Distributed Utility Estimation with Heterogeneous Relative Information

Original
Distributed Utility Estimation with Heterogeneous Relative Information / Menci, M.; Oliva, G.; Papi, M.; Setola, R.; Zoppello, M.. - In: IEEE CONTROL SYSTEMS LETTERS. - ISSN 2475-1456. - 2:2(2018), pp. 248-253. [10.1109/LCSYS.2018.2819964]

Availability:
This version is available at: 11583/2984484 since: 2023-12-13T09:53:32Z

Publisher:
Institute of Electrical and Electronics Engineers Inc.

Published
DOI:10.1109/LCSYS.2018.2819964

(Article begins on next page)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2819964, IEEE Control Systems Letters

Distributed Utility Estimation with Heterogeneous Relative Information

M. Menci[1], G. Oliva[1*], M. Papi[1], R. Setola[1] and M. Zoppello[2]

***Abstract*—In this paper we consider a scenario where a set of agents, interconnected by a network topology, aim at computing an estimate of their own utility, importance or value, based on pairwise relative information having heterogeneous nature. In more detail, the agents are able to measure the difference between their value and the value of some their neighbors, or have an estimate of the ratio between their value and the value the remaining neighbors. This setting may find application in problems involving information provided by heterogeneous sensors (e.g., differences and ratios), as well as in scenarios where estimations provided by humans have to be merged with sensor measurements. Specifically, we develop a distributed algorithm that lets each agent asymptotically compute a utility value. To this end, we first characterize the task at hand in terms of a least-squares minimum problem, providing a necessary and sufficient condition for the existence of a unique global minimum, and then we show that the proposed algorithm asymptotically converges to a global minimum. The paper is concluded by numerical analyses that corroborate the theoretical findings.**

***Index Terms*—Optimization, Optimization algorithms, Sensor fusion**

I. INTRODUCTION

MAKING judgements is a core task in decision making processes involving humans or artificial intelligent systems. Such a task can either be carried out in an absolute or relative way. However, as noted by Blumenthal in his seminal work [1], also absolute judgements require some sort of comparison; indeed "to make the judgement, a person must compare an immediate impression with impression in memory of similar stimuli" [1]. Therefore, it is of paramount importance to be able to make judgements based on relative information. We point out that, in some cases, it might be impossible to take absolute judgements; this is due to technological reasons, e.g., the available sensors are not able to measure absolute information, or "psychological" reasons, e.g., the decision-maker might not be confident in assessing absolute information, or might feel more comfortable in handling relative information. In recent years, a large body of scientific literature has been aimed at endowing networked agents with the ability to distributedly compute absolute information based on relative measurements. A relevant example in this sense is *Sensor Network Localization*, where networked sensors aim at computing their location based on relative information such as bearings [2], [3], distances [4], presence within the sensing range [5], or combinations of distance and presence information [6], [7]. Other examples include *Formation Control* and *Distributed Analytic Hierarchy Process*. Within Formation Control problems [8]–[11], networked mobile agents aim at occupying locations that satisfy prescribed relative positions (e.g., in a least-squares sense as done in [8], [9] or exactly, under the assumption that the network is rigid, as done in [10]). Conversely, within distributed Analytic Hierarchy Process algorithms, the nodes in the network aim at computing their own utility or importance value based on the knowledge of perturbed utility ratios [12]–[15].

[1]Unit of Automatic Control, Department of Engineering, Università Campus Bio-Medico di Roma, via Àlvaro del Portillo 21, 00128, Rome, Italy. [2] Department of mathematics "Tullio Levi Civita" - University of Padova, Via Trieste, 63 - 35121 Padova, Italy. * Corresponding author, email: g.oliva@unicampus.it

A. Contribution and Novelty

To the best of our knowledge, current approaches in the literature operate based on homogeneous information. However, there are situations where one can improve the quality of the estimate by mixing heterogeneous pieces of information. For instance, consider a scenario where humans and machines cooperate; in this case, while sensors might be able to provide measurements of the difference between two quantities, while humans might be able to provide ratio information, e.g., assessing how many times one light or sound source is brighter or louder than another (see for instance [16]). Another example is the fusion of the information provided by sensors of heterogeneous nature, e.g., some able to measure distances [4], some able to measure ratios, such as signal strength ratios [17] or hop-count ratios [18]. To overcome some of the limitations of previous works, in this paper we consider a hybrid scenario where networked agents aim at computing their own utility, position or importance, based on heterogeneous pairwise relative information. Specifically, each agent knows: (i) an estimate of the differences between its utility and the utility of some of its neighbors; (ii) an estimate of the utility ratio with respect to the remaining neighbors. Based on such heterogeneous relative information, the agents cooperate in order to compute the absolute utility of each agent. To this end, we first define a least-squares minimization problem, by characterizing its global minima and providing a necessary and sufficient condition for the existence of a unique global minimum. Then, we develop a synchronous continuous-time distributed algorithm and we show that its dynamics converges to a global minimum, discussing the conditions guaranteeing that such a problem admits a unique or several global minima. We point out that the proposed problem setting is a mixture of the formation control approach, where agents are equipped with sensors able to measure relative positions, and the case where just ratios are considered. However, due

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2819964, IEEE Control Systems Letters

to the presence of information of heterogeneous nature, there is no straightforward way[1] to apply either of the above methodologies, thus calling for a different approach.

B. Paper Outline

The paper outline is as follows: in Section II we give some preliminary definitions while in Section III we formally describe our problem setting and the proposed distributed algorithm; Sections IV and V provide, respectively, an optimization problem whose solution corresponds to the unknown utilities for the nodes and a necessary and sufficient condition for the existence of a unique global minimum; Section VI characterizes the convergence properties of the proposed distributed algorithm; Section VII contains a simulation campaign aimed at numerically demonstrate the effectiveness of the proposed algorithm, while Section VIII collects some conclusive remarks and future work directions.

II. Preliminaries

We denote vectors by boldface lowercase letters and matrices with uppercase letters. We refer to the (i,j)-th entry of a matrix A by A_{ij}. We represent by $\mathbf{0}_n$ and $\mathbf{1}_n$ vectors with n components, all equal to zero and to one, respectively.

A. Graph Theory

Let $G=\{V,E\}$ be a *graph* with n nodes $V=\{v_1,v_2,\ldots,v_n\}$ and e edges $E\subseteq V\times V\setminus\{(v_i,v_i)\,|\,i\in V\}$, where $(v_i,v_j)\in E$ captures the existence of a link from node v_i to node v_j. A weighted graph is a graph $G=\{V,E\}$, together with a set of weights W such that $w_{ij}\in W$ represents the weight of each edge $(v_i,v_j)\in E$. A *bidirectional graph* is a graph such that $(v_i,v_j)\in E$ whenever $(v_j,v_i)\in E$; note that, in general, for weighted bidirectional graphs the weights w_{ij} and w_{ji} can be different. Let the *in-neighborhood* $\mathcal{N}_i^{in}$ of a node v_i be the set of nodes v_j such that $(v_j,v_i)\in E$, while the out-neighborhood $\mathcal{N}_i^{out}$ is the set of nodes v_j such that $(v_i,v_j)\in E$. The *in-degree* d_i^{in} of a node v_i is the number of its incoming edges, i.e., $d_i^{in}=|\mathcal{N}_i^{in}|$, while the *out-degree* d_i^{out} is the number of its outgoing edges, i.e., $d_i^{out}=|\mathcal{N}_i^{out}|$. A graph is *connected* if each node can be reached from each other node by using the edges in E, regardless of their orientation, while it is *strongly connected* if each node can be reached from each other node by using the edges in E, considering their orientation. Clearly, a bidirectional connected graph is also strongly connected. Given a graph $G=\{V,E\}$, let us define the set of matrices compatible with G as $\mathbb{A}_G=\left\{M\in\mathbb{R}^{|V|\times|V|}\,|\,M_{ij}=0,\quad\forall(v_j,v_i)\notin E, i\neq j\right\}$; in other words, a matrix $M\in\mathbb{A}_G$ has nonzero off-diagonal entries M_{ij} if and only if $(v_j,v_i)\in E$, while it can have nonzero diagonal entries.

III. Problem Statement

Let us consider a bidirectional (strongly) connected graph $G=\{V,E\}$ with n nodes, where each node $v_i\in V$ represents an agent and each link (v_i,v_j) captures the existence of a communication channel from agent v_i to agent v_j. Each agent $v_i\in V$ has the task to compute a value $f_i>0$ (e.g., its utility, position or importance[2]), based on relative information with respect to its neighboring agents; we assume that such an information has an heterogeneous nature, as discussed next. In more detail, for each agent v_i the in-neighborhood $\mathcal{N}_i^{in}$ over G is partitioned into two mutually exclusive[3] sets $\mathscr{D}_i^{in}$ and $\mathscr{R}_i^{in}$, i.e., $\mathscr{D}_i^{in}\bigcap\mathscr{R}_i^{in}=\emptyset$ and $\mathcal{N}_i^{in}=\mathscr{D}_i^{in}\bigcup\mathscr{R}_i^{in}$. The set $\mathscr{D}_i^{in}$ contains the in-neighbors of i for which relative information on the *difference* of the values is available; in other words, for all $v_j\in\mathscr{D}_i^{in}$ the agent v_i knows the value d_{ij} for the difference f_i-f_j. The set $\mathscr{R}_i^{in}$ contains the in-neighbors of i for which relative *ratios* are available; in other words, for all $v_j\in\mathscr{R}_i^{in}$ the agent v_i knows the value $r_{ij}>0$ for the ratio f_i/f_j. Note that, for simplicity, we assume that $v_j\in\mathscr{D}_i^{in}$ whenever $v_i\in\mathscr{D}_j^{in}$ and $v_j\in\mathscr{R}_i^{in}$ whenever $v_i\in\mathscr{R}_j^{in}$. Moreover, for each available difference d_{ij} it holds $d_{ji}=-d_{ij}$, while for each available ratio r_{ij} it holds $r_{ji}=1/r_{ij}$. Note that we can express E as $E=E^d\bigcup E^r$, where $E^d=\{(v_i,v_j)\in E\,|\,v_i\in\mathscr{D}_j^{in}\}$ and $E^r=\{(v_i,v_j)\in E\,|\,v_i\in\mathscr{R}_j^{in}\}$; clearly, it holds $E^d\bigcap E^r=\emptyset$.

In this paper we provide a distributed algorithm to let each agent asymptotically estimate its utility. To this end, we first formulate a least-squares optimization problem; then, we prove that our distributed algorithm asymptotically converges to a global optimal solution of the least-squares optimization problem.

IV. Optimization Problem

In this section we consider the problem of finding an $\boldsymbol{x}^*\in\mathbb{R}^n$ that satisfies all the distance and ratio constraints in an optimal least-squares sense. To this end, we consider a function $g:\mathbb{R}^n\to\mathbb{R}$ defined as

$$
\begin{aligned}
g(\boldsymbol{x}) &= \frac{1}{2}\sum_{(v_i,v_j)\in E^d}(x_i-x_j-d_{ij})^2 \\
&+\frac{1}{2}\sum_{(v_i,v_j)\in E^r}\left(\frac{r_{ji}}{1+r_{ji}}x_i-\frac{r_{ij}}{1+r_{ij}}x_j\right)^2.
\end{aligned}
\tag{1}
$$

In order to solve the problem at hand in this paper, we look for a global minimum $\boldsymbol{x}^*$ of $g(\cdot)$, i.e., we aim at finding $\boldsymbol{x}^*$ that satisfies $g(\boldsymbol{x}^*)=\min_{\boldsymbol{x}\in\mathbb{R}^n}\{g(\boldsymbol{x})\}$. It is immediate to recognize that $g(\boldsymbol{x})\geq 0$ for all $\boldsymbol{x}\in\mathbb{R}^n$ and that $g(\boldsymbol{x})=0$ if and only if for all $(v_i,v_j)\in E^d$ it holds $x_i-x_j=d_{ij}$ and for all $(v_i,v_j)\in E^r$ it holds $x_i/x_j=r_{ij}$; hence, in order to solve the problem at hand in this paper, we seek a global minimum $\boldsymbol{x}^*$ for $g(\cdot)$.

Let us now characterize the structure of the optimal solutions of the above problem. By straightforward computations, and since by assumption $j\in\mathscr{D}_i^{in}$ whenever $i\in\mathscr{D}_j^{in}$ and

[1]Indeed, we point out that a simple replacement of the ratios by their logarithm, with the aim to resort to an approach able to handle just distances, would not be an effective choice. In fact, there would be the need to introduce additional variables and constraints (i.e., constraints in the form $z_i=\log(x_i)$), which would need to be carefully handled; this represents an interesting direction that we leave for future research.

[2]In the following, we refer to the i-th value f_i simply as utility.

[3]Note that the proposed approach can be easily extended to the case where the graph is a *multigraph* with at most two links connecting any pair of nodes, i.e., a node v_j may belong to both sets. In this way it would be possible to handle situations where both difference and ratio information is provided for the same link.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2819964, IEEE Control Systems Letters

$j \in \mathscr{R}_i^{in}$ whenever $i \in \mathscr{R}_j^{in}$, it follows that the first order partial derivative of $g(\cdot)$ with respect to x_i is given by

$$\frac{\partial g(\boldsymbol{x})}{\partial x_i} = \sum_{j \in \mathscr{D}_i^{in}} (x_i - x_j) - \sum_{j \in \mathscr{D}_i^{in}} d_{ij} + \sum_{j \in \mathscr{R}_i^{in}} \frac{r_{ji}}{1+r_{ji}} \left(\frac{r_{ji}}{1+r_{ji}} x_i - \frac{r_{ij}}{1+r_{ij}} x_j \right). \quad (2)$$

Again, by simple computations, it can be shown that the $n \times n$ Hessian matrix $H(\cdot)$ associated to $g(\cdot)$ is such that

$$H_{ij}(\boldsymbol{x}) = \frac{\partial^2 g(\boldsymbol{x})}{\partial x_i \partial x_j} = \begin{cases} |\mathscr{D}_i^{in}| + \sum_{j \in \mathscr{R}_i^{in}} \frac{r_{ji}^2}{(1+r_{ji})^2}, & \text{if } i = j \\ -1 & \text{if } v_j \in \mathscr{D}_i^{in} \\ -\frac{1}{(1+r_{ij})(1+r_{ji})} & \text{if } v_j \in \mathscr{R}_i^{in} \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Let us now collect some observations about $g(\cdot)$ and $H(\cdot)$.

Remark 1: Since $g(\cdot)$ represents a nonnegative second order polynomial, H is constant, symmetric and positive semidefinite [19]. We point out that, since H is positive semidefinite, the function $g(\cdot)$ is convex [4] (see, for instance [20], Chapter 2). Finally, we notice that for $i \neq j$ it holds $H_{ij} \neq 0$ if and only if $(v_j, v_i) \in E$; hence, $H \in \mathbb{A}_G$.

Remark 2: Since $g(\cdot)$ is convex, any of its global minima $\boldsymbol{x}^*$ satisfies

$$\left. \frac{\partial g(\boldsymbol{x})}{\partial x_i} \right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0, \quad \forall i \in \{1, \ldots, n\}.$$

Stacking Eq. (2) for all $i \in \{1, \ldots, n\}$, setting $\delta_i = \sum_{j \in \mathscr{D}_i^{in}} d_{ij}$, $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_n]^T$ and evaluating at $\boldsymbol{x} = \boldsymbol{x}^*$, we conclude that the global minima $\boldsymbol{x}^*$ of $g(\cdot)$ satisfy $H\boldsymbol{x}^* = \boldsymbol{\delta}$. Therefore, we observe that $g(\cdot)$ has a unique global minimum $\boldsymbol{x}^* = H^{-1}\boldsymbol{\delta}$ if and only if $rank(H) = n$. Otherwise, it holds $rank(H) = m < n$ and the set of global minima of $g(\cdot)$ is a subspace of $\mathbb{R}^n$ with dimension equal to $n - m$.

V. Existence of a Unique Global Minimum

In this section, we provide a necessary and sufficient condition that guarantees the existence of the unique global optimal solution to the minimization problem. To this end, let us now provide a necessary and sufficient condition for a vector $\boldsymbol{x} \in \mathbb{R}^n$ to belong to the kernel of the Hessian matrix H of $g(\cdot)$.

Proposition 1: Let $G = \{V, E^d \bigcup E^r\}$ be a connected bidirectional graph with n nodes, where E^d and E^r reflect, respectively, the difference and ratio information available; moreover, let $g(\cdot)$ be defined as in Eq (1). A vector $\boldsymbol{x} \in \mathbb{R}^n$ satisfies $H\boldsymbol{x} = \boldsymbol{0}_n$, where H is the Hessian matrix associated to $g(\cdot)$, if and only if it holds

$$\sum_{(v_i,v_j) \in E^d} (x_i - x_j)^2 + \sum_{(v_i,v_j) \in E^r} \frac{(x_i - r_{ij}x_j)^2}{(1+r_{ij})^2} = 0. \quad (4)$$

Proof: To establish the result we notice that, being H symmetric, it holds $H\boldsymbol{x} = \boldsymbol{0}_n$ if and only if $\boldsymbol{x}^T H \boldsymbol{x} = 0$. Moreover, we have that $H = H' + H''$, where H' and H'' are symmetric matrices having entries given by

$$H'_{ij} = \begin{cases} |\mathscr{D}_i^{in}|, & \text{if } i = j \\ -1 & \text{if } v_j \in \mathscr{D}_i^{in} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

and

$$H''_{ij} = \begin{cases} \sum_{j \in \mathscr{R}_i^{in}} \frac{r_{ji}^2}{(1+r_{ji})^2}, & \text{if } i = j \\ -\frac{1}{(1+r_{ij})(1+r_{ji})} & \text{if } v_j \in \mathscr{R}_i^{in} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

In other words, we have that $H\boldsymbol{x} = \boldsymbol{0}_n$ if and only if it holds

$$\boldsymbol{x}^T H \boldsymbol{x} = \boldsymbol{x}^T H' \boldsymbol{x} + \boldsymbol{x}^T H'' \boldsymbol{x} = 0. \quad (7)$$

We observe that

$$\boldsymbol{x}^T H' \boldsymbol{x} = \sum_{i=1}^{n} x_i \sum_{j \in \mathscr{D}_i^{in}} (x_i - x_j) = \sum_{(v_i,v_j) \in E^d} (x_i - x_j)^2, \quad (8)$$

and that

$$\boldsymbol{x}^T H'' \boldsymbol{x} = \sum_{i=1}^{n} x_i \sum_{j \in \mathscr{R}_i^{in}} \left(\frac{r_{ji}^2}{(1+r_{ji})^2} x_i - \frac{1}{(1+r_{ij})(1+r_{ji})} x_j \right).$$

At this point, we notice that, by some algebra, it holds

$$\boldsymbol{x}^T H'' \boldsymbol{x} = \sum_{(v_i,v_j) \in E^r} \frac{(x_i - r_{ij}x_j)^2}{(1+r_{ij})^2}. \quad (9)$$

The proof follows. ■

We now show that $rank(H) \geq n - 1$.

Lemma 1: Let $G = \{V, E^d \bigcup E^r\}$ be a connected bidirectional graph with n nodes, where E^d and E^r reflect, respectively, the difference and ratio information available; moreover, let $g(\cdot)$ be defined as in Eq (1). The Hessian matrix H associated to $g(\cdot)$ is such that $rank(H) \geq n - 1$.

Proof: Let $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}_n\}$ be such that $H\boldsymbol{x} = \boldsymbol{0}_n$. By Proposition 1, $\boldsymbol{x}$ satisfies Eq. (4). Therefore, it must hold $x_i - x_j = 0$, for all $(v_i, v_j) \in E^d$ and $x_i - r_{ij}x_j = 0$, for all $(v_i, v_j) \in E^r$. The above conditions can be rearranged as $x_i/x_j = 1$, for all $(v_i, v_j) \in E^d$ and $x_i/x_j = r_{ij}$, for all $(v_i, v_j) \in E^r$. In other words, for each link $(v_i, v_j) \in E$, the ratio of the entries x_i and x_j of the vector $\boldsymbol{x}$ must be equal to a given w_{ij}, where $w_{ij} = r_{ij}$ for links corresponding to ratio information and $w_{ij} = 1$ for links corresponding to difference information. Let $W \in \mathbb{A}_G$ be the $n \times n$ matrix collecting such ratios, i.e., $W_{ij} = w_{ij}$ if $(v_j, v_i) \in E^d \bigcup E^r$ and $W_{ij} = 0$, otherwise. In [14], the authors demonstrate that, when a matrix W collecting sparse ratio information has the same structure as a connected bidirectional graph, a necessary and sufficient condition for the existence of a vector $\boldsymbol{x}$ such that $W_{ij} = x_i/x_j$ for all $W_{ij} \neq 0$ is that the product of the entries W_{ij} along any cycle of the graph is equal to one; otherwise, no solution exists. When such a condition is satisfied, we observe that the ratios x_i/x_j are defined up to a scaling factor. Hence, the kernel of H has dimension one and $rank(H) = n - 1$. In the latter case, no solution exists (other than the trivial one) and therefore the kernel of H coincides with $\{\boldsymbol{0}_n\}$ and $rank(H) = n$. The proof is complete. ■

As a consequence of Lemma 1, we can state a necessary

[4]The positive semidefiniteness of H implies convexity but not strict convexity, i.e., $g(\cdot)$ might have multiple global minima.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2819964, IEEE Control Systems Letters

and sufficient condition for $rank(H)$ to be equal to n.

Proposition 2: Let H be the Hessian matrix associated to $g(\cdot)$ and let us assign a weight $w_{ij} = 1$ to all $(v_i, v_j) \in E^d$ and a weight $w_{ij} = r_{ij}$ to all $(v_i, v_j) \in E^r$. It holds $rank(H) = n$ if and only if there is a cycle $c = \{(v_1, v_2), \ldots, (v_m, v_1)\}$ over $G = \{V, E^d \bigcup E^r\}$ such that $\prod_{(v_i,v_j)\in c} w_{ij} \neq 1$.

Proof: By Lemma 1, it holds $rank(H) = n-1$ if and only if all cycles over G satisfy $\prod_{(v_i,v_j)\in c} w_{ij} = 1$, otherwise $rank(H) = n$. The proof follows. ■

VI. Proposed Algorithm

If all the information can be collected and processed in a centralized way then, as noted in Remark 1, a solution to the problem at hand in this paper is to find $\boldsymbol{x}^*$ that satisfies $H\boldsymbol{x}^* = \boldsymbol{\delta}$. In several situations it might be impossible to solve the problem by means of a centralized supervisory entity; in those cases, each agent aims at computing its own utility in a distributed way. Specifically, based on the information regarding its neighbors, each agent i aims at computing a value x_i^* such that, overall, the vector $\boldsymbol{x}^*$ satisfies all difference and ratio constraints in a least-squares sense; in other words, the agents aim at computing a vector $\boldsymbol{x}^*$ that is a global minimum for $g(\cdot)$. Within the proposed algorithm, each agent i executes the following continuous-time and synchronous update algorithm

$$\begin{aligned}\dot{x}_i(t) = \alpha &\sum_{j\in\mathscr{D}_i^{in}} (x_j - x_i) \\ &+ \alpha \sum_{j\in\mathscr{R}_i^{in}} \frac{r_{ij}}{1+r_{ij}} \left(\frac{r_{ij}}{1+r_{ij}} x_j - \frac{r_{ji}}{1+r_{ji}} x_i \right) + \alpha\delta_i,\end{aligned} \tag{10}$$

where $\alpha > 0$ and $\delta_i = \sum_{j\in\mathscr{D}_i^{in}} d_{ij}$.

Let us now show that the proposed distributed algorithm asymptotically converges to a global minimum[5] $\boldsymbol{x}^*$ for $g(\cdot)$.

Theorem 1: Let us consider a connected bidirectional graph $G = \{V, E^d \bigcup E^r\}$ with n nodes, where E^d and E^r reflect, respectively, the difference and ratio information available to the agents. Let the agents execute the synchronous update rule in Eq. (10), with initial condition $x_i(0) > 0$ and $\alpha > 0$. It holds $\lim_{t\to\infty} x_i(t) = x_i^*$, where $\boldsymbol{x}^* = [x_1^*, \ldots, x_n^*]^T$ is a global minimum for $g(\cdot)$.

Proof: Stacking Eq. (10) for all the agents and setting $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)]^T$ and $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_n]^T$, we get

$$\dot{\boldsymbol{x}}(t) = -\alpha H \boldsymbol{x}(t) + \alpha\boldsymbol{\delta}. \tag{11}$$

As noted in Remark 1, we have that H is positive semidefinite and by Lemma 1 it holds $rank(-\alpha H) \geq n-1$. Hence, matrix $-\alpha H$ is stable in the continuous-time sense and the system $\dot{\boldsymbol{x}}(t) = -\alpha H\boldsymbol{x}(t)$ converges to an equilibrium point. We point out that the presence of the constant input $\alpha\boldsymbol{\delta}$ does not affect stability; hence, also the dynamics in Eq. (11) converges to an equilibrium point $\boldsymbol{x}_{eq}$, which satisfies $\boldsymbol{0}_n = -\alpha H\boldsymbol{x}_{eq} + \alpha\boldsymbol{\delta}$, that is, $H\boldsymbol{x}_{eq} = \boldsymbol{\delta}$. Therefore, using the same reasoning as in Remark 2, we conclude that the equilibrium reached corresponds to a global minimum of $g(\cdot)$. The proof is complete. ■

[5]As discussed in the previous section, the solution is unique if and only if the condition in Proposition 2 is satisfied.

Remark 3: The parameter α can be used to arbitrarily increase the speed of convergence of the proposed algorithm, e.g., by letting each agent choose the same $\alpha \gg 1$. However, to select a specific (e.g., instance-dependent) value of α, some form of distributed coordination or agreement (e.g., distributed consensus [9]) is required before the execution of the proposed algorithm.

We now characterize the structure of the particular global minimum of $g(\cdot)$ computed by the proposed algorithm.

Theorem 2: Let us consider a connected bidirectional graph $G = \{V, E^d \bigcup E^r\}$ with n nodes, where E^d and E^r reflect, respectively, the difference and ratio information available to the agents. Let the agents execute the synchronous update rule in Eq. (10), with initial condition $x_i(0) > 0$ and $\alpha > 0$. Without loss of generality, let λ_i be the i-th smallest eigenvalue of the Hessian matrix H of $g(\cdot)$ and let $\boldsymbol{z}_i$ be the corresponding eigenvector such that the set $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$ represents an orthonormal basis for H. The state of the agents asymptotically converges to

$$\boldsymbol{x}_{eq} = \begin{cases} \left(\boldsymbol{z}_1^T \boldsymbol{x}(0)\right)\boldsymbol{z}_1 + \sum_{i=2}^n \frac{1}{\lambda_i}\left(\boldsymbol{z}_i^T\boldsymbol{\delta}\right)\boldsymbol{z}_i, & \text{if } \lambda_1 = 0, \\ \sum_{i=1}^n \frac{1}{\lambda_i}\left(\boldsymbol{z}_i^T\boldsymbol{\delta}\right)\boldsymbol{z}_i, & \text{otherwise.} \end{cases} \tag{12}$$

Proof: As noted in Remark 1 and Lemma 1, the Hessian matrix H is symmetric and it has at most one eigenvalue equal to zero, while all other eigenvalues are positive. Hence, we diagonalize H by writing $H = Z\Lambda Z^{-1}$, where Λ is a diagonal $n \times n$ matrix with $\Lambda_{ii} = \lambda_i$ and $Z = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n]$; since $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$ represents an orthonormal basis, we have that it holds $Z^{-1} = Z^T$. The state of the agents at time t is given by $\boldsymbol{x}(t) = e^{-\alpha Ht}\boldsymbol{x}(0) + \int_0^t e^{-\alpha H(t-\tau)}\alpha\boldsymbol{\delta}d\tau$ and can be rearranged as

$$\boldsymbol{x}(t) = Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{x}(0) + \alpha Ze^{-\alpha\Lambda t}Z^{-1}\int_0^t Ze^{\alpha\Lambda\tau}Z^{-1}\boldsymbol{\delta}d\tau$$

Let us define $\boldsymbol{\eta} = Z^{-1}\boldsymbol{\delta}$, so that it holds $\eta_i = \boldsymbol{z}_i^T\boldsymbol{\delta}$ for all $i \in \{1, \ldots, n\}$. Notice that $H\boldsymbol{x}_{eq} = \boldsymbol{\delta}$; hence, when $\lambda_1 = 0$ it holds $\eta_1 = \boldsymbol{z}_1^T\boldsymbol{\delta} = \boldsymbol{z}_1^T H\boldsymbol{x}_{eq} = 0$. Since $e^{\alpha\Lambda\tau}$ is diagonal and $Z^{-1} = Z^T$, it holds

$$\begin{aligned}\boldsymbol{x}(t) &= Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{x}(0) + \alpha Ze^{-\alpha\Lambda t}Z^{-1}\int_0^t \sum_{i=1}^n \boldsymbol{z}_i e^{\alpha\lambda_i\tau}\eta_i d\tau \\ &= Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{x}(0) + \alpha\sum_{i=1}^n Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{z}_i\eta_i\int_0^t e^{\alpha\lambda_i\tau}d\tau.\end{aligned}$$

Let $\boldsymbol{e}_i$ be the i-th vector in the canonical base in $\mathbb{R}^n$; since $Z^{-1} = Z^T$, we have that

$$Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{z}_i = Ze^{-\alpha\Lambda t}\boldsymbol{e}_i = Ze^{-\alpha\lambda_i t}\boldsymbol{e}_i = e^{-\alpha\lambda_i t}Z\boldsymbol{e}_i = e^{-\alpha\lambda_i t}\boldsymbol{z}_i;$$

hence, it holds

$$\boldsymbol{x}(t) = Ze^{-\alpha\Lambda t}Z^{-1}\boldsymbol{x}(0) + \alpha\sum_{i=1}^n \eta_i e^{-\alpha\lambda_i t}\boldsymbol{z}_i\int_0^t e^{\alpha\lambda_i\tau}d\tau.$$

At this point we notice that, when $\lambda_1 = 0$, it holds $\eta_1 = 0$; hence,

$$\boldsymbol{x}(t) = \sum_{i=1}^n e^{-\alpha\lambda_i t}\left(\boldsymbol{z}_i^T\boldsymbol{x}(0)\right)\boldsymbol{z}_i + \sum_{i=2}^n \left(\boldsymbol{z}_i^T\boldsymbol{\delta}\right)\frac{1-e^{-\alpha\lambda_i t}}{\lambda_i}\boldsymbol{z}_i$$

and therefore $\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}_{eq}$, where $\boldsymbol{x}_{eq}$ corresponds to the first case in Eq. (12). Conversely, when $\lambda_1 > 0$, we have that

$$\boldsymbol{x}(t) = \sum_{i=1}^{n} e^{-\alpha\lambda_i t} \left(\boldsymbol{z}_i^T \boldsymbol{x}(0)\right) \boldsymbol{z}_i + \sum_{i=1}^{n} \left(\boldsymbol{z}_i^T \boldsymbol{\delta}\right) \frac{1 - e^{-\alpha\lambda_i t}}{\lambda_i} \boldsymbol{z}_i$$

and therefore $\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}_{eq}$, where this time $\boldsymbol{x}_{eq}$ corresponds to the second case in Eq. (12). This completes our proof. ■

A few remarks are now in order.

Remark 4: Notice that, when $\lambda_1 = 0$ the set of global minima of $g(\cdot)$ correspond to a subspace of $\mathbb{R}^n$ of dimension equal to one. Conversely, when $\lambda_1 > 0$ the problem admits a unique global minimum. In particular, as shown in Eq. (12), in the first case the solutions coincide with an affine space of the eigenspace spanned by $\boldsymbol{z}_1$ (the particular value computed by the agents depends on the initial condition $\boldsymbol{x}(0)$ and on the complete sets of eigenvalues and eigenvectors of H), while in the latter case the solution is unique, and it depends on all the eigenvalues and eigenvectors of H but is independent on the initial condition.

Remark 5: Note that, although Eq. (12) provides a closed-form solution for the global minima of $g(\cdot)$, its structure depends on the entire set of eigenvalues and eigenvectors of H; a distributed algorithm to compute such information has a remarkably higher computational burden for the agents (e.g., see [21]) with respect to the proposed algorithm, thus justifying the adoption of our approach in a distributed computing scenario.

VII. Simulation Results

In this section we provide numerical evidence to corroborate the theoretical findings. Let us take into account two small scale instances such that $g(\boldsymbol{x}^*) = 0$, i.e., such that the available information is perfectly consistent. Specifically, we consider two graphs with $|V| = 5$ nodes and $|E| = 12$ edges (i.e., six pairs of bidirectional edges); the graphs and the available differences/ratios are reported, respectively, in Figure 1a and Figure 1c. Let us now discuss the first example. Note that the information associated to the example in Figure 1a satisfies the necessary and sufficient condition in Proposition 2, hence it can be shown that $\boldsymbol{g}(\cdot)$ has a unique global minimum at $\boldsymbol{x}^* = \sum_{i=1}^{n} \frac{1}{\lambda_i} \left(\boldsymbol{z}_i^T \boldsymbol{\delta}\right) \boldsymbol{z}_i = [1,2,2,8,1]^T$, thus numerically corroborating Eq. (12). Figure 1b shows the evolution of the proposed distributed algorithm when $\alpha = 1$; it can be noted that the state $x_i(t)$ of each agent v_i effectively converges to x_i^*. Let us now discuss the example in Figure 1c; Figure 1d shows that, for $\boldsymbol{x}(0) = [0.5768, 0.0259, 0.4465, 0.6463, 0.5212]^T$, the state of the agents converges to an $\boldsymbol{x}_{eq}$ corresponding to the first case in Eq. (12), i.e., $\boldsymbol{x}_{eq} = [-2.5790, -1.5790, 4.4210, 2.2105, 0.5526]^T$, thus numerically validating Eq. (12). It can be easily shown that for any $\varepsilon \in \mathbb{R}$ the vector $\boldsymbol{x}^*(\varepsilon) = [1+\varepsilon, 2+\varepsilon, 8+\varepsilon, 4+\varepsilon/2, 1+\varepsilon/8]^T$ is a global minimum, since it holds $g(\boldsymbol{x}^*(\varepsilon)) = 0$ (in our example, we have that $\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}^*(-3.579)$); note that, as demonstrated in Theorem 2, the particular value of ε associated to the asymptotic solution found depends on the initial condition $\boldsymbol{x}(0)$.

In order to assess the effect of perturbations on the available information, and to compare with standard formation control and AHP approaches, in Figure 2 we consider a graph where $|V| = 100$ nodes are sampled uniformly at random in the unit square $[0,1]^2$ and a pair of nodes v_i, v_j is connected by an edge provided that their Euclidean distance is smaller than $\rho = 0.2$; the resulting graph has $|E| = 1038$ links (i.e., 519 distinct pairs); the graph is reported in Figure 2a. Moreover, we consider a scenario where the utility of the i-th agent is $x_i^* = \frac{2i}{n(n+1)}$ (so that, overall, it holds $\mathbf{1}_n^T \boldsymbol{x}^* = 1$) and we partition the links of E into the sets E^d (black solid lines in Figure 2a) and E^r (blue dotted lines in Figure 2a), which correspond to difference and ratio information, respectively. Specifically, in order to guarantee that the problem can be solved based on just differences or ratios, we first calculate two edge-disjoint spanning trees over G and we assign their links to E^d and E^r, respectively; then, we randomly partition the remaining links in E, assigning them to the sets E^d and E^r with equal probability; as a result we obtain $|E^d| = 554$ (277 distinct pairs) and $|E^r| = 484$ (242 distinct pairs). In order to evaluate the effectiveness of the proposed methodology, we consider multiplicative errors affecting the available ratios and additive perturbations affecting the available differences. In more detail, we consider ratios affected by log-normal random perturbations (as typically done in the AHP literature, see for instance [22]), i.e., we set $r_{ij} = exp(\mathcal{N}(0,\sigma)) x_i^*/x_j^*$, where $\mathcal{N}(0,\sigma)$ is a normal random number with zero average and standard deviation σ. Then, we select random additive perturbations for the difference information which are comparable to the magnitude of the multiplicative ones. To this end, we observe that if $x_i^*/x_j^* = r_{ij}e^{\sigma}$, then $x_i^* - x_j^* = (r_{ij}e^{\sigma} - 1)x_j^* = x_i^* e^{\sigma} - x_j^*$; therefore, if we seek for a perturbation γ such that $x_i^* - x_j^* = d_{ij} + \gamma$ we have that $\gamma = x_i^* e^{\sigma} - x_j^* - d_{ij} = x_i^*(e^{\sigma} - 1)$. For the above reason, we set $d_{ij} = x_i^* - x_j^* + \mathcal{N}(0, x_i^*(e^{\sigma} - 1))$. In Figure 2b we compare the results achieved by considering only difference information via formation control (blue dashed line), only ratios via the AHP approach in [22]) (green dotted line) and the performance of the proposed algorithm when we consider both differences and ratios (red solid lines); for all curves we show the results in terms of average and standard deviation over $M = 100$ runs with the same choice of σ. Specifically, we plot against σ the *Kendall's Tau Distance* [23] τ between the nominal and perturbed ranking of the agents; such a distance is such that $\tau \in [0,1]$, where $\tau = 0$ means that the ranking is the same and $\tau = 1$ means that the rankings are in reverse order. As shown by Figure 2b, the proposed approach is remarkably more robust to the perturbations; for instance, $\tau \leq 0.01$ for $\sigma \leq 0.1$ (while using only differences or ratios we get $\tau \approx 0.07$ and $\tau \approx 0.04$, respectively). The difference in the result of the three approaches widens as σ grows, and for $\sigma = 0.3$ we have that the proposed approach yields $\tau \approx 0.07$, while the cases of using only differences and ratios yield $\tau \approx 0.31$ and $\tau \approx 0.13$, respectively.

VIII. Conclusions and Future Work

In this paper we develop a novel distributed decision making technique that endows a network of agents with the capability

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2819964, IEEE Control Systems Letters

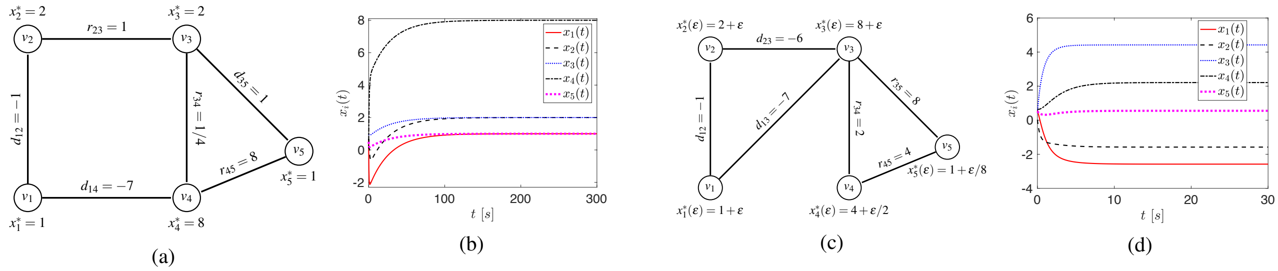

Figure 1. Examples with $|V| = 5$ nodes and $|E| = 12$ edges (i.e., six pairs of bidirectional edges). In panel 1a the condition in Proposition 2 is satisfied and there is a unique global minimum for $g(\cdot)$. In panel 1c the condition in Proposition 2 is violated and there are several global minima for $g(\cdot)$ (we show them as a function of the parameter ε). Panels 1b and 1d show the evolution of the proposed algorithm for $\alpha = 1$, considering the instance in panels 1c and 1c, respectively.

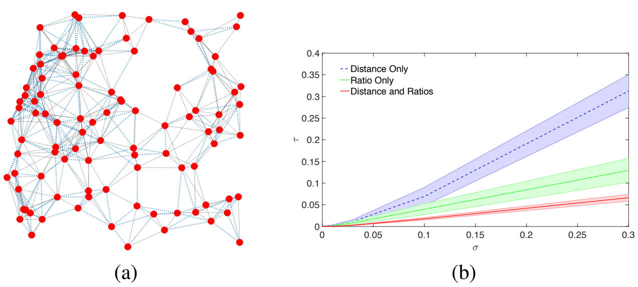

Figure 2. Left panel: graph considered in the simulation. Right panel: comparison of the proposed approach (red line) with formation control (blue dashed line) and AHP (green dotted line) for growing perturbations.

to compute a quantity that represents their own utility or importance, based on the knowledge of pairwise relative information of heterogeneous nature, i.e., the differences and ratios of the utilities of a node with respect to its neighbors. Specifically, we frame the problem in terms of a least-squares minimization problem and we characterize the structure of the global minima of such problem, providing a necessary and sufficient condition that guarantees the existence of a unique solution. Moreover, we develop a distributed continuous-time algorithm that lets the agents asymptotically find a global minimum. Future work will follow four main directions: (i) extending the framework to directed graphs; (ii) introducing constraints in the formulation; (iii) including in the framework different typologies of nonlinear functions describing the relative information available; (iv) extending the approach to wireless sensor network localization, considering a scenario where some sensors are able to measure distances while other sensors are able to estimate of the ratio between their distance from pairs of neighbors.

References

[1] A. L. Blumenthal, *The process of cognition.* Prentice Hall/Pearson Education, 1977.

[2] K. Briechle and U. D. Hanebeck, “Localization of a mobile robot using relative bearing measurements,” *IEEE Transactions on Robotics and Automation*, vol. 20, no. 1, pp. 36–44, 2004.

[3] A. N. Bishop, B. D. Anderson, B. Fidan, P. N. Pathirana, and G. Mao, “Bearing-only localization using geometrically constrained optimization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 1, pp. 308–320, 2009.

[4] J. Aspnes, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, Y. R. Yang, B. D. Anderson, and P. N. Belhumeur, “A theory of network localization,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, pp. 1663–1678, 2006.

[5] C. Laoudias, M. P. Michaelides, and C. Panayiotou, “Fault tolerant fingerprint-based positioning,” in *Communications (ICC), 2011 IEEE International Conference on.* IEEE, 2011, pp. 1–5.

[6] G. Oliva, C. Laoudias, F. Pascucci, R. Setola, and C. G. Panayiotou, “Hybrid sensor networks localization dealing with range-capable and range-free nodes,” in *Indoor Positioning and Indoor Navigation (IPIN), 2014 International Conference on*. IEEE, 2014, pp. 281–290.

[7] G. Oliva, S. Panzieri, F. Pascucci, and R. Setola, “Sensor networks localization: Extending trilateration via shadow edges,” *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2752–2755, 2015.

[8] J. A. Fax and R. M. Murray, “Information flow and cooperative control of vehicle formations,” *IEEE transactions on automatic control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[9] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[10] B. D. Anderson, C. Yu, J. M. Hendrickx *et al.*, “Rigid graph control architectures for autonomous formations,” *IEEE Control Systems*, vol. 28, no. 6, 2008.

[11] K.-K. Oh, M.-C. Park, and H.-S. Ahn, “A survey of multi-agent formation control,” *Automatica*, vol. 53, pp. 424–440, 2015.

[12] P. T. Harker, “Alternative modes of questioning in the analytic hierarchy process,” *Mathematical Modelling*, vol. 9, no. 3-5, pp. 353–360, 1987.

[13] ——, “Incomplete pairwise comparisons in the analytic hierarchy process,” *Mathematical Modelling*, vol. 9, no. 11, pp. 837–848, 1987.

[14] G. Oliva, R. Setola, and A. Scala, “Sparse and distributed analytic hierarchy process,” *Automatica*, vol. 85, pp. 211–220, 2017.

[15] M. Menci, G. Oliva, M. Papi, R. Setola, and A. Scala, “A suite of distributed methodologies to solve the sparse analytic hierarchy process problem,” in *IEEE European Control Conference 2018*, 2018 (to appear).

[16] J. C. Middlebrooks and D. M. Green, “Sound localization by human listeners,” *Annual review of psychology*, vol. 42, no. 1, pp. 135–159, 1991.

[17] S.-Y. Jung, C.-K. Choi, S. H. Heo, S. R. Lee, and C.-S. Park, “Received signal strength ratio based optical wireless indoor localization using light emitting diodes for illumination,” in *Consumer Electronics (ICCE), 2013 IEEE International Conference on*. IEEE, 2013, pp. 63–64.

[18] S. Yang, J. Yi, and H. Cha, “Hcrl: A hop-count-ratio based localization in wireless sensor networks,” in *Sensor, Mesh and Ad Hoc Communications and Networks, 2007. SECON’07. 4th Annual IEEE Communications Society Conference on*. IEEE, 2007, pp. 31–40.

[19] L. V. Ahlfors, “Complex analysis: an introduction to the theory of analytic functions of one complex variable,” *New York, London*, p. 177, 1953.

[20] A. Beck, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*. SIAM, 2014.

[21] H. Strakova and W. N. Gansterer, “A distributed eigensolver for loosely coupled networks,” in *Parallel, Distributed and Network-Based Processing (PDP), 2013 21st Euromicro International Conference on*. IEEE, 2013, pp. 51–57.

[22] J. Barzilai, W. D. Cook, and B. Golany, “Consistent weights for judgements matrices of the relative importance of alternatives,” *Operations research letters*, vol. 6, no. 3, pp. 131–134, 1987.

[23] M. G. Kendall, “A new measure of rank correlation,” *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.