2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

# Network Clustering: A Dynamical Systems and Saddle-Point Perspective

Mathias Bürger, Daniel Zelazo, and Frank Allgöwer

***Abstract*— This paper studies a class of cooperative networks that exhibit clustering in their steady-state behavior. We consider a collection of agents with heterogeneous dynamics and a bounded interaction rule between neighboring systems. We relate the steady state-behavior of the dynamical network to a static saddle-point problem. The saddle-point description of the system allows for a precise characterization of clustering. We show that the graph forms clusters along edges that are *saturated* and the corresponding cluster values depend only on these edges and the objective functions of each agent. We then provide a Lyapunov stability proof connecting the steady-state behavior of the dynamic system to the solution of the static saddle-point problem.**

## I. Introduction

Clustering is the phenomenon that in a dynamical network of interacting agents, the network partitions into several groups and all agents within the same group agree upon a common state. It is important to understand the mechanisms that lead to network clustering and to develop analytic tools providing information about where the network is most likely to split.

The literature relevant to this work can be partitioned into two broad groups. On one hand, research on the dynamic behavior and the stability of cooperative networks is relevant. Within the many contributions in this field, the works [1] and [2] are of particular interest due to the similarities in the network model we use for this work. On the other hand, research related to the general area of clustering behavior in dynamic networks is also relevant. The work [3] studies the aggregation of cooperative dynamic networks where the notion of time-scale separation between strongly connected groups is the main analytic tool. However, the heterogeneity of the agent dynamics is not taken into account to explain partitioning of the network. The contributions [4], [5] and [6] focus more on this aspect. However, the literature in this field is by far not as extensive as in the first group, and up to now clustering in dynamical networks seems to not be fully explained.

The contributions of this paper are as follows. We propose a class of dynamic network models that exhibit clustering in their steady-state behavior. We consider a collection of agents with heterogeneous dynamics that are coupled over a graph. A distinguishing feature in the model we adapt is that the interaction rule between neighboring agents are bounded. We establish a relation of the dynamical network model to a specific static saddle-point problem. The static saddle-point problem allows us to explicitly characterize network clustering. In particular, we show that the solution of the saddle-point problem leads to clustering behavior if a sub-set of the variables associated with the edges in the graph are *saturated*. The notion of saturation relates constraints that are active to the flow space of the graph. This result also allows us to explicitly characterize the value that each cluster obtains. We then show, using a Lyapunov stability proof, that the trajectories of the dynamics system will always converge to a solution of the saddle-point problem. This result allows to analyze the clustering behavior of the dynamical system by studying a static saddle point problem, and in this way to relate the properties of the nodes and of the network structure to the resulting clustering pattern.

The authors thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart and the Priority Programme 1305 “Control Theory of Digitally Networked Dynamical Systems”.

M. Bürger, D. Zelazo, and F. Allgöwer are with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, 70550 Stuttgart, Germany, {mathias.buerger, daniel.zelazo, frank.allgower}@ist.uni-stuttgart.de.

The remainder of the paper is organized as follows. In §II the notion of clustering is defined. The dynamical network is presented in §III. We then present in §IV a static saddle-point problem that explains network clustering. The connection between the dynamical network and the static saddle point problem is finally established in §V.

*Notation:* For a vector $x \in \mathbb{R}^n$, its transpose is given by $x'$ and the $i$th component by $x_i$; The $ij$th element of a matrix $A$ is denoted $[A]_{ij}$. The inner-product of two vectors is denoted $\langle x, y\rangle = x'y$; the standard Euclidean norm is $\|x\| = \langle x, x\rangle^{1/2}$. The null space and range space of a matrix is denoted as $\mathcal{N}(A)$ and $\mathcal{R}(A)$ respectively. The boundary of a set $\Gamma$ is denoted as $\partial\Gamma$ and the interior by $\mathrm{int}\Gamma$. The vector $\mathbb{1}$ is the vector of all ones.

## II. Preliminaries

Throughout this paper we consider systems defined over graphs [7]. A *graph* $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, consists of a set of *nodes*, $\mathbf{V} = \{v_1, \ldots, v_n\}$, and a set of *edges*, $\mathbf{E} = \{e_1, \ldots, e_m\}$ describing the incidence relation between pairs of nodes. The notation $v_i \sim v_j$ denotes that node $v_i$ is connected (or adjacent) to node $v_j$. Equivalently, $e = (v_i, v_j) \in \mathbf{E}$ is the *directed* edge connecting $v_i$ and $v_j$. A simple *path* in a graph is a sequence of distinct nodes such that consecutive nodes are adjacent to each other in the graph and each node is used once.A simple *cycle* in a graph is a path where the initial and terminal nodes are the same. A graph is *connected* if there exists a path between any pair of nodes; otherwise the graph is called *disconnected*. We also use the convention that an isolated vertex is a connected graph.

A graph $\mathcal{G}' = (\mathbf{V}', \mathbf{E}')$ is a *subgraph* of $\mathcal{G}$ if $\mathbf{V}' \subseteq \mathbf{V}$ and $\mathbf{E}' \subseteq \mathbf{E}$; equivalently, we write $\mathcal{G}' \subseteq \mathcal{G}$. Subgraphs can be induced by either a node set or an edge set. For example, the subgraph $\mathcal{P} \subseteq \mathcal{G}$ induced by the node set $\mathbf{P} \subseteq \mathbf{V}$ is the graph $\mathcal{P} = (\mathbf{P}, \mathbf{E}')$, with $\mathbf{E}' = \{e = (v_i, v_j) \,|\, v_i, v_j \in \mathbf{P}, e \in \mathbf{E}\}$. Similarly, the subgraph $\mathcal{Q} \subseteq \mathcal{G}$ induced by the edge set $\mathbf{Q} \subseteq \mathbf{E}$ is the graph $\mathcal{Q} = (\mathbf{V}', \mathbf{Q})$, with $\mathbf{V}' \subseteq \mathbf{V}$ the set of all nodes incident to the edges in $\mathbf{Q}$. A disconnected graph can be expressed as the union of connected subgraphs; each connected subgraph is referred to as a *component* of $\mathcal{G}$. Throughout this paper we follow the convention that bold-faced capital letters refer to sets, as in $\mathbf{V}$, and the script notation for graphs, as in $\mathcal{Q}$.

The *incidence matrix* $E(\mathcal{G}) \in \mathbb{R}^{|\mathbf{V}|\times|\mathbf{E}|}$ of the graph $\mathcal{G}$, is a $\{0, \pm1\}$ matrix with the rows and columns indexed by the vertices and edges of $\mathcal{G}$ such that $[E(\mathcal{G})]_{ik}$ has value ‘+1’ if node $i$ is the initial node of edge $k$, ‘-1’ if it is the terminal node, and ‘0’ otherwise. At times we will refer to the *flow space* and the *cut space* of the incidence matrix, defined as $\mathcal{N}(E(\mathcal{G}))$ and $\mathcal{R}(E(\mathcal{G})')$ respectively [7]. The cycles in a graph provide an important characterization of the flow space.

*Definition 2.1:* A *signed path vector* $\zeta \in \mathbb{R}^{|\mathbf{E}|}$ of a connected graph $\mathcal{G}$ corresponds to a path such that the $i$-th element of $\zeta$ takes the value '+1' if edge $i$ is traversed positively, '-1' if traversed negatively, and '0' if the edge is not used in the path.

*Theorem 2.2 ([7]):* Given a connected graph $\mathcal{G}$ with arbitrary orientation, the flow space $\mathcal{N}(E(\mathcal{G}))$ is spanned by all the linearly independent signed path vectors corresponding to the cycles in $\mathcal{G}$.

We now provide some definitions related to graph partitioning and clustering.

*Definition 2.3:* A *cluster* $\mathcal{P}$ is a connected subgraph of $\mathcal{G}$ induced by a node set $\mathbf{P} \subseteq \mathbf{V}$.

*Definition 2.4:* A *p-Partition* of the graph $\mathcal{G}$ is a collection of node sets $\mathbb{P} = \{\mathbf{P}_1, \ldots, \mathbf{P}_p\}$ with $\mathbf{P}_i \subseteq \mathbf{V}$, $\cup_{i=1}^{p} \mathbf{P}_i = \mathbf{V}$, and $\mathbf{P}_i \cap \mathbf{P}_j = \emptyset$ for all $\mathbf{P}_i, \mathbf{P}_j \in \mathbb{P}$, such that each subgraph $\mathcal{P}_i$ induced by the node sets $\mathbf{P}_i$ is connected.

Note that each subgraph $\mathcal{P}_i$ induced by a $p$-partition is also a cluster. At times we will also refer to the *p-cluster* of a graph to mean the set of subgraphs induced by a p-partition. For a connected graph $\mathcal{G}$, the union of all clusters induced by a partition will not reconstruct the original graph; that is, $\cup_{i=1}^{p} \mathcal{P}_i \subset \mathcal{G}$. This is formalized by the definition of a *cut-set*.

*Definition 2.5:* A *cut-set* of the graph $\mathcal{G}$ is a set of edges whose deletion leads to an increase in the number of connected components in $\mathcal{G}$.

Any $p$-partition of a graph will induce a cut-set (possibly empty, if $p = 1$). In this case, the cut-set is defined as

$$\mathbf{Q} = \{(v_i, v_j) \in \mathbf{E} | v_i \in \mathbf{P}_k, v_j \in \mathbf{P}_l, \forall \mathbf{P}_k, \mathbf{P}_l \in \mathbb{P}, k \neq l\}.$$

Similarly, a cut-set can be used to create p-partitions. Figure 1(a) is an example of a connected graph with cycles. Figure 1(b) shows the resulting induced clusters formed by the 3-partition $\mathbb{P} = \{\{v_1, v_2, v_3, v_4\}, \{v_5, v_6, v_7\}, \{v_8\}\}$. The cut set induced by this partition are the set of edges between each cluster, visualized as dotted edges in the graph.

Throughout this paper we associate scalar variables with each node and edge in a graph. For example, each component $x_i$ of the vector $x \in \mathbb{R}^{|\mathbf{V}|}$ is associated with a node $v_i \in \mathbf{V}$. Similarly, each component $z_i$ of a vector $z \in \mathbb{R}^{|\mathbf{E}|}$ is associated with an edge $e_i \in \mathbf{E}$. This can be used to provide an additional characterization of clusters and partitions of a graph.

*Definition 2.6:* A cluster $\mathcal{P}$ is in *$\epsilon$-agreement* if

$$\|x_k - x_l\| \leq \epsilon, \quad \text{for all } v_k, v_l \in \mathbf{P}.$$

The cluster is in an (exact) *agreement* with a *cluster value* $\beta$ if it is in $\epsilon$-agreement for $\epsilon = 0$.

For a vector $x \in \mathbb{R}^{|\mathbf{V}|}$ defined on the nodes of $\mathcal{G}$ and a subgraph $\mathcal{P} = (\mathbf{P}, \mathbf{E}') \subseteq \mathcal{G}$, we write $x(\mathcal{P}) \in \mathbb{R}^{|\mathbf{P}|}$ to denote the vector of all components $x_j$ associated with the nodes $v_j \in \mathbf{P}$; a similar notation is adopted for vectors defined on the edges.

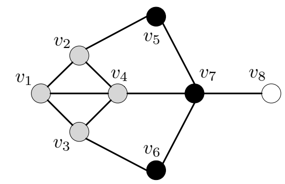
Using this notation, we can express the values of a cluster in exact agreement as $x(\mathcal{P}) = \beta \mathbb{1}_{|\mathbf{P}|}$.

## III. A Dynamic Network Model

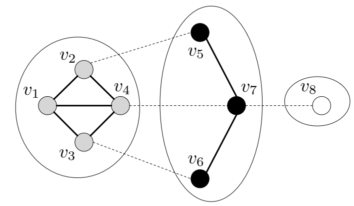
In the following, we consider a dynamical network defined on a graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ with $\mathbf{V} = \{v_1, \ldots, v_n\}$. Each node $v_i \in \mathbf{V}$ is associated with a dynamic *node state* $x_i(t) \in \mathbb{R}$ that evolves according to the scalar dynamical system,

$$\Sigma_i : \quad \dot{x}_i(t) = -\nabla J_i(x_i(t)) + u_i(t), \tag{1}$$

where $u_i(t) \in \mathbb{R}$ is an external input, and the nonlinear function $\nabla J_i(x_i(t))$ is the gradient of a *strongly convex objective function*



(a) A graph on 8 nodes.



(b) A 3-cluster of the graph.

Fig. 1. A connected graph on 8 nodes. The 3-partition $\mathbb{P} = \{\{v_1, v_2, v_3, v_4\}, \{v_5, v_6, v_7\}, \{v_8\}\}$ induces 3 clusters, shown in (b). The cut-set induced by this partition is $\mathbf{Q} = \{(v_2, v_5), (v_4, v_7), (v_3, v_6), (v_7, v_8)\}$.

$J_i(x_i(t))$. Strong convexity of $J_i(x_i(t))$ implies that there exists a $\eta_i > 0$ such that

$$(\nabla J_i(x_i) - \nabla J_i(\tilde{x}_i))(x_i - \tilde{x}_i) \geq \eta_i (x_i - \tilde{x}_i)^2, \forall x_i, \tilde{x}_i \in \mathbb{R}.$$

Coupling between each agent is realized through their control input. Each agent has access to relative state information for its control, defined over the edges of the graph $\mathcal{G}$. In this direction, we define the *network output* $y \in \mathbb{R}^{|\mathbf{E}|}$ of the system as,

$$y(t) = E(\mathcal{G})' x(t). \tag{2}$$

Associated with each edge $e_k \in \mathbf{E}$ is an *edge state* $z_k(t) \in \mathbb{R}$ that evolves according to the integrator dynamics driven by the network output,

$$\Pi_k \quad : \quad \begin{cases} \dot{z}_k(t) &= \quad y_k(t) \\ w_k(t) &= \quad \alpha_k \psi_k(z_k(t)) \end{cases} \tag{3}$$

The normalized nonlinear functions $\psi_k(z_k(t))$ vanish at the origin ($\psi_k(0) = 0$), are monotonically increasing ($z_k > \tilde{z}_k \Rightarrow \psi_k(z_k) > \psi_k(\tilde{z}_k)$), and are bounded,

$$\lim_{z_k \to \infty} \psi_k(z_k) = +1 \text{ and } \lim_{z_k \to -\infty} \psi_k(z_k) = -1. \tag{4}$$

The parameter $\alpha_k > 0$ can be interpreted as an *edge capacity*, and plays an important role in the clustering behavior of this network. The output of each edge dynamical system is used to generate the control for each agent, distributed by the incidence matrix as

$$u(t) = -E(\mathcal{G}) w(t). \tag{5}$$

Note that since $w(t)$ is bounded, the driving force $u(t)$ will also always be bounded.

Combining equations (1)-(5), we obtain the following closed-loop dynamical system,

$$\dot{x}(t) = -\nabla \mathbf{J}(x(t)) - E(\mathcal{G}) W \boldsymbol{\psi}(z(t)) \tag{6a}$$

$$\dot{z}(t) = E(\mathcal{G})' x(t), \tag{6b}$$

with $\nabla \mathbf{J}(x(t)) = [\nabla J_1(x_1(t)), \ldots, \nabla J_n(x_n(t))]'$, $\boldsymbol{\psi}(z) = [\psi_1(z_1(t)), \ldots, \psi_m(z_{|\mathbf{E}|}(t))]'$ and the diagonal matrix $W = \text{diag}([\alpha_1, \ldots, \alpha_{|\mathbf{E}|}])$ containing all the edge capacities. The considered network is depicted in Figure 2.

This model represents a broad class of coupled dynamical systems, and is in the general spirit of the class of problems studied in, for example, [1], [2]. An important distinction for the model used here is each agent has different dynamics and even distinct equilibria for $u_i(t) = 0$. Furthermore, the coupling functions $\psi_k$ are bounded. This implies that the attraction between neighboring nodes cannot be arbitrarily large. For the heterogeneous network (6) with unbounded coupling functions (e.g., $\alpha_k = \infty$), we have

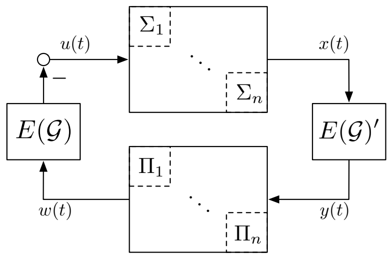
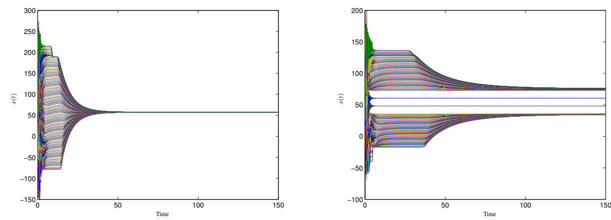



Fig. 2. Structure of the network model.



(a) . (b)

Fig. 3. Simulation of the node trajectories $x(t)$ for a random problem with different edge capacities: (a) trajectories converge to a 1-partition; (b) trajectories converge to a 4-partition.

shown in a previous work [8] that the nodes reach an agreement on a common value for the node states.

As mentioned above, the edge capacities play an important role for the behavior of the system (6). As a motivating preview, Figure 3 shows the trajectories of 100 agents, connected over a complete graph, running the protocol (6). All agents have randomly chosen quadratic objective functions $J_i(x_i)$ with minima at distinct points. The coupling functions are chosen as $\psi_k = \tanh(z_k)$. All initial conditions are chosen randomly and the edge capacities are equal for all edges of the graph, $W = \alpha I$. The two simulations differ only in the choice of the edge capacities $\alpha$. For a sufficiently large edge capacity $\alpha$, Figure 3(a) shows all agents reaching agreement (e.g., they form an exact 1-cluster). Reducing the edge capacity leads to a different steady-state behavior, as shown in Figure 3(b). In this case, we observe that the agents from an exact 4-cluster each with different cluster values, comprised of two large clusters and two isolated agents. Having qualitatively observed the clustering behavior of the proposed model we aim to characterize it more precisely.

The above example suggests that clustering behavior will depend on three parameters of the system (6): i) the interaction graph $\mathcal{G}$, (ii) the local objective functions $J_i$ of each agent, and (iii) the edge capacities $\alpha_k$. These observations motivate the main goals of this paper. Namely, we would like to answer the following questions:

- For which edge capacities $\alpha_k$ will the network achieve agreement (will the network form a 1-partition)?
- If the network forms a p-partition (p>1), along which edges will the network split?
- What will the cluster value $\beta_i$ for each cluster be?

The main analytic machinery we use to address these questions surprisingly come from a corresponding *static saddle-point problem*. In the sequel, we will discuss how a certain class of saddle-point problems lead to a clustering behavior of the optimization variables as a function of the constraint sets. Within the static problem set-up we are able to characterize certain properties of clusters. This will then lead to the main result of this work, connecting the static problem to the dynamic system (6).

## IV. Saddle-Point Problems and Static Network Clustering

The clustering phenomena in networks can be described by a static optimization problem that is intimately connected to the dynamical network (6). In particular, we examine the following static *max-min problem*, referred to as a *saddle-point problem*:

$$\max_{\mu\in\Gamma}\ \min_{\mathrm{x}}\ L(\mathrm{x},\mu) = \sum_{i=1}^{n} J_i(\mathrm{x}_i) + \mu' E(\mathcal{G})'\mathrm{x}, \qquad (7)$$

where $\mathrm{x} = [\mathrm{x}_1, \dots, \mathrm{x}_n]' \in \mathbb{R}^n$ are decision variables associated with each node in the graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, and $\mu = [\mu_1, \dots, \mu_{|\mathbf{E}|}]' \in \mathbb{R}^{|\mathbf{E}|}$ are variables associated with the edges in $\mathbf{E}$. The objective functions $J_i(\mathrm{x}_i)$ are the integral functions of the gradients $\nabla J_i$ appearing in the dynamics (1). We will sometimes abbreviate the notation writing $J(\mathrm{x}) := \sum_{i=1}^{n} J_i(\mathrm{x}_i)$. Note that the vector x here is a static vector while the state vector $x(t)$ of the system (6) is dynamic.

The *constraint set* $\Gamma = \Gamma_1 \times \cdots \times \Gamma_{|\mathbf{E}|}$, with $\Gamma_k = [-\alpha_k, \alpha_k]$, for some $\alpha_k > 0$, is a box constraint. Throughout this paper, except if explicitly stated otherwise, we assume that $0 < \alpha_k < \infty$ and thus $\Gamma$ is a compact and convex set. This notation is intentionally introduced to show a correspondence with the edge capacities defined for the dynamic network problem.

A special instance of the problem (7) occurs when the variables $\mu_k$ are *unconstrained*, corresponding to $\Gamma = \mathbb{R}^{|\mathbf{E}|}$. In this case, (7) corresponds to the *dual problem* of a corresponding network optimization problem. The primal problem in this case can be written as

$$\min_{\mathrm{x}}\ \sum_{i=1}^{n} J_i(\mathrm{x}_i) \quad \text{s.t. } E(\mathcal{G})'\mathrm{x} = 0. \qquad (8)$$

Denote by $(\mathrm{x}^*, \mu^*)$ a primal and dual solution to the problem (8). The $\mu_k^*$ variables are the dual variables of (8). The optimal solution $\mathrm{x}^*$ of (8) will always form a 1-cluster in agreement, i.e. $E(\mathcal{G})'\mathrm{x}^* = 0 \ \Rightarrow \mathrm{x}^* = \beta\mathbb{1}$, for some $\beta \in \mathbb{R}$. A optimal dual solution can then be obtained by the first order optimality condition $\nabla \mathbf{J}(\mathrm{x}^*) + E(\mathcal{G})\mu^* = 0$. This problem falls under a broad class of network optimization problems and consequently can be solved efficiently using a variety of methods (see, for example, [9]).

There is a game-theoretic interpretation of the saddle point problem. A decision maker in each node $v_i$ aims to minimize its individual objective function $J_i$; simultaneously, another decision maker, attached to an edge, penalizes any deviation between the decision variables of its incident nodes. For the problem (8), the dual variables associated with the constraints will force the decision makers on the nodes to reach an exact agreement on their values. However, the saddle-point problem formulation (7) does not permit the edge decision makers to arbitrarily penalize the deviation between neighboring agents, e.g. the penalty variables $\mu$ are restricted to be contained in the set $\Gamma$. This additional constraint has a strong impact on the structure of the primal solution x.

### A. Saddle Points

We provide here some properties of the saddle-points associated with (7).

*Definition 4.1:* A point $(\bar{x}, \bar{\mu})$ is a *saddle-point* of (7) if $\bar{\mu} \in \Gamma$ and $L(\bar{x}, \mu) \leq L(\bar{x}, \bar{\mu}) \leq L(x, \bar{\mu})$, for all $x \in \mathbb{R}^n, \mu \in \Gamma$.

In general there can be more than one saddle-point. We will denote the *set of all saddle-points* in the following by $\mathbb{X} \times \mathbb{M}$.

*Lemma 4.2:* The set of all saddle-points $\mathbb{X} \times \mathbb{M}$ for (7) is nonempty.

*Proof:* The set $\Gamma$ is nonempty, convex, and compact. The function $L : \mathbb{R}^n \times \Gamma \to \mathbb{R}$ is convex for each fixed $\mu \in \Gamma$, and concave for each $x \in \mathbb{R}^n$. Furthermore, for some $\bar{\mu} \in \Gamma$ and $\beta \in \mathbb{R}$ the level sets $\{x \in \mathbb{R}^n | \sum_{i=1}^n J_i(x_i) + \bar{\mu}' E(\mathcal{G})' x \leq \beta\}$ are nonempty and compact since each $J_i$ is a strongly convex function. The statement follows from the *Saddle Point Theorem* [10, Proposition 4.7]. ■

Saddle-points also admit some first-order optimality conditions. Let $(\bar{x}, \bar{\mu})$ be a saddle point, then ([11])

$$\nabla \mathbf{J}(\bar{x}) + E(\mathcal{G})\bar{\mu} = 0, \text{ and } \bar{x}' E(\mathcal{G})(\mu - \bar{\mu}) \leq 0, \; \forall \mu \in \Gamma. \quad (9)$$

*Lemma 4.3:* Let $(\bar{x}, \bar{\mu})$ be a saddle point of (7), then $\mathbb{X} = \{\bar{x}\}$ and $\mathbb{M} = \{\mu \in \Gamma \,|\, \mu = \bar{\mu} + \nu, \, \nu \in \mathcal{N}(E(\mathcal{G}))\}$

*Proof:* We first show the uniqueness of the saddle-point in the x-coordinate. Suppose $(\bar{x}, \bar{\mu})$ and $(\tilde{x}, \tilde{\mu})$ are both saddle-points with $\bar{x} \neq \tilde{x}$. Then $L(\bar{x}, \bar{\mu}) = L(\tilde{x}, \tilde{\mu})$ and furthermore, $J(\bar{x}) - J(\tilde{x}) = \tilde{\mu}' E(\mathcal{G})' \tilde{x} - \bar{\mu}' E(\mathcal{G})' \bar{x}$. The first-order optimality conditions state that $\tilde{\mu}' E(\mathcal{G})' \tilde{x} \geq \bar{\mu}' E(\mathcal{G})' \tilde{x}$ and $E(\mathcal{G})\bar{\mu} = -\nabla \mathbf{J}(\bar{x})$, implying that

$$\begin{aligned} J(\bar{x}) - J(\tilde{x}) &= \tilde{\mu}' E(\mathcal{G})' \tilde{x} - \bar{\mu}' E(\mathcal{G})' \bar{x} \\ &\geq \bar{\mu} E(\mathcal{G})' \tilde{x} - \bar{\mu}' E(\mathcal{G})' \bar{x} = -\nabla \mathbf{J}(\bar{x})(\tilde{x} - \bar{x}). \end{aligned}$$

On the other hand, due to the strong convexity of $J_i$, we have

$$J(\tilde{x}) - J(\bar{x}) \geq \nabla \mathbf{J}(\bar{x})'(\tilde{x} - \bar{x}) + \frac{m}{2}\|\tilde{x} - \bar{x}\|^2, \quad m > 0.$$

This leads to a contradiction, proving that $\bar{x} = \tilde{x}$ and thus $\mathbb{X} = \{\bar{x}\}$.

Let $(\bar{x}, \bar{\mu})$ be a saddle-point of (7). If $\mathcal{N}(E(\mathcal{G}))$ is non-trivial, then for any vector $\nu \in \mathcal{N}(E(\mathcal{G}))$, one has $L(\bar{x}, \bar{\mu}) = L(\bar{x}, \bar{\mu} + \nu)$. Any vector $\mu = \bar{\mu} + \nu \in \Gamma$ satisfies the saddle-point and first-order optimality conditions. ■

The result states that there is a unique vector $\bar{x}$ at which a saddle-point can be attained. However, the set of all saddle-points depends on the structure of the graph, and in particular its flow space.

*Lemma 4.4:* For $\Gamma = \mathbb{R}^m$, the set $\mathbb{M}$ contains more than one point if and only if $\mathcal{G}$ contains at least one cycle.

*Proof:* From Theorem 2.2, the flow space of $E(\mathcal{G})$ is non-trivial if and only if $\mathcal{G}$ contains at least one cycle. ■

### B. Network Clustering

Having established the existence and uniqueness properties of the saddle-points for (7), we now show how these solutions lead to clusters in the graph $\mathcal{G}$. First, we introduce the notion of a *saturated edge* in the graph.

*Definition 4.5:* An edge $e_k \in \mathbf{E}$ is said to be *saturated* if for all $\bar{\mu} \in \mathbb{M}$, $\bar{\mu}_k \in \partial\Gamma_k$ (e.g., $|\bar{\mu}_k| = \alpha_k$).

Note, however, that $\bar{\mu}_k \in \partial\Gamma_k$ for a particular $\bar{\mu}$ does not imply the edge is saturated. For an edge to be saturated, the constraint associated with that edge must be active for *all* possible saddle-points in the set $\mathbb{M}$. The following lemma connects the definition of saturated edges to graph properties.

*Lemma 4.6:* Any cycle in $\mathcal{G}$ contains either none or at least two saturated edges.

*Proof:* The statement is proven by contradiction. Assume that edge $e_k$ is the *only* saturated edge contained in a cycle with a corresponding signed path vector $\zeta$. Then $\zeta_k \neq 0$ and from Theorem 2.2, $\zeta \in \mathcal{N}(E(\mathcal{G}))$. From Lemma 4.3, there exists a $\delta \in \mathbb{R}$ sufficiently small such that $\tilde{\mu} = \bar{\mu} + \delta\zeta \in \mathbb{M}$ and $\tilde{\mu}_k \in \text{int}\Gamma_k$. But this is a contradiction to the definition of a saturated edge. Therefore, $e_k$ *cannot* be saturated. This implies that if a cycle contains a saturated edge, it must contain at least two saturated edges. ■

We now show that if the set $\mathbb{M}$ contains saturated edges, then there is a corresponding cut-set for the graph comprised of those edges.

*Lemma 4.7:* The set of saturated edges in $\mathbb{M}$ forms a cut-set for the graph.

*Proof:* First, assume that $\mathcal{G}$ is a spanning tree. Then $\mathbb{M} = \{\bar{\mu}\}$. Assume that at least one edge constraint is active, then every $\bar{\mu}_k \in \partial\Gamma_k$ is saturated, and its deletion results in an increase in the number of components, thus forming a cut-set.

Next, let $\mathcal{G}$ contain cycles. Assume that edge $e_k$ is saturated and is *not* contained in any cycle. Then its deletion results in an increase in the number of components in $\mathcal{G}$, and is included in a cut-set for the graph.

Now assume that a saturated edge $e_k$ is contained in one or more cycles. Then by Lemma 4.6 any cycle contains at least one other saturated edge. The deletion of two or more edges from a cycle results in an increase in the number of components in the graph, and thus each saturated edge in a cycle is included in a cut-set. ■

Lemma 4.7 makes a strong connection between the saddle-points of (7), saturated edges, and cut-sets. We are now able to state the main result of this section, relating clustering to saddle-points.

*Theorem 4.8:* Let $\mathbb{X} \times \mathbb{M}$ be the saddle-points of (7), and let $\mathbf{Q} \subseteq \mathbf{E}$ be the set of saturated edges. Then $\mathbf{Q}$ induces a $p$-partition $\mathbb{P} = \{\mathbf{P}_1, \ldots, \mathbf{P}_p\}$ and each cluster $\mathcal{P}_i$ induced by the set $\mathbf{P}_i$ is in exact agreement.

*Proof:* Let $(\bar{x}, \bar{\mu}) \in \mathbb{X} \times \mathbb{M}$ be a saddle-point with $\bar{\mu}_k \in \text{int}\Gamma_k$ for all non-saturated edges. Note that stating that cluster $\mathcal{P}_i$ is in agreement is equivalent to $E(\mathcal{P}_i)'\bar{x}(\mathcal{P}_i) = 0$. Assume, in order to arrive at a contradiction, that there exists some $P_i \in \mathbb{P}$ such that $E(\mathcal{P}_i)'\bar{x}(\mathcal{P}_i) \neq 0$. Denote by $\mathcal{Q}$ the subgraph induced by $\mathbf{Q}$. The function (7) can be written as

$$\begin{aligned} L(\bar{x}, \bar{\mu}) = &\sum_{i=1}^n J_i(\bar{x}_i) + \sum_{j=1}^p \bar{\mu}(\mathcal{P}_j)' E(\mathcal{P}_j)' \bar{x}(\mathcal{P}_j) + \\ &\bar{\mu}(\mathcal{Q})' E(\mathcal{Q})' \bar{x}(\mathcal{Q}). \end{aligned} \quad (10)$$

Since all clusters except $\mathcal{P}_i$ are assumed to be in agreement, the second term of (10) can be written as $\sum_{e_i = (v_k, v_l) \in \mathcal{P}_i} \bar{\mu}_i(\bar{x}_k - \bar{x}_l)$. Assume without loss of generality that only the edge $e_k = (v_i, v_j)$ in $\mathcal{P}_i$ connects two nodes that are not in agreement with a positive difference (e.g., $\bar{x}_i - \bar{x}_j > 0$). Then there exists an $\epsilon > 0$ such that $\bar{\mu}_k + \epsilon \in \Gamma_k$ and $\bar{\mu}_k(\bar{x}_i - \bar{x}_j) < (\bar{\mu}_k + \epsilon)(\bar{x}_i - \bar{x}_j)$. Let $\tilde{\mu}$ be the edge value after adding $\epsilon$ to only edge value $\bar{\mu}_k$ as described above. Then $L(\bar{x}, \bar{\mu}) \leq L(\bar{x}, \tilde{\mu})$, contradicting the assumption that $(\bar{x}, \bar{\mu})$ is a saddle-point. Therefore, each cluster $\mathcal{P}_i$ must be in agreement. ■

With this theorem, we already answered the second question posed in §III. That is, the critical edges along which the network will partition are related to the saturated edges in the static problem (7). More precisely, the network will partition along the saturated edges contained in the saddle points of (7).

Theorem 4.8 can be used to express any saddle-point $\bar{x}$ in the form $x(\mathcal{P}_i) = \beta_i \mathbb{1}, \quad i \in \{1, \ldots, p\}$, for some $\beta_i \in \mathbb{R}$. The clusters $\mathcal{P}_i$ are the connected components of the graph $\mathcal{G}$ after deleting all the saturated edges.

We are now also prepared to answer the remaining two original questions on network clustering for the static case. In the following, denote $\mathbb{X} \times \mathbb{M}$ as the set of saddle points for (7) and $\mathbb{X}^* \times \mathbb{M}^*$ as the primal and dual optimal solution sets to the network optimization problem (8).

*Lemma 4.9:* The solution $\bar{\mathrm{x}} \in \mathbb{X}$ of the saddle point problem (7) forms a 1-cluster in agreement if and only if $\Gamma \cap \mathbb{M}^* \neq \emptyset$.

*Proof:* If $\Gamma \cap \mathbb{M}^* \neq \emptyset$, there exists a $\mu^* \in \mathbb{M}^*$ which is also contained in $\Gamma$. Since $\mu^*$ is the optimal dual solution to (8), we know that the corresponding optimal primal solution is $\mathrm{x}^* = \beta \mathtt{1}$. This solution $(\mathrm{x}^*, \mu^*)$ satisfies the first order optimality conditions for (7). Thus $(\mathrm{x}^*, \mu^*)$ is also a saddle point of (7) and therefore $\bar{\mathrm{x}} = \beta \mathtt{1}$. Now suppose that $\bar{\mathrm{x}} = \beta \mathtt{1}$. There exists $\bar{\mu} \in \Gamma$ such that $\nabla \mathbf{J}(\beta \mathtt{1}) + E(\mathcal{G})\bar{\mu} = 0$. Since $\bar{\mathrm{x}}$ satisfies also the second condition of optimality for (8), $E(\mathcal{G})'\bar{\mathrm{x}} = \beta E(\mathcal{G})'\mathtt{1} = 0$. Thus $(\bar{x}, \bar{\mu})$ is also a saddle point of (8) and therefore $\bar{\mu} \in \mathbb{M}^*$. This shows that $\Gamma \cap \mathbb{M}^* \neq \emptyset$ and concludes the proof. ∎

The importance of Lemma 4.9 is that it provides a condition for achieving a 1-cluster for (7) in terms of the solution of an unconstrained network optimization problem. Considering that there are many efficient algorithms for solving (8), checking if the saddle-point problem achieves a 1-cluster is equivalent to solving the unconstrained problem.

We can now characterize the agreement values of the clusters. Assume that the network forms a $p$-cluster $(\mathcal{P}_1, \ldots, \mathcal{P}_p)$. We have already shown that $\bar{\mathrm{x}}(\mathcal{P}_i) = \beta_i \mathtt{1}$. For notational simplicity, define

$$J_{\mathcal{P}_i}(\beta_i) = \sum_{j \in \mathcal{P}_i} J_j(\beta_i) \text{ and } \nabla J_{\mathcal{P}_i}(\beta_i) = \sum_{j \in \mathcal{P}_i} \nabla J_j(\beta_i).$$

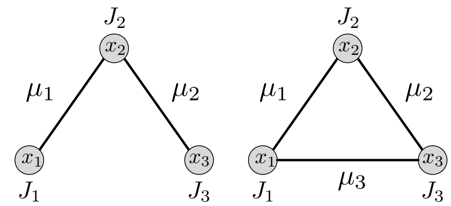
Note that both $J_{\mathcal{P}_i}$ and $\nabla J_{\mathcal{P}_i}$ are functions mapping $\mathbb{R}$ to $\mathbb{R}$. The notation $\nabla J_{\mathcal{P}_i}$ denotes the summation over all gradients within one cluster. The notation is introduced to prepare an alternative representation of the first order optimality condition (9) in terms of the network partitions. With this in mind, we also point out that the summation of $E(\mathcal{G})\bar{\mu}$ over one partition is $\mathtt{1}'E(\mathcal{P}_i)\bar{\mu}(\mathcal{P}_i) = 0$.

Recall also that each edge in the cut-set $\mathbf{Q}$ that induces the $p$-cluster is saturated; each component of the vector $\bar{\mu}(\mathcal{Q})$ is either $+\alpha_k$ or $-\alpha_k$. We define therefore the vector $y_{\mathcal{P}_i} = \nabla_{\mathcal{P}_i}\bar{\mu}(\mathcal{Q})'E(\mathcal{Q})'\bar{\mathrm{x}}(\mathcal{Q})$ where the notation $\nabla_{\mathcal{P}_i}$ means "take the gradient with respect to only the nodes in $\mathcal{P}_i$." This vector has a special structure. In particular, for all edges in $\mathbf{Q}$ that are not incident to any nodes in $\mathbf{P}_i$, the corresponding value of $y_{\mathcal{P}_i}$ is 0. Otherwise, the corresponding value of $y_{\mathcal{P}_i}$ is $\pm\alpha_k$ for an edge $k$ that is in the cut-set and incident to nodes in $\mathbf{P}_i$.

*Lemma 4.10:* The agreement value for the cluster $\mathcal{P}_i$ is given by $\beta_i = \nabla J^*_{\mathcal{P}_i}(-y'_{\mathcal{P}_i}\mathtt{1})$ where $J^*_{\mathcal{P}_i}$ is the convex conjugate of the function $J_{\mathcal{P}_i}$.

*Proof:* The statement follows from the first order conditions of optimality. We can rewrite the standard conditions of optimality using the cluster notation as $\nabla J_{\mathcal{P}_i}(\beta_i) + \bar{\mu}(\mathcal{P}_i)'E(\mathcal{P}_i)\mathtt{1} + y'_{\mathcal{P}_i}\mathtt{1} = 0, \quad i \in \{1, \ldots, p\}$ Note that the previous equation can be derived by taking the sum of all first order optimality conditions, which correspond to the nodes in the cluster $\mathcal{P}_i$. Considering now that $E(\mathcal{P}_i)'\mathtt{1} = 0$ and that the gradient of the convex conjugate function is the inverse of the gradient of the function ([12]), the statement follows directly. ∎

This result highlights the important property that the cluster agreement value $\beta_i$ depends only on the objective functions of the cluster nodes $J_{\mathcal{P}_i}$ and the edge capacities $\alpha_k$ of the edges separating the cluster from neighboring clusters. It is independent of the distribution of the dual variables within the cluster. Having given the saturated edges and the saturation bounds, one can easily predict the agreement value of the cluster.




(a) A path graph on 3 nodes. (b) A cycle graph on 3 nodes.

Fig. 4. Example illustrating the role of cycle edges in saddle-point solution.

We now present a simple example illustrating the implications of Lemma 4.7 and Theorem 4.8.

*Example 4.11:* We consider the two graphs, the path graph $\mathcal{G}_l$ in Figure 4(a) and the cycle graph $\mathcal{G}_c$ in Figure 4(b). The objective functions for each node are chosen as $J_i = \frac{1}{2}(x_i - \xi_i)^2$, with $\xi_1 = 10$, $\xi_2 = 5$, and $\xi_3 = 15$.

Let us first consider the two problems for $\Gamma = \mathbb{R}^{|\mathbf{E}|}$ to gain more insight into the solution. Denote by $(\mathrm{x}^*, \mu^*)$ the solution of the unconstrained problem. For the path graph, one can quickly compute $\mathrm{x}^* = 10$, $\mu_1^* = 0$, and $\mu_2^* = -5$. This solution is unique since $E(\mathcal{G}_l)$ has no cycles. The cycle graph $\mathcal{G}_c$ has the same solution $\mathrm{x}^* = 10$, $\mu_1^* = 0$, and $\mu_2^* = -5$ but with the additional edge value $\mu_3^* = 0$. The cycle graph, however, contains one cycle. In particular, the flow space can be expressed as $\mathcal{N}(E(\mathcal{G}_c)) = \beta[1\,1\,1]'$, $\beta \in \mathbb{R}$. Thus, the saddle points for $\mathcal{G}_c$ are $\mathrm{x}^* = 10$ and $\mu^* \in \mathbb{M}$ with

$$\mathbb{M} = \left\{ \mu \in \Gamma \mid \mu = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \beta \in \mathbb{R} \right\}.$$

Consider now the same problem with $\Gamma_i = [-\alpha, \alpha]$ for each edge, with $\alpha > 0$ (each edge has identical box constraints). Denote the solution to the constrained problem as $(\bar{\mathrm{x}}, \bar{\mu})$. Observe that for $\alpha \geq 5$, the constrained solution is identical to the unconstrained solution (for the cycle graph, the set of saddle-points is smaller). For $\alpha < 5$, edge $e_2$ in the path graph $\mathcal{G}_l$ will saturate and form the 2-partition $\mathbb{P} = \{\{v_1, v_2\}, \{v_3\}\}$.

In contrast, consider now the cycle graph $\mathcal{G}_c$ and choose $\alpha = 4.5$. Then there exists a $\tilde{\mu} \in \mathbb{M}$ (e.g., $\tilde{\mu} = [4, -1, 4]'$) for which no variable is on the boundary of $\Gamma$. Note that since $E(\mathcal{G}_c)\bar{\mu} = E(\mathcal{G}_c)\tilde{\mu}$, both $(\bar{x}, \bar{\mu})$ and $(\bar{x}, \tilde{\mu})$ with $\bar{x} = 10\,[1\,1\,1]'$ are saddle points. If $\alpha \leq 2.5$, then for the cycle graph, $\mathbb{M}$ collapses to the point $[2.5\ -2.5\,2.5]'$ and all three edges are considered saturated. The network will split into the 3-partition $\mathbb{P} = \{\{v_1\}, \{v_2\}, \{v_3\}\}$.

From this simple example, one can infer that the cycle graph is somehow more "robust" to capacity reductions on the edges. That is, the path graph will result in clustering before the cycle graph if the edge capacities are gradually reduced.

## V. Dynamic Network Clustering and Saddle-Point Problems

It remains now to connect the solution of the static problem (7) to the behavior of the dynamic system (6). A main contribution of this paper is the observation that the asymptotic behavior of the dynamic network (6) is directly connected to the saddle points of (7). The following theorem summarizes this result.

*Theorem 5.1:* Let $\mathbb{X} \times \mathbb{M}$ be the set of saddle points of problem (7). The trajectories $x(t)$ of (6a) and $w(t)$ of (6b) remain bounded and $\lim_{t\to\infty} x(t) \to \mathbb{X}, \quad \lim_{t\to\infty} w(t) \to \mathbb{M}$.

*Proof:* The proof relies on a Lyapunov-type argument. Let throughout the proof $(\bar{\mathrm{x}}, \bar{\mu})$ be a particular saddle point in $\mathbb{X} \times$

$\mathbb{M}$. Without loss of generality, assume that the solution of (7) forms a $p$-partition. First, consider the steady-state behavior of the dynamics (6a) $\dot{x}(t) = 0 = -\nabla \mathbf{J}(x(t)) - E(\mathcal{G})w(t)$. Note that this equilibrium has the same form as the first-order optimality condition for the saddle-point problem (7). Thus we know that $(\bar{\mathrm{x}}, \bar{\mu})$ is an equilibrium point for this dynamics in the sense $x(t) = \bar{\mathrm{x}}$ and $w(t) \in \mathbb{M}$. Note that such a point always exists since $w_k(t) \in [-\alpha_k, \alpha_k]$. Following the argumentation, we have for (6b) $\dot{z}(t) = 0 = E(\mathcal{G})'x(t)$.Contrary to the dynamics (6a), a saddle point $\bar{\mathrm{x}}$ in general does not give rise to an equilibrium for these dynamics. In particular, since $\bar{\mathrm{x}}$ corresponds to some $p$-partition, it holds in general that $E(\mathcal{G})'\bar{\mathrm{x}} \neq 0$.

We must therefore show that the solution $(x(t), z(t))$, even if $z(t)$ might be unbounded, converges such that $(x(t), w(t)) \to \mathbb{X} \times \mathbb{M}$. We will show this by the construction of an integral Lyapunov-like function. Define the variables $\bar{\mathrm{z}}_k = \psi_k^{-1}(\bar{\mu}_k/\alpha_k)$ if $|\mu_k| < \alpha_k$, $\bar{\mathrm{z}}_k = -\infty$ if $\bar{\mu}_k = -\alpha_k$, and $\bar{\mathrm{z}}_k = \infty$ if $\bar{\mu}_k = \alpha_k$ for $k \in \{1, \dots, |\mathbf{E}|\}$. Consider the following Lyapunov function candidate,

$$V = \frac{1}{2}\|x(t) - \bar{\mathrm{x}}\|^2 + \sum_{k=1}^{|\mathbf{E}|} \int_{\bar{\mathrm{z}}_k}^{z_k(t)} (\alpha_k \psi_k(s) - \bar{\mu}_k)\, ds. \qquad (11)$$

To begin, we show that $V(x, z)$ is positive-definite and vanishes at $x(t) = \bar{\mathrm{x}}$ and $z(t) = \bar{\mathrm{z}}$. Positive-definiteness of the first term is obvious and we must only verify it for the second summand.

Define $h_k(z_k) := \int_{\bar{\mathrm{z}}_k}^{z_k(t)} (\alpha_k \psi_k(s) - \bar{\mu}_k)\, ds$. It is clear that $h(\bar{z}_k) = 0$. Furthermore, the first derivative vanishes at $\bar{z}$, i.e. $\frac{\partial h_k}{\partial z_k}|_{z_k = \bar{\mathrm{z}}_k} = (\alpha_k \psi_k(\bar{\mathrm{z}}_k) - \bar{\mu}_k) = 0$. Finally, the second derivative is everywhere non-negative, $\frac{\partial^2 h_k}{\partial z_k^2} = \frac{\partial \psi_k}{\partial z_k} \geq 0$ due to the monotonic property of $\psi_k(z_k)$. Now we can conclude that $h_k(z_k)$ is a positive-definite function attaining its minimum at $z_k = \bar{\mathrm{z}}_k$. The function $V(x, z)$ is therefore a suitable Lyapunov function candidate.

To analyze the behavior of the system, we consider the directional derivative of $V$ along the trajectories

$$\dot{V} = (x - \bar{\mathrm{x}})'\dot{x} + \sum_{k=1}^{m} (\alpha_k \psi_k(z_k) - \bar{\mu}_k)\dot{z}$$
$$= (x - \bar{\mathrm{x}})'(-\nabla \mathbf{J}(x) - E(\mathcal{G})W\boldsymbol{\psi}(z)) + (W\boldsymbol{\psi}(z) - \bar{\mu})'E(\mathcal{G})'x.$$

We now add zero to obtain

$$\dot{V} = (x - \bar{\mathrm{x}})'(-\nabla \mathbf{J}(x) + \nabla \mathbf{J}(\bar{\mathrm{x}}) - \nabla \mathbf{J}(\bar{\mathrm{x}}) - E(\mathcal{G})'W\boldsymbol{\psi}(z)) + (W\boldsymbol{\psi}(z) - \bar{\mu})'E(\mathcal{G})'x.$$

Using the first order condition of optimality $\nabla \mathbf{J}(\bar{\mathrm{x}}) = -E(\mathcal{G})\bar{\mu}$ we can write this as $\dot{V} = -(x - \bar{\mathrm{x}})'(\nabla \mathbf{J}(x) - \nabla \mathbf{J}(\bar{\mathrm{x}}))$
$+ (x - \bar{\mathrm{x}})'(E(\mathcal{G})\bar{\mu} - E(\mathcal{G})W\boldsymbol{\psi}(z)) + (W\boldsymbol{\psi}(z) - \bar{\mu})'E(\mathcal{G})'x$.

Due to strong convexity of the objective functions, $(x - \bar{\mathrm{x}})'(\nabla \mathbf{J}(x) - \nabla \mathbf{J}(\bar{\mathrm{x}})) \geq \eta(x - \bar{\mathrm{x}})'(x - \bar{\mathrm{x}})$ one gets

$$\dot{V} \leq -\eta(x - \bar{\mathrm{x}})'(x - \bar{\mathrm{x}})$$
$$+(x - \bar{\mathrm{x}})'E(\mathcal{G})(\bar{\mu} - W\boldsymbol{\psi}(z)) + (W\boldsymbol{\psi}(z) - \bar{\mu})'E(\mathcal{G})'x,$$

and thus $\dot{V} \leq -\eta(x - \bar{\mathrm{x}})'(x - \bar{\mathrm{x}}) - \bar{\mathrm{x}}'E(\mathcal{G})(\bar{\mu} - W\boldsymbol{\psi}(z))$.

We now note that $\bar{x}'E(\mathcal{G})$ is a vector with one entry for each edge of $\mathcal{G}$. The entry $[\bar{x}'E(\mathcal{G})]_k$ of this vector is nonzero if and only if the edge $e_k$ is saturated ($|\bar{\mu}| = \alpha_k$). Thus, we can write

$$\dot{V} \leq -\bar{\mathrm{x}}(\mathcal{Q})'E(\mathcal{Q})(\bar{\mu}(\mathcal{Q}) - W(\mathcal{Q})\boldsymbol{\psi}(\mathcal{Q})) \qquad (12)$$

with all entries on the vector $\bar{\mu}(\mathcal{Q})$ having the value '+$\alpha_k$' or '-$\alpha_k$'.

We show now that the right hand side of (12) is non-positive. Note therefore that the entries of the vector $(\bar{\mu}(\mathcal{Q}) - W(\mathcal{Q})\boldsymbol{\psi}(\mathcal{Q}))$ have the same sign as the corresponding entries of the vector $\bar{\mu}(\mathcal{Q})$. In particular if $\bar{\mu}_k = +\alpha_k$ ($\bar{\mu}_k = -\alpha_k$) then $\bar{\mu}_k - \alpha_k \psi_k(z_k) > 0$ ($\bar{\mu}_k - \alpha_k \psi_k(z_k) < 0$) for all $z_k$.

We also note that the entries of the vector $E(\mathcal{Q})'\bar{\mathrm{x}}(\mathcal{Q})$ have the same sign as $\bar{\mu}(\mathcal{Q})$. This is a condition of optimality. If $\bar{\mu}_k[E(\mathcal{G})'\bar{\mathrm{x}}]_k \not\geq 0$ the solution $\bar{\mu}$ can not be optimal since a simple change in the sign would increase the value of $L(\bar{\mathrm{x}}, \bar{\mu})$. Taking these observations into account, one knows that each component

$$\left[\bar{\mathrm{x}}(\mathcal{Q})'E(\mathcal{Q})\right]_k \left[(\bar{\mu}(\mathcal{Q}) - W\boldsymbol{\psi}(\mathcal{Q}))\right]_k \geq 0$$

for all $z$ and consequently $\dot{V} \leq \eta(x - \bar{\mathrm{x}})'(x - \bar{\mathrm{x}})$.

All solutions converge to the set $\{x : x \equiv \bar{\mathrm{x}}\}$ [13]. It has already been shown that the saddle point $\bar{\mathrm{x}}$ is unique. By LaSalle's Invariance Principle [13] we can now conclude that all solutions converge to the largest invariant set contained in $\{x : x \equiv \bar{\mathrm{x}}\}$. Thus, all solutions approach the set $\{(x, z) : x = \bar{\mathrm{x}}, W\boldsymbol{\psi}(z) = \bar{\mu} + \nu,\ \nu \in \mathcal{N}(E)\}$. This proves the theorem. ■

The result establishes a direct connection between the dynamic network (6) and the static saddle point problem (7). We have already used the properties of the static problem to analyze the clustering behavior. Performing the analysis in the static problem is significantly easier. We have then directly used the properties of the static network clustering to characterize the behavior of the dynamic network. Naturally, all the results on the properties of the network clustering which are presented in §IV hold in the same way for the limiting behavior of the dynamic network.

## VI. Conclusions

This work examined the phenomena of clustering within a coupled dynamical system. We considered a network model and showed that its trajectories converge to the solution of an associated saddle-point problem. This result provides an explicit connection between static optimization problems and dynamic systems in a networked setup. This relation allows for a direct analysis of possible clustering properties of the dynamic network.

## References


[1] D. Zelazo and M. Mesbahi, "Edge Agreement: Graph-theoretic Performance Bounds and Passivity Analysis," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 554–555, 2010.

[2] M. Arcak, "Passivity as a design tool for group coordination," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1380–1390, 2007.

[3] E. Biyik and M. Arcak, "Area aggregation and time scale modeling for sparse nonlinear networks," *Systems and Control Letters*, vol. 57, no. 2, pp. 142–149, 2008.

[4] F. D. Smet and D. Aeyeles, "Clustering in a network of non-identical and mutually interacting agents," *Proceedings of the Royal Society A*, vol. 465, pp. 745–768, 2009.

[5] ——, "Cluster transition in a multi-agent clustering model," in *Proc. of the 48th IEEE Conf. on Dec. and Contr.*, 2009, pp. 4778–4784.

[6] W. Xia and M. Cao, "Cluster synchronization algorithms," in *American Control Conference*, Baltimore, USA, June 2010.

[7] C. Godsil and G. Royle, *Algebraic Graph Theory*. Springer, 2001.

[8] M. Bürger, G. S. Schmidt, and F. Allgöwer, "Preference Based Group Agreement in Cooperative Control," in *Proc. of the 8th IFAC Symposium on Nonlinear Control Systems*, 2010, pp. 149–154.

[9] R. T. Rockafellar, *Network Flows and Monotropic Optimization*. Belmont, Massachusetts: Athena Scientific, 1998.

[10] D. Bertsekas, "Min common / max crossing duality: A geometric view of conjugacy in convex optimization," Lab. for Information and Decision Systems, MIT, Tech. Rep. Report LIDS-P-2796, 2009.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2003.

[12] R. Rockafellar, *Convex Analysis*. Princeton University Press, 1997.

[13] H. Khalil, *Nonlinear Systems*. Upper Saddle River, New Jersey: Prentice Hall, 2002.