23rd International Symposium on Mathematical Theory of Networks and Systems
Hong Kong University of Science and Technology, Hong Kong, July 16-20, 2018

Distributed Multi-Agent Optimization for Pareto Optimal Problem over Unbalanced Networks via Exact Penalty Methods with Equality and Inequality Constraints

Izumi Masubuchi[1†], Takayuki Wada[2], Yasumasa Fujisaki[2], and Fabrizio Dabbene[3]

***Abstract*— This paper proposes a distributed multi-agent optimization protocol to solve a Pareto optimal problem. The protocol only requires local communications between agents to exchange decision variables and the graph representing the communications has to be only strongly connected but does not need to be balanced. This extends the implementability of the protocol to real-world applications. The protocol is based on exact penalty methods and can handle inequality and equality constraints. The computation is executed without disclosing objective and constraint functions.**

***Index Terms*—Distributed multi-agent optimization, networked systems, exact penalty method, equality and inequality constraints. AMS subject classification: 90C25.**

I. Introduction

Distributed multi-agent optimization is one of the key methodologies fundamental for control and decision making over large-scale network. Various protocols have been proposed to realize consensus between agents via linear protocols [1], [2], which are extended to distributed convex optimization [3], [4], [5], [6], [7], [8], [9], [10], where the agents do not need to disclose their objective and constraint functions to other agents. This is a major advantage of distributed optimization for agents, who are connected via a network to exchange limited data but can not disclose information contained in those objective and constraint functions.

In distributed optimization methods to minimize sum of the objective functions of the agents based on the linear consensus protocol, the graph representing the availability of decision variables from other agents is needed to be balanced, or the weight matrix to compute convex combination of the decision variables of neighboring agents has to be doubly stochastic. In practice, it can be difficult to set up such a doubly stochastic matrix since that agents need to negotiate to determine the elements.

It has been shown in [4], however, that the consensus-protocol-based protocol solves a weighted sum of the objective functions of the agents, where the weight is given as the left eigenvector of the weight matrix which is only right stochastic. Wada and Fujisaki [11] proved that the protocol attains a Pareto optimum with the set of objective functions and the results are extended to randomly varying unbalanced [13]. Also Xie et al. [12] investigated distributed optimization over time-varying unbalanced graphs with inequality constraints.

This paper extends previous results of constrained distributed optimization over networks with unbalanced graphs representing the agents' communications so that equality conditions can be included as local constraints. The protocol proposed in this paper is a generalization of the protocol [9], [10] based on exact penalty methods [14] and the linear protocol for consensus [1], [2] Under appropriate choices of step length on the gradient terms, the protocol attains a consensus at a Pareto optimal solution satisfying equality and inequality conditions. Numerical examples are provided to illustrate the convergence to a Pareto optimal solution over unbalanced network. Similarly to [12], our protocol can be applied to distributed min-max optimization [5], for which we also show a numerical example.

The rest of the paper is organized as follows. Section II formulates a distributed multi-agent optimization problem with equality and inequality constraint to make consensus at a Pareto optimal solution and a protocol to solve this problem over unbalanced network is presented in Section III. The proof of the consensus and consensus is provided in Section IV, following which we show numerical examples of the proposed protocol with application to a min-max problem. Section VI concludes the paper.

Notation

Let $\mathbf{R}^n$ and $\mathbf{R}^{m\times n}$ denote the sets of real (column) n-vectors and $m\times n$-matrices, respectively. $\|\cdot\|$ is the Euclidean norm for vectors and $\langle\cdot,\cdot\rangle$ is the standard inner product of $\mathbf{R}^n$. Let $\|\cdot\|_\infty$ denote the infinity norm. Inequalities on real vectors are elementwise. Let $\mathbf{1}_N = [\,1 \quad 1 \quad \cdots \quad 1\,]^\top \in \mathbf{R}^N$ and I_N be the identity matrix of size N. The diagonal matrix whose diagonal entries are $d_1, d_2, \ldots, d_n$ is represented as $\mathrm{diag}\{d_1, d_2, \ldots, d_n\}$. Let $\overline{\mathrm{B}}(R; x)$ be the closed ball whose center is x and whose radius is R. For matrices A, B, $A\otimes B$ stands for the Kronecker product.

II. Problem formulation

A. Multi-agent Pareto optimization problem

We consider a multi-agent optimization problem on a network of agents. Let $\mathcal{A} = \{1, 2, \ldots, N\}$ be the set of agents. Suppose that each agent i has functions $f^i : \mathbf{R}^n \to$

[1]Izumi Masubuchi is with Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan.

[2]Takayuki Wada and Yasumasa Fujisaki are with Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan.

[3]Fabrizio Dabbene is with Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy (CNR-IEIIT), Torino, Italy.

[†]Corresponding author. E-mail: msb@harbor.kobe-u.ac.jp

$\mathbf{R}$, $g^i : \mathbf{R}^n \to \mathbf{R}$, $h^i : \mathbf{R}^n \to \mathbf{R}^{m^i}$, and that the benefit of agent i is to minimize $f^i(x^i)$ subject to inequality constraint $g^i(x^i) \le 0$ and equality constraint $h^i(x^i) = 0$. Meanwhile, variable x^i is common to the network and hence the consensus between agents, i.e., the coincidence of variables x^i is needed. Accordingly, we formulate the following problem:

Problem 1: Find a Pareto optimal solution for

$$\min_{x^1,x^2,\dots,x^N\in\mathbf{R}^n} f^1(x^1), f^2(x^2), \dots, f^N(x^N) \tag{1}$$

subject to constraints:

$$g^i(x^i) \le 0, \quad h^i(x^i) = 0, \quad i \in \mathcal{A} \tag{2}$$

and the consensus of decision variables, namely

$$x^1 = x^2 = \cdots = x^N. \tag{3}$$

To realize the consensus, the agents can communicate with each other to exchange variables x^i. Let $\mathcal{J}^i$ be a subset of $\mathcal{A} \setminus \{i\}$ and suppose that agent i can receive variable x^j from agents j that belong to $\mathcal{J}^i$. By this we define a graph $\mathcal{G} = (\mathcal{A}, \mathcal{E})$, where $\mathcal{E} = \{(j, i) \in \mathcal{A} \times \mathcal{A} : j \in \mathcal{J}^i\}$.

Assumption 1: Graph $\mathcal{G}$ is strongly connected.
We assume the following on functions f^i, g^i, h^i.

Assumption 2: Functions $f^i, g^i, i \in \mathcal{A}$ are proper subdifferentiable convex functions on $\mathbf{R}^n$ and h^i, $i \in \mathcal{A}$ are affine functions on $\mathbf{R}^n$.

Let $\partial f^i(x)$ denote the subdifferential of f^i at $x \in \mathbf{R}^n$. It holds for all $x \in \mathbf{R}^n$ and for all $v \in \partial f^i(x_0)$ that $f^i(x) \ge f^i(x_0) + \langle v, x - x_0\rangle$. A subdifferential is a closed set and by Assumption 2 $\partial f^i(x)$ and $\partial g^i(x)$ are nonempty.

Assumption 3: Functions $f^i, g^i, i \in \mathcal{A}$ are globally Lipschitz continuous on $\mathbf{R}^n$.

Since h^i's are affine, they are also Lipschitz continuous. Let L be a common upper bound of the Lipschitz constants of all these functions, namely $|f^i(x) - f^i(y)| \le L\|x - y\|$, $|g^i(x)-g^i(y)| \le L\|x-y\|$, and $\|h^i(x)-h^i(y)\| \le L\|x-y\|$ hold for all $x, y \in \mathbf{R}^n$ and for all $i \in \mathcal{A}$. Let $\mathbf{P}^N$ be a subset of $\mathbf{R}^N$ such that $\pi = \begin{bmatrix} \pi^1 & \pi^2 & \cdots & \pi^N \end{bmatrix}^\top \in \mathbf{P}^N$ satisfy $\pi^i > 0$, $i = 1, 2, \dots, N$ and $\sum_{i=1}^N \pi^i = 1$.

A solution to a Pareto optimal problem with convex functions is solved through the following problem with any vector $\pi \in \mathbf{P}^N$:

Problem 2: Find $x \in \mathbf{R}^n$ that minimizes

$$\sum_{i\in\mathcal{A}} \pi^i f^i(x^i) \tag{4}$$

subject to (2) and (3).

More precisely, any Pareto optimal solution is an optimal solution to Problem 2 for some $\pi \in \mathbf{P}^N$. Let x_* give a Pareto optimal solution of Problem 1, i.e., $x^1 = x^2 = \cdots = x^N = x_*$ attains a Pareto optimum. Then, for some vector $\pi \in \mathbf{P}^N$, $f_{*\pi} = \sum_{i\in\mathcal{A}} \pi^i f^i(x_*)$, $g^i(x_*) \le 0$, and $h^i(x_*) = 0$, $i \in \mathcal{A}$ and $f_{*\pi}$ is the minimum of Problem 2. Let $X_{*\pi}$ be the set of such x_* with vector π.

Assumption 4: $X_{*\pi} \ne \emptyset$.

If the objective functions are common over the network, i.e., if $f^1(x) = f^2(x) = \cdots = f^N(x) = f(x)$, Problem 1 provides the optimum of $\min f(x)$ with $g^i(x) \le 0$, $h^i(x) = 0$, $i \in \mathcal{A}$ since $\sum_{i\in\mathcal{A}} \pi^i = 1$, which is explored in [12]. An important application of this is the distributed min-max optimization:

Problem 3: $\min_{\tilde{x}^1,\tilde{x}^2,\dots,\tilde{x}^N} \max\{\tilde{f}^i(\tilde{x}^i) : i \in \mathcal{A}\}$ subject to $\tilde{g}^i(\tilde{x}^i) \le 0$, $\tilde{h}^i(\tilde{x}^i) = 0$ and consensus $\tilde{x}^1 = \tilde{x}^2 = \cdots = \tilde{x}^N$.

This problem is embedded in Problem 1 by defining $x^i = \begin{bmatrix} (\tilde{x}^i)^\top & \tilde{y}^i \end{bmatrix}^\top$, $f^i(x^i) = \tilde{y}^i$, $g^i(x^i) = \max\{\tilde{f}^i(\tilde{x}^i) - \tilde{y}^i, \tilde{g}^i(\tilde{x}^i)\}$, $h^i(x^i) = \tilde{h}^i(\tilde{x}^i)$, where $\tilde{y}^i \in \mathbf{R}$. A distributed min-max optimization with inequality constraints has been also considered in [5].

III. Distributed optimization protocol

For Problem 1, we provide a protocol by which the agents in $\mathcal{A}$ achieve consensus at a Pareto optimum of Problem 1. Let $W = (w^{ij}) \in \mathbf{R}^{N\times N}$ be an arbitrary matrix that meets the following assumptions:

Assumption 5: 1) $w^{ij} > 0$ if $i = j$ or $(j, i) \in \mathcal{E}$, and $w^{ij} = 0$ otherwise. 2) $W\mathbf{1}_N = \mathbf{1}_N$.

Notice that W is only right stochastic.

Define $G^i(x) = \max\{g^i(x), \|h^i(x)\|_\infty\}$ and let $N_{\rm iter}$ be a positive integer.

Protocol 1: Each agent i executes the following:

1) Let $x_1^i \in \mathbf{R}^n$, $i \in \mathcal{A}$ be arbitrary and $t := 1$.
2) Receive x_t^j, $j \in \mathcal{J}^i$ and compute $\xi_t^i = \sum_{j\in\mathcal{A}} w^{ij} x_t^j$.
3) Update x^i as

$$x_{t+1}^i := \xi_t^i - a_t u_t^i - b_t \tilde{v}_t^i, \tag{5}$$

where $u_t^i \in \partial f^i(\xi_t^i)$, $v_t^i \in \partial G^i(\xi_t^i)$ and

$$\tilde{v}_t^i := \begin{cases} v_t^i & \text{if } G^i(\xi_t^i) > c_t, \\ 0 & \text{if } G^i(\xi_t^i) \le c_t. \end{cases} \tag{6}$$

4) If $t < N_{\rm iter}$, then let $t := t + 1$ and go to 2.

In Protocol 1, step-length parameters a_t, b_t and c_t need to satisfy the following assumption.

Assumption 6: We assume that parameters a_t, b_t and c_t are positive numbers that converge to 0 as $t \to \infty$ and

$$\sum_{t=1}^\infty a_t = \infty, \quad \sum_{t=1}^\infty b_t^2 < \infty, \tag{7}$$

$$\lim_{t\to\infty} \frac{a_t}{b_t} = 0, \quad \lim_{t\to\infty} \frac{b_t^2}{a_t} = 0, \quad \lim_{t\to\infty} \frac{a_t}{b_t c_t} = 0. \tag{8}$$

Assumption 6 implies that there exist an integer $t_0 \ge 1$ and positive numbers r_1 such that

$$a_t \le r_1 b_t \quad \forall t \ge t_0. \tag{9}$$

Note that a_t is also square summable from (9). For example, $a_t = a_1/t^k$, $b_t = b_1/t^l$ and $c_t = c_1/t^m$ satisfy Assumption 6 if $0.5 < l < l + m < k \le 1$ and $a_1, b_1, c_1 > 0$.

Remark 1: The update law (5) can be written as

$$x_{t+1}^i := \xi_t^i - a_t \left(u_t^i - \frac{b_t}{a_t}\tilde{v}_t^i\right).$$

We remark that a standard exact penalty method is a gradient method with composite gradient of the objective and constraint functions, where the latter is multiplied by a constant positive number large enough [14]. The proposed protocol replaces such a constant with a sequence that goes to ∞ and also in the iterations guarantees the consensus and the convergence.

We make the following technical assumption on the feasible set of the problem, which is not restrictive in practice.

Assumption 7: For any $y \geq 0$ and for any $\pi \in \mathbf{P}^N$, the following set is bounded:

$$\mathcal{S}_\pi(y) = \Big\{ x \in \mathbf{R}^n : \sum_{i \in \mathcal{A}} \pi^i f^i(x) - f_{*\pi} \leq y, \\ g^i(x) \leq y,\ \|h^i(x)\|_\infty \leq y,\ i \in \mathcal{A} \Big\}. \quad (10)$$

Now we can state the main result of the paper.

Theorem 1: (i) Let Assumptions 1-7 hold. Then for any initial value $(x_1^1, x_1^2, \ldots, x_1^N)$, the sequence $\{(\xi_t^1, \xi_t^2, \ldots, \xi_t^N)\}$ generated by Protocol 1 satisfies

$$\lim_{t\to\infty} (\xi_t^i - \xi_t^j) = 0, \quad i, j \in \mathcal{A}, \quad (11)$$

$$\liminf_{t\to\infty} g^i(\xi_t^i) \leq 0, \quad i \in \mathcal{A}, \quad (12)$$

$$\liminf_{t\to\infty} \|h^i(\xi_t^i)\|_\infty = 0, \quad i \in \mathcal{A}, \quad (13)$$

$$\liminf_{t\to\infty} \sum_{i \in \mathcal{A}} \pi^i f^i(\xi_t^i) = f_{*\pi}, \quad (14)$$

where $\pi \in \mathbf{P}^N$ is a left eigenvector of W corresponding to eigenvalue 1[1].

Proof: The proof is shown in Section IV. ■

IV. Proof of Theorem 1

A. Proof of consensus

Let us prove the consensus (11) of Theorem 1. Define

$$\bar{x}_t := \sum_{i \in \mathcal{A}} \pi^i x_t^i, \quad \bar{u}_t := \sum_{i \in \mathcal{A}} \pi^i u_t^i,$$

$$\bar{v}_t := \sum_{i \in \mathcal{A}} \pi^i \tilde{v}_t^i, \quad x_t^\delta := \begin{bmatrix} x_t^1 - x_t^2 \\ \vdots \\ x_t^{N-1} - x_t^N \end{bmatrix}. \quad (15)$$

Then $\sum_{i \in \mathcal{A}} \pi^i \xi_t^i = \bar{x}_t$ and $\bar{x}_{t+1} = \bar{x}_t - a_t \bar{u}_t - b_t \bar{v}_t$. Moreover, for some $d > 0$, it holds that

$$\|x_t^i - \bar{x}_t\| \leq d\|x_t^\delta\|, \quad \|\xi_t^i - \bar{x}_t\| \leq d\|x_t^\delta\| \quad (16)$$

for $i \in \mathcal{A}$. Define

$$u_t^\delta := \begin{bmatrix} u_t^1 - u_t^2 \\ \vdots \\ u_t^{N-1} - u_t^N \end{bmatrix}, \quad \tilde{v}_t^\delta := \begin{bmatrix} \tilde{v}_t^1 - \tilde{v}_t^2 \\ \vdots \\ \tilde{v}_t^{N-1} - \tilde{v}_t^N \end{bmatrix}.$$

Lemma 1: The sequence $\{x_t^\delta\}$ is square summable.

Proof: Define

$$Z = \begin{bmatrix} 1 & -1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 1 & -1 \end{bmatrix} \in \mathbf{R}^{N \times (N-1)}$$

[1]Since $\mathcal{G}$ is connected, such a π uniquely exists.

and $T = [\ Z^\top \quad \pi\ ]^\top$, $\Pi = \mathrm{diag}\{\pi^1, \ldots, \pi^N\}$. Then it is easy to see that

$$T^{-1} = [\ \Pi^{-1} Z^\top (Z \Pi^{-1} Z^\top)^{-1} \quad \mathbf{1}_N\ ]$$

and

$$T^{-1} W T = \begin{bmatrix} W_0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$W_0 = Z W \Pi^{-1} Z^\top (Z \Pi^{-1} Z^\top)^{-1},$$

where the magnitude of the eigenvalues of W_0 is less than 1. Moreover, x_t^δ satisfies

$$x_{t+1}^\delta = (W_0 \otimes I_n) x_t^\delta - a_t u_t^\delta - b_t \tilde{v}_t^\delta. \quad (17)$$

From the assumptions in Section III, a_t and b_t are square summable and u_t^δ and $\tilde{v}_t^\delta$ are bounded. Hence (17) is a stable linear discrete-time system with l_2 inputs, which implies that x_t^δ is square summable. ■

From this lemma and (16), $\xi_t^i - \bar{x}_t$, $i \in \mathcal{A}$ is also square summable. Therefore $\lim_{t\to\infty}(\xi_t^i - \xi_t^j) = 0$, $i, j \in \mathcal{A}$.

B. Boundedness

To prove the boundedness of the sequences $\{x_t^i\}$ and $\{\xi_t^i\}$, let us see some preliminaries first. Define

$$H_\pi(x) = \max\Big\{ \sum_{i \in \mathcal{A}} \pi^i f^i(x) - f_{*\pi},\, G^1(x), \ldots, G^N(x) \Big\}.$$

It holds that $H_\pi(x) \geq 0$ for all $x \in \mathbf{R}^n$ and $x \in X_{*\pi}$ iff $H_\pi(x) = 0$. Below let $x_{*\pi} \in X_{*\pi}$.

Lemma 2: For any $\pi \in \mathbf{P}^N$, there exist $p > 0$ and $q \in \mathbf{R}$ for which

$$H_\pi(x) \geq p\|x - x_{*\pi}\| - q \quad (18)$$

holds for any $x \in \mathbf{R}^n$.

Proof: Fix $\pi \in \boldsymbol{P}^N$. From Assumptions 4 and 7, $\in X_{*\pi} = \mathcal{S}_\pi(0) = \{x \in \mathbf{R}^n : H_\pi(x) = 0\}$ is not empty and bounded. Noticing that $x_{*\pi} \in X_{*\pi}$, we see that there exist $\rho_0 > 0$ such that $X_{*\pi} \subset \overline{\mathrm{B}}(\rho_0; x_{*\pi})$. Similarly, $\mathcal{S}_\pi(1) = \{x \in \mathbf{R}^n : H_\pi(x) \leq 1\}$ is bounded and hence there exist $\rho_1 > \rho_0$ such that $\{x \in \mathbf{R}^n : H_\pi(x) \leq 1\} \subset \overline{\mathrm{B}}(\rho_1; x_{*\pi})$. Then it is easy to see that $p = 1/(\rho_1 - \rho_0)$ and $q = \rho_1/(\rho_1 - \rho_0)$ satisfy (18). ■

In the following, Lemma 2 is extended to the sequence $\{\xi_t^i\}$:

Lemma 3: For any $\pi \in \mathbf{P}^N$, there exist $p > 0$ and $q_1 \in \mathbf{R}$ such that

$$\max\Big\{ \sum_{i \in \mathcal{A}} \pi^i f^i(\xi_t^i) - f_{*\pi},\, G^1(\xi_t^1), \ldots, G^N(\xi_t^N) \Big\} \\ \geq p\|\bar{x}_t - x_{*\pi}\| - q_1, \quad t = 1, 2, \ldots. \quad (19)$$

Proof: From Lemma 2, there exists $p > 0$ and q such that $H_\pi(x) \geq p\|x - x_{*\pi}\| - q$ holds for all $x \in \mathbf{R}^n$. From (16),

$$\max\Big\{ \sum_{i \in \mathcal{A}} \pi^i f^i(\xi_t^i) - f_{*\pi},\, G^1(\xi_t^1), \ldots, G^N(\xi_t^N) \Big\} \\ \geq H_\pi(\bar{x}_t) - Ld\|x_t^\delta\| \\ \geq p\|\bar{x}_t - x_{*\pi}\| - q - Ld \sup_{t \geq 1} \|x_t^\delta\|, \quad t = 1, 2, \ldots.$$

From Lemma 1, x_t^δ is square summable. Hence $\sup_{t\ge 1}\|x_t^\delta - x_{*\pi}\| < \infty$. This implies that $q_1 = q + Ld\sup_{t\ge 1}\|x_t^\delta\|$ completes the proof. ∎

Let $\mathcal{A}_t^+$ denote the set of $i \in \mathcal{A}$ for which $G^i(\xi_t^i) > c_t$. From (5) and (9), the following inequality holds for all $t \ge t_0$:

$$\begin{aligned}\sum_{i\in\mathcal{A}}&\pi^i\|x_{*\pi} - x_{t+1}^i\|^2\\ &= \sum_{i\in\mathcal{A}}\pi^i\left\|x_{*\pi} + a_t u_t^i + b_t\tilde{v}_t^i - \xi_t^i\right\|^2\\ &= \sum_{i\in\mathcal{A}}\pi^i\Big\{\|x_{*\pi} - \xi_t^i\|^2 + 2a_t\left\langle u_t^i, x_{*\pi} - \xi_t^i\right\rangle\\ &\qquad +2b_t\left\langle \tilde{v}_t^i, x_{*\pi} - \xi_t^i\right\rangle + \|a_t u_t^i + b_t\tilde{v}_t^i\|^2\Big\}\\ &\le \sum_{i\in\mathcal{A}}\pi^i\Big\{\|x_{*\pi} - x_t^i\|^2 + 2a_t(f^i(x_{*\pi}) - f^i(\xi_t^i))\Big\}\\ &\qquad + \sum_{i\in\mathcal{A}_t^+} 2\pi^i b_t(G^i(x_{*\pi}) - G^i(\xi_t^i)) + L_2 b_t^2 \end{aligned} \tag{20}$$

where $L_2 := 2L^2(1 + r_1^2)$. This inequality will be exploited below. Define $D_t = \sum_{i\in\mathcal{A}}\pi^i\|x_{*\pi} - x_t^i\|^2$.

Lemma 4: The sequences $\{x_t^i\}$ and $\{\xi_t^i\}$ are bounded.

Proof: First, consider positive integers $t \ge t_0$ for which x_t^i satisfies the following inequality:

$$\sum_{i\in\mathcal{A}}\pi^i\|x_{*\pi} - x_t^i\|^2 \le \sum_{i\in\mathcal{A}}\pi^i\|x_{*\pi} - x_{t+1}^i\|^2. \tag{21}$$

We obtain an inequality on the average $\bar{x}_t$ for such t's. If t satisfies the condition (21), it holds from (20) that

$$\begin{aligned}0 \le &\sum_{i\in\mathcal{A}} 2\pi^i a_t(f^i(x_{*\pi}) - f^i(\xi_t^i))\\ &+ \sum_{i\in\mathcal{A}_t^+} 2\pi^i b_t(G^i(x_{*\pi}) - G^i(\xi_t^i)) + L_2 b_t^2.\end{aligned}$$

Since $G^i(x_{*\pi}) = 0$, we have

$$2a_t\left(\sum_{i\in\mathcal{A}}\pi^i f^i(\xi_t^i) - f_{*\pi}\right) + \sum_{i\in\mathcal{A}} 2\pi^i b_t G^i(\xi_t^i) \le L_2 b_t^2. \tag{22}$$

From Lemma 3,

$$\text{LHS of (22)} \ge 2(a_t + b_t)(p\|\bar{x}_t - x_{*\pi}\| - q_1)$$

and hence

$$\begin{aligned}p\|\bar{x}_t - x_{*\pi}\| - q_1 &\le \frac{L_2 b_t^2}{2(a_t + b_t)} \le \frac{L_2}{4}\max\left\{\frac{b_t^2}{a_t},\, b_t\right\}\\ \to 0 &\quad \text{as} \quad t \to \infty.\end{aligned}$$

This implies that $\bar{x}_t$ is bounded. Since $\lim_{t\to\infty}(x_t^i - \bar{x}_t) \to 0$ holds from the consensus proved above, we can see that there exists a constant $c > 0$ such that $\|x_t^i - x_{*\pi}\| \le \alpha$ for any t that satisfies (21). Therefore, for any $t \ge 1$, it holds that $D_t > D_{t+1}$ or $D_t \le \alpha^2$. This implies the boundedness of $\{x_t^i\}$, by which the boundedness of $\{\xi_t^i\}$ is obvious. ∎

C. *Proof of convergence*

To prove (12)–(14), assume that there exist $\varepsilon > 0$ and integer $t_1 \ge t_0$ such that for all $t \ge t_1$ it holds that

$$\sum_{i\in\mathcal{A}}\pi^i f^i(\xi_t^i) > f_{*\pi} + \varepsilon \quad \text{or} \quad \max_{i\in\mathcal{A}} G^i(\xi_t^i) > c_t \tag{23}$$

and lead to a contradiction. Let $t \ge t_1$.

First, suppose that $\mathcal{A}_t^+$ is empty. Then it holds that $G^i(\xi_t^i) \le c_t$ for all $i \in \mathcal{A}$. and hence $\sum_{i\in\mathcal{A}}\pi^i f^i(\xi_t^i) > f_{*\pi} + \varepsilon$ from (23). Hence

$$D_{t+1} < D_t - 2\varepsilon a_t + L_2 b_t^2. \tag{24}$$

Next, suppose that the set $\mathcal{A}_t^+$ is not empty. Recall Lemma 4 to see $M := \sup_{t\ge 1}\max_{i\in\mathcal{A}}\|x_{*\pi} - \xi_t^i\| < \infty$. Setting $\underline{\pi} = \min_{1\le i\le N}\pi^i (> 0)$, observe that

$$\begin{aligned}&\text{RHS of (20)}\\ &= \sum_{i\in\mathcal{A}}\pi^i\left\{\|x_{*\pi} - x_t^i\|^2 + 2a_t(f^i(x_{*\pi}) - f^i(\xi_t^i))\right\}\\ &\qquad + \sum_{i\in\mathcal{A}_t^+} 2\pi^i b_t(G^i(x_{*\pi}) - G^i(\xi_t^i)) + L_2 b_t^2\\ &< D_t + \sum_{i\in\mathcal{A}} 2\pi^i a_t L\|x_{*\pi} - x_t^i\| - 2\underline{\pi} b_t c_t + L_2 b_t^2\\ &\le D_t + 2LMa_t - 2\underline{\pi} b_t c_t + L_2 b_t^2. \end{aligned} \tag{25}$$

Since $b_t c_t / a_t \to \infty$ as $t \to \infty$ from Assumption 6, there exists an integer $t_2 \ge t_1$ such that

$$2LM + 2\varepsilon \le 2\underline{\pi}\frac{b_t c_t}{a_t} \quad \forall t \ge t_2,$$

which implies

$$\text{RHS of (25)} \le D_t - 2\varepsilon a_t + L_2 b_t^2 \quad \forall t \ge t_2. \tag{26}$$

From (24) and (26), it can be seen that

$$\varepsilon < \frac{D_{t_2} + L_2\sum_{\tau=t_2}^t b_\tau^2}{2\sum_{\tau=t_2}^t a_\tau} \tag{27}$$

is true for all $t \ge t_2$. Recall that b_t is square summable and a_t is not summable. Therefore the RHS of (27) converges to zero as $t \to \infty$ and contradicts $\varepsilon > 0$. Thus (23) is found false and hence

$$\begin{aligned}\forall\varepsilon > 0 \quad \forall l \ge t_2 \quad \exists t_l \ge l \quad &\sum_{i\in\mathcal{A}}\pi^i f^i(\xi_{t_l}^i) \le f_{*\pi} + \varepsilon\\ \text{and} \quad &G^i(\xi_{t_l}^i) \le c_{t_l}, \quad i \in \mathcal{A}.\end{aligned}$$

This completes the proof.

V. Numerical example

Consider a directed graph of five agents with $\mathcal{E} = \{(1,2),(2,3),(2,4),(3,2),(3,4),(4,1),(4,5),(5,1),(5,3)\}$. The graph is not balanced. We set

$$W = \begin{bmatrix} 0.2 & 0 & 0 & 0.4 & 0.4\\ 0.4 & 0.2 & 0.4 & 0 & 0\\ 0 & 0.4 & 0.6 & 0 & 0\\ 0 & 0.4 & 0.4 & 0.2 & 0\\ 0 & 0 & 0 & 0.4 & 0.6 \end{bmatrix},$$

which is only right stochastic. Each $f^i(x)$ is a linear function of $x \in \mathbf{R}^5$ and each inequality $g^i(x) \le 0$ consists of three affine inequalities. The step-length parameters are set as $a_t = 10/t$, $b_t = 10/t^{0.7}$, and $c_t = 0.001/t^{0.2}$. Agent 1 has a one-dimensional affine equality constraint and the other agents do not have equality constraints. Protocol 1 is executed as shown in Figs. 1–5, which show ξ_t^i, the weighted sum of $f^i(\xi_t^i)$, $g^i(\xi_t^i)$, $h^1(\xi_t^1)$ and the error of the solution, respectively. The error is the norm of the difference between ξ_t^i and the optimal solution of Problem 1 with centralized setting (i.e., $x^1, x^2, \ldots, x^N$ is replaced with common x), where $x_{*\pi} = [\, 0.3560 \quad -0.1290 \quad 0.3044 \quad 2.0178 \quad -1.2263 \,]^\top$ and $f_{*\pi} = 5.4779$. Figs. 1–5 show that Under Protocol 1, which does not depend on π, ξ_t^i converges to the optimal solution with satisfying the equality and inequality constraints.

Next, we solved a minimax problem for the same objective and constraint functions. The results of Protocol 1 adopted to Problem 3 are shown in Figs. 6–10 for ξ_t^i, $f^i(\xi_t^i)$, $g^i(\xi_t^i)$, $h^1(\xi_t^1)$ and the error, respectively, where the error is computed in the same way as the first example. The solution of the centralized problem is $x_{*\pi} = [\, -0.4262 \quad 0.2623 \quad 0.8934 \quad -0.0902 \quad 0.2951 \,]^\top$ and $f_{*\pi} = 9.2377$. Figs. 6–10 verify the convergence to the optimal solution.

VI. Conclusions

In this paper, we proposed a distributed multi-agent optimization protocol to solve a Pareto optimal problem with equality and inequality constraints. The graph of the network does not need to be balanced but is needed to be a strongly connected undirected graph. The protocol is based on exact penalty methods and extends the previous results of [9], [10] to the Pareto optimal problem. Equality constraints can be included in the proposed protocol. In particular, if the agents have a common objective function, the protocol attains the optimal of the objective function [12], which is applied to the min-max optimization. Numerical examples demonstrated that the proposed protocol attains a consensus at a Pareto optimal solution and solves a distributed min-max optimization problem over an unbalanced network.

Acknowledgment

This work was supported by COOPS, International Bilateral Joint CNR Laboratories, and JST CREST Grant Number JPMJCR15K2, Japan.

References

[1] A. Jadbabaie, J. Lin and A. S. Morse: Coordination of groups of mobile autonomous agents using nearest neighbor rules, *IEEE Transactions on Automatic Control*, Vol. 48, No. 6, pp. 988-1001 (2003)

[2] R. Olfati-Saber and R. M. Murray: Consensus problems in networks of agents with switching topology and time-delays, *IEEE Transactions on Automatic Control*, Vol. 49, No. 9, pp. 1520-1533 (2004)

[3] A. Nedić and A. Ozdaglar: Distributed subgradient methods for multi-agent optimization, *IEEE Transactions on Automatic Control*, Vol. 54, No. 1, pp. 48-61 (2009)

[4] A. Nedić, A. Ozdaglar and P. A. Parrilo: Constrained consensus and optimization in multi-agent networks, *IEEE Transactions on Automatic Control*, Vol. 55, No. 4, pp. 922-938 (2010)

[5] K. Srivastava, A. Nedić and D. Stipanović: Distributed min-max optimization in networks, *Proceedings of the 17th Digital Signal Processing Conference* (2011)

[6] M. Zhu and S. Martínez: On distributed convex optimization under inequality and equality constraints, *IEEE Transactions on Automatic Control*, Vol. 57, No. 1, pp. 151-164 (2012)

[7] J. C. Duchi, A. Agarwal and M. J. Wainwright: Dual averaging for distributed optimization: convergence analysis and network scaling, *IEEE Transactions on Automatic Control*, Vol. 57, No. 3, pp. 592-606 (2012)

[8] T.-H. Chang, A. Nedić and A. Scaglione: Distributed constrained optimization by consensus-based primal-dual perturbation method, *IEEE Transactions on Automatic Control*, Vol. 59, No. 6, pp. 1524-1538 (2014)

[9] I. Masubuchi, T. Wada, N. T. Asai, L. T. H. Nguyen, Y. Ohta and Y. Fujisaki: Distributed constrained optimization protocol via an exact penalty method, *Proceedings of the 14th European Control Conference*, pp. 1480-1485 (2015)

[10] I. Masubuchi, T. Wada, N. T. Asai, L. T. H. Nguyen, Y. Ohta and Y. Fujisaki: Distributed multi-agent optimization based on an exact penalty method with equality and inequality constraints, *SICE Journal of Control, Measurement, and System Integration*, Vol. 9, No. 4, pp. 179-186 (2016)

[11] T. Wada and Y. Fujisaki: Distributed optimization over directed unbalanced networks, *Proceedings of the SICE International Symposium on Control Systems 2016*, SY0003/16/0000-870 (2016)

[12] P. Xie, K. You, R. Tempo, S. Song, and C. Wu: Distributed convex optimization with inequality constraints over time-varying unbalanced digraphs, *arXive:* 1612.09029v1 [cs.DC] (2016)

[13] T. Wada, I. Masubuchi, K. Hanada, T. Asai, and Y. Fujisaki: Distributed multi-objective optimization over randomly varying unbalanced network, *Proceedings of the 20th IFAC World Congress*, pp. 2403-2408 (2017)

[14] D. P. Bertsekas: *Nonlinear Programming*, Athena Scientific (1999)

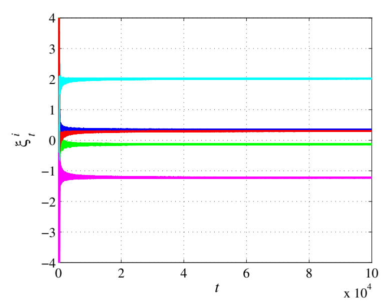

Fig. 1. ξ_t^i

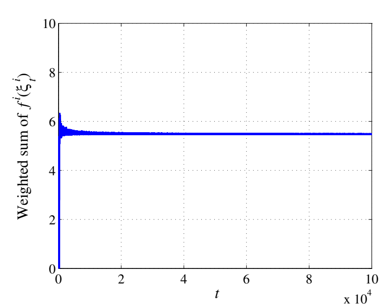

Fig. 2. Weighted sum of objective functions

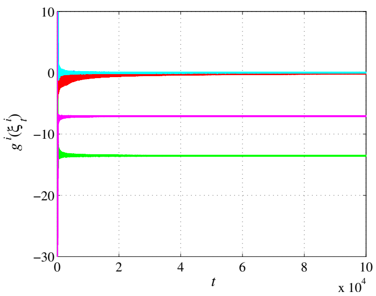

Fig. 3. Inequality constraints

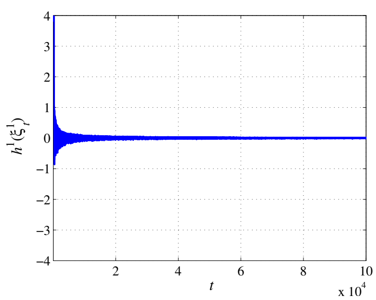

Fig. 4. Equality constraint

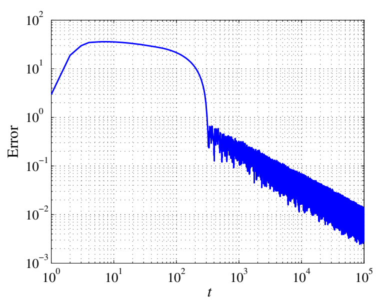

Fig. 5. Error of solution

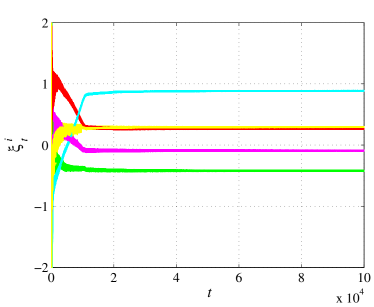

Fig. 6. ξ_t^i

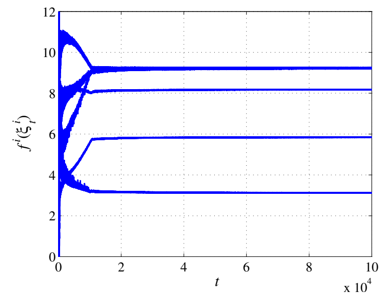

Fig. 7. Objective functions

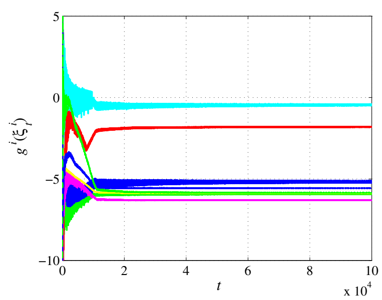

Fig. 8. Inequality constraints

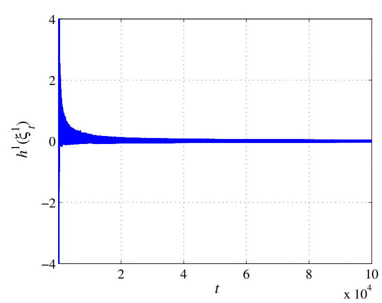

Fig. 9. Equality constraint

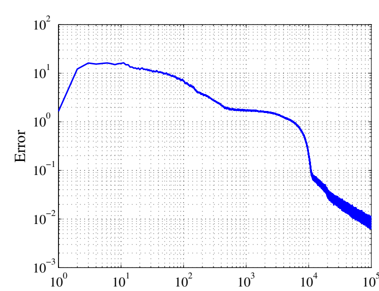

Fig. 10. Error of solution