial\lambda_i}\Delta\dot{\lambda}_i + u_i\right) \\ &\leq -\frac{\eta_i\alpha}{l_i}\|\Delta\dot{\lambda}_i\|^2 + \eta_i\Delta\dot{\lambda}_i^T u_i, \\ \frac{d(-\Delta\lambda_i^T\Delta\gamma_i)}{dt} &= -\Delta\dot{\lambda}_i^T\Delta\gamma_i + \Delta\lambda_i^T u_i. \end{aligned}$$

Recall that $\nabla J_i(\lambda_i) = h_i(\lambda_i) - d_i$, and it follows that

$$\begin{aligned} & \alpha\cdot\frac{d\left(J_i(\lambda_i^*) - J_i(\lambda_i) + (h_i(\lambda_i^*) - d_i)^T\Delta\lambda_i\right)}{dt} \\ =\; & \alpha\left(-\nabla J_i(\lambda_i) + (h_i(\lambda_i^*) - d_i)\right)^T\Delta\dot{\lambda}_i \\ =\; & -(\alpha B_{\lambda_i}\Delta\lambda_i)^T\Delta\dot{\lambda}_i. \end{aligned}$$

By combining the above equations, one has that

$$\begin{aligned} \dot{V}_i \;=\; & \frac{\eta_i}{2}\cdot\frac{d\|\Delta\dot{\lambda}_i\|^2}{dt} + \frac{d(-\Delta\lambda_i^T\Delta\gamma_i)}{dt} + \\ & \alpha\cdot\frac{d\left(J_i(\lambda_i^*) - J_i(\lambda_i) + (h_i(\lambda_i^*) - d_i)^T\Delta\lambda_i\right)}{dt} \\ \leq\; & -\frac{\eta_i\alpha}{l_i}\|\Delta\dot{\lambda}_i\|^2 + \eta_i\Delta\dot{\lambda}_i^T u_i + \Delta\lambda_i^T u_i \\ & -(\alpha B(\lambda_i)\Delta\lambda_i + \Delta\gamma_i)^T\Delta\dot{\lambda}_i \\ =\; & \left(-\frac{\eta_i\alpha}{l_i} + 1\right)\|\Delta\dot{\lambda}_i\|^2 + \eta_i\Delta\dot{\lambda}_i^T u_i + \Delta\lambda_i^T u_i \end{aligned} \tag{28}$$

with $-\frac{\eta_i\alpha}{l_i} + 1 < 0$. Since

$$\left(-\frac{\eta_i\alpha}{l_i} + 1\right)\|\Delta\dot{\lambda}_i\|^2 + \eta_i\Delta\dot{\lambda}_i^T u_i \leq \frac{\eta_i^2}{4\left(\frac{\eta_i\alpha}{l_i} - 1\right)}u_i^T u_i,$$

it follows that

$$\dot{V}_i \leq \Delta\lambda_i^T u_i + \frac{\eta_i^2}{4\left(\frac{\eta_i\alpha}{l_i} - 1\right)}u_i^T u_i.$$

Finally, let us prove $\nu_i \geq -\frac{l_i^2}{\alpha^2}$. To this end, consider the following optimization problem

$$\min_{\eta_i > \frac{l_i}{\alpha}} \; \frac{\eta_i^2}{4\left(\frac{\eta_i\alpha}{l_i} - 1\right)},$$

and it can be verified that the optimal solution is given by $\eta_i^* = \frac{2l_i}{\alpha}$ and the corresponding minimum value of the above objective function is $\frac{l_i^2}{\alpha^2}$.

Thus, it can be summarized that $\dot{V}_i \leq \Delta\lambda_i^T u_i + \frac{l_i^2}{\alpha^2} u_i^T u_i$, which completes the proof.

PROOF OF THEOREM 2

Recall the storage function defined in (26) for individual system, and consider the Lyapunov function $V = \sum_{i=1}^N V_i$ for the overall distributed algorithm. Denote $u = \text{col}(u_1, \ldots, u_N)$, $\Delta\lambda = \text{col}(\Delta\lambda_1, \ldots, \Delta\lambda_N)$, and it follows from (12) that $u = -\beta\left(L(t) \otimes I_m\right)\Delta\lambda$. Based on the result in Theorem 1, one has

$$\begin{aligned} \dot{V} &\leq \textstyle\sum_{i=1}^N \Delta\lambda_i^T u_i + \frac{l_i^2}{\alpha^2} u_i^T u_i \\ &= -\beta\Delta\lambda^T\left(L(t) \otimes I_m\right)\Delta\lambda + \beta^2\Delta\lambda^T\left(L(t)^T \otimes I_m\right) \times \\ &\quad \left(\text{diag}\left(\tfrac{l_i^2}{\alpha^2}\right) \otimes I_m\right)\left(L(t) \otimes I_m\right)\Delta\lambda \\ &= \Delta\lambda^T\left(M \otimes I_m\right)\Delta\lambda \end{aligned}$$

with

$$M = -\frac{\beta}{2}\left(L(t) + L(t)^T\right) + \beta^2\left(L(t)^T \text{diag}\left(\frac{l_i^2}{\alpha^2}\right) L(t)\right).$$

Since a weight-balanced digraph $\mathcal{G}$ is strongly connected if and only if it is weakly connected (Lemma 1 in [17]), any weight-balanced digraph can be decomposed into a set of strongly connected balanced digraphs. For a strongly connected balanced digraph, it is apparent that its Laplacian L has the same null space with L^T, which is span$\{1_N\}$. Then, for a weight-balanced digraph, its Laplacian L and L^T have the same null space. Therefore, for a time-varying weight-balanced digraph, Null$(L(t)+L(t)^T)$ is the same with Null$(L(t)^T\text{diag}\left(l_i^2\right)L(t))$ at any time t. Besides, since $\mathcal{G}(t)$ is weight-balanced for all t, it can be easily verified that $L(t) + L(t)^T \geq 0$ and $L(t)^T\text{diag}\left(l_i^2\right)L(t) \geq 0$. Since the above two matrices are both positive semi-definite and have the same null space, it can be implied from the min-max theorem that if the condition in (13) holds, then

$$\alpha^2\left(L(t) + L(t)^T\right) \geq 2\beta L(t)^T\text{diag}\left(l_i^2\right)L(t). \tag{29}$$

Thus, it can be concluded that $M \leq 0$, which leads to $\dot{V} \leq 0$. Note that at any time t, M has the same null space with $L(t)$'s, so $\dot{V}(t) = 0$ only if the nodes belonging to the same strongly connected subgraph reach output consensus. According to LaSalle's invariance principle, the trajectory $\Delta\lambda$ tends to the largest invariant set of $\{\Delta\lambda | \dot{V}(t) = 0\}$. Moreover, since the graph $\mathcal{G}(t)$ is infinitely jointly strongly connected, one has that $\Delta\lambda$ will converge to the set $\{\Delta\lambda | \Delta\lambda_1 = \ldots = \Delta\lambda_N\}$.

According to (27), $V \geq 0$ and V is radially unbounded, i.e., $V \to \infty$ as $\|(\Delta\lambda^T, \Delta\gamma^T)^T\| \to \infty$. Since $\dot{V} \leq 0$, then V is non-increasing, and the state is bounded, i.e., λ, γ are bounded. Let us recall that $\Lambda_i \triangleq \text{range}(\nabla f_i(x_i))$ with $x_i \in \mathbb{R}^m$, and $h_i(\nabla f_i(x_i)) = x_i$. Let $\bar{\Lambda}_i$ be the boundary of the set Λ_i. Since $x_i \in \mathbb{R}^m$ is unbounded in our Problem (5) and f_i is strictly convex, then $\|h_i(\lambda_i)\| \to \infty$ when $\lambda_i \to \bar{\Lambda}_i$. From the first line of (9), this yields that $\|\dot{\lambda}_i\| \to \infty$ when $\lambda_i \to \bar{\Lambda}_i$ since γ_i is bounded. Consequently, based on (26), $V \to \infty$, which contradicts the fact that V is non-increasing. Therefore, for all $i \in \mathcal{I}$, the set Λ_i is a positively invariant set of λ_i.

Next, let us show that $\dot{V} = 0 \Rightarrow \Delta\dot{\lambda}_1 = \ldots = \Delta\dot{\lambda}_N = \mathbf{0}$. Since the inequality in (13) is strict, it follows that there exists a small enough scalar $\epsilon > 0$ such that

$$0 < \beta < \frac{\alpha^2\sigma_{min}^+(L(t) + L(t)^T)}{2\sigma_N\left(L(t)^T\text{diag}\left(l_i^2 + \epsilon\right)L(t)\right)}. \tag{30}$$

By substituting η_i with $\eta_i^* = \frac{2l_i}{\alpha}$ in (28), we have

$$\dot{V}_i \leq -\|\Delta\dot{\lambda}_i\|^2 + \frac{2l_i}{\alpha}\Delta\dot{\lambda}_i^T u_i + \Delta\lambda_i^T u_i.$$

By completing the square, we further have $-\|\Delta\dot{\lambda}_i\|^2 + \frac{2l_i}{\alpha}\Delta\dot{\lambda}_i^T u_i \leq -\frac{\epsilon}{(l_i^2/\alpha^2 + \epsilon)}\|\Delta\dot{\lambda}_i\|^2 + \left(\frac{l_i^2}{\alpha^2} + \epsilon\right)u_i^T u_i$. Hence,

$$\dot{V}_i \leq -\frac{\epsilon}{\left(\frac{l_i^2}{\alpha^2} + \epsilon\right)}\|\Delta\dot{\lambda}_i\|^2 + \left(\frac{l_i^2}{\alpha^2} + \epsilon\right)u_i^T u_i + \Delta\lambda_i^T u_i. \tag{31}$$

Hence, by similar argument before, it follows that $\dot{V} \leq \Delta\lambda^T\left(\hat{M} \otimes I_m\right)\Delta\lambda - \sum_{i=1}^N \frac{\epsilon}{(l_i^2/\alpha^2 + \epsilon)}\|\Delta\dot{\lambda}_i\|^2$ where $\hat{M} = -\frac{\beta}{2}\left(L(t) + L(t)^T\right) + \beta^2 L(t)^T\text{diag}\left(\frac{l_i^2}{\alpha^2} + \epsilon\right)L(t)$ and $\hat{M} \leq 0$. As a consequence, it can be concluded that $\dot{V} \leq 0$ and $\dot{V} = 0$ only if $\Delta\dot{\lambda}_1 = \ldots = \Delta\dot{\lambda}_N = \mathbf{0}$.

Because of the LaSalle's invariance principle, we have that $\Delta\dot{\lambda} \to \mathbf{0}$ and $\Delta\lambda \to 1_N \otimes s$ for some $s \in \mathbb{R}^m$ as $t \to \infty$. Furthermore, by (12), one has $\Delta\dot{\gamma} \to \mathbf{0}$ as $t \to \infty$. Thus, the states λ, γ under algorithm (9) will converge to an equilibrium point. With the initial condition $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$, it follows from Lemma 2 that algorithm (9) will converge to the optimal solution of the problem (7).

PROOF OF COROLLARY 1

Define a vector variable $x = \text{col}(x_1, \ldots, x_N)^T \in \mathbb{R}^{mN}$ and it can be observed that $x^T(\mathbf{L}(t) + \mathbf{L}(t)^T)x(t) = 2\sum_{i=1}^N x_i \sum_{j=1}^N a_{ij}(t)(x_i - x_j) = \sum_{i=1}^N \sum_{j=1}^N a_{ij}(t)(x_i - x_j)^2$ where the second equality follows from the balance of the graph $\mathcal{G}(t)$. Suppose condition (14) holds, i.e., $\alpha^2 > 2\beta 2 l_i^2 d_{in}^i(t)$ for all $i \in \mathcal{I}$. Then, one has

$$\begin{aligned} \alpha^2 x^T(\mathbf{L}(t) + \mathbf{L}(t)^T)x(t) &= \alpha^2\sum_{i=1}^N\sum_{j=1}^N a_{ij}(t)(x_i - x_j)^2 \\ &\geq 2\beta\sum_{i=1}^N l_i^2 d_{in}^i(t)\sum_{j=1}^N a_{ij}(t)(x_i - x_j)^2. \end{aligned}$$

Since $d_{in}^i(t) = \sum_{j=1}^N a_{ij}(t)$, it follows from Cauchy-Schwartz inequality that $d_{in}^i(t)\sum_{j=1}^N a_{ij}(t)(x_i - x_j)^2 \geq \left(\sum_{j=1}^N a_{ij}(t)(x_i - x_j)\right)^2$. This yields that

$$\begin{aligned} &\sum_{i=1}^N l_i^2 d_{in}^i(t)\sum_{j=1}^N a_{ij}(t)(x_i - x_j)^2 \\ \geq\ &\sum_{i=1}^N l_i^2\left(\sum_{j=1}^N a_{ij}(t)(x_i - x_j)\right)^2 \\ =\ &x^T\mathbf{L}(t)^T\text{diag}(l_i^2)\mathbf{L}(t)x(t). \end{aligned}$$

Hence, we have for all $x \in \mathbb{R}^{mN}$, $\alpha^2 x^T(\mathbf{L}(t)+\mathbf{L}(t)^T)x(t) \geq 2\beta x^T \mathbf{L}(t)^T \mathrm{diag}(l_i^2)\mathbf{L}(t)x(t)$, which is equivalent to (29). Following the same reasoning after (29) will complete the proof.

Proof of Theorem 3

Let us consider a revised storage function $\bar{V}_i = \frac{1}{T_s}\left(V_i + \kappa\|z_i\|^2\right)$ with V_i defined in (26) and the coefficient $\kappa > 0$ will be designed later. The positive definiteness of $\bar{V}_i$ can be easily verified since V_i is positive definite according to the proof of Theorem 1 and $\kappa\|z_i\|^2 \geq 0$.

Consider the difference of $\bar{V}_i$ between two consecutive sampling instants, kT_s and $(k+1)T_s$ for any $k \in \mathbb{N}$, we have

$$\int_{kT_s}^{(k+1)T_s} \dot{\bar{V}}_i dt = \bar{V}_i((k+1)T_s) - \bar{V}_i(kT_s) = \frac{1}{T_s}\left(\int_{kT_s}^{(k+1)T_s} \dot{V}_i dt + \kappa\|z_i((k+1)T_s)\|^2 - \kappa\|z_i(kT_s)\|^2\right).$$

It is proved by Theorem 1 that $\dot{V}_i \leq \Delta\lambda_i^T u_i + \frac{l_i^2}{\alpha^2}u_i^T u_i$. By expressing $\Delta\lambda_i(t)$ as $\Delta\bar{\lambda}_i(k) + \left(\Delta\lambda_i(t) - \Delta\bar{\lambda}_i(k)\right)$, one has

$$\begin{aligned} & \int_{kT_s}^{(k+1)T_s} \dot{V}_i dt \\ \leq\; & \int_{kT_s}^{(k+1)T_s} \Delta\bar{\lambda}_i(k)^T u_i dt + \int_{kT_s}^{(k+1)T_s} \\ & \left(\Delta\lambda_i(t) - \Delta\bar{\lambda}_i(k)\right)^T u_i dt + \frac{l_i^2}{\alpha^2}\int_{kT_s}^{(k+1)T_s} u_i^T u_i dt \\ \leq\; & T_s \Delta\bar{\lambda}_i(k)^T \bar{u}_i(k) + T_s\frac{l_i^2}{\alpha^2}\|\bar{u}_i(k)\|^2 \\ & + \int_{kT_s}^{(k+1)T_s}\left(\frac{1}{2\theta}\|\Delta\lambda_i(t) - \Delta\bar{\lambda}_i(k)\|^2 + \frac{\theta}{2}\|\bar{u}_i(k)\|^2\right)dt \end{aligned}$$

where θ can be any positive scalar, and the second inequality holds since $u_i(t)$ is set to be a piecewise signal due to the zero order holder (16). Lemma 4 provides

$$\begin{aligned} & \int_{kT_s}^{(k+1)T_s} \|\Delta\lambda_i(t) - \Delta\bar{\lambda}_i(k)\|^2 dt \\ \leq\; & T_s^3\frac{l_i^2}{\alpha^2}\|\bar{u}_i\|^2 + T_s^2\frac{l_i}{\alpha}\left(\|z_i(kT_s)\|^2 - \|z_i((k+1)T_s)\|^2\right) \end{aligned}$$

which follows that

$$\begin{aligned} & \int_{kT_s}^{(k+1)T_s} \dot{V}_i dt \\ \leq\; & T_s\Delta\bar{\lambda}_i(k)^T\bar{u}_i(k) + \frac{T_s l_i^2}{\alpha^2}\|\bar{u}_i(k)\|^2 + \left(\frac{T_s\theta}{2} + \frac{T_s^3 l_i^2}{2\theta\alpha^2}\right)\cdot \\ & \|\bar{u}_i(k)\|^2 + \frac{T_s^2 l_i}{2\theta\alpha}\left(\|z_i(kT_s)\|^2 - \|z_i((k+1)T_s)\|^2\right). \end{aligned}$$

By selecting θ to minimize the value of $\left(\frac{T_s\theta}{2} + \frac{T_s^3 l_i^2}{2\theta\alpha^2}\right)$, it can be easily obtained that

$$\theta^* = T_s\frac{l_i}{\alpha} \text{ and } \min\left(\frac{T_s\theta}{2} + \frac{T_s^3}{2\theta}\frac{l_i^2}{\alpha^2}\right) = T_s^2\frac{l_i}{\alpha}$$

Now, let us choose $\theta = T_s\frac{l_i}{\alpha}$ and $\kappa = \frac{T_s}{2}$. It follows that

$$\begin{aligned} & \bar{V}_i((k+1)T_s) - \bar{V}_i(kT_s) \\ =\; & \frac{1}{T_s}\left(\int_{kT_s}^{(k+1)T_s}\dot{V}_i dt + \kappa\|z_i((k+1)T_s)\|^2 - \kappa\|z_i(kT_s)\|^2\right) \\ \leq\; & \Delta\bar{\lambda}_i(k)^T\bar{u}_i(k) + \left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\|\bar{u}_i(k)\|^2. \end{aligned}$$

Thus, it can be observed that the sampled system $\bar{\Psi}_i$ is IFP($\bar{\nu}_i$) from $\bar{u}_i$ to $\Delta\bar{\lambda}_i$ with IFP index $\bar{\nu}_i \geq -\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)$.

Proof of Theorem 4

First, let us consider the equilibrium point of (23) with initial condition satisfying $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$ whose compact form is represented as

$$\begin{aligned} \dot{\lambda^*} &= -\alpha(h(\lambda^*) - d) - \gamma^* = \mathbf{0} \\ \dot{\gamma^*} &= \beta\mathbf{L}(k)\hat{\lambda}^* = \mathbf{0}. \end{aligned} \tag{32}$$

By similar reasoning in Lemma 2, we can obtain that $\sum_{i=1}^N \gamma_i(t) = \mathbf{0}$ for any $t > 0$ and $\nabla J(\lambda^*) = 0$. Besides, $\dot{\gamma^*} = \beta\mathbf{L}(t)\hat{\lambda}^* = \mathbf{0}$ leads to $\hat{\lambda}_i^* = \hat{\lambda}_j^*, \forall i,j \in \mathcal{I}$. Due to the triggering condition (25), we have $\|\lambda_i^* - \hat{\lambda}_i^*\| = 0$, indicating $\lambda^* = \hat{\lambda}^*$ and $\lambda_i^* = \lambda_j^*, \forall i,j \in \mathcal{I}$. Under Assumption 1, the equilibrium (λ^*, γ^*) is unique with λ^* being the optimal solution of (7).

Next, the error dynamics in each individual subsystem is obtained by comparing (23) and (32) as

$$\hat{\Psi}_i : \begin{cases} \Delta\dot{\lambda}_i &= -\alpha\left(h_i(\lambda_i) - h_i(\lambda^*)\right) - \Delta\gamma_i \\ \Delta\dot{\gamma}_i &= -u_i \\ \hat{u}_i(k) &= \beta\sum_{j=1}^N a_{ij}(k)(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)) \end{cases}$$

with $\Delta\hat{\lambda}_i = \hat{\lambda}_i - \lambda_i^*$. Since the dynamic from input u_i to output $\Delta\bar{\lambda}_i$ is the same with that in (17) and $u_i(t) = \hat{u}_i(k), \forall t \in [kT_s, (k+1)T_s)$, it follows from Theorem 3 that

$$\begin{aligned} & \bar{V}_i((k+1)T_s) - \bar{V}_i(kT_s) \\ \leq\; & \Delta\bar{\lambda}_i(k)^T\hat{u}_i(k) + \left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\|\hat{u}_i(k)\|^2, \forall i \in \mathcal{I} \end{aligned}$$

with $\bar{V}_i$ defined in the proof of Theorem 3. Consider the Lyapunov function $\bar{V} = \sum_{i=1}^N \bar{V}_i$, and it yields that

$$\begin{aligned} & \bar{V}(k+1) - \bar{V}(k) \\ \leq & \sum_{i=1}^N \Delta\bar{\lambda}_i(k)^T\hat{u}_i(k) + \left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\|\hat{u}_i(k)\|^2 \\ = & \sum_{i=1}^N \beta\Delta\bar{\lambda}_i^T(k)\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right) \\ & + \sum_{i=1}^N \beta^2\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\left\|\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right)\right\|^2 \\ = & \sum_{i=1}^N \beta\left(\Delta\hat{\lambda}_i(k) + e_i(k)\right)^T\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right) \\ & + \sum_{i=1}^N \beta^2\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\left\|\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right)\right\|^2 \\ = & \beta\sum_{i=1}^N\sum_{j=1}^N e_i(k)^T a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right) + \beta\sum_{i=1}^N\sum_{j=1}^N \\ & a_{ij}(k)\Delta\hat{\lambda}_i(k)^T\Delta\hat{\lambda}_j(k) - \beta\sum_{i=1}^N\sum_{j=1}^N a_{ij}(k)\Delta\hat{\lambda}_i(k)^T\Delta\hat{\lambda}_i(k) \\ & + \sum_{i=1}^N \beta^2\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\left\|\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right)\right\|^2 \end{aligned}$$

where the second equality holds since $e_i(k) = \bar{\lambda}_i(k) - \hat{\lambda}_i(k) = \Delta\bar{\lambda}_i(k) - \Delta\hat{\lambda}_i(k)$. It can be derived from $\mathcal{G}(t)$ being balanced that

$$\begin{aligned}&\textstyle\sum_{i=1}^N \sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_i(k)^T\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)^T\Delta\hat{\lambda}_i(k)\right)\\ &= -\tfrac{1}{2}\textstyle\sum_{i=1}^N \sum_{j=1}^N a_{ij}(k)\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2.\end{aligned}$$

Let us observe that for all $\tau_i > 0$

$$\begin{aligned}& e_i(k)^T a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right)\\ \le\ & a_{ij}(k)\left(\tfrac{1}{2\tau_i}\|e_i(k)\|^2 + \tfrac{\tau_i}{2}\|\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2\right)\end{aligned}$$

and by Cauchy-Schwartz inequality we have that

$$\begin{aligned}&\left\|\textstyle\sum_{j=1}^N a_{ij}(k)\left(\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\right)\right\|^2\\ \le\ & d_{in}^i(k)\textstyle\sum_{j=1}^N a_{ij}(k)\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2.\end{aligned}$$

With the above equations, we can now have for any $\tau_i > 0$

$$\begin{aligned}&\bar{V}(k+1) - \bar{V}(k)\\ \le\ & -\frac{\beta}{2}\sum_{i=1}^N\sum_{j=1}^N a_{ij}(k)\left(\left(1 - \tau_i - 2\beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\right)\cdot\right.\\ &\left.\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2 - \frac{\|e_i(k)\|^2}{\tau_i}\right)\\ =\ & -\frac{\beta}{2}\sum_{i=1}^N\left(\left(1 - \tau_i - 2\beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\right)\sum_{j=1}^N\right.\\ &\left. a_{ij}(k)\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2 - d_{in}^i(k)\frac{\|e_i(k)\|^2}{\tau_i}\right).\end{aligned}$$

By letting $\tau_i = \frac{1}{2} - \beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)$, it can be verified by (22) that $\tau_i > 0$, and the above inequality becomes

$$\begin{aligned}&\bar{V}(k+1) - \bar{V}(k)\\ \le\ & -\tfrac{\beta}{2}\sum_{i=1}^N\left(\left(\frac{1}{2} - \beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\right)\sum_{j=1}^N a_{ij}(k)\cdot\right.\\ &\left.\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2 - \frac{d_{in}^i(k)\|e_i(k)\|^2}{\left(\frac{1}{2} - \beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\right)}\right)\end{aligned}$$

Suppose condition (25) holds. Then it follows that

$$\begin{aligned}&\bar{V}(k+1) - \bar{V}(k)\\ \le\ & -\tfrac{\beta}{2}(1 - c_i)\textstyle\sum_{i=1}^N\left(\frac{1}{2} - \beta d_{in}^i(k)\left(\frac{l_i^2}{\alpha^2} + T_s\frac{l_i}{\alpha}\right)\right)\\ &\textstyle\sum_{j=1}^N a_{ij}(k)\|\Delta\hat{\lambda}_j(k) - \Delta\hat{\lambda}_i(k)\|^2.\end{aligned}$$

Since $0 < c_i < 1$, it leads to $\bar{V}(k+1) - \bar{V}(k) \le 0$. Under Assumption 3, the largest invariant set of $\{\Delta\hat{\lambda}|\bar{V}(k+1) - \bar{V}(k) = 0\}$ is $\{\Delta\hat{\lambda}|\Delta\hat{\lambda}_1 = \ldots = \Delta\hat{\lambda}_N\}$. Therefore, according to the discrete-time LaSalle's invariance principle [44], we have that $\Delta\hat{\lambda}_i(k) - \Delta\hat{\lambda}_j(k) \to 0, \forall i, j \in \mathcal{I}$ as $k \to \infty$. Then, it can be indicated from (25) that $\lim_{k\to\infty} e_i(k) = \mathbf{0}$, and hence, $\lim_{k\to\infty}\Delta\bar{\lambda}_i(k) = \lim_{t\to\infty}\Delta\hat{\lambda}_i(k), \forall i \in \mathcal{I}$. It follows from (23) that $\lim_{t\to\infty}\dot{\gamma} = \mathbf{0}$.

Next, since the inequalities of (22) and $c_i < 1$ are strict, by following (31) with similar argument after (31) in the proof of Theorem 2, it can be proved that $\bar{V}(k+1) - \bar{V}(k) = 0 \Rightarrow \Delta\dot{\lambda}_1 = \ldots = \Delta\dot{\lambda}_N = \mathbf{0}$.

Based on the result that $\lim_{t\to\infty}\Delta\dot{\lambda} = \mathbf{0}, \lim_{t\to\infty}\dot{\gamma} = \mathbf{0}$, and $\lim_{t\to\infty}\Delta\lambda = 1_N \otimes s$ for some $s \in \mathbb{R}^m$, it can be concluded that the states λ and γ under algorithm (23) with the triggering condition (25) will converge to an equilibrium point (λ^*, γ^*), and λ^* is identical to the optimal solution of (7) if the initial condition satisfies $\sum_{i=1}^N \gamma_i(0) = \mathbf{0}$.

References

[1] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[2] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151–164, 2011.

[3] B. Gharesifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2013.

[4] W. Shi, Q. Ling, G. Wu, and W. Yin, "Extra: An exact first-order algorithm for decentralized consensus optimization," *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.

[5] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *2011 50th IEEE conference on decision and control and European control conference*. IEEE, 2011, pp. 3800–3805.

[6] 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.

[7] Y. Zhao, Y. Liu, G. Wen, and G. Chen, "Distributed optimization for linear multiagent systems: Edge-and node-based adaptive designs," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3602–3609, 2017.

[8] A. Cherukuri and J. Cortés, "Distributed generator coordination for initialization and anytime optimization in economic dispatch," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 226–237, 2015.

[9] P. Yi, Y. Hong, and F. Liu, "Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems," *Automatica*, vol. 74, pp. 259–269, 2016.

[10] Z. Deng, S. Liang, and Y. Hong, "Distributed continuous-time algorithms for resource allocation problems over weight-balanced digraphs," *IEEE transactions on cybernetics*, vol. 48, no. 11, pp. 3116–3125, 2017.

[11] S. S. Kia, "Distributed optimal in-network resource allocation algorithm design via a control theoretic approach," *Systems & Control Letters*, vol. 107, pp. 49–57, 2017.

[12] L. Ding, G. Y. Yin, W. X. Zheng, Q.-L. Han *et al.*, "Distributed energy management for smart grids with an event-triggered communication scheme," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 5, pp. 1950–1961, 2018.

[13] Y. Zhu, W. Ren, W. Yu, and G. Wen, "Distributed resource allocation over directed graphs via continuous-time algorithms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.

[14] C. Li, X. Yu, W. Yu, T. Huang, and Z.-W. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Transactions on Industrial informatics*, vol. 12, no. 5, pp. 1775–1785, 2015.

[15] X. Shi, Y. Wang, S. Song, and G. Yan, "Distributed optimisation for resource allocation with event-triggered communication over general directed topology," *International Journal of Systems Science*, vol. 49, no. 6, pp. 1119–1130, 2018.

[16] T. T. Doan and C. L. Beck, "Distributed resource allocation over dynamic networks with uncertainty," *IEEE Transactions on Automatic Control*, 2020.

[17] N. Chopra and M. W. Spong, "Passivity-based control of multi-agent systems," in *Advances in robot control*. Springer, 2006, pp. 107–134.

[18] M. Li, L. Su, and G. Chesi, "Consensus of heterogeneous multi-agent systems with diffusive couplings via passivity indices," *IEEE Control Systems Letters*, vol. 3, no. 2, pp. 434–439, 2019.

[19] H. Yu and P. J. Antsaklis, "Output synchronization of networked passive systems with event-driven communication," *IEEE transactions on automatic control*, vol. 59, no. 3, pp. 750–756, 2013.

[20] Y. Yan, L. Su, V. Gupta, and P. Antsaklis, "Analysis of two-dimensional feedback systems over networks using dissipativity," *IEEE Transactions on Automatic Control*, 2019.

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

[22] Y. Wang, M. Xia, V. Gupta, and P. J. Antsaklis, "On feedback passivity of discrete-time nonlinear networked control systems with packet drops," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2434–2439, 2014.

[23] H. Zakeri and P. J. Antsaklis, "Recent advances in analysis and design of cyber-physical systems using passivity indices," in *2019 27th Mediterranean Conference on Control and Automation (MED)*. IEEE, 2019, pp. 31–36.

[24] Q. Lü and H. Li, "Event-triggered discrete-time distributed consensus optimization over time-varying graphs," *Complexity*, vol. 2017, 2017.

[25] Y. Kajiyama, N. Hayashi, and S. Takai, "Distributed subgradient method with edge-based event-triggered communication," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2248–2255, 2018.

[26] C. Liu, H. Li, Y. Shi, and D. Xu, "Distributed event-triggered gradient method for constrained convex minimization," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 778–785, 2019.

[27] C. Liu, H. Li, and Y. Shi, "Resource-aware exact decentralized optimization using event-triggered broadcasting," *IEEE Transactions on Automatic Control*, 2020.

[28] S. S. Kia, J. Cortés, and S. Martínez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254–264, 2015.

[29] W. Chen and W. Ren, "Event-triggered zero-gradient-sum distributed consensus optimization over directed networks," *Automatica*, vol. 65, pp. 90–97, 2016.

[30] W. Yu, Z. Deng, H. Zhou, and Y. Hong, "Distributed resource allocation optimization with discrete-time communication and application to economic dispatch in power systems," in *13th IEEE Conference on Automation Science and Engineering (CASE)*. IEEE, 2017, pp. 1226–1231.

[31] W. Du, X. Yi, J. George, K. H. Johansson, and T. Yang, "Distributed optimization with dynamic event-triggered mechanisms," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 969–974.

[32] A. Wang, X. Liao, and T. Dong, "Event-triggered gradient-based distributed optimisation for multi-agent systems with state consensus constraint," *IET Control Theory & Applications*, vol. 12, no. 10, pp. 1515–1519, 2018.

[33] X. Yi, L. Yao, T. Yang, J. George, and K. H. Johansson, "Distributed optimization for second-order multi-agent systems with dynamic event-triggered communication," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 3397–3402.

[34] J. Liu, W. Chen, and H. Dai, "Event-triggered zero-gradient-sum distributed convex optimisation over networks with time-varying topologies," *International Journal of Control*, vol. 92, no. 12, pp. 2829–2841, 2019.

[35] X. Shi, Z. Lin, T. Yang, and X. Wang, "Distributed dynamic event-triggered algorithm with minimum inter-event time for multi-agent convex optimisation," *International Journal of Systems Science*, pp. 1–12, 2020.

[36] S. Liu, L. Xie, and D. E. Quevedo, "Event-triggered quantized communication-based distributed convex optimization," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 167–178, 2016.

[37] Y. Tang, Y. Hong, and P. Yi, "Distributed optimization design based on passivity technique," in *2016 12th IEEE International Conference on Control and Automation (ICCA)*. IEEE, 2016, pp. 732–737.

[38] 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.

[39] J. Bao and P. L. Lee, *Process control: the passive systems approach*. Springer Science & Business Media, 2007.

[40] G. J. Minty *et al.*, "On the monotonicity of the gradient of a convex function." *Pacific Journal of Mathematics*, vol. 14, no. 1, pp. 243–247, 1964.

[41] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-dynamic programming*. Athena Scientific Belmont, MA, 1996, vol. 5.

[42] A. Ruszczyński, *Nonlinear optimization*. Princeton university press, 2006, vol. 13.

[43] P. Wan and M. D. Lemmon, "Event-triggered distributed optimization in sensor networks," in *Proceedings of the 2009 International Conference on Information Processing in Sensor Networks*. IEEE Computer Society, 2009, pp. 49–60.

[44] W. Mei and F. Bullo, "LaSalle invariance principle for discrete-time dynamical systems: A concise and self-contained tutorial," *arXiv preprint arXiv:1710.03710*, 2017.

Lanlan Su is a Lecturer in the School of Engineering at the University of Leicester. She received her B.E. degree in Electrical Engineering from Zhejiang University, China, in 2014, and her Ph.D degree from the University of Hong Kong, in 2018. Before she joined the University of Leicester in 2019, she was a postdoctoral research associate in University of Notre Dame. She also is an awardee of the Hong Kong Ph.D. Fellowship Scheme established by the Research Grants Council of Hong Kong. Her research interests include multiplier-based analysis, networked control system and optimization.

Mengmou Li is currently a research associate with the Control Group, University of Cambridge. He received his B.S. degree in Physics from Zhejiang University, China, in 2016, and the Ph.D. degree in Electrical and Electronic Engineering from the University of Hong Kong, in 2020. From October 2020 to January 2021, he was a Post-Doctoral Fellow with Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology. His research interests include distributed optimization, consensus and synchronization, passivity.

Vijay Gupta is a Professor in the Department of Electrical Engineering at the University of Notre Dame, having joined the faculty in January 2008. He received his B. Tech degree at Indian Institute of Technology, Delhi, and his M.S. and Ph.D. at California Institute of Technology, all in Electrical Engineering. Prior to joining Notre Dame, he also served as a research associate in the Institute for Systems Research at the University of Maryland, College Park. He received the 2018 Antonio Ruberti Award from IEEE Control Systems Society, the 2013 Donald P. Eckman Award from the American Automatic Control Council and a 2009 National Science Foundation (NSF) CAREER Award. His research and teaching interests are broadly in distributed decision making.

Graziano Chesi is a Professor at the Department of Electrical and Electronic Engineering of the University of Hong Kong. He received the Laurea in Information Engineering (summa cum laude and encomium) and the Best Student Award of the Faculty of Engineering from the University of Florence in 1997, and the PhD in Systems Engineering from the University of Bologna in 2001. He joined the University of Siena in 2000, and the University of Hong Kong in 2006. He served as Associate Editor for various journals, including Automatica, the European Journal of Control, the IEEE Control Systems Letters, the IEEE Transactions on Automatic Control, the IEEE Transactions on Computational Biology and Bioinformatics, and Systems and Control Letters. He also served as Guest Editor for the IEEE Transactions on Automatic Control, for the International Journal of Robust and Nonlinear Control, and for Mechatronics. He founded and served as chair of the Technical Committee on Systems with Uncertainty of the IEEE Control Systems Society. He also served as chair of the Best Student Paper Award Committees for the IEEE Conference on Decision and Control and the IEEE Multi-Conference on Systems and Control. He is author of the books "Homogeneous Polynomial Forms for Robustness Analysis of Uncertain Systems" (Springer 2009) and "Domain of Attraction: Analysis and Control via SOS Programming" (Springer 2011). He is a Fellow of the IEEE.