Distributed Controllers for Multi-Agent Coordination Via Gradient-Flow Approach

Kazunori Sakurama, *Member, IEEE*, Shun-ichi Azuma, *Member, IEEE*, and Toshiharu Sugie, *Fellow, IEEE*

***Abstract*—This paper provides a unified solution for a general distributed control problem of multi-agent systems based on the gradient-flow approach. First, a generalized coordination is presented as a control objective which represents a wide range of coordination tasks (e.g., consensus, formation and pattern decision) in a unified manner. Second, a necessary and sufficient condition for the gradient-based controllers to be distributed is derived. It turns out that the notion of clique (i.e., complete subgraph) plays a crucial role to obtain any distributed controllers. Furthermore, all such controllers are explicitly characterized with free design parameters. Third, it is shown how to choose an optimal controller in a systematic way among all distributed ones, where an optimality measure is introduced for the generalized coordination. Finally, the effectiveness of the proposed method is demonstrated through simulations, where a distributed pattern decision is discussed as an example of the generalized coordination.**

***Index Terms*—Distributed controllers, generalized coordination, gradient-flow method, multi-agent systems.**

I. Introduction

COOPERATIVE and distributed control of multi-agent systems have attracted a lot of attention in the recent years [1]. In the control systems, a group of components called agents interact with each other based on their local information through network communications and/or local sensing [2]. In practical applications, the agents represent vehicles, robots, satellites, sensors and so on, which cooperatively perform a control task [3]. In intelligent transportation systems, a large number of vehicles are controlled so as to drive safely and efficiently [4]. As more fundamental tasks, consensus [5], [6], formation [7]–[9], flocking [10], [11], coverage [12], [13], purchase [14], [15] and attitude synchronization [16], [17] have been vigorously investigated.

One of the most powerful tools to facilitate cooperative motions in multi-agent systems is the gradient-flow method [18]. This method is based on an objective function (or a potential function), which encodes a control task. The way to design objective functions is generally stated as follows: Consider n agents whose dynamics are given by $\dot{x}_i = u_i$ for $i \in \{1, 2, \ldots, n\}$ with the state $x_i \in \mathbb{R}^d$ and input $u_i \in \mathbb{R}^d$ of agent i. These

Manuscript received June 17, 2013; revised February 20, 2014, August 25, 2014, and November 4, 2014; accepted November 14, 2014. Date of publication November 26, 2014; date of current version May 21, 2015. This work was supported by MEXT/JSPS KAKENHI (24760336). Recommended by Associate Editor S. Zampieri.

K. Sakurama is with the Graduate School of Engineering, Tottori University, Tottori 680-8552, Japan (e-mail: sakurama@mech.tottori-u.ac.jp).

S-i. Azuma and T. Sugie are with the Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan (e-mail: sazuma@i.kyoto-u.ac.jp; sugie@i.kyoto-u.ac.jp).

Digital Object Identifier 10.1109/TAC.2014.2374951

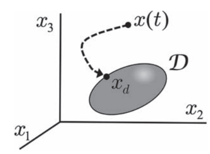

Fig. 1. Geometric illustration of the multi-agent coordination.

agents can communicate with and/or sense each other over a network graph $\mathcal{G}$. A target set $\mathcal{D} \subset \mathbb{R}^{nd}$ is given so as to describe a control task, which is achieved if $x \in \mathcal{D}$ for $x := [x_1^\top\, x_2^\top \cdots x_n^\top]^\top \in \mathbb{R}^{nd}$. Then, find an objective function $V(x)$ such that

S1) $\partial V/\partial x_i(x)$ depends only on the states x_j of the agents who can communicate to agent i over the graph $\mathcal{G}$,

S2) $V(x)$ takes a minimum for $x \in \mathcal{D}$.

A solution to this problem provides the gradient-based controller $u_i = -\partial V/\partial x_i(x)$ for $i \in \{1, 2, \cdots, n\}$. From S1), this controller is *distributed*, namely uses only local information. From S2), the resulting dynamics $\dot{x}_i = -\partial V/\partial x_i(x)$ achieves the convergence $x(t) \to x_d$ for a point $x_d \in \mathcal{D}$, where the given control task is attained. See Fig. 1 for the illustration of this problem for three agents on a one-dimensional space ($n = 3$, $d = 1$).

Typical coordination tasks for multi-agent systems are listed in Table I with the corresponding target sets $\mathcal{D}$ and objective functions $V(x)$, where $\mathcal{E}$ is the edge set of the graph $\mathcal{G}$. Fig. 2 illustrates the target sets $\mathcal{D}$. For example, the consensus problem is the case of $\mathcal{D} = \{x \in \mathbb{R}^{nd} | x_i = x_j\ \forall\, i, j \in \{1, 2, \ldots, n\}\}$, where all the states x_i reach a consensus. This $\mathcal{D}$ represents a line. Pattern decision introduced in Section VI is a control task to decide and form a suitable formation pattern from given multiple patterns. This task is formulated with the set $\mathcal{D}$ of multiple lines, each of which corresponds to a given pattern. Hence, the gradient-flow method provides solutions for a wide range of multi-agent problems. In this research, a common objective function $V(x)$ is employed to share a task between agents. This is ordinary in the gradient-flow approach [18], but is not in many potential games [19].

On the other hand, each existing work addresses a specific $\mathcal{D}$ as shown in Table I. In other words, each problem has been addressed independently so far. Hence, it contributes to clarify appropriate network structures ($\mathcal{G}$) to each task ($\mathcal{D}$), while it is not clear what kind of tasks ($\mathcal{D}$) are achievable over a given network ($\mathcal{G}$). In order to clarify this point, it is important to derive solutions of the above-mentioned problem for a general class of $\mathcal{D}$. Furthermore, such solutions enable us to design distributed controllers in a unified way. In this sense, this problem is fairly fundamental in multi-agent problems, but it is still open.

TABLE I
CONTROL TASKS ACHIEVABLE WITH THE GRADIENT-FLOW METHOD

Control task	Target set $\mathcal{D}$	Objective function $V(x)$	Reference
Consensus	Homeomorphic to $\mathbb{R}^d$	$\sum_{(i,j)\in\mathcal{E}} \Vert x_i - x_j\Vert^2$	[5], [6]
Distance-based formation	Homeomorphic to $\mathbb{SE}(d)$	$\sum_{(i,j)\in\mathcal{E}} (\Vert x_i - x_j\Vert - d_{ij})^2$	[8], [9]
Flocking	α-lattice	Distance-based formation + velocity consensus	[10], [11]
Coverage	Multiple points	$\sum_{i=1}^{n}\int_{\mathcal{V}_i} f(\Vert q - x_i\Vert)\phi(q)\mathrm{d}q$	[12], [13]
Pattern decision	Homeomorphic to $\mathbb{R}^d \times \{1,2,\ldots,p\}$	$\sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})}\prod_{l=1}^{p}\sum_{i\in\mathcal{I}}\left\Vert \sum_{j\in\mathcal{I}\setminus\{i\}}(x_i - x_j - \xi_{ij}^l)\right\Vert^2$	Sec. VI

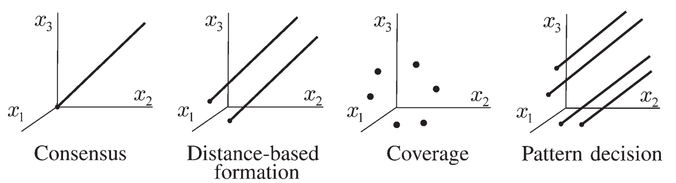

Fig. 2. Target sets $\mathcal{D}$ for several control tasks.

This paper provides a unified solution for the general distributed control problem based on the gradient-flow method. First, we give a complete characterization of all distributed controllers over a given $\mathcal{G}$, which corresponds to S1). Note that, as seen in Table I, the conventional gradient-based controllers are derived from $V(x)$ of the form $\sum_{(i,j)\in\mathcal{E}} W_{ij}(x_i, x_j)$, namely it is based on edges. However, it turns out that not edges but cliques (i.e., complete subgraphs) are the critical units for distributed controllers of multi-agent systems. Second, based on the characterization, we provide a unified solution for the general distributed control problem, which corresponds to both S1) and S2). To do so, we introduce a clique space, that is, a subspace spanned by the states x_i of the agents belonging to a clique. The key is to use a projection of $\mathcal{D}$ onto each clique space in the characterization of the distributed controllers. This solution will be a distributed controller achieving S2) if it exists. Even if it does not exist due to the limited local information, it provides us the best approximate solution.

One of the contributions of this paper is to derive a necessary and sufficient condition for gradient-based controllers to be distributed. This result gives a complete characterization of all such controllers for any given network by exploiting clique-based functions, and yields a systematic procedure to design distributed controllers for given performance specifications. Hence, if there exists a controller which satisfies the specifications, we can always find it in the set. While, if there is no controller satisfying the specifications in the set, we can guarantee that this problem is infeasible. Instead, we can find a best approximation of the solution from the set. This is possible due to the necessity part of this condition. Note that some existing papers (e.g., [20]) have already employed an optimization method based on cliques. Thus, the sufficient part is rather easy. However, the necessity part is critical from the above viewpoint and it is proved for the first time in this paper.

In addition, the stability of the multi-agent system with the proposed distributed controller is investigated. Since an equilibrium is not given by a point, but is given by a set for the general distributed control problem from S2), the Lyapunov stability theorem is extended to the equilibrium set. We show that the proposed controller achieves the asymptotic stability for a certain class of target sets $\mathcal{D}$ and the global asymptotic stability for convex $\mathcal{D}$. Since the paper considers the multi-agent control in the framework of the gradient-flow approach, the global asymptotic stability is not always guaranteed. This is the limitation inherent to the general gradient-flow approach (not only for multi-agent systems but general dynamical systems). Nevertheless, the gradient-flow approach plays a crucial role in the multi-agent control. In fact, as listed in Table I, many multi-agent coordination tasks are based on this approach. Note that this paper mainly considers fixed graphs. Thus, in Table I, consensus, formation and pattern decision are optimally solved in a unified way, but flocking and coverage on unfixed graphs are not included. In spite of the limitation, the results of the distributed controller characterization and optimal controller design could be of a fundamental importance for multi-agent systems.

This paper is organized as follows. Section II formulates the generalized coordination problem for a multi-agent system. In Section III, an optimal distributed controller which gives the best approximate solution is designed. In Section IV, a result of characterization of all gradient-based distributed controllers is proved, which is used in Section III. In Section V, the stability is analyzed for the multi-agent system with the proposed distributed controller. In Section VI, the effectiveness of the proposed method is demonstrated through an application to the pattern decision problem. In Section VII, these results are expanded to the case of non-differentiable objective functions. Section VIII is the conclusion.

Notations: Let $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}^d$ be the sets of all real numbers, non-negative numbers and d-dimensional real vectors, respectively. The Kronecker product of vectors is denoted by $\otimes$. The notations $\langle\cdot,\cdot\rangle$ and $\Vert\cdot\Vert$ denote the inner product and the Euclidean norm of vectors, respectively. The notations $cl(\cdot)$ and $co(\cdot)$ represent the closure and convex hull of a set, respectively. Let $|\cdot|$ and $pow(\cdot)$ be the cardinality (i.e., the number of the elements) and the power set (i.e., the family of the subsets) of a set. The empty set is denoted by $\emptyset$. Consider n real vectors $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$ and a set $\mathcal{I} = \{i_1, i_2, \ldots, i_m\}$ of m natural numbers such that $i_j \neq i_k$ and $i_j \leq n$ for $j,k \in \{1,2,\ldots,m\}, j \neq k$. Then, the set $\{x_{i_1}, x_{i_2}, \ldots, x_{i_m}\}$ of the m real vectors corresponding to the index set $\mathcal{I}$ is denoted by $x_{\mathcal{I}}$. Following [21],

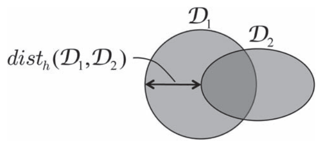

Fig. 3. Directed Hausdorff distance.

we do not distinguish the set $x_{\mathcal{I}}$ of the vectors and the corresponding collective vector $[x_{i_1}^\top\ x_{i_2}^\top\ \cdots\ x_{i_m}^\top]^\top \in \mathbb{R}^{md}$. Let $I_n \in \mathbb{R}^{n\times n}$, $\mathbf{1}_n \in \mathbb{R}^n$ and $e_{ni} \in \mathbb{R}^n$ be the identity matrix, the vector which has all its entries 1, and the unit vector whose i-th component is 1, respectively.

Let class $\mathcal{C}^p$ be the set of functions whose i-th derivative exists and is continuous for $i \in \{1,2,\ldots,p\}$. Let $\mathcal{L}_{\mathrm{loc}}$ be the set of locally Lipschitz continuous functions. The operator $Z(\cdot)$ from a function $f:\mathbb{R}^n \to \mathbb{R}^m$ to a subset in $\mathbb{R}^n$ yields the zero set of f, namely $Z(f) := \{x \in \mathbb{R}^n | f(x) = 0\}$. For sets $\mathcal{A}$, $\mathcal{D} \subset \mathbb{R}^n$, $[\mathcal{A}]_{\mathcal{D}}$ is the subset of $\mathcal{A}$ whose components are connected to $\mathcal{D}$ as

$$[\mathcal{A}]_{\mathcal{D}} := \left\{ x \in \mathcal{A} \,\middle|\, \begin{array}{l} \text{there is a path } \pi:[0,1]\to\mathcal{A} \\ \text{s.t. } \pi(0) = x \text{ and } \pi(1) \in \mathcal{D} \end{array} \right\} \tag{1}$$

where the path means the continuous map from the unit closed interval to $\mathcal{A}$. The distance function from a point $x \in \mathbb{R}^n$ to a set $\mathcal{D} \subset \mathbb{R}^n$ is defined as $dist(x,\mathcal{D}) := \inf_{y\in\mathcal{D}} \|x-y\|$. A non-negative function $\rho(x,\mathcal{D})$ satisfying the following is called a semi-distance function between x and $\mathcal{D}$

$$\rho(x,\mathcal{D}) = 0 \Leftrightarrow x \in cl(\mathcal{D}). \tag{2}$$

The directed Hausdorff distance from a set $\mathcal{D}_1 \subset \mathbb{R}^n$ to $\mathcal{D}_2 \subset \mathbb{R}^n$ is defined as

$$dist_h(\mathcal{D}_1,\mathcal{D}_2) := \sup_{x\in\mathcal{D}_1} dist(x,\mathcal{D}_2) \tag{3}$$

which evaluates how $\mathcal{D}_1$ is different from $\mathcal{D}_2$ as illustrated in Fig. 3.

II. Problem Formulation

A. System Description

Consider a group of n agents governed by the homogeneous first-order differential equations

$$\dot{x}_i = u_i,\ i \in \mathcal{N} \tag{4}$$

where $x_i \in \mathbb{R}^d$ and $u_i \in \mathbb{R}^d$ denote the state and the input of agent i in the d-dimensional space, and $\mathcal{N} := \{1,2,\ldots,n\}$ is the agent set. Let $x := x_{\mathcal{N}}$ and $u := u_{\mathcal{N}}$ be the collections of all the states x_i and inputs u_i. Agents exchange information through communication and/or local sensing. Their network is represented by the set $\mathcal{E} \subset \mathcal{N}\times\mathcal{N}$. If $(i,j) \in \mathcal{E}$, agent i can communicate with agent j. We say that agent j is adjacent to agent i if $(i,j) \in \mathcal{E}$. Let $\mathcal{G} = (\mathcal{N},\mathcal{E})$ be the graph for the node set $\mathcal{N}$ and the edge set $\mathcal{E}$. Assume that $\mathcal{G}$ is time-invariant.

The graph $\mathcal{G}$ represents the information available to the control input of each agent. Then, the controller u_i which can be embedded in agent i should be of the form

$$u_i = f_i\left(x_i, x_{\mathcal{N}_i}\right) \tag{5}$$

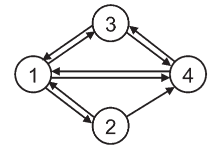

Fig. 4. Example of a graph.

for some function $f_i : \mathbb{R}^d \times \mathbb{R}^{|\mathcal{N}_i|d} \to \mathbb{R}^d$, where $\mathcal{N}_i \subset \mathcal{N}$ is the neighbor set of agent i, i.e.,

$$\mathcal{N}_i := \{j \in \mathcal{N} | (i,j) \in \mathcal{E}\}. \tag{6}$$

A controller of the form (5) is said to be *distributed* (or *decentralized*) over the graph $\mathcal{G}$.

The subgraph induced by a node subset $\mathcal{I} \subset \mathcal{N}$ is denoted by $\mathcal{G}[\mathcal{I}] = (\mathcal{I},\mathcal{E}')$, where $\mathcal{E}' \subset \mathcal{E}$ contains any edges $(i,j) \in \mathcal{E}$ such that $i,j \in \mathcal{I}$, i.e.,

$$\mathcal{E}' = \{(i,j) \in \mathcal{E} | i,j \in \mathcal{I}\}.$$

A node subset $\mathcal{I} \subset \mathcal{N}$ is called a *clique* if and only if its induced subgraph $\mathcal{G}[\mathcal{I}]$ is complete. The number $|\mathcal{I}|$ of the nodes in the clique is called its size. Let $\mathcal{S}_{clq}(\mathcal{G}) \subset pow(\mathcal{N})$ be the family of all cliques in the graph $\mathcal{G}$, namely

$$\mathcal{S}_{clq}(\mathcal{G}) := \{\mathcal{I} \subset \mathcal{N} | (i,j) \in \mathcal{E} \quad \forall\, i,j \in \mathcal{I} \text{ s.t. } i \neq j\}. \tag{7}$$

Subsets of $\mathcal{S}_{clq}(\mathcal{G})$ concerning particular agents are defined as follows:

$$\mathcal{S}_{clq,i}(\mathcal{G}) := \{\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G}) | i \in \mathcal{I}\} \tag{8}$$

$$\mathcal{S}_{clq,ij}(\mathcal{G}) := \{\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G}) | i,j \in \mathcal{I}\}. \tag{9}$$

Example 1: Consider the network depicted in Fig. 4. The node and edge sets of the corresponding graph and the neighbor set of each agent are obtained as follows:

$$\begin{aligned} \mathcal{N} &= \{1,2,3,4\} \\ \mathcal{E} &= \{(1,2),(1,3),(1,4),(2,1),(2,4),(3,1),(3,4), \\ &\qquad (4,1),(4,3)\} \\ \mathcal{N}_1 &= \{2,3,4\},\ \mathcal{N}_2 = \{1,4\},\ \mathcal{N}_3 = \{1,4\},\ \mathcal{N}_4 = \{1,3\}. \end{aligned}$$

The family of all cliques and its subsets concerning agents 1 and 3 are given as follows:

$$\begin{aligned} \mathcal{S}_{clq}(\mathcal{G}) &= \{\{1,3,4\},\{1,2\},\{1,3\},\{1,4\},\{3,4\}, \\ &\qquad \{1\},\{2\},\{3\},\{4\},\emptyset\} \\ \mathcal{S}_{clq,1}(\mathcal{G}) &= \{\{1,3,4\},\{1,2\},\{1,3\},\{1,4\},\{1\}\} \\ \mathcal{S}_{clq,13}(\mathcal{G}) &= \{\{1,3,4\},\{1,3\}\}. \end{aligned}$$

■

B. Control Objective

For the multi-agent system (4) over the graph $\mathcal{G}$, consider a target set $\mathcal{D} \subset \mathbb{R}^{nd}$. The set $\mathcal{D}$ is assumed to be non-empty, but not necessarily to be connected. The set $\mathcal{D}$ describes our control task, and $x \in \mathcal{D}$ implies that the task is achieved. Hence, we design a controller to achieve

$$\lim_{t\to\infty} dist\left(x(t),\mathcal{D}\right) = 0 \tag{10}$$

where $x(t) \in \mathbb{R}^{nd}$ is the solution of the system (4). This control task is called the *generalized coordination*.

Example 2: The consensus problem can be formulated by the generalized coordination with the target set

$$\mathcal{D} = \{x \in \mathbb{R}^{nd} | x_i = x_j \quad \forall i, j \in \mathcal{N}\} \tag{11}$$

with which (10) implies that $x_i(t) - x_j(t) \to 0$ for all $i, j \in \mathcal{N}$. The set (11) is equivalent to

$$\mathcal{D} = \{x \in \mathbb{R}^{nd} | \exists \theta \in \mathbb{R}^d \text{ s.t. } x = \mathbf{1}_n \otimes \theta\} \tag{12}$$

where the vector θ represents the consensus point. Thus, as depicted in Fig. 2, for $d = 1$, the shape of $\mathcal{D}$ is the straight line in the space $\mathbb{R}^n$ with slope 1 which passes through the origin. ■

Remark 1: In the problem formulation, $\mathcal{D}$ is given in advance, which is global information. However, as revealed in Section III, each agent does not need to know the whole information on $\mathcal{D}$, but needs local information concerning its neighbors. For example, in the consensus problem, agent i does not have to know the global goal given in (11), but should know the local requirement $x_i = x_j$ for its neighbors $j \in \mathcal{N}_i$. See Example 4 for details. ■

The gradient-flow method is one of the most powerful tools to design controllers for multi-agent systems. We just have to design an objective function $V : \mathbb{R}^{nd} \to \mathbb{R}_+$ to evaluate the achievement of the control task represented by $\mathcal{D}$. Then, the controller for agent $i \in \mathcal{N}$ is obtained as the corresponding components of the gradient of $V(x)$ as

$$u_i = -\frac{\partial V}{\partial x_i}(x). \tag{13}$$

By using this controller, from (4) and (13)

$$\dot{V}(x) = \sum_{i=1}^{n} \left\langle \frac{\partial V}{\partial x_i}(x),\ \dot{x}_i \right\rangle = -\sum_{i=1}^{n} \left\| \frac{\partial V}{\partial x_i}(x) \right\|^2 \leq 0 \tag{14}$$

is derived and $V(x)$ is monotonically decreasing. Then, as long as $Z(V) \cap \mathcal{D}$ is non-empty, it is expected that $x(t)$ locally approaches the set $[Z(V)]_{\mathcal{D}}$ which is a subset of the minimum set of $V(x)$ connected to $\mathcal{D}$.

In addition, we have to design $V(x)$ such that the derived controller (13) is of the form (5) to be distributed for the implementation. In this sense, we say a class $\mathcal{C}^1$ function $V : \mathbb{R}^{nd} \to \mathbb{R}$ is *gradient-distributed* over the graph $\mathcal{G}$ if and only if the gradient-based controller (13) is distributed, that is, there exist n continuous functions $f_i : \mathbb{R}^d \times \mathbb{R}^{|\mathcal{N}_i|d} \to \mathbb{R}^d$ such that

$$\frac{\partial V}{\partial x_i}(x) = -f_i\left(x_i, x_{\mathcal{N}_i}\right) \tag{15}$$

for all $i \in \mathcal{N}$.

Two important classes of objective functions are defined. First, let $\mathcal{F}_{gd}(\mathcal{G}) \subset \mathcal{C}^1$ be the set of all gradient-distributed functions over the graph $\mathcal{G}$. Second, let $\mathcal{F}_z(\mathcal{D}) \subset \mathcal{C}^0$ be the set of all functions $V : \mathbb{R}^{nd} \to \mathbb{R}_+$ such that $\mathcal{D} \subset Z(V)$. Note that $Z(V) \cap \mathcal{D}$ is always non-empty for $V \in \mathcal{F}_z(\mathcal{D})$.

Now, we formulate the problem which we tackle in this paper as follows.

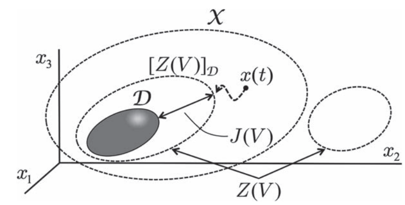

Fig. 5. Meaning of the performance index $J(V)$.

Problem 1: For a graph $\mathcal{G}$ and a non-empty set $\mathcal{D} \subset \mathbb{R}^{nd}$, find a function $V(x)$ of class $\mathcal{F}_{gd}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D})$ minimizing the performance index

$$J(V) := dist_h\left([Z(V)]_{\mathcal{D}}, \mathcal{D}\right). \tag{16}$$

■

Let $x(t) \in \mathbb{R}^{nd}$ be the solution of the system (4) with the gradient-based controller (13). Then, Fig. 5 illustrates the meaning of the performance index $J(V)$, which tells us the worst case distance between $\mathcal{D}$ and the convergence point of the state $x(t)$ to $[Z(V)]_{\mathcal{D}}$. Indeed, if and only if $J(V) = 0$ is obtained, $[Z(V)]_{\mathcal{D}} = \mathcal{D}$ holds and the generalized coordination (10) is locally achieved. However, this is not always the case due to the limit of the network connections in the graph $\mathcal{G}$. Even then, a solution to Problem 1 provides the best objective function $V(x)$ over the graph $\mathcal{G}$ in terms of the performance $J(V)$. The gradient-based controller (13) with such $V(x)$ is called *J-optimal*. Actually, the J-optimal controller gives the smallest value of the worst case distance between $\mathcal{D}$ and $x(t)(t \to \infty)$ from initial states $x(0) \in \mathcal{X}$ for a certain neighborhood $\mathcal{X}$ of $\mathcal{D}$.

In order to complete this discussion, the convergence of $x(t)$ to $[Z(V)]_{\mathcal{D}}$ should be verified. For this purpose, the stability is analyzed for the system (4) with the controller (13). Note that $[Z(V)]_{\mathcal{D}}$ is closed and is an equilibrium set of this system. Then, $[Z(V)]_{\mathcal{D}}$ is said to be *stable* if for each open set $\mathcal{O}_1 \supset [Z(V)]_{\mathcal{D}}$, there exists an open set $\mathcal{O}_2 \supset [Z(V)]_{\mathcal{D}}$ such that

$$x(0) \in \mathcal{O}_2 \Rightarrow x(t) \in \mathcal{O}_1 \quad \forall t \geq 0. \tag{17}$$

Moreover, $[Z(V)]_{\mathcal{D}}$ is said to be *asymptotically stable* if it is stable and there exists an open set $\mathcal{O} \supset [Z(V)]_{\mathcal{D}}$ such that

$$x(0) \in \mathcal{O} \Rightarrow \lim_{t \to \infty} dist\left(x(t), [Z(V)]_{\mathcal{D}}\right) = 0. \tag{18}$$

In addition, $[Z(V)]_{\mathcal{D}}$ is said to be *globally asymptotically stable* if $\mathcal{O} = \mathbb{R}^{nd}$ in (18).

The problem on the stability analysis is given as follows.

Problem 2: For a graph $\mathcal{G}$ and a non-empty set $\mathcal{D} \subset \mathbb{R}^{nd}$, consider the multi-agent system (4) over $\mathcal{G}$ with the gradient-based controller (13). Find a solution $V(x)$ to Problem 1 such that the equilibrium set $[Z(V)]_{\mathcal{D}}$ is asymptotically stable. ■

Remark 2: This paper mainly focuses on the local convergence similarly to the existing papers based on the gradient-flow approach as [18] because of the following reasons. i) Many important coordination problems including consensus are formulated with convex sets $\mathcal{D}$. Then, the objective function proposed later achieves the global asymptotic stability. ii) Although undesired local minima are often unavoidable for non-convex $\mathcal{D}$, we can achieve the global convergence via various methods based on gradients, e.g., [22]. The results in this paper are still important because these methods can be used in a distributed manner with our proposed methods. ■

III. J-OPTIMAL CONTROLLER

We design a J-optimal distributed controller based on the gradient-flow method for the generalized coordination to solve Problem 1. In this section, we particularly consider $\mathcal{D}$ given by a finite union of convex sets because of the differentiability of objective functions. The general case will be discussed in Section VII.

First, we characterize all gradient-distributed functions over the graph $\mathcal{G}$. This characterization will be a powerful tool to design distributed controllers.

Theorem 1: For a graph $\mathcal{G}$, a function $V : \mathbb{R}^{nd} \to \mathbb{R}$ is of class $\mathcal{F}_{gd}(\mathcal{G})$ if and only if there exist class $\mathcal{C}^1$ functions $W^{\mathcal{I}} : \mathbb{R}^{|\mathcal{I}|d} \to \mathbb{R}$ for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$ such that

$$V(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} W^{\mathcal{I}}(x_{\mathcal{I}}). \tag{19}$$

In particular, if $V(x)$ is non-negative, $W^{\mathcal{I}}(x_{\mathcal{I}})$ can be chosen to be non-negative for all $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$.

Proof: The sufficiency is shown. The partial derivatives of both sides in (19) are calculated as

$$\frac{\partial V}{\partial x_i}(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} \frac{\partial W^{\mathcal{I}}}{\partial x_i}(x_{\mathcal{I}}). \tag{20}$$

Consider a clique $\mathcal{I} \in \mathcal{S}_{clq,i}(\mathcal{G})$. For a node $j \in \mathcal{I}$, either of $j \in \mathcal{N}_i$ or $j = i$ holds from (8), which implies that

$$\bigcup_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} \mathcal{I} \subset \mathcal{N}_i \cup \{i\}. \tag{21}$$

Then, from (21), the right-hand side of (20) depends on $x_{\mathcal{N}_i\cup\{i\}}$, which can be represented as (15) with a function $f_i(x_i, x_{\mathcal{N}_i})$. This holds for any $i \in \mathcal{N}$, and $V(x)$ is of class $\mathcal{F}_{gd}(\mathcal{G})$.

The rest of the proof is shown in Section IV. ■

Theorem 1 implies that *all* gradient-distributed functions are characterized as (19) via free parameters $W^{\mathcal{I}}(x_{\mathcal{I}})$, each of which depends on the states of agents in one clique. Thanks to the necessity part of this theorem, when designing objective functions, we just have to consider the form (19) because, otherwise, the gradient-based controller (13) is not distributed. Therefore, all we have to do in solving Problem 1 is to find $W^{\mathcal{I}}(x_{\mathcal{I}})$ in (19) which minimizes $J(V)$. The key idea here is to use the projections of x and $\mathcal{D}$ onto the clique $\mathcal{I}$'s space. These projections are given by $x_{\mathcal{I}}$ and $P^{\mathcal{I}}(\mathcal{D})$ with the projection operator $P^{\mathcal{I}} : pow(\mathbb{R}^{nd}) \to pow(\mathbb{R}^{|\mathcal{I}|d})$ defined as

$$P^{\mathcal{I}}(\mathcal{D}) := \left\{ y \in \mathbb{R}^{|\mathcal{I}|d} | \exists x \in \mathcal{D} \text{ s.t. } y = x_{\mathcal{I}} \right\}. \tag{22}$$

Using these projections, the function

$$W^{\mathcal{I}}(x_{\mathcal{I}}) = \rho^{\mathcal{I}}\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D})\right) \tag{23}$$

yields the desired result, where $\rho^{\mathcal{I}}$ is a semi-distance function between $x_{\mathcal{I}}$ and $P^{\mathcal{I}}(\mathcal{D})$. Equation (23) quantifies the discrepancy between x and $\mathcal{D}$ on the clique $\mathcal{I}$'s space, whose geometric interpretation is given in Fig. 6 for $\mathcal{I} = \{1, 2\}$. See Example 3 for more details. Then, from Theorem 1, the following function is given as a candidate for solution to Problem 1.

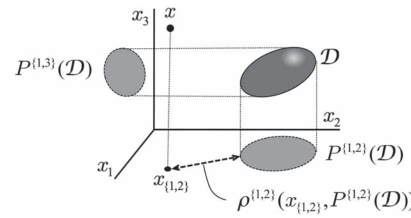

Fig. 6. Projections of x and $\mathcal{D}$ onto a clique's space.

Lemma 1: For a graph $\mathcal{G}$ and a non-empty set $\mathcal{D} \subset \mathbb{R}^{nd}$, the function

$$V_*(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \rho^{\mathcal{I}}\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D})\right) \tag{24}$$

with semi-distance functions $\rho^{\mathcal{I}}(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$ is of class $\mathcal{F}_z(\mathcal{D})$ and satisfies

$$J(V_*) \le J(V) \quad \forall V \in \mathcal{F}_{gd}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D}). \tag{25}$$

Moreover, if all semi-distance functions $\rho^{\mathcal{I}}(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ are of class $\mathcal{C}^1$, then $V_*(x)$ is of class $\mathcal{F}_{gd}(\mathcal{G})$.

Proof: For the simplicity of the proof, we assume that $P^{\mathcal{I}}(\mathcal{D})$ is closed for all $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$. It is easy to consider the other case.

First, the following inclusion relation is shown:

$$Z(V_*) \subset Z(V) \quad \forall V \in \mathcal{F}_{gd}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D}). \tag{26}$$

Consider a vector $y \in Z(V_*)$. Then, from (2) and (24), $y_{\mathcal{I}} \in P^{\mathcal{I}}(\mathcal{D})$ holds for any cliques $\mathcal{I}$. From (22), there exists a vector $z \in \mathcal{D}$ satisfying $z_{\mathcal{I}} = y_{\mathcal{I}}$ for each $\mathcal{I}$. Consider a function $V \in \mathcal{F}_{gd}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D})$. From Theorem 1, $V \in \mathcal{F}_{gd}(\mathcal{G})$ is of the form (19) with certain continuous functions $W^{\mathcal{I}}(x_{\mathcal{I}})$. From Theorem 1 and $V \in \mathcal{F}_z(\mathcal{D})$, every $W^{\mathcal{I}}(x_{\mathcal{I}})$ is non-negative, and $W^{\mathcal{I}}(x_{\mathcal{I}}) = 0$ holds for any $x \in \mathcal{D}$. Therefore

$$W^{\mathcal{I}}(y_{\mathcal{I}}) = W^{\mathcal{I}}(z_{\mathcal{I}}) = 0 \tag{27}$$

holds. Because (27) holds for any cliques $\mathcal{I}$, $V(y) = 0$, namely, $y \in Z(V)$ is satisfied from (19). Thus, (26) is proved.

Next, we show the inclusion relation

$$[Z(V_*)]_{\mathcal{D}} \subset [Z(V)]_{\mathcal{D}} \quad \forall V \in \mathcal{F}_{gd}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D}) \tag{28}$$

which is sufficient for (25). Consider $x \in [Z(V_*)]_{\mathcal{D}}$, then there is a path $\pi : [0, 1] \to Z(V_*)$ from $\pi(0) = x$ to $\pi(1) \in \mathcal{D}$. From (26), the range of the path π can be extended to $Z(V)$, and π is a path from $\pi(0) = x$ to $\pi(1) \in \mathcal{D}$ on $Z(V)$. Thus, $x \in [Z(V)]_{\mathcal{D}}$ holds, which indicates (28).

From (22) and (24), $V_* \in \mathcal{F}_z(\mathcal{D})$ holds because $V_*(x) = 0$ for all $x \in \mathcal{D}$. From Theorem 1, $V_* \in \mathcal{F}_{gd}(\mathcal{G})$ holds if all $\rho^{\mathcal{I}}(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ are of class $\mathcal{C}^1$. The proof is completed. ■

The optimal objective function $V_*(x)$ in (24) is not uniquely defined because of the options of the functions $\rho^{\mathcal{I}}$. However, the value of $J(V_*)$ is unique. Actually, from the definition of the semi-distance in (2)

$$\begin{aligned}Z(V_*) &= \bigcap_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \left\{x\in\mathbb{R}^{nd} | \rho^{\mathcal{I}}\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D})\right)=0\right\} \\ &= \bigcap_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \left\{x\in\mathbb{R}^{nd} | x_{\mathcal{I}}\in cl\left(P^{\mathcal{I}}(\mathcal{D})\right)\right\} \quad (29)\end{aligned}$$

is derived, and the performance index is calculated from (16) as

$$J(V_*) = dist_h\left(\left[\bigcap_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \left\{x\in\mathbb{R}^{nd} | x_{\mathcal{I}}\in cl\left(P^{\mathcal{I}}(\mathcal{D})\right)\right\}\right]_{\mathcal{D}}, \mathcal{D}\right)$$

which is independent from $\rho^{\mathcal{I}}$. The options of $\rho^{\mathcal{I}}$ can be used to achieve a desired convergence property of $x(t)$.

Example 3: Consider the elliptic target set $\mathcal{D}$ in Fig. 6 and the graph $\mathcal{G}=(\mathcal{N},\mathcal{E})$ with $\mathcal{N}=\{1,2,3\}$ and $\mathcal{E}=\{(1,2), (1,3), (2,1), (3,1)\}$. The set of cliques is given by $\mathcal{S}_{clq}(\mathcal{G}) = \{\{1,2\},\{1,3\},\{1\},\{2\},\{3\},\emptyset\}$. The coordinate $x_{\{1,2\}} = (x_1, x_2)$ and the ellipse $P^{\{1,2\}}(\mathcal{D})$ are the projection of the state x and the target set $\mathcal{D}$ onto the clique $\{1,2\}$'s space (i.e., the x_1-x_2 plane). The dashed arrow describes the discrepancy between x and $\mathcal{D}$ after the projection, which can be quantified by the semi-distance function $\rho^{\{1,2\}}(x_{\{1,2\}}, P^{\{1,2\}}(\mathcal{D}))$. In the same manner, $\rho^{\{1,3\}}(x_{\{1,3\}}, P^{\{1,3\}}(\mathcal{D}))$ is considered for $\{1,3\}$. From (29), the set $[Z(V_*)]_{\mathcal{D}}$ of the objective function (24) is the intersection of two elliptic cylinders whose bottoms are $P^{\{1,2\}}(\mathcal{D})$ and $P^{\{1,3\}}(\mathcal{D})$. This zero set is the closest to $\mathcal{D}$ than those of all other functions $V\in\mathcal{F}_{gd}(\mathcal{G})\cap\mathcal{F}_z(\mathcal{D})$ in the sense of the performance index $J(V)$.

From Lemma 1, the objective function $V_*(x)$ in (24) is a solution to Problem 1 if semi-distance functions $\rho^{\mathcal{I}}(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ are of class $\mathcal{C}^1$. The following lemma presents such a semi-distance function.

Lemma 2: For a convex set $\mathcal{D}\subset\mathbb{R}^n$ and a vector $x\in\mathbb{R}^n$

$$\rho(x,\mathcal{D}) = \left(dist(x,\mathcal{D})\right)^2 \quad (30)$$

is a semi-distance function between x and $\mathcal{D}$ and is of class $\mathcal{C}^1$. Its derivative is given by

$$\frac{\partial\rho(x,\mathcal{D})}{\partial x} = 2(x-\bar{y}) \quad (31)$$

where $\bar{y}$ is the closest point of $cl(\mathcal{D})$ from x, that is, satisfies $\|x-\bar{y}\|\le\|x-y\|$ for all $y\in cl(\mathcal{D})$.

Proof: See Appendix A. ■

The following lemma gives the convex properties of the sets projected by (22).

Lemma 3: For a convex set $\mathcal{D}\subset\mathbb{R}^{nd}$ and $\mathcal{I}\subset\mathcal{N}$, the sets $P^{\mathcal{I}}(\mathcal{D})$ and $\{x\in\mathbb{R}^{nd}|x_{\mathcal{I}}\in P^{\mathcal{I}}(\mathcal{D})\}$ are convex.

Proof: For $y,\hat{y}\in P^{\mathcal{I}}(\mathcal{D})$, we have to show that $sy+(1-s)\hat{y}\in P^{\mathcal{I}}(\mathcal{D})$ holds for any $s\in[0,1]$. From (22), there exist $x,\hat{x}\in\mathcal{D}$ such that $x_{\mathcal{I}}=y, \hat{x}_{\mathcal{I}}=\hat{y}$. Because $\mathcal{D}$ is convex, $sx+(1-s)\hat{x}\in\mathcal{D}$ holds for any $s\in[0,1]$. Then, $(sx+(1-s)\hat{x})_{\mathcal{I}}\in P^{\mathcal{I}}(\mathcal{D})$ from (22). Thus

$$(sx+(1-s)\hat{x})_{\mathcal{I}} = sx_{\mathcal{I}}+(1-s)\hat{x}_{\mathcal{I}} = sy+(1-s)\hat{y}\in P^{\mathcal{I}}(\mathcal{D})$$

is obtained, which completes the proof. ■

Consequently, a solution to Problem 1 is derived from (24) by appropriately choosing differentiable semi-distance functions.

Theorem 2: For a graph $\mathcal{G}$ and a non-empty set $\mathcal{D}\subset\mathbb{R}^{nd}$ given by a finite union of convex sets, the function

$$V_*(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \frac{k^{\mathcal{I}}}{2}\prod_{l=1}^{p}\left(dist\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}_l)\right)\right)^2 \quad (32)$$

is of class $\mathcal{F}_{gd}(\mathcal{G})\cap\mathcal{F}_z(\mathcal{D})$ and satisfies (25), where $k^{\mathcal{I}}$ are positive constants for $\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})$ and $\mathcal{D}_l\subset\mathbb{R}^{nd}$ are non-empty convex sets for $l\in\{1,2,\ldots,p\}$ such that $\mathcal{D}=\bigcup_{l=1}^{p}\mathcal{D}_l$.

Proof: Note the equations

$$P^{\mathcal{I}}(\mathcal{D}) = P^{\mathcal{I}}\left(\bigcup_{l=1}^{p}\mathcal{D}_l\right) = \bigcup_{l=1}^{p}P^{\mathcal{I}}(\mathcal{D}_l) \quad (33)$$

where $P^{\mathcal{I}}(\mathcal{D}_l)$ is convex for the convex set $\mathcal{D}_l$ from Lemma 3. Then, $\prod_{l=1}^{p}(dist(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}_l)))^2$ is a semi-distance function between $x_{\mathcal{I}}$ and $P^{\mathcal{I}}(\mathcal{D})$, and is of class $\mathcal{C}^1$ from Lemma 2. Thus, from Lemma 1, $V_*(x)$ in (32) satisfies (25) and is of class $\mathcal{F}_{gd}(\mathcal{G})\cap\mathcal{F}_z(\mathcal{D})$. The proof is completed. ■

Now, from Theorem 2, a J-optimal controller is derived. Let $\bar{y}^{\mathcal{I}l}$ be the closest point of $cl(P^{\mathcal{I}}(\mathcal{D}_l))$ from $x_{\mathcal{I}}$ as

$$\|x_{\mathcal{I}}-\bar{y}^{\mathcal{I}l}\|\le\|x_{\mathcal{I}}-y\| \quad \forall y\in cl\left(P^{\mathcal{I}}(\mathcal{D}_l)\right) \quad (34)$$

which is uniquely determined because $cl(P^{\mathcal{I}}(\mathcal{D}_l))$ is convex from Lemma 3. Then, the gradient of (32) is calculated as

$$\begin{aligned}\frac{\partial V_*}{\partial x}(x) &= \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})}\frac{k^{\mathcal{I}}}{2}\sum_{j=1}^{p}\frac{\partial\left(dist\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}_j)\right)\right)^2}{\partial x} \\ &\quad\times\prod_{l=1,l\ne j}^{p}\left(dist\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}_l)\right)\right)^2 \\ &= \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} k^{\mathcal{I}}\sum_{j=1}^{p}\bar{E}^{\mathcal{I}}(x_{\mathcal{I}}-\bar{y}^{\mathcal{I}j})\prod_{l=1,l\ne j}^{p}\|x_{\mathcal{I}}-\bar{y}^{\mathcal{I}l}\|^2\end{aligned} \quad (35)$$

from Lemma 2, where $\bar{E}^{\mathcal{I}} = E^{\mathcal{I}}\otimes I_d$ for the matrix $E^{\mathcal{I}}\in\mathbb{R}^{n\times|\mathcal{I}|}$ whose (i,k)-th component is given by

$$(E^{\mathcal{I}})_{ik} = \begin{cases}1, & \text{if } i=i_k\\ 0, & \text{otherwise}\end{cases} \quad (36)$$

for $i_1, i_2, \ldots, i_{|\mathcal{I}|}\in\mathcal{N}$ such that $\mathcal{I}=\{i_1, i_2, \ldots, i_{|\mathcal{I}|}\}$ and $i_1 < i_2 < \cdots < i_{|\mathcal{I}|}$. Then, the J-optimal controller of agent i is given from (13) for $V(x)=V_*(x)$ as

$$u_i = -\sum_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} k^{\mathcal{I}}\sum_{j=1}^{p}\left(x_i - \left(e_{|\mathcal{I}|k}^{\top}\otimes I_d\right)\bar{y}^{\mathcal{I}j}\right) \times\prod_{l=1,l\ne j}^{p}\|x_{\mathcal{I}}-\bar{y}^{\mathcal{I}l}\|^2 \quad (37)$$

where k is a positive integer such that $i=i_k$. This controller requires the information on x_i, $x_{\mathcal{I}}$ and $P^{\mathcal{I}}(\mathcal{D}_l)$ for cliques $\mathcal{I}$ which agent i belongs to. Since any agents in $\mathcal{I}$ are adjacent to agent i from (7), the subset $P^{\mathcal{I}}(\mathcal{D}_l)$ represents the desired coordination between agent i and its neighbors. Thus, (37) can be computed based on only local information.

Example 4: We show that a J-optimal distributed controller is derived in a systematic way. Consider the consensus problem in Example 2, where the target set $\mathcal{D}$ is given by the convex set (12), hence $p=1$ and $\mathcal{D}=\mathcal{D}_1$. Note that $\|x_{\mathcal{I}}-y\|$ is minimized for $y \in P^{\mathcal{I}}(\mathcal{D}_1)$ at

$$\bar{y}^{\mathcal{I}1} = \frac{1}{|\mathcal{I}|}\mathbf{1}_{|\mathcal{I}|} \otimes \left(\mathbf{1}_{|\mathcal{I}|}^{\top} \otimes I_d\right) x_{\mathcal{I}} = \frac{1}{|\mathcal{I}|}\mathbf{1}_{|\mathcal{I}|} \otimes \sum_{j\in\mathcal{I}} x_j. \quad (38)$$

Then, the J-optimal controller (37) is calculated as

$$\begin{aligned} u_i &= -\sum_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} k^{\mathcal{I}}\left(x_i - \left(e_{|\mathcal{I}|k}^{\top}\otimes I_d\right)\bar{y}^{\mathcal{I}1}\right) \\ &= -\sum_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} \frac{k^{\mathcal{I}}}{|\mathcal{I}|}\sum_{j\in\mathcal{I}\setminus\{i\}}(x_i - x_j) = -\sum_{j\in\mathcal{N}_i} a_{ij}(x_i - x_j) \end{aligned}$$

where $a_{ij} = \sum_{\mathcal{I}\in\mathcal{S}_{clq,ij}(\mathcal{G})} k^{\mathcal{I}}/|\mathcal{I}|$. ■

IV. Proof of Theorem 1

In this section, we show the rest of the proof of Theorem 1 after some preliminaries.

First, a mathematical operation, called *partial difference*, is introduced. Consider a function $g(x)\in\mathbb{R}^m$ for $x=[x_1^{\top}\ x_2^{\top}\ \cdots\ x_n^{\top}]^{\top}$ where $x_i\in\mathbb{R}^d$ for $i\in\{1,2,\ldots,n\}$. For a constant vector $c\in\mathbb{R}^d$, the partial difference $\Delta_c^{x_i}:\mathcal{C}^0\to\mathcal{C}^0$ of $g(x)$ with respect to x_i is defined as

$$\Delta_c^{x_i} g(x) := g(x) - g(x)|_{x_i=c} \quad (39)$$

where $g(x)|_{x_i=c} = g(x_1,x_2,\ldots,x_{i-1},c,x_{i+1},\ldots,x_n)$. The partial difference enables us to check the independence of a function from a variable without partial differentiation. Now, $g(x)$ is said to be *independent from* x_i if and only if there exists a function $\hat{g}:\mathbb{R}^{(n-1)d}\to\mathbb{R}^m$ such that

$$g(x) \equiv \hat{g}(x_1,x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_n). \quad (40)$$

Then, the following holds.

Lemma 4: A function $g(x)\in\mathbb{R}^m$ is independent from x_i if and only if $\Delta_c^{x_i} g(x)\equiv 0$ holds for a constant vector $c\in\mathbb{R}^d$.

Proof: Assume that $g(x)$ is independent from x_i and satisfies (40). Then, both of $g(x)$ and $g(x)|_{x_i=c}$ are equivalent to $\hat{g}(x_1,x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$, which leads to $\Delta_c^{x_i}g(x)\equiv 0$ from (39). Conversely, assume that $\Delta_c^{x_i}g(x)\equiv 0$ is satisfied. Then, $g(x)\equiv g(x)|_{x_i=c}$ holds from (39). We can substitute $\hat{g}(x_1,x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$ for $g(x)|_{x_i=c}$, and (40) is derived. The proof is completed. ■

The partial difference is commutative as follows.

Lemma 5: For constants vectors $c_i, c_j\in\mathbb{R}^d$, we obtain

$$\Delta_{c_i}^{x_i}\left(\Delta_{c_j}^{x_j} g\right)(x) = \Delta_{c_j}^{x_j}\left(\Delta_{c_i}^{x_i} g\right)(x). \quad (41)$$

Proof: From (39), the following equations hold

$$\begin{aligned} &\Delta_{c_i}^{x_i}\left(\Delta_{c_j}^{x_j} g\right)(x) = \Delta_{c_i}^{x_i}\left(g(x) - g(x)|_{x_j=c_j}\right) \\ &= \left(g(x) - g(x)|_{x_j=c_j}\right) - \left(g(x)|_{x_i=c_i} - g(x)|_{x_j=c_j,x_i=c_i}\right) \\ &= \left(g(x) - g(x)|_{x_i=c_i}\right) - \left(g(x)|_{x_j=c_j} - g(x)|_{x_i=c_i,x_j=c_j}\right) \\ &= \Delta_{c_j}^{x_j}\left(g(x) - g(x)|_{x_i=c_i}\right) = \Delta_{c_j}^{x_j}\left(\Delta_{c_i}^{x_i} g\right)(x) \end{aligned}$$

The proof is completed. ■

From the commutativity, we can define the high-order partial difference of the function $g(x)$ with respect to $x_{\mathcal{J}} = [x_{j_1}^{\top}\ x_{j_2}^{\top}\ \cdots\ x_{j_q}^{\top}]^{\top}$ for a same-dimensional constant vector $c_{\mathcal{J}} = [c_{j_1}^{\top}\ c_{j_2}^{\top}\ \cdots\ c_{j_q}^{\top}]^{\top}$ as

$$\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}} g(x) := \Delta_{c_{j_q}}^{x_{j_q}}\left(\Delta_{c_{j_{q-1}}}^{x_{j_{q-1}}}\cdots\left(\Delta_{c_{j_1}}^{x_{j_1}} g\right)\cdots\right)(x) \quad (42)$$

where $\mathcal{J}=\{j_1,j_2,\ldots,j_q\}$ and $q=|\mathcal{J}|$. Because of the commutativity, the order of the partial differences in (42) can be freely changed.

If the function $g(x)$ is scalar and of class $\mathcal{C}^1$, the partial difference can be described in the integral form as

$$\Delta_{\mu}^{x_{ik}} g(x) = \int_{\mu}^{x_{ik}} \frac{\partial g}{\partial x_{ik}}(x)\mathrm{d}x_{ik} \quad (43)$$

for a constant $\mu\in\mathbb{R}$ and $k\in\{1,2,\ldots,d\}$, where x_{ik} is the k-th component of x_i. Conversely, the partial derivative is described by the partial difference as

$$\frac{\partial g}{\partial x_{ik}}(x) = \lim_{\mu\to x_{ik}} \frac{\Delta_{\mu}^{x_{ik}} g(x)}{x_{ik}-\mu}. \quad (44)$$

Moreover, the commutativity between these two operations is valid as follows.

Lemma 6: For a class $\mathcal{C}^1$ scalar function $g(x)$ and a constant vector $c\in\mathbb{R}^d$, the following holds:

$$\Delta_c^{x_j}\frac{\partial g}{\partial x_i}(x) = \frac{\partial}{\partial x_i}\Delta_c^{x_j} g(x). \quad (45)$$

Proof: The k-th component of the left-hand side of (45) is calculated as

$$\begin{aligned} \Delta_c^{x_j}\frac{\partial g}{\partial x_{ik}}(x) &= \Delta_c^{x_j}\left(\lim_{\mu\to x_{ik}}\frac{\Delta_{\mu}^{x_{ik}} g(x)}{x_{ik}-\mu}\right) \\ &= \lim_{\mu\to x_{ik}}\frac{\Delta_{\mu}^{x_{ik}}\left(\Delta_c^{x_j} g\right)(x)}{x_{ik}-\mu} \\ &= \frac{\partial}{\partial x_{ik}}\Delta_c^{x_j} g(x) \end{aligned}$$

from (41) and (44). By taking this operation for all $k\in\{1,2,\ldots,d\}$, (45) is obtained. ■

Now, an important decomposition form of a scalar function is derived, which is valid for any function. Each term in the decomposition has a certain property described by the partial difference.

Lemma 7: For a function $V:\mathbb{R}^{nd}\to\mathbb{R}$, there exists a set family $\mathcal{S}_{\mathrm{com}}(V)\subset pow(\mathcal{N})$ with some functions $W^{\mathcal{I}}:\mathbb{R}^{|\mathcal{I}|d}\to\mathbb{R}$ for $\mathcal{I}\in\mathcal{S}_{\mathrm{com}}(V)$ satisfying

$$V(x) = \sum_{\mathcal{I}\in\mathcal{S}_{\mathrm{com}}(V)} W^{\mathcal{I}}(x_{\mathcal{I}}) \quad (46)$$

and

$$\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}} W^{\mathcal{I}}(x_{\mathcal{I}}) \begin{cases} \not\equiv 0, & \mathcal{I}=\mathcal{J} \\ \equiv 0, & \mathcal{I}\neq\mathcal{J} \end{cases} \quad (47)$$

for any $\mathcal{J}\in\mathcal{S}_{\mathrm{com}}(V)$. In particular, if $V(x)$ is of class $\mathcal{C}^1$ and/or non-negative, $W^{\mathcal{I}}(x_{\mathcal{I}})$ can be chosen to be of class $\mathcal{C}^1$ and/or non-negative for all $\mathcal{I}\in\mathcal{S}_{\mathrm{com}}(V)$.

Proof: See Appendix B. ■

Example 5: For $n=4$ and $d=1$, consider the function

$$V(x)=x_1x_2x_3+x_3x_4+x_2x_3+x_3.$$

Equations (46) and (47) hold with $\mathcal{S}_{\rm com}(V)=\{\{1,2,3\},\{3,4\}\}$ and the functions

$$W^{\{1,2,3\}}(x_1,x_2,x_3)=x_1x_2x_3+x_2x_3$$
$$W^{\{3,4\}}(x_3,x_4)=x_3x_4+x_3.$$

■

Rest of the Proof of Theorem 1: The necessity part is shown. We prove that a function $V\in\mathcal{F}_{gd}(\mathcal{G})$ is of the form (19). From Lemma 7, $V(x)$ is of the form (46) with class $\mathcal{C}^1$ functions $W^{\mathcal{I}}(x_{\mathcal{I}})$ satisfying (47) for $\mathcal{I}\in\mathcal{S}_{\rm com}(V)$. Then, we just have to show the relation

$$\mathcal{S}_{\rm com}(V)\subset\mathcal{S}_{clq}(\mathcal{G}) \tag{48}$$

which reduces (46) to (19).

To show (48) by contradiction, we assume that there exists a set $\mathcal{J}\subset\mathcal{N}$ such that $\mathcal{J}\in\mathcal{S}_{\rm com}(V)$ and $\mathcal{J}\notin\mathcal{S}_{clq}(\mathcal{G})$. From the assumption $\mathcal{J}\in\mathcal{S}_{\rm com}(V)$ and (47)

$$\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}}V(x)=\sum_{\mathcal{I}\in\mathcal{S}_{\rm com}(V)}\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}}W^{\mathcal{I}}(x_{\mathcal{I}})=\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}}W^{\mathcal{J}}(x_{\mathcal{J}})\not\equiv 0 \tag{49}$$

is derived. From the assumption $\mathcal{J}\notin\mathcal{S}_{clq}(\mathcal{G})$ and (7), there exist $j_a,j_b\in\mathcal{J}$ such that $(j_a,j_b)\notin\mathcal{E}$ and $j_a\neq j_b$. Then, from (6), $j_b\notin\mathcal{N}_{j_a}$ holds. Because we assume that $V(x)$ is of class $\mathcal{F}_{gd}(\mathcal{G})$, (15) holds for $i=j_a$. The partial difference of this equation with respect to x_{j_b} is reduced to

$$\Delta_{c_{j_b}}^{x_{j_b}}\frac{\partial V}{\partial x_{j_a}}=-\Delta_{c_{j_b}}^{x_{j_b}}f_{j_a}\left(x_{j_a},x_{\mathcal{N}_{j_a}}\right)\equiv 0 \tag{50}$$

from Lemma 4 because $f_{j_a}(x_{j_a},x_{\mathcal{N}_{j_a}})$ is independent from x_{j_b} for $j_b\notin\mathcal{N}_{j_a}$. By taking the integral of the k-th component of (50) with respect to x_{j_ak} for $k\in\{1,2,\dots,d\}$,we obtain

$$\int_{c_{j_ak}}^{x_{j_ak}}\Delta_{c_{j_b}}^{x_{j_b}}\frac{\partial V}{\partial x_{j_ak}}(x)\mathrm{d}x_{j_ak}=\int_{c_{j_ak}}^{x_{j_ak}}\frac{\partial}{\partial x_{j_ak}}\Delta_{c_{j_b}}^{x_{j_b}}V(x)\mathrm{d}x_{j_ak}$$
$$=\Delta_{c_{j_ak}}^{x_{j_ak}}\Delta_{c_{j_b}}^{x_{j_b}}V(x)\equiv 0 \tag{51}$$

where the first two equations are from (45) and (43), respectively. By taking (51) for all $k\in\{1,2,\dots,d\}$, $\Delta_{c_{j_a}}^{x_{j_a}}\Delta_{c_{j_b}}^{x_{j_b}}V(x)\equiv 0$ is derived. Then, from (41) and (42), the equations

$$\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}}V(x)=\Delta_{c_{\mathcal{J}\setminus\{j_a,j_b\}}}^{x_{\mathcal{J}\setminus\{j_a,j_b\}}}\Delta_{c_{j_a}}^{x_{j_a}}\Delta_{c_{j_b}}^{x_{j_b}}V(x)\equiv 0 \tag{52}$$

hold because $j_a,j_b\in\mathcal{J}$. Consequently, (49) and (52) contradict. Thus, (48) holds.

Finally, the property of the differentiability is satisfied from Lemma 7, which completes the proof. ■

V. Stability Analysis

The stability is analyzed for the multi-agent system (4) with the J-optimal controller (37) to solve Problem 2. In general, by using the gradient-based controller (13) with an objective function $V\in\mathcal{F}_z(\mathcal{D})$, the following result is achieved for the stability of $[Z(V)]_{\mathcal{D}}$. This is directly obtained from [23].

Lemma 8: For a non-empty set $\mathcal{D}\subset\mathbb{R}^{nd}$, consider the system (4) with the gradient-based controller (13), where $V(x)$ is of class $\mathcal{F}_z(\mathcal{D})\cap\mathcal{C}^1$. Assume that $[Z(V)]_{\mathcal{D}}$ is compact. If there exists an open set $\mathcal{O}_a\supset[Z(V)]_{\mathcal{D}}$ such that

$$V(x)>0 \text{ for } x\in\mathcal{O}_a\setminus[Z(V)]_{\mathcal{D}} \tag{53}$$

then $[Z(V)]_{\mathcal{D}}$ is stable. Moreover, if there exists an open set $\mathcal{O}_b\supset[Z(V)]_{\mathcal{D}}$ such that

$$\frac{\partial V}{\partial x}(x)\neq 0 \text{ for } x\in\mathcal{O}_b\setminus[Z(V)]_{\mathcal{D}} \tag{54}$$

then $[Z(V)]_{\mathcal{D}}$ is asymptotically stable. In addition, if (53) and (54) hold for $\mathcal{O}_a=\mathcal{O}_b=\mathbb{R}^{nd}$ and

$$V(x)\to\infty \text{ as } \|x\|\to\infty \tag{55}$$

is satisfied, then $[Z(V)]_{\mathcal{D}}$ is globally asymptotically stable. ■

Now, we consider the J-optimal controller (37) based on the objective function $V_*(x)$ in (32) and investigate the stability of $[Z(V_*)]_{\mathcal{D}}$. First, a structure of $Z(V_*)$ is given as follows.

Lemma 9: For a graph $\mathcal{G}$ and non-empty closed convex sets $\mathcal{D}_l\subset\mathbb{R}^{nd}$ for $l\in\{1,2,\dots,p\}$, the zero set of $V_*(x)$ in (32) with positive constants $k^{\mathcal{I}}$ for $\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})$ is described as

$$Z(V_*)=\bigcup_{\lambda\in\bar{\mathcal{P}}}\mathcal{K}_\lambda \tag{56}$$

for $\lambda=[\lambda_1\ \ \lambda_2\cdots\lambda_m]^\top\in\mathbb{R}^m$, where $m=|\mathcal{S}_{clq}(\mathcal{G})|$ is the number of the cliques in $\mathcal{G}$, $\mathcal{K}_\lambda\subset\mathbb{R}^{nd}$ is the closed convex set defined as

$$\mathcal{K}_\lambda:=\bigcap_{k=1}^{m}\left\{x\in\mathbb{R}^{nd}|x_{\mathcal{I}_k}\in cl\left(P^{\mathcal{I}_k}\left(\mathcal{D}_{\lambda_k}\right)\right)\right\} \tag{57}$$

for $\mathcal{S}_{clq}(\mathcal{G})=\{\mathcal{I}_1,\mathcal{I}_2,\dots,\mathcal{I}_m\}$, and $\bar{\mathcal{P}}$ is defined as

$$\bar{\mathcal{P}}:=\{\lambda\in\mathcal{P}^m|\mathcal{K}_\lambda\neq\emptyset\} \tag{58}$$

for $\mathcal{P}:=\{1,2,\dots,p\}$.

Proof: From (33) and the property of closed sets, the equations

$$cl\left(P^{\mathcal{I}}(\mathcal{D})\right)=cl\left(\bigcup_{l=1}^{p}P^{\mathcal{I}}(\mathcal{D}_l)\right)=\bigcup_{l=1}^{p}cl\left(P^{\mathcal{I}}(\mathcal{D}_l)\right)$$

hold. Define the set $\mathcal{S}_{kl}:=\{x\in\mathbb{R}^{nd}|x_{\mathcal{I}_k}\in cl(P^{\mathcal{I}_k}(\mathcal{D}_l))\}$, and from (29), the equations

$$Z(V_*)=\bigcap_{k=1}^{m}\left\{x\in\mathbb{R}^{nd}|x_{\mathcal{I}_k}\in\bigcup_{l=1}^{p}cl\left(P^{\mathcal{I}_k}(\mathcal{D}_l)\right)\right\}=\bigcap_{k=1}^{m}\bigcup_{l=1}^{p}\mathcal{S}_{kl}$$
$$=\bigcup_{\lambda_1=1}^{p}\mathcal{S}_{1\lambda_1}\cap\bigcup_{\lambda_2=1}^{p}\mathcal{S}_{2\lambda_2}\cdots\cap\bigcup_{\lambda_m=1}^{p}\mathcal{S}_{m\lambda_m}$$
$$=\bigcup_{\lambda_1,\lambda_2,\dots,\lambda_m=1}^{p}\bigcap_{k=1}^{m}\mathcal{S}_{k\lambda_k}=\bigcup_{\lambda\in\mathcal{P}^m}\mathcal{K}_\lambda \tag{59}$$

are obtained. Remove λ such that $\mathcal{K}_\lambda = \emptyset$ from $\mathcal{P}^m$ in (59), then (56) is achieved. Note that $cl(P^{\mathcal{I}_k}(\mathcal{D}_{\lambda_k}))$ is closed and convex from Lemma 3, and $\mathcal{K}_\lambda$ is closed and convex too from (57). ■

Then, $[Z(V_*)]_\mathcal{D}$ is proved to be stable.

Lemma 10: For a graph $\mathcal{G}$ and a set $\mathcal{D} \subset \mathbb{R}^{nd}$ given by a finite union of non-empty compact convex sets $\mathcal{D}_l \subset \mathbb{R}^{nd}$ for $l \in \{1, 2, \ldots, p\}$, consider the system (4) with the J-optimal controller (37) where $k^\mathcal{I}$ are positive constants for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$. Then, the equilibrium set $[Z(V_*)]_\mathcal{D}$ is stable for $V_*(x)$ given by (32).

Proof: The set $[Z(V_*)]_\mathcal{D}$ is an equilibrium set of the system because $\partial V_*/\partial x(x) = 0$ for $x \in [Z(V_*)]_\mathcal{D}$, hence the input (37) is zero from (35).

From Lemma 9, $Z(V_*)$ is of the form (56). Because $\mathcal{K}_\lambda$ in (57) is closed and convex, each $\mathcal{K}_\lambda$ is contained by either of $[Z(V_*)]_\mathcal{D}$ or $Z(V_*) \setminus [Z(V_*)]_\mathcal{D}$ in (56). Otherwise, there are points $x_a, x_b \in \mathcal{K}_\lambda$ such that $x_a \in [Z(V_*)]_\mathcal{D}$ and $x_b \in Z(V_*) \setminus [Z(V_*)]_\mathcal{D}$. Then, because $\mathcal{K}_\lambda$ is convex, there is a path from x_b to x_a such that $\pi(s) = sx_b + (1-s)x_a \in \mathcal{K}_\lambda$ for $s \in [0, 1]$. Then, for $x_a \in [Z(V_*)]_\mathcal{D}$, there is a path from $\mathcal{D}$ to x_a. Thus, there is a path from $\mathcal{D}$ to x_b, which contradicts $x_b \in Z(V_*) \setminus [Z(V_*)]_\mathcal{D}$. Therefore, there exist sets $\bar{\mathcal{P}}_a, \bar{\mathcal{P}}_b \subset \bar{\mathcal{P}}$ such that

$$[Z(V_*)]_\mathcal{D} = \bigcup_{\lambda \in \bar{\mathcal{P}}_a} \mathcal{K}_\lambda \quad (60)$$

$$Z(V_*) \setminus [Z(V_*)]_\mathcal{D} = \bigcup_{\lambda \in \bar{\mathcal{P}}_b} \mathcal{K}_\lambda \quad (61)$$

$$\bar{\mathcal{P}}_a \cap \bar{\mathcal{P}}_b = \emptyset,\ \bar{\mathcal{P}}_a \cup \bar{\mathcal{P}}_b = \bar{\mathcal{P}} \quad (62)$$

where $\mathcal{K}_{\lambda_a} \cap \mathcal{K}_{\lambda_b} = \emptyset$ holds for $\lambda_a \in \bar{\mathcal{P}}_a$ and $\lambda_b \in \bar{\mathcal{P}}_b$. Thus, the sets $[Z(V_*)]_\mathcal{D}$ and $Z(V_*) \setminus [Z(V_*)]_\mathcal{D}$ are disjoint.

Since $\{i\} \in \mathcal{S}_{clq}(\mathcal{G})$ holds for any $i \in \mathcal{N}$

$$Z(V_*) \subset \bigcap_{i \in \mathcal{N}} \left\{ x \in \mathbb{R}^{nd} | x_i \in cl\left(P^{\{i\}}(\mathcal{D})\right) \right\} \quad (63)$$

is derived from (29). From the assumption, $\mathcal{D}$ is compact, hence $P^{\{i\}}(\mathcal{D})$ is bounded for any $i \in \mathcal{N}$. Then, the right-hand side of (63) is bounded, and $Z(V_*)$ is bounded too. Moreover, from (60) and (61), the two sets $[Z(V_*)]_\mathcal{D}$ and $Z(V_*) \setminus [Z(V_*)]_\mathcal{D}$ are closed for closed sets $\mathcal{K}_\lambda$. Thus, the two sets are compact.

The two compact disjoint sets can be covered by two disjoint open sets $\mathcal{O}_a, \mathcal{O}_c \subset \mathbb{R}^{nd}$ as

$$[Z(V_*)]_\mathcal{D} \subset \mathcal{O}_a,\ \ Z(V_*) \setminus [Z(V_*)]_\mathcal{D} \subset \mathcal{O}_c,\ \ \mathcal{O}_a \cap \mathcal{O}_c = \emptyset. \quad (64)$$

Then, $(Z(V_*) \setminus [Z(V_*)]_\mathcal{D}) \cap \mathcal{O}_a = \emptyset$ holds. Therefore, $V_*(x) \neq 0$ holds for $x \in \mathcal{O}_a \setminus [Z(V_*)]_\mathcal{D}$, and (53) is satisfied for $V(x) = V_*(x)$. Thus, Lemma 8 guarantees that $[Z(V_*)]_\mathcal{D}$ is stable. ■

Now, $[Z(V_*)]_\mathcal{D}$ is proved to be asymptotically stable.

Theorem 3: For a graph $\mathcal{G}$ and a set $\mathcal{D} \subset \mathbb{R}^{nd}$ given by a finite union of non-empty compact convex sets $\mathcal{D}_l \subset \mathbb{R}^{nd}$ for $l \in \{1, 2, \ldots, p\}$, consider the system (4) with the J-optimal controller (37) where $k^\mathcal{I}$ are positive constants for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$. Then, the equilibrium set $[Z(V_*)]_\mathcal{D}$ is asymptotically stable for $V_*(x)$ given by (32).

Proof: In the proof of Lemma 10, $\bar{\mathcal{P}}_a$ is non-empty from (60) because $[Z(V_*)]_\mathcal{D}$ is non-empty for $V_* \in \mathcal{F}_z(\mathcal{D})$. Therefore, we can consider $z \in [Z(V_*)]_\mathcal{D} = \bigcup_{\lambda \in \bar{\mathcal{P}}_a} \mathcal{K}_\lambda$. Then, a non-empty set $\hat{\mathcal{P}} \subset \bar{\mathcal{P}}_a$ is defined as the set satisfying

$$z \in \bigcap_{\lambda \in \hat{\mathcal{P}}} \mathcal{K}_\lambda \text{ and } z \notin \bigcup_{\lambda \in \bar{\mathcal{P}}_a \setminus \hat{\mathcal{P}}} \mathcal{K}_\lambda. \quad (65)$$

We define the neighborhood of z as

$$\mathcal{U}_z := \{x \in \mathbb{R}^{nd} | \ \|x - z\| < \epsilon\} \setminus \bigcup_{\lambda \in \bar{\mathcal{P}}_a} \mathcal{K}_\lambda \quad (66)$$

where $\epsilon > 0$ is small enough to satisfy

$$\mathcal{U}_z \cap \bigcup_{\lambda \in \bar{\mathcal{P}}_b} \mathcal{K}_\lambda = \emptyset. \quad (67)$$

This is possible from (61), (64) and (65). Let $\Lambda_k \subset \mathcal{P}$ for $k \in \{1, 2, \ldots, m\}$ be the set of the k-th components λ_k of all $\lambda \in \hat{\mathcal{P}}$ as

$$\Lambda_k := \{l \in \mathcal{P} | \exists \lambda \in \hat{\mathcal{P}} \text{ s.t. } l = \lambda_k\}. \quad (68)$$

Because $\hat{\mathcal{P}}$ is non-empty, the following holds for any $k \in \{1, 2, \ldots, m\}$

$$|\Lambda_k| \geq 1. \quad (69)$$

Hereafter, we consider a certain $x \in \mathcal{U}_z$. From (62), (66) and (67), x satisfies

$$x \notin \mathcal{K}_\lambda \ \ \forall \lambda \in \bar{\mathcal{P}}. \quad (70)$$

Let $\mathcal{M} \subset \{1, 2, \ldots, m\}$ be the set

$$\mathcal{M} := \{k \in \{1, 2, \ldots, m\} | \ \left\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k l}\right\| \neq 0 \quad \forall\, l \in \mathcal{P}\} \quad (71)$$

where $\bar{y}^{\mathcal{I}l} \in cl(P^\mathcal{I}(\mathcal{D}_l))$ is the closest point of $cl(P^\mathcal{I}(\mathcal{D}_l))$ from $x_\mathcal{I}$, namely satisfies (34). $\mathcal{M}$ is non-empty because otherwise for any $k \in \{1, 2, \ldots, m\}$, there exists $\lambda_k^* \in \mathcal{P}$ such that $x_{\mathcal{I}_k} = \bar{y}^{\mathcal{I}_k \lambda_k^*}$. Then, for $\lambda^* = [\lambda_1^*\ \lambda_2^* \cdots \lambda_m^*]^\top \in \bar{\mathcal{P}}$, $x \in \mathcal{K}_{\lambda^*}$ holds from (57), which contradicts (70).

From (57) and (65), the inclusion $z_{\mathcal{I}_k} \in cl(P^{\mathcal{I}_k}(\mathcal{D}_{\lambda_k}))$ holds for any $\lambda \in \hat{\mathcal{P}}$ and $k \in \{1, 2, \ldots, m\}$. Thus, from (34)

$$\left\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k}\right\| \leq \|x_{\mathcal{I}_k} - z_{\mathcal{I}_k}\| \leq \|x - z\| < \epsilon \quad (72)$$

is derived. Moreover, from the hyperplane separation theorem

$$\begin{aligned} 0 &\geq \left\langle z_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k}, x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k} \right\rangle \\ &= \left\langle x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k} - (x_{\mathcal{I}_k} - z_{\mathcal{I}_k}), x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k} \right\rangle \\ &= \left\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k}\right\|^2 - \left\langle x_{\mathcal{I}_k} - z_{\mathcal{I}_k}, x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k} \right\rangle \end{aligned}$$

is obtained. Therefore, the following inequality holds:

$$\left\langle x_{\mathcal{I}_k} - z_{\mathcal{I}_k}, x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k} \right\rangle \geq \left\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k \lambda_k}\right\|^2. \quad (73)$$

Without the loss of generality, assume that $k^\mathcal{I} = 1$ for all $\mathcal{I}$. Then, from (35) and (71)

$$\frac{\partial V_*}{\partial x}(x) = \sum_{k \in \mathcal{M}} \sum_{j=1}^{p} \bar{E}^{\mathcal{I}_k} \alpha_k(x) \frac{x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}}{\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}\|^2} \quad (74)$$

is derived where

$$\alpha_k(x) = \prod_{l=1}^{p} \left\| x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k l} \right\|^2 > 0, k \in \mathcal{M}. \tag{75}$$

Now, we obtain

$$\begin{aligned} &\left\langle x - z, \frac{\partial V_*}{\partial x}(x) \right\rangle \\ &= \sum_{k\in\mathcal{M}} \alpha_k(x) \sum_{j=1}^{p} \left\langle x - z, \bar{E}^{\mathcal{I}_k} \frac{x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}}{\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}\|^2} \right\rangle \\ &= \sum_{k\in\mathcal{M}} \alpha_k(x) \sum_{j=1}^{p} \frac{\left\langle x_{\mathcal{I}_k} - z_{\mathcal{I}_k}, x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j} \right\rangle}{\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}\|^2} \\ &= \sum_{k\in\mathcal{M}} \alpha_k(x) \left(\sum_{j\in\Lambda_k} \frac{\left\langle x_{\mathcal{I}_k} - z_{\mathcal{I}_k}, x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j} \right\rangle}{\|x_{\mathcal{I}_k} - \bar{y}^{\mathcal{I}_k j}\|^2} + O(\epsilon) \right) \\ &\geq \sum_{k\in\mathcal{M}} \alpha_k(x) \left(\sum_{j\in\Lambda_k} 1 + O(\epsilon) \right) \\ &= \sum_{k\in\mathcal{M}} \alpha_k(x) \left(|\Lambda_k| + O(\epsilon) \right) \end{aligned} \tag{76}$$

where $O(\epsilon)$ is the Landau symbol. In (76), the third equation is from (68) and (72), and the inequality is from (73). From (69) and (75), the right-hand side of (76) is positive for a sufficiently small $\epsilon > 0$, which implies that

$$\frac{\partial V_*}{\partial x}(x) \neq 0 \quad \forall\, x \in \mathcal{U}_z. \tag{77}$$

Consider the set

$$\mathcal{O}_b = [Z(V_*)]_{\mathcal{D}} \cup \bigcup_{z\in[Z(V_*)]_{\mathcal{D}}} \mathcal{U}_z \tag{78}$$

which is an open set containing $[Z(V_*)]_{\mathcal{D}}$ from (60) and (66). Then, (54) holds for $V(x) = V_*(x)$ from (77). From Lemma 10, $[Z(V_*)]_{\mathcal{D}}$ is stable. Consequently, Lemma 8 guarantees that $[Z(V_*)]_{\mathcal{D}}$ is asymptotically stable. ■

If $\mathcal{D}$ is convex, the global asymptotic stability is guaranteed.

Theorem 4: In Theorem 3, if $p = 1$, then $[Z(V_*)]_{\mathcal{D}}$ is globally asymptotically stable.

Proof: First, from Lemma 9, $Z(V_*)$ is given by a convex set for $p=1$. Thus, $[Z(V_*)]_{\mathcal{D}} = Z(V_*)$ holds. Therefore, (53) is satisfied for $\mathcal{O}_a = \mathbb{R}^{nd}$. Second, in the proof of Theorem 3, all the relations in (76) hold without $O(\epsilon)$ for $p=1$. Thus, (54) holds for $\mathcal{O}_b = \mathbb{R}^{nd}$. Finally, consider a sequence $x^1, x^2, \ldots \in \mathbb{R}^{nd}$ such that $\|x^k\| \to \infty$ as $k \to \infty$. Then, there exists $i \in \mathcal{N}$ such that $\|x_i^k\| \to \infty$ where $x^k = [(x_1^k)^\top\ (x_2^k)^\top\ \cdots\ (x_n^k)^\top]^\top \in \mathbb{R}^{nd}$ and $x_j^k \in \mathbb{R}^d$ for $j \in \mathcal{N}$. From the assumption that $\mathcal{D}_1$ is compact, $P^{\{i\}}(\mathcal{D}_1)$ is bounded. Thus, $dist(x_i^k, P^{\{i\}}(\mathcal{D}_1)) \to \infty$ holds as $k \to \infty$. Because $\{i\} \in \mathcal{S}_{clq}(\mathcal{G})$, the inequality $V_*(x^k) \geq k^{\mathcal{I}}(dist(x_i^k, P^{\{i\}}(\mathcal{D}_1)))^2/2$ holds from (32). Thus, (55) is satisfied. The proof is completed from Lemma 8. ■

Remark 3: In Theorems 3 and 4, the compactness of $\mathcal{D}_l$ is assumed. If this assumption is not satisfied, we can consider a workspace for each agent without the loss of utility. Let a compact convex set $\mathcal{W}_i \subset \mathbb{R}^d$ be a large enough workspace of agent $i \in \mathcal{N}$, and consider the new objective function $V_*(x) + \sum_{i\in\mathcal{N}} k_i(dist(x_i, \mathcal{W}_i))^2$ for $k_i > 0$. Then, the stability properties are satisfied. ■

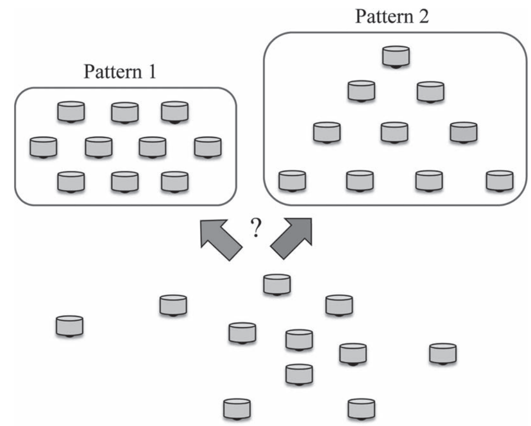

Fig. 7. Distributed pattern decision.

VI. Numerical Example

In this section, we illustrate a way to design a J-optimal distributed controller for an application to a coordination task called *distributed pattern decision*. The effectiveness of the designed controller is demonstrated though simulations.

Distributed pattern decision is a control task for the group of agents to decide and form a suitable formation pattern from given multiple patterns. Fig. 7 illustrates this task, where two formation patterns are prepared for rendezvous (Pattern 1) and travel (Pattern 2) while the agents are scattered around at the initial time. The agents can use sensors which measure the relative locations from the neighbors. There is no leader to determine the final pattern, nor communication devices to spread the information about which pattern is going to be formed.

We formulate this control task with the generalized coordination (10). Let $p \geq 2$ be the number of the prepared formation patterns. The set $\mathcal{P} = \{1, 2, \ldots, p\}$ denotes the set of the pattern indexes. For each index $l \in \mathcal{P}$, the desired relative locations between agents i and j are given by $\xi_{ij}^l \in \mathbb{R}^d$, and one of the patterns from $l \in \mathcal{P}$ is expected to be realized. Moreover, we expect that each pattern is possibly achieved from some initial states. Our control task is formulated as follows: For each $l \in \mathcal{P}$, there exists a set $\mathcal{X}_l \subset \mathbb{R}^{nd}$ whose measure (volume) is nonzero such that

$$x(0) \in \mathcal{X}_l \Rightarrow \lim_{t\to\infty} (x_i(t) - x_j(t)) = \xi_{ij}^l \quad \forall\, i, j \in \mathcal{N}. \tag{79}$$

This is equivalent to the local achievement of (10) with the target set

$$\mathcal{D} = \bigcup_{l\in\mathcal{P}} \mathcal{D}_l,\ \mathcal{D}_l = \{x \in \mathbb{R}^{nd} | x_i - x_j = \xi_{ij}^l \ \ \forall\, i, j \in \mathcal{N}\}. \tag{80}$$

The proposed objective function (32) is given as follows for the target set (80):

$$V_*(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} \frac{k^{\mathcal{I}}}{2|\mathcal{I}|^{2p}} \prod_{l=1}^{p} \sum_{h\in\mathcal{I}} \left\| \sum_{j\in\mathcal{I}\setminus\{h\}} \left(x_h - x_j - \xi_{hj}^l \right) \right\|^2 \tag{81}$$

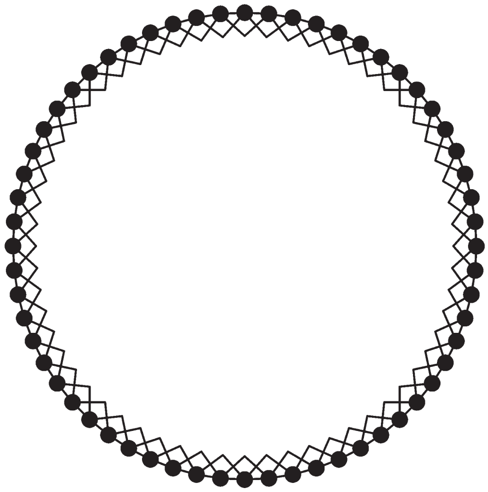

Fig. 8. 4-regular graph.

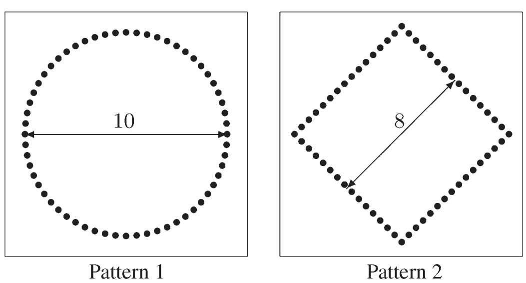

Fig. 9. Formation patterns in the simulations.

for $k^{\mathcal{I}} > 0$. Then, the gradient-based distributed controller (37) is derived as follows:

$$u_i = -\sum_{\mathcal{I}\in\mathcal{S}_{clq,i}(\mathcal{G})} \frac{k^{\mathcal{I}}}{|\mathcal{I}|^{2p-1}} \sum_{k=1}^{p} \sum_{j\in\mathcal{I}\setminus\{i\}} \left(x_i - x_j - \xi_{ij}^{k}\right) \times \prod_{l=1,l\neq k}^{p} \sum_{h\in\mathcal{I}} \left\| \sum_{j\in\mathcal{I}\setminus\{h\}} \left(x_h - x_j - \xi_{hj}^{l}\right) \right\|^2. \quad (82)$$

We demonstrate the effectiveness of the designed controller (82) through simulations. Let $n = 60$, then the node set is given by $\mathcal{N} = \{1, 2, \ldots, 60\}$ with the 4-regular graph depicted in Fig. 8. Let $d = 2$ be the dimension of the space, $p = 2$ be the number of the formation patterns. The two formation patterns are depicted in Fig. 9: the circle with a diameter of 10 and the square with sides of 8. These patterns define the corresponding relative locations $\xi_{ij}^1, \xi_{ij}^2 \in \mathbb{R}^d$ for $i, j \in \mathcal{N}$. Two simulations are performed from two different initial positions for the system (4) with the J-optimal controller (82) for $k^{\mathcal{I}} = |\mathcal{I}|^{2p-1}$. Figs. 10 and 11 show the simulation results of the transitions of the agents' positions from the two initial positions. From these figures, it is observed that the agents form the different formation patterns, circle and square assigned in Fig. 9 depending on the initial positions. It should be stressed that the agents share no global information including the patterns which will be formed, but can measure the relative positions of the neighbors over the network. From each individual motion, one of the assigned patterns is attained in each simulation.

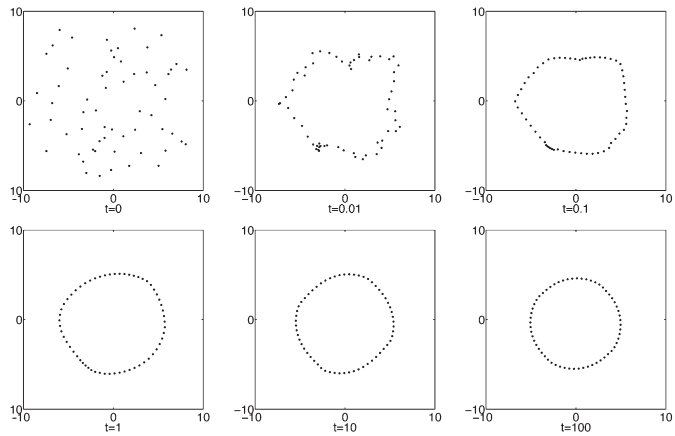

Fig. 10. Transitions of agents' positions from the first initial positions.

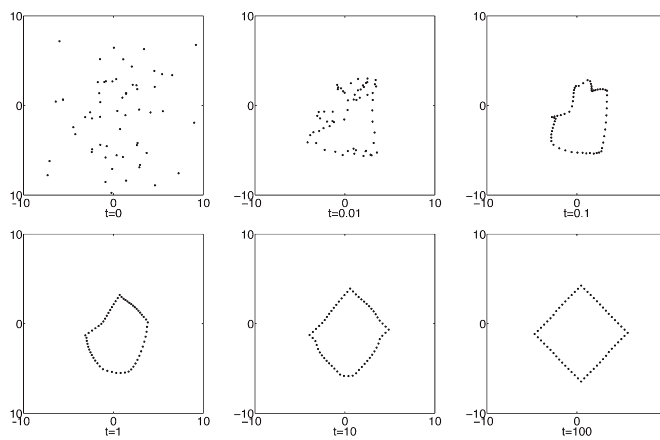

Fig. 11. Transitions of agents' locations from the second initial positions.

Remark 4: The distributed pattern decision is more difficult than the ordinary formation [7] because any patterns are acceptable and each agent has to guess which pattern will be formed. Mathematically, this difficulty is caused by the non-convexity of the target set (80). To overcome this difficulty, we propose the best objective function (81) out of all gradient-distributed functions, which minimizes the performance index $J(V)$. The state $x(t)$ still might be trapped in undesired minima, where some patterns are mixed. This is because only the local asymptotic stability is guaranteed for the non-convex $\mathcal{D}$ from Theorem 3. Nevertheless, compared with edge-based objective functions such as

$$V(x) = \sum_{(i,j)\in\mathcal{E}} k_{ij} \prod_{l=1}^{p} \left\| x_i - x_j - \xi_{ij}^{l} \right\|^2$$

the function (81) performs much better as shown in [24] because it includes less undesired zeros (i.e., $Z(V) \setminus \mathcal{D}$). Note that (81) is clique-based because it includes the products $\prod_{i\in\mathcal{I}} x_i$ of the agent states belonging to each clique $\mathcal{I}$. ■

VII. Extension to Non-Differentiable Objective Functions

In this section, the results in Section III are extended to the case of non-differentiable objective functions of class $\mathcal{L}_{\mathrm{loc}}$ (i.e., locally Lipschitz continuous). Then, the target set $\mathcal{D}$ does not necessarily have to be a union of convex sets.

Consider an objective function $V(x)$ of class $\mathcal{L}_{\rm loc}$. The solution $x(t)$ of the system (4) with (13) is regarded as the solution of the differential inclusion $\dot{x} \in -\partial V(x)$ in the sense of Filippov [25]. The set-valued function $\partial V : \mathbb{R}^{nd} \to pow(\mathbb{R}^{nd})$ called *generalized gradient* [26] is defined as

$$\partial V(x) := co\left(\left\{ \lim_{k\to\infty} \frac{\partial V}{\partial x}(x^k) \middle| \{x^k\}_{k=1,2,\ldots} \subset \mathbb{R}^{nd} \setminus \mathcal{U} \right.\right.$$
$$\left.\left. \text{s.t. } \lim_{k\to\infty} x^k = x \right\}\right)$$

for an arbitrary set $\mathcal{U} \subset \mathbb{R}^{nd}$ of measure zero including the points where $V(x)$ is not differentiable. We assume that the solution $x(t)$ exists in this sense for all $t \geq 0$ for each initial state $x(0)$. The function $V(x)$ of class $\mathcal{L}_{\rm loc}$ is called *gradient-distributed* over the graph $\mathcal{G}$ if there exist n functions $f_i : \mathbb{R}^d \times \mathbb{R}^{|\mathcal{N}_i|d} \to \mathbb{R}^d$ such that

$$\frac{\partial V}{\partial x_i}(x) = -f_i\left(x_i, x_{\mathcal{N}_i}\right) \quad \forall\, x \in \mathbb{R}^{nd} \setminus \mathcal{U}$$

for all $i \in \mathcal{N}$. Let $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G}) \subset \mathcal{L}_{\rm loc}$ be the set of all gradient-distributed functions over the graph $\mathcal{G}$ of class $\mathcal{L}_{\rm loc}$.

Now, we consider the class $\mathcal{L}_{\rm loc}$ version of Problem 1, that is, find a function $V(x)$ minimizing the performance index $J(V)$ in (16) among all functions in $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D})$. First, the characterization in Theorem 1 is compatible with class $\mathcal{L}_{\rm loc}$ functions as follows.

Theorem 5: For a graph $\mathcal{G}$, a function $V : \mathbb{R}^{nd} \to \mathbb{R}$ is of class $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G})$ if and only if there exist class $\mathcal{L}_{\rm loc}$ functions $W^{\mathcal{I}} : \mathbb{R}^{|\mathcal{I}|d} \to \mathbb{R}$ for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$ satisfying (19).

Proof: This can be proved by the same techniques as the proof of Theorem 1 thought the Lebesgue integration. ■

Then, a solution to Problem 1 for class $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G})$ functions is derived from Lemma 1. This result is available for any non-empty set $\mathcal{D}$ in contrast to Theorem 2.

Theorem 6: For a graph $\mathcal{G}$ and a non-empty set $\mathcal{D} \subset \mathbb{R}^{nd}$, the function

$$V_*(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G})} k^{\mathcal{I}} dist\left(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D})\right) \tag{83}$$

with positive constants $k^{\mathcal{I}}$ for $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$ is of class $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D})$ and satisfies

$$J(V_*) \leq J(V) \quad \forall\, V \in \mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G}) \cap \mathcal{F}_z(\mathcal{D}). \tag{84}$$

Proof: Because the set $\mathcal{D}$ is non-empty, so is $P^{\mathcal{I}}(\mathcal{D})$ for any $\mathcal{I} \in \mathcal{S}_{clq}(\mathcal{G})$. Thus, the function $dist(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ is of class $\mathcal{L}_{\rm loc}$ [27]. Then, from Theorem 5, $V_*(x)$ in (83) is of class $\mathcal{F}_{gd,\mathcal{L}_{\rm loc}}(\mathcal{G})$. Because $dist(x_{\mathcal{I}}, P^{\mathcal{I}}(\mathcal{D}))$ is a semi-distance function, $V_*(x)$ is of class $\mathcal{F}_z(\mathcal{D})$. Moreover, (84) is proved in the same way as Lemma 1 because the differentiability of $V_*(x)$ is not assumed in its proof. ■

The stability of the set $[Z(V)]_{\mathcal{D}}$ for $V \in \mathcal{L}_{\rm loc}$ can be analyzed as follows, which is corresponding to Lemma 8. This is a straightforward consequence of [28].

Theorem 7: For a non-empty set $\mathcal{D} \subset \mathbb{R}^{nd}$, consider the system (4) with the gradient-based controller (13), where $V(x)$ is of class $\mathcal{F}_z(\mathcal{D}) \cap \mathcal{L}_{\rm loc}$. Assume that $[Z(V)]_{\mathcal{D}}$ is compact. If there exists an open set $\mathcal{O}_a \supset [Z(V)]_{\mathcal{D}}$ satisfying (53), then $[Z(V)]_{\mathcal{D}}$ is stable. Moreover, if there exists an open set $\mathcal{O}_b \supset [Z(V)]_{\mathcal{D}}$ such that

$$0 \notin \partial V(x) \text{ for } x \in \mathcal{O}_b \setminus [Z(V)]_{\mathcal{D}}$$

then $[Z(V)]_{\mathcal{D}}$ is asymptotically stable. ■

Remark 5: The result of the characterization in Theorem 5 can be directly extended to time-varying graphs. For example, consider the state-dependent graph $\mathcal{G}(x) = \mathcal{G}_k$ for $x \in \mathcal{D}_k$, where $\mathcal{G}_k$ are graphs of the node set $\mathcal{N}$ and $\mathcal{D}_k \subset \mathbb{R}^{nd}$ are open subsets such that $\bigcup_{k=1,2,\ldots} cl(\mathcal{D}_k) = \mathbb{R}^{nd}$ and $\mathcal{D}_k \cap \mathcal{D}_l = \emptyset$ for $k, l = 1, 2, \ldots, k \neq l$. Then, a class $\mathcal{L}_{\rm loc}$ objective function $V(x)$ is gradient-distributed over the graph $\mathcal{G}(x)$ if and only if it is of the form

$$V(x) = \sum_{\mathcal{I}\in\mathcal{S}_{clq}(\mathcal{G}_k)} W^{\mathcal{I}}(x_{\mathcal{I}})$$

for $x \in \mathcal{D}_k$ for any $k = 1, 2, \ldots$. ■

VIII. Conclusion

This paper provided a unified solution for the general distributed control problem of multi-agent systems based on the gradient-flow approach. First, all gradient-based distributed controllers were characterized with a family of functions each of which depends on a clique. This paper showed that the complete characterization cannot be obtained by using the edges only, but it is possible by introducing the cliques. This essential feature was revealed by this paper for the first time. Next, we designed a J-optimal distributed controller for the generalized coordination, where the projection into the clique's space is the key to minimize the performance measure. This result enables us to design J-optimal distributed controllers in a systematic way from the network structure $\mathcal{G}$ and the target set $\mathcal{D}$. Moreover, the asymptotic stability of the system with the J-optimal distributed controller was shown. The effectiveness of the proposed method was demonstrated through the application to the distributed pattern decision. Because we focus on the gradient-flow approach and its local performance, we can reach the above fundamental results on multi-agent problems. The future work includes controller design enhancing the local performance and investigation of global performance by considering alternative stability analysis. Moreover, the design procedure of a J-optimal distributed controller for time-varying graphs will be developed through the characterization shown in Remark 5.

Appendix A
Proof of Lemma 2

For the simplicity, we denote $f(x) = \rho(x, \mathcal{D})$ in (30). The directional derivative of $f(x)$ at a point $x \in \mathbb{R}^n$ along a unit vector $v \in \mathbb{R}^n$ is defined as

$$(D_v f)(x) := \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(x+hv) - f(x)}{h}. \tag{85}$$

Consider $x \notin cl(\mathcal{D})$, and $x + hv \notin cl(\mathcal{D})$ holds for a sufficiently small h. Then

$$dist(x+hv, \mathcal{D}) = \|x - \bar{y}\| + \left\langle hv, \frac{x-\bar{y}}{\|x-\bar{y}\|} \right\rangle + O(h^2)$$

holds from [27]. Thus, from (85) and the fact that $dist(x, \mathcal{D}) = \|x - \bar{y}\|$

$$\begin{aligned}(D_{e_{ni}} f)(x) &= \lim_{\substack{h\to 0\\ h\neq 0}} \frac{2\|x-\bar{y}\| \left\langle he_{ni}, \frac{x-\bar{y}}{\|x-\bar{y}\|} \right\rangle + O(h^2)}{h} \\ &= 2\langle e_{ni}, x - \bar{y}\rangle = 2(x_i - \bar{y}_i) \end{aligned} \tag{86}$$

is achieved, where x_i and $\bar{y}_i$ are the i-th components of x and $\bar{y}$, respectively. For $x \in cl(\mathcal{D})$, $(D_{e_{ni}} f)(x) = 0$ holds from (85). Note that $x = \bar{y}$ holds for $x \in cl(\mathcal{D})$. Then, with (86), $\partial f / \partial x_i(x) = (D_{e_i} f)(x) = 2(x_i - \bar{y}_i)$ is achieved for any $x \in \mathbb{R}^n$, which is equivalent to (31). The proof is completed.

Appendix B
Proof of Lemma 7

The following two lemmas are given for preliminaries.

Lemma 11: Consider sets $\mathcal{I}, \mathcal{J}, \mathcal{K} \subset \mathcal{N}$, and a function $W(x_{\mathcal{I}}) \in \mathbb{R}$. If the relations $\mathcal{J} \not\subset \mathcal{K}$ and $\mathcal{I} \subset \mathcal{K}$ hold, the equation $\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}} W(x_{\mathcal{I}}) \equiv 0$ holds for any constant vectors $c_{\mathcal{J}} \in \mathbb{R}^{|\mathcal{J}|d}$.

Proof: From $\mathcal{I} \subset \mathcal{K}$ and $\mathcal{J} \not\subset \mathcal{K}$, the relation $\mathcal{J} \not\subset \mathcal{I}$ holds. Then, there exists an index $j \in \mathcal{J} \setminus \mathcal{I}$, and $\Delta_{c_j}^{x_j} W(x_{\mathcal{I}}) \equiv 0$ holds from Lemma 4. Then, from (41) and (42), the equations

$$\Delta_{c_{\mathcal{J}}}^{x_{\mathcal{J}}} W(x_{\mathcal{I}}) = \Delta_{c_{\mathcal{J}\setminus\{j\}}}^{x_{\mathcal{J}\setminus\{j\}}} \left(\Delta_{c_j}^{x_j} W\right)(x_{\mathcal{I}}) \equiv 0$$

hold, which completes the proof. ∎

Lemma 12: For a function $V : \mathbb{R}^{nd} \to \mathbb{R}$, there exists a set family $\mathcal{S}_{\text{com}}(V) \subset pow(\mathcal{N})$ with some functions $W^{\mathcal{I}} : \mathbb{R}^{|\mathcal{I}|d} \to \mathbb{R}$ for $\mathcal{I} \in \mathcal{S}_{\text{com}}(V)$ satisfying (46) and

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} W^{\mathcal{I}}(x_{\mathcal{I}}) \not\equiv 0 \quad \forall \mathcal{I} \in \mathcal{S}_{\text{com}}(V) \tag{87}$$

$$\mathcal{J} \not\subset \mathcal{I} \quad \forall \mathcal{I},\ \mathcal{J} \in \mathcal{S}_{\text{com}}(V) \text{ s.t. } \mathcal{I} \neq \mathcal{J}. \tag{88}$$

In particular, if $V(x)$ is of class $\mathcal{C}^1$ and/or non-negative, $W^{\mathcal{I}}(x_{\mathcal{I}})$ can be chosen to be of class $\mathcal{C}^1$ and/or non-negative for all $\mathcal{I} \in \mathcal{S}_{\text{com}}(V)$.

Proof: This lemma is proved later. ∎

Proof of Lemma 7: Lemma 7 is straightforward from Lemma 12 because the upper part of (47) is equivalent to (87) and the lower one is from (88) and Lemma 11 for $\mathcal{K} = \mathcal{I}$. ∎

The following is required to prove Lemma 12.

Lemma 13: For a set $\mathcal{I} = \{i_1, i_2, \ldots, i_m\}$ where $m = |\mathcal{I}|$, a function $V : \mathbb{R}^{nd} \to \mathbb{R}$ satisfies

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} V(x) \equiv 0 \tag{89}$$

for a constant vector $c_{\mathcal{I}} \in \mathbb{R}^{md}$ if and only if there exist functions $W_l : \mathbb{R}^{(n-1)d} \to \mathbb{R}$ for $l \in \{1, 2, \ldots, m\}$ satisfying

$$V(x) = \sum_{l=1}^{m} W_l\left(x_{\mathcal{N}\setminus\{i_l\}}\right). \tag{90}$$

In particular, if $V(x)$ is of class $\mathcal{C}^1$ and/or non-negative, $W_l(x_{\mathcal{N}\setminus\{i_l\}})$ can be chosen to be of class $\mathcal{C}^1$ and/or non-negative for all $l \in \{1, 2, \ldots, m\}$.

Proof: To show the sufficiency of the first part, consider $V(x)$ of the form (90). The partial difference of each term in the summation of (90) with respect to $x_{\mathcal{I}}$ is calculated as

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} W_l\left(x_{\mathcal{N}\setminus\{i_l\}}\right) = \Delta_{c_{\mathcal{I}\setminus\{i_l\}}}^{x_{\mathcal{I}\setminus\{i_l\}}} \Delta_{c_{i_l}}^{x_{i_l}} W_l\left(x_{\mathcal{N}\setminus\{i_l\}}\right) \equiv 0$$

from (41), (42) and Lemma 4. Thus, (89) holds.

Second, to show the necessity, we use the mathematical induction with respect to m. The case of $m = 1$ is directly from Lemma 4. Next, assume that this lemma holds for $m - 1$ instead of m, and we consider the case of m. Equation (89) is reduced to

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} V(x) = \Delta_{c_{\mathcal{I}\setminus\{i_m\}}}^{x_{\mathcal{I}\setminus\{i_m\}}} \Delta_{c_{i_m}}^{x_{i_m}} V(x) \equiv 0. \tag{91}$$

Note that $|\mathcal{I} \setminus \{i_m\}| = m - 1$, and the assumption of the mathematical induction is available for $\Delta_{c_{i_m}}^{x_{i_m}} V(x)$ in (91). Thus

$$\Delta_{c_{i_m}}^{x_{i_m}} V(x) = \sum_{l=1}^{m-1} W_l\left(x_{\mathcal{N}\setminus\{i_l\}}\right) \tag{92}$$

is obtained for certain functions $W_l : \mathbb{R}^{(n-1)d} \to \mathbb{R}$ for $l \in \{1, 2, \ldots, m-1\}$. Then, from (39), (92) is reduced to

$$V(x) = V(x)|_{x_{i_m} = c_{i_m}} + \sum_{l=1}^{m-1} W_l\left(x_{\mathcal{N}\setminus\{i_l\}}\right). \tag{93}$$

Substitute $W_m(x_{\mathcal{N}\setminus\{i_m\}})$ for $V(x)|_{x_{i_m}=c_{i_m}}$, and (90) is obtained. In this substitution, $W_m(x_{\mathcal{N}\setminus\{i_m\}}) = V(x)|_{x_{i_m}=c_{i_m}}$ is of class $\mathcal{C}^1$ and/or non-negative if $V(x)$ is class $\mathcal{C}^1$ and/or non-negative. The proof is completed. ∎

Proof of Lemma 12: We use the mathematical induction according to the number n. Assume that $V(x)$ is of class $\mathcal{C}^1$ and/or non-negative.

First, consider the case of $n = 1$. If $\Delta_{c_1}^{x_1} V(x) \not\equiv 0$ holds, $\mathcal{S}_{\text{com}}(V) = \{\{1\}\}$ satisfies (46), (87) and (88) with $W^{\mathcal{I}}(x_{\mathcal{I}}) = V(x)$. Otherwise, $V(x)$ is constant from Lemma 4, and $\mathcal{S}_{\text{com}}(V) = \{\emptyset\}$ satisfies (46), (87) and (88) with $W^{\mathcal{I}}(x_{\mathcal{I}}) = V(x)$. It is obvious that $W^{\mathcal{I}}(x_{\mathcal{I}})$ is of class $\mathcal{C}^1$ and/or non-negative.

Next, assume that Lemma 12 holds for $1, 2, \ldots, n-1$ instead of n, and show it for n. To this end, we show that for a repetition number $r \in \{1, 2, \ldots\}$, there exist set families $\mathcal{S}_\alpha^r$, $\mathcal{S}_\beta^r \subset pow(\mathcal{N})$ satisfying

$$\mathcal{J} \not\subset \mathcal{I} \quad \forall \mathcal{I},\ \mathcal{J} \in \mathcal{S}_\alpha^r \cup \mathcal{S}_\beta^r \text{ s.t. } \mathcal{I} \neq \mathcal{J} \tag{94}$$

$$\left|\mathcal{S}_\beta^r\right| \leq \left|\mathcal{S}_\beta^{r-1}\right| - 1,\ r \geq 2 \tag{95}$$

$$\forall \mathcal{I} \in \mathcal{S}_\beta^r, \exists \mathcal{J} \in \mathcal{S}_\beta^{r-1} \text{ s.t. } |\mathcal{I}| \leq |\mathcal{J}|,\ r \geq 2 \tag{96}$$

and class $\mathcal{C}^1$ and/or non-negative functions $W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}}), W_{\beta\mathcal{I}}^r(x_{\mathcal{I}})$ satisfying

$$V(x) = \sum_{\mathcal{I}\in\mathcal{S}_\alpha^r} W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}}) + \sum_{\mathcal{I}\in\mathcal{S}_\beta^r} W_{\beta\mathcal{I}}^r(x_{\mathcal{I}}) \tag{97}$$

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}}) \not\equiv 0 \quad \forall \mathcal{I} \in \mathcal{S}_\alpha^r \tag{98}$$

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}} W_{\beta\mathcal{I}}^r(x_{\mathcal{I}}) \equiv 0 \quad \forall \mathcal{I} \in \mathcal{S}_\beta^r. \tag{99}$$

It is obvious that from (95), there exists a natural number $\bar{r}$ such that $|\mathcal{S}_\beta^{\bar{r}}|=0$. Then, (94), (97) and (98) are corresponding to (88), (46), and (87) with $\mathcal{S}_{\rm com}(V)=\mathcal{S}_\alpha^{\bar{r}}$ and $W^{\mathcal{I}}(x_{\mathcal{I}})=W_{\alpha\mathcal{I}}^{\bar{r}}(x_{\mathcal{I}})$.

We show the above statements by the mathematical induction with respect to r. Consider $r=1$. If $\Delta_{c_{\mathcal{N}}}^{x_{\mathcal{N}}}V(x)\not\equiv 0$ holds, (94), (97), (98) and (99) are fulfilled with

$$\mathcal{S}_\alpha^1=\{\mathcal{N}\},\ \mathcal{S}_\beta^1=\{\emptyset\},\ W_{\alpha\mathcal{N}}(x_{\mathcal{N}})=V(x),\ W_{\beta\emptyset}(x_\emptyset)=0.$$

Otherwise, from Lemma 13, there exist class $\mathcal{C}^1$ and/or non-negative functions $W_l(x_{\mathcal{N}\setminus\{l\}})$ satisfying (90). Then, (94), (97), (98) and (99) are fulfilled with

$$\begin{aligned}
&\mathcal{S}_\alpha^1=\left\{\{\mathcal{N}\setminus\{l\}\}\,|\Delta_{c_{\mathcal{N}\setminus\{l\}}}^{x_{\mathcal{N}\setminus\{l\}}}W_l\left(x_{\mathcal{N}\setminus\{l\}}\right)\not\equiv 0\right\}\\
&\mathcal{S}_\beta^1=\left\{\{\mathcal{N}\setminus\{l\}\}\,|\Delta_{c_{\mathcal{N}\setminus\{l\}}}^{x_{\mathcal{N}\setminus\{l\}}}W_l\left(x_{\mathcal{N}\setminus\{l\}}\right)\equiv 0\right\}\\
&W_{\alpha\mathcal{N}\setminus\{l\}}=W_{\beta\mathcal{N}\setminus\{l\}}=W_l(x_{\mathcal{N}\setminus\{l\}}).
\end{aligned}$$

The relation (94) is satisfied because of the relation $\mathcal{N}\setminus\{l_1\}\not\subset\mathcal{N}\setminus\{l_2\}$ for $l_1,l_2\in\mathcal{N}$ such that $l_1\neq l_2$. Here, note that

$$|\mathcal{I}|=|\mathcal{N}\setminus\{l\}|=n-1,\ \mathcal{I}\in\mathcal{S}_\beta^1. \tag{100}$$

Assume that there exist set families $\mathcal{S}_\alpha^r,\mathcal{S}_\beta^r\subset pow(\mathcal{N})$ satisfying (94), (95) and (96) and functions $W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})$, $W_{\beta\mathcal{I}}^r(x_{\mathcal{I}})$ satisfying (97), (98) and (99). Now, we show that these conditions are fulfilled for $r+1$.

Consider a set $\mathcal{I}_0\in\mathcal{S}_\beta^r$. From (96) and (100), $|\mathcal{I}_0|\le n-1$ is satisfied. Therefore, from the assumption of the mathematical induction with respect to n, there exists a set family $\mathcal{S}_{\rm com}(W_{\beta\mathcal{I}_0}^r)\subset pow(\mathcal{I}_0)$ such that there exist class $\mathcal{C}^1$ and/or non-negative functions $W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}})$ satisfying (46), (87) and (88), i.e.,

$$W_{\beta\mathcal{I}_0}^r(x_{\mathcal{I}_0})=\sum_{\mathcal{I}\in\mathcal{S}_{\rm com}\left(W_{\beta\mathcal{I}_0}^r\right)}W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}}) \tag{101}$$

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}})\not\equiv 0\quad\forall\mathcal{I}\in\mathcal{S}_{\rm com}\left(W_{\beta\mathcal{I}_0}^r\right) \tag{102}$$

$$\mathcal{J}\not\subset\mathcal{I}\quad\forall\mathcal{I}\mathcal{J}\in\mathcal{S}_{\rm com}\left(W_{\beta\mathcal{I}_0}^r\right)\ \text{s.t.}\ \mathcal{I}\neq\mathcal{J}. \tag{103}$$

Categorize the elements of the set family $\mathcal{S}_{\rm com}(W_{\beta\mathcal{I}_0}^r)$ into

$$\mathcal{S}_{\mathcal{I}_0\mathcal{I}}^r:=\left\{\mathcal{J}\in\mathcal{S}_{\rm com}\left(W_{\beta\mathcal{I}_0}^r\right)|\mathcal{J}\subset\mathcal{I}\right\}\ \text{for}\ \mathcal{I}\in\mathcal{S}_\alpha^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\} \tag{104}$$

$$\mathcal{S}_{\mathcal{I}_0}^r:=\mathcal{S}_{\rm com}\left(W_{\beta\mathcal{I}_0}^r\right)\setminus\cup_{\mathcal{I}\in\mathcal{S}_\alpha^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\}}\mathcal{S}_{\mathcal{I}_0\mathcal{I}}^r. \tag{105}$$

Then, from (101), (97) is reduced to

$$V(x)=\sum_{\mathcal{I}\in\mathcal{S}_\alpha^r}\hat{W}_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})+\sum_{\mathcal{I}\in\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\}}\hat{W}_{\beta\mathcal{I}}^r(x_{\mathcal{I}})+\sum_{\mathcal{I}\in\mathcal{S}_{\mathcal{I}_0}^r}W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}}) \tag{106}$$

where $\hat{W}_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})$ and $\hat{W}_{\beta\mathcal{I}}^r(x_{\mathcal{I}})$ represent the class $\mathcal{C}^1$ and/or non-negative functions

$$\hat{W}_{\iota\mathcal{I}}^r(x_{\mathcal{I}})=W_{\iota\mathcal{I}}^r(x_{\mathcal{I}})+\sum_{\mathcal{J}\in\mathcal{S}_{\mathcal{I}_0\mathcal{I}}^r}W_{\beta\mathcal{I}_0\mathcal{J}}^r(x_{\mathcal{J}}) \tag{107}$$

where $\iota=\alpha,\beta$. Then, the $r+1$ version of (97) is given by (106) with

$$\mathcal{S}_\alpha^{r+1}=\mathcal{S}_\alpha^r\cup\mathcal{S}_{\mathcal{I}_0}^r\quad\text{for}\quad\mathcal{S}_\beta^{r+1}=\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\} \tag{108}$$

$$W_{\iota\mathcal{I}}^{r+1}(x_{\mathcal{I}})=\hat{W}_{\iota\mathcal{I}}^r(x_{\mathcal{I}})\quad\text{for}\quad\mathcal{I}\in\mathcal{S}_\iota^r\setminus\{\mathcal{I}_0\} \tag{109}$$

$$W_{\alpha\mathcal{I}}^{r+1}(x_{\mathcal{I}})=W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}})\quad\text{for}\quad\mathcal{I}\in\mathcal{S}_{\mathcal{I}_0}^r. \tag{110}$$

It is obvious that $W_{\iota\mathcal{I}}^{r+1}(x_{\mathcal{I}})$ is of class $\mathcal{C}^1$ and/or non-negative. In the rest of the proof, we will prove that (94)–(99) hold for $r+1$ instead of r.

First, the $r+1$ versions of (95) and (96) are directly from (108).

Next, consider the $r+1$ version of (98), which is given by

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\alpha\mathcal{I}}^{r+1}(x_{\mathcal{I}})=\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}\hat{W}_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})\not\equiv 0\ \text{for}\ \mathcal{I}\in\mathcal{S}_\alpha^r \tag{111}$$

$$\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\alpha\mathcal{I}}^{r+1}(x_{\mathcal{I}})=\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\beta\mathcal{I}_0\mathcal{I}}^r(x_{\mathcal{I}})\not\equiv 0\ \text{for}\ \mathcal{I}\in\mathcal{S}_{\mathcal{I}_0}^r \tag{112}$$

from (108), (109) and (110). Inequation (111) is obtained from the calculation of the partial difference (107) as

$$\begin{aligned}
\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}\hat{W}_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})&=\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})+\sum_{\mathcal{J}\in\mathcal{S}_{\mathcal{I}_0\mathcal{I}}^r}\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\beta\mathcal{I}_0\mathcal{J}}^r(x_{\mathcal{J}})\\
&=\Delta_{c_{\mathcal{I}}}^{x_{\mathcal{I}}}W_{\alpha\mathcal{I}}^r(x_{\mathcal{I}})\not\equiv 0.
\end{aligned} \tag{113}$$

The inequality is from (98), and the second equation is from Lemma 11 because $\mathcal{I}\not\subset\mathcal{I}_0$ and $\mathcal{J}\subset\mathcal{I}_0$ hold, where the first relation is from $\mathcal{I}\in\mathcal{S}_\alpha^r$, $\mathcal{I}_0\in\mathcal{S}_\beta^r$ and (94), and the second one is from $\mathcal{J}\in\mathcal{S}_{\mathcal{I}_0\mathcal{I}}^r\subset\mathcal{S}_{\rm com}(W_{\beta\mathcal{I}_0}^r)\subset pow(\mathcal{I}_0)$. On the other hand, (112) is directly from (102) and (105).

The proof of the condition corresponding to (99) is omitted because it is verified in the same way.

Consider the condition corresponding to (94) for $r+1$, that is

$$\mathcal{J}\not\subset\mathcal{I}\ \text{for}\ \mathcal{I},\mathcal{J}\in\left(\mathcal{S}_\alpha^{r+1}\cup\mathcal{S}_\beta^{r+1}\right)=\left(\mathcal{S}_\alpha^r\cup\mathcal{S}_{\mathcal{I}_0}^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\}\right)\ \text{s.t.}\ \mathcal{I}\neq\mathcal{J} \tag{114}$$

from (108). From (94) and (103), we just have to show the following two relations in (114):

$$\mathcal{J}\not\subset\mathcal{I}\ \text{for}\ \mathcal{I}\in\mathcal{S}_\alpha^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\},\mathcal{J}\in\mathcal{S}_{\mathcal{I}_0}^r \tag{115}$$

$$\mathcal{J}\not\subset\mathcal{I}\ \text{for}\ \mathcal{I}\in\mathcal{S}_{\mathcal{I}_0}^r,\mathcal{J}\in\mathcal{S}_\alpha^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\}. \tag{116}$$

The relation (115) is directly from (104) and (105). Because $\mathcal{I}\in\mathcal{S}_{\mathcal{I}_0}^r\subset\mathcal{S}_{\rm com}(W_{\beta\mathcal{I}_0}^r)\subset pow(\mathcal{I}_0)$ holds from (105), the relation $\mathcal{I}\subset\mathcal{I}_0$ holds. From (94) and $\mathcal{I}_0\in\mathcal{S}_\beta^r$, the relation $\mathcal{J}\not\subset\mathcal{I}_0$ holds for $\mathcal{J}\in\mathcal{S}_\alpha^r\cup\mathcal{S}_\beta^r\setminus\{\mathcal{I}_0\}$. The two relations are reduced to (116), which completes the proof.

Acknowledgment

The authors would like to thank the associate editor and the anonymous reviewers who helped improve the quality of the paper.

REFERENCES

[1] J. Shamma, *Cooperative Control of Distributed Multi-Agent Systems*. New York, NY, USA: Wiley-Interscience, 2008.

[2] N. A. Lynch, *Distributed Algorithms*. Burlington, MA, USA: Morgan Kaufmann, 1996.

[3] R. M. Murray, "Recent research in cooperative control of multivehicle systems," *J. Dynamic Syst., Meas. Control*, vol. 129, no. 5, pp. 571–583, 2007.

[4] B. Chen and H. H. Cheng, "A review of the applications of agent technology in traffic and transportation systems," *IEEE Trans. Intell. Transport. Syst.*, vol. 11, no. 2, pp. 485–497, 2010.

[5] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[6] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[7] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[8] R. Olfati-Saber and R. M. Murray, "Distributed cooperative control of multiple vehicle formations using structural potential functions," in *Proc. IFAC World Congress Autom. Control*, 2002, pp. 1–6.

[9] R. Olfati-Saber and R. M. Murray, "Graph rigidity and distributed formation stabilization of multi-vehicle systems," in *Proc. 41st IEEE Conf. Decision Control*, 2002, pp. 2965–2971.

[10] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, 2006.

[11] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 863–868, 2007.

[12] J. Cortés, S. Martínez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 243–255, 2004.

[13] J. Cortés, S. Martínez, and F. Bullo, "Spatially-distributed coverage optimization and control with limited-range interactions," *ESAIM: Control, Optim. Calculus Variations*, vol. 11, pp. 691–719, 2005.

[14] J. A. Marshall, M. E. Broucke, and B. A. Francis, "Formations of vehicles in cyclic pursuit," *IEEE Trans. Autom. Control*, vol. 49, no. 11, pp. 1963–1974, 2004.

[15] T.-H. Kim and T. Sugie, "Cooperative control for target-capturing task based on a cyclic pursuit strategy," *Automatica*, vol. 43, no. 8, pp. 1426–1431, 2007.

[16] Y. Igarashi, T. Hatanaka, M. Fujita, and M. W. Spong, "Passivity-based attitude synchronization in $SE(3)$," *IEEE Trans. Control Syst. Technol.*, vol. 17, no. 5, pp. 1119–1134, 2009.

[17] W. Ren, "Distributed cooperative attitude synchronization and tracking for multiple rigid bodies," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 2, pp. 383–392, 2010.

[18] S. Martínez, J. Cortés, and F. Bullo, "Motion coordination with distributed information," *IEEE Control Syst. Mag.*, vol. 27, no. 4, pp. 75–88, 2007.

[19] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst., Man. Cybern. B: Cybernetics*, vol. 39, no. 6, pp. 1393–1407, 2009.

[20] S. Bolognani and S. Zampieri, "A gossip-like distributed optimization algorithm for reactive power flow control," in *Proc. IFAC World Congress*, 2011, pp. 1–14.

[21] J. Cortés, S. Martínez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1289–1298, 2006.

[22] F. Aluffi-Pentini, V. Parisi, and F. Zirilli, "Global optimization and stochastic differential equations," *J. Optim. Theory Appl.*, vol. 47, no. 1, pp. 1–16, 1985.

[23] W. M. Haddad and V. Chellaboina, *Nonlinear Dynamical Systems and Control: A Lyapunov-based Approach*. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[24] K. Sakurama, S. Azuma, and T. Sugie, "Parameterization of all distributed controllers based on gradient-flow method for networked multi-agent systems," in *Proc. 51st IEEE Conf. Decision Control*, 2012, pp. 5465–5470.

[25] A. F. Filippov, *Differential Equations With Discontinuous Right-Hand Sides*. Norwell, MA, USA: Kluwer Academic Publishers, 1988.

[26] F. H. Clarke, "Generalized gradients and applications," *Trans. Amer. Math. Society*, vol. 205, pp. 247–262, 1975.

[27] M. Giaquinta and G. Modica, *Mathematical Analysis: Foundations and Advanced Techniques for Functions of Several Variables*. Boston, MA, USA: Birkhäuser, 2012.

[28] A. Bacciotti and F. Ceragioli, "Stability and stabilization of discontinuous systems and nonsmooth lyapunov functions," *ESAIM: Control, Optim. Calculus Variations*, vol. 4, pp. 361–376, 1999.

Kazunori Sakurama (M'12) received the B.S. degree in engineering, and the M.S. and Ph.D. degrees in informatics from Kyoto University, Kyoto, Japan in 1999, 2001, and 2004, respectively.

He was a Research Fellow of the Japan Society for the Promotion of Science from 2003 to 2004, an Assistant Professor at the University of Electro-Communications, Tokyo, Japan from 2004 to 2011 and a Program-Specific Researcher at Kyoto University in 2011. He is currently an Associate Professor with the Graduate School of Engineering, Tottori University, Tottori, Japan. His research interests include control of multi-agent systems, networked systems and nonlinear systems.

Dr. Sakurama received Control Division Conference Award in 2014 from the Society of Instrument and Control Engineers (SICE) in 2014. He is a member of the SICE and the Institute of Systems, Control and Information Engineers (ISCIE).

Shun-ichi Azuma (S'03–M'05) was born in Tokyo, Japan, in 1976. He received the B.Eng. degree in electrical engineering from Hiroshima University, Higashi Hiroshima, Japan, in 1999, and the M.Eng. and Ph.D. degrees in control engineering from Tokyo Institute of Technology, Tokyo, Japan, in 2001 and 2004, respectively.

He was a Research Fellow of the Japan Society for the Promotion of Science from 2004 to 2005 and an Assistant Professor in the Department of Systems Science, Graduate School of Informatics, Kyoto University, Uji, Japan, from 2005 to 2011. He is currently an Associate Professor at Kyoto University. He held visiting positions at Georgia Institute of Technology, Atlanta GA, USA, from 2004 to 2005 and at University of Pennsylvania, Philadelphia PA, USA, from 2009 to 2010. He became an Associate Editor of *IFAC Journal Automatica* in 2014. His research interests include analysis and control of hybrid systems.

Dr. Azuma was an Associate Editor of the IEEE CSS Conference Editorial Board in 2011 and the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS since 2013.

Toshiharu Sugie (F'07) received the B.E., M.E., and Ph.D. degrees in engineering form Kyoto University, Japan, in 1976, 1978 and 1985, respectively.

From 1978 to 1980, he was a research member of Musashino Electric Communication Laboratory, NTT, Musashino, Japan. From 1984 to 1988, he was a Research Associate, Department of Mechanical Engineering, University of Osaka Prefecture, Osaka. In 1988, he joined Kyoto University, where he is currently a Professor with the Department of Systems Science. He is an area editor for *Automatica*. His research interests are in robust control, identification, and control application to mechanical systems.

Dr. Sugie is a Fellow of the SICE.