2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

Distributed Algorithm for Achieving Minimum l_1 Norm Solutions of Linear Equation

Jingqiu Zhou, Wang Xuan, Shaoshuai Mou, and Brian. D. O. Anderson

***Abstract*— This paper proposes a distributed algorithm for multi-agent networks to achieve a minimum l_1-norm solution to a linear equation $Ax = b$ where A has full row rank. When the underlying network is undirected and fixed, it is proved that the proposed algorithm drive all agents' individual states to converge in finite-time to the same minimum l_1-norm solution. Numerical simulations are also provided as validation of the proposed algorithms.**

I. Introduction

A significant amount of effort in the control community has recently been given to distributed algorithms for solving linear equations over multi-agent networks, in which each agent only knows part of the equation and controls a state vector that can be looked at as an estimate of the solution of the overall linear equations [1]–[5]. Numerous extensions along this direction include achieving solutions with the minimum Euclidean norm [6], [7], elimination of the initialization step [8], reduction of state vector dimension by utilizing the sparsity of the linear equation [9] and achieving least square solutions [10]–[15]. All these algorithms yield asymptotic convergence, but require an infinite number of sensing or communication events.

Solutions to underdetermined linear equations with the minimum l_1 norm are perhaps the most important in many engineering applications including earthquake location detection [16], analysis of statistical data [17], solving biomeganetic inverse problems [18], and so on. One most intriguing case among these applications is compressive sensing, which enables transmission of sparse data in a very efficient way [19]. The decoding process of compressive sensing requires solving of linear equations with a minimum number of non-zero entries of the solution vectors, which, however, is an NP-hard problem and usually computationally costly [20]. Thus researchers usually turn to achieve solutions with minimum l_1 norm instead for which the function to be minimized is convex [21], [22]. Most existing results for achieving minimum l_1 norm solutions are based on the idea of Lasso [23] including Alternating Direction Method of Multipliers (ADMM) [24], the Primal-Dual Interior-Point Method [25], [26], Gradient Projection Methods [27], Homotopy Methods [28], Iterative Shrinkage-Thresholding Methods [29] and Proximal Gradient Methods [30]. Interesting as these results are, they either achieve limited accuracy dominated by a threshold parameter, involve solving a much larger linear equation, or lead to high computational complexity.

In this paper we aim to develop distributed algorithms for multi-agent networks to achieve in **finite time** a solution of linear equations with the minimum l_1 norm. By distributed is meant that each agent only knows part of the overall linear equation and can communicate with only its nearby neighbors. The problem of interest is formulated in Section II. We introduce in section III the concepts to be employed in the paper including Filippov set-valued maps, Filippov solutions, generalized Lie derivatives, based on which a preliminary result is achieved. In Section IV, we will first propose a centralized update for achieving a solution with the minimum l_1 norm. Motivated by the projection-consensus flow proposed in [13] and the finite-time gradient flow for consensus devised in [31]–[34], we utilize a combination of the finite time consensus flow and the proposed centralized algorithm to develop a distributed linear equation solver for achieving a minimum l_1 norm solution, which is shown to converge in finite time. We provide simulations in Section V and concluding remarks in Section VI.

Notation: Let r denote an arbitrary positive integer. Let $\mathbf{1}_r$ denote the vector in $\mathbb{R}^r$ with all entries equal to 1s. Let I_r denote the $r \times r$ identity matrix. We let $\mathrm{col}\,\{A_1, A_2, \cdots, A_r\}$ be a stack of matrices A_i possessing the same number of columns with the index in a top-down ascending order, $i = 1, 2, \cdots, r$. Let $\mathrm{diag}\,\{A_1, A_2, \cdots, A_r\}$ denote a block diagonal matrix with A_i the ith diagonal block entry, $i = 1, 2, \cdots, r$. By $M > 0$ and $M \geq 0$ are meant that the square matrix M is positive definite and positive semi-definite, respectively. By M' is meant the transpose of a matrix M. Let $\ker M$ and $\mathrm{image}\, M$ denote the kernel and image of a matrix M, respectively. Let $\otimes$ denote the Kronecker product. Let $\|\cdot\|_1$ denote the l_1 norm of a vector in $\mathbb{R}^r$.

II. Problem Formulation

Consider a network of m agents, $i = 1, 2, ..., m$; inside this network, each agent can observe states of certain other agents called its *neighbors*. Let $\mathcal{N}_i$ denote the set of agent i's neighbors. We assume that the neighbor relation is symmetric, that is, $j \in \mathcal{N}_i$ if and only if $i \in \mathcal{N}_j$. Then all

This work was supported by a funding from Northrop Grumman Cooperation. J. Zhou, X. Wang and S. Mou are with the School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47906 USA (e-mail: zhou745@purdue.edu, wang3156@purdue.edu, mous@purdue.edu). B. D. O. Anderson is with Hangzhou Dianzi University, Hangzhou, China, The Australian National University and Data-61 CSIRO (formerly NICTA), Canberra, ACT 2600 Australia, e-mail: Brian.Anderson@anu.edu.au; his work is supported by Data-61, CSIRO, and by the Australian Research Council's Discovery Projects DP-130103610 and DP-160104500. Corresponding Author: Shaoshuai Mou.

Proofs of the assertions in this paper are available from the first author upon request and will appear in a full length version of this paper.

these neighbor relations can be described by an m-node-$\bar{m}$-edge undirected graph $\mathbb{G}$ such that there is an undirected edge connecting i and j if and only if i and j are neighbors. In this paper we only consider the case in which $\mathbb{G}$ is connected, fixed and undirected.

Suppose that each agent i knows $A_i \in \mathbb{R}^{n_i \times n}$ and $b_i \in \mathbb{R}^{n_i}$ and controls a state vector $y_i(t) \in \mathbb{R}^n$. Then all these A_i and b_i can be stacked into an overall equation $Ax = b$, where $A = \text{col}\,\{A_1, A_2, \cdots, A_m\}$, $b = \text{col}\,\{b_1, b_2, \cdots, b_m\}$. Without loss of generality for the problems of interest to us, we assume A to have full row-rank. Let $\bar{x}^*$ denote a *minimum l_1-norm solution* to $Ax = b$, that is,

$$\bar{x}^* = \arg\min_{Ax=b} \|x\|_1 \tag{1}$$

The problem of interest in this paper is to develop distributed algorithms for each agent i to update its state vector $y_i(t)$ by only using its neighbors' states such that all $y_i(t)$ to converge in finite time to a common minimum l_1-norm solution $\bar{x}^*$.

III. Key Concepts and Preliminary Results

Before proceeding, we introduce some key concepts and preliminary results for future derivation and analysis. key references for the background we summarize are [35] and [36].

A. Filippov set-valued maps and Filippov Solutions

By a *Filippov set-valued map* $F[f] : \mathbb{R}^r \to \mathcal{B} \subset \mathbb{R}^r$ associated with a function $f : \mathbb{R}^r \to \mathbb{R}^r$ is meant

$$F[f](x) \triangleq \bigcap_{\delta>0} \bigcap_{\mu(\mathcal{S})=0} \overline{\text{co}}\{f(B(x,\delta))/\mathcal{S}\} \tag{2}$$

Here $B(x,\delta)$ stands for the open ball on $\mathbb{R}^r$, whose center is at x and has a radius of δ; $\mu(\mathcal{S})$ denotes the Lebesgue measure of $\mathcal{S}$; and $\overline{\text{co}}$ stands for the convex closure. Let $\text{sgn}\,(x) : \mathbb{R}^r \to \mathbb{R}^r$ be a function with the kth entry, $k = 1, 2, ..., r$, defined as

$$(\text{sgn}\,(x))_k = \begin{cases} 1, & (x)_k > 0; \\ -1, & (x)_k < 0; \\ 0, & (x)_k = 0. \end{cases} \tag{3}$$

It follows that the Filippov set-valued map $F[\text{sgn}\,](x)$ for $x \in \mathbb{R}^r$ is defined entrywise as:

$$(F[\text{sgn}\,](x))_k = \begin{cases} 1 & (x)_k > 0 \\ [-1,1] & (x)_k = 0 \\ -1 & (x)_k < 0 \end{cases} \tag{4}$$

for $k = 1, 2, ..., r$. Note that even if $(x)_i = (x)_j = 0$, the ith and jth entries of a vector in $F[\text{sgn}\,](x)$ may not necessarily be equal since each of them could be chosen as arbitrary values in the interval $[-1,1]$. From the definition of $F[\text{sgn}\,](x)$, one can verify that

$$q'x = \|x\|_1, \quad \forall q \in F[\text{sgn}\,](x) \tag{5}$$

While for any $w \in \mathbb{R}^r$, there holds

$$q'w \leq \|w\|_1 \tag{6}$$

.

By *a Filippov solution* for $\dot{x} \in F[f](x)$ is meant a Caratheodory solution $x(t)$ such that $\dot{x} \in F[f](x)$ for almost all t, $x(t)$ is absolutely continuous and can be written in the form of an indefinite integral. The following two lemmas treat existence of such a Filippov solution.

Lemma 1: (Proposition 3 in [35]) If $f : \mathbb{R}^r \to \mathbb{R}^r$ is measurable and locally bounded, then for any initial point $x_0 \in \mathbb{R}^r$, there exists a Filippov solution[1] for $\dot{x} \in F[f](x)$.

Lemma 2: (Theorem 8 in page 85 of [36]) Let a vector-valued $f(t,x)$ be defined almost-everywhere in the domain G of time space (t,x). With $f(t,x)$ measurable and locally bounded almost-everywhere in an open domain G, let

$$F[f](t,x) \triangleq \bigcap_{\delta>0} \bigcap_{\mu(\mathcal{S})=0} \overline{\text{co}}\{f(t, B(x,\delta)/\mathcal{S})\}. \tag{7}$$

Then for any point $(t_0, x_0) \subset G$, there exists a Filippov solution of $\dot{x} \in F[f](t,x)$ with $x(t_0) = x_0$.

Note that Lemma 2 establishes the existence of a solution for time-varying systems. This is more general than Lemma 1, which only guarantees the existence of solutions to time-invariant systems.

B. Generalized Gradients and Generalized Lie Derivatives

For a locally Lipschitz function $w : \mathbb{R}^r \to \mathbb{R}$, the *generalized gradient* of w is

$$\partial w(x) \triangleq \text{co}\{\lim_{i\to\infty} \nabla w(x_i) : x_i \to x, x_i \notin \mathcal{S} \cup \Omega_w\} \tag{8}$$

where $\mathcal{S} \subset \mathbb{R}^r$ is an arbitrarily chosen set of measure zero, Ω_w denotes the set of points at which w is not differentiable, and co denotes convex hull. Specially, for the function $||x||_1$, one computes the kth element of its generalized gradient to be:

$$(\partial ||x||_1)_k = \begin{cases} 1 & x_k > 0 \\ [-1,1] & x_k = 0 \\ -1 & x_k < 0 \end{cases} \tag{9}$$

It follows from this and the definition of $F[\text{sgn}\,](x)$ in (4) that

$$F[\text{sgn}\,](x) = \partial \|x\|_1 \tag{10}$$

For a set-valued map $\mathcal{F} : \mathbb{R}^r \to \mathcal{B}(\mathbb{R}^r)$, the *generalized Lie derivative* of w is defined as

$$\begin{aligned} \tilde{\mathcal{L}}_{\mathcal{F}} w(x) = \{ & q \in \mathbb{R} : \text{there exists } \alpha \in \mathcal{F}(x) \\ & \text{such that } \forall \beta \in \partial w(x), q = \beta'\alpha\} \end{aligned} \tag{11}$$

The above definition of generalized Lie derivative implies that for each α in $\mathcal{F}(x)$, we check if the inner product $\beta'\alpha$ is a fixed value for all $\beta \in \partial w(x)$. If so, this inner product is an element in $\tilde{\mathcal{L}}_{\mathcal{F}} w(x)$, but note that the set $\tilde{\mathcal{L}}_{\mathcal{F}} w(x)$ may be empty. Moreover, for locally Lipschitz and regular(see [37],p.39 and [38], p.3 for detailed discussion of regular function[2]) functions $w(x)$, one has the following lemma:

[1]There is no implication that the solution exists on an infinite interval

[2]That a function $w(x) : \mathbb{R}^n \to \mathbb{R}$ is called regular at $x \in \mathbb{R}^n$ if

1) for all $v \in \mathbb{R}^n$ there exists the usual right directional derivative $w'_+(x,v)$.
2) for all $v \in \mathbb{R}^n$, $w'_+(x,v) = w^o(x,v)$.

Lemma 3: (Proposition 10 in [35]) Let $x : [0, t_1] \to \mathbb{R}^r$ be a solution for $\dot{x} \in \mathcal{F}(x(t))$, where $\mathcal{F}$ is any set-valued map. Let $w(x) : \mathbb{R}^r \to \mathbb{R}$ be locally Lipschitz and regular. Then $w(x(t))$ is differentiable at almost all $t \in [0, t_1]$; The derivative of $w(x(t))$ satisfies $\frac{dw(x(t))}{dt} \in \tilde{\mathcal{L}}_{\mathcal{F}} w(x(t))$ for almost all $t \in [0, t_1]$.

Lemma 3 guarantees the existence of generalized Lie derivatives for functions that are locally Lipschitz and regular. If one focuses on a specific solution, one can show that α in (11) is a special vector as summarized in the following lemma.

Lemma 4: (See Proof of Lemma 1 in [38]) Let $x(t)$ denote a specific solution of a differential enclosure. Suppose $w(x)$ is locally Lipschitz and regular. Let $\mathcal{I} \subset [0, \infty)$ denote the time interval for which $\dot{x}(t)$ exists. Then

$$\frac{dw(x(t)))}{dt} = \beta' \dot{x}(t) \tag{12}$$

where β is any vector in $\partial w(x)$.

C. *Preliminary Results*

For any positive semi-definite matrix $M \in \mathbb{R}^{r \times r}$, $M \neq 0$ one can define

$$\Phi(M) = \{q \in \mathbb{R}^r | \, \exists \phi \in F[\text{sgn } (q)], M\phi(q) = 0\} \tag{13}$$

and its compliment

$$\Phi_c(M) = \{q \in \mathbb{R}^r | \, \forall \phi \in F[\text{sgn } (q)], M\phi(q) \neq 0\}. \tag{14}$$

We impose a further requirement on M, namely that $\Phi_c(M)$ is nonempty. This is easily ensured, see below. Let

$$\Lambda(M) = \{\phi | \phi \in F[\text{sgn }](q), q \in \Phi_c(M)\} \tag{15}$$

Now $F[\text{sgn }](q)$ is a closed set for any fixed q; also note that $F[\text{sgn }](q)$ can only be one of a finite number of different sets; hence it is easy to check for a given M whether $\Phi_c(M)$ is nonempty, (and in a later use of the result, it proves easy to check). It further follows that $\Lambda(M)$ is also a closed set. Consequently, the continuous function $f(\phi) = \phi' M \phi$ has a nonzero minimum on $\Lambda(M)$. We denote

$$\lambda(M) = \min_{\phi \in \Lambda(M)} f(\phi). \tag{16}$$

From the definition of $\Phi_c(M)$ and $\Lambda(M)$, one has $\lambda(M) > 0$. To summarize, one has the following lemma:

Lemma 5: For any nonnegative-definite matrix M, we let $\Phi_c(M)$, $\Lambda(M)$ and $\lambda(M)$ defined as above. Suppose that $\Phi_c(M)$ is nonempty. Then $\lambda(M)$ is a positive constant.

For the m-node-$\bar{m}$-edge graph $\mathbb{G}$, we label all its nodes as $1, 2, ..., m$ and all its edges as $1, 2, ..., \bar{m}$. Assign an arbitrary direction to each edge in $\mathbb{G}$. Then the incidence matrix of $\mathbb{G}$ denoted by $H = [h_{ik}]_{m \times \bar{m}}$ is defined as follows

$$h_{ik} = \begin{cases} 1, & i \text{ is the head of the } k\text{th edge;} \\ -1, & i \text{ is the tail of the } k\text{th edge;} \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Since $\mathbb{G}$ is connected, then $\ker H'$ is the span of $\mathbf{1}_m$ [39]. Moreover, one has the following lemma:

Lemma 6: Suppose A has full-row rank and $\mathbb{G}$ is connected. Let $\bar{P} = \text{diag}\,\{P_1, P_2, ..., P_m\}$ where each P_i is the projection matrix to $\ker A_i$. Let $\bar{H} = H \otimes I_n$ with H the incidence matrix of $\mathbb{G}$. Then one has

$$\text{image } \bar{H} \cap \ker \bar{P} = 0 \tag{18}$$

and

$$\text{image } \bar{H}' \cap \Phi(\bar{H}' \bar{P} \bar{H}) = 0 \tag{19}$$

For any $q \in \text{image } \bar{H}' \cap \Phi(\bar{H}' \bar{P} \bar{H})$, there exists a vector p such that

$$q = \bar{H}' p \tag{20}$$

and a vector $\phi \in F[\text{sgn }](q)$ such that $\bar{H}' \bar{P} \bar{H} \phi = 0$. Note that $\bar{P}$ is a projection matrix; then

$$\bar{P} \bar{H} \phi = 0. \tag{21}$$

From (18) one then has

$$\bar{H} \phi = 0. \tag{22}$$

From $\phi \in F[\text{sgn }](q)$ and (5), one has

$$||q||_1 = \phi' q$$

which together with (20) and (22) implies $||q||_1 = 0$. Then one has $q = 0$ and (19) is true. ■

Now consider the system

$$\dot{x} \in -MF[\text{sgn }](x) \tag{23}$$

with any positive semi-definite matrix $M \in \mathbb{R}^{r \times r}$. The existence of a Filippov solution to (23) can be guaranteed by Lemma 1. The existence interval is $t \in [0, \infty)$ because of the global bound on the right-hand side to (23). Let $x(t)$ denote such a Filippov solution for any given $x(0)$. Note that the function $\|x\|_1$ is locally Lipschitz and regular(The word was introduced early with reference). Then by Lemma 3, the time derivative of $\|x(t)\|_1$ exists for almost all $t \in [0, \infty)$ and is in the set of generalized Lie derivatives. In other words, there exists a set $\mathcal{I} = [0, \infty) \backslash \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that

$$\frac{d\|x(t)\|_1}{dt} \text{ exists for all } t \in \mathcal{I} \tag{24}$$

Proposition 1: Let $x(t)$ denote a Filippov solution to (23) for any given $x(0) \in \mathbb{R}^r$. Then

- $$\frac{d\|x(t)\|_1}{dt} \leq 0, \quad t \in \mathcal{I}; \tag{25}$$
- there exists a finite time T such that
$$x(T) \in \Phi(M); \tag{26}$$
- if it is further true that
$$\frac{d\|x(t)\|_1}{dt} = 0, \quad t \in [T, \infty) \setminus \mathcal{T}, \tag{27}$$
one has
$$x(t) = x(T) \quad t \in [T, \infty) \tag{28}$$

Remark 1: Note that the right-hand side of (23) is a projection of the gradient flow of the potential function $\|x(t)\|_1$.

It is also standard that the gradient law of a real analytic function with a lower bound converges to a single point which is both a local minimum and a critical point of the potential function [40]. However, if the real analytic property does not hold, the convergence result may fail. Indeed, the function $\|x(t)\|_1$ here is obviously not real analytic, and one cannot immediately assert that (23) will drive $\|x(t)\|_1$ to its minimum, not to mention the finite time result in Proposition 1. Thus Proposition 1 is nontrivial and will serve as the foundation for devising finite-time distributed linear equation solvers in this paper.

IV. ALGORITHMS AND MAIN RESULTS

In this section, we first present a centralized update for minimum l_1-norm solutions to $Ax = b$ and then devise a distributed one for all agents in the network to achieve the same minimum l_1-norm solution in finite time.

A. Finite-Time Centralized Update for Minimum l_1-norm Solution

In this subsection, we will propose a centralized update for achieving a minimum l_1-norm solution to $Ax = b$. By noting that $\|x\|_1$ is convex, we conceive of using a negative gradient flow of $\|x\|_1$ subject to x remaining on the manifold $Ax = b$ in order to achieve $\bar{x}^* = \arg\min_{Ax=b} \|x\|_1$. This leads us to the following update:

$$\dot{y} \in -PF[\text{sgn}\,](y), \quad Ay(0) = b. \tag{29}$$

where P denotes the projection matrix onto the kernel of A. Again by Lemma 1 one has there exists a Filippov solution to system (29) for $t \in [0, \infty)$, which we denote by $y(t)$. By Lemma 3, there exists a set $\mathcal{I} = [0, \infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of measure 0 such that $\frac{d\|y(t)\|_1}{dt}$ exists for all $t \in \mathcal{I}$. Moreover, we have the following main theorem:

Theorem 1: With A of full row rank, the Filippov solution $y(t)$ to (29) converges in finite time to a constant, which is a minimum l_1-norm solution to $Ax = b$.

B. Finite-Time Distributed Update for Minimum l_1-norm Solutions

In this subsection we will develop a distributed update for a multi-agent network to achieve a minimum l_1-norm solution to $Ax = b$ in finite time. Of course, A is assumed to have full row rank, but not necessarily be square. Motivated by study a combination of the finite-time consensus flow [34] and the finite-time centralized update for minimum l_1-norm solutions in (29), we propose the following update for agent i, $i = 1, 2, ..., m$:

$$\dot{y}_i = -k(t)P_i\phi_i - P_i \sum_{j\in\mathcal{N}_i} \phi_{ij} \tag{30}$$

where $\phi_i \in F[\text{sgn}\,](y_i)$, $\phi_{ij} \in F[\text{sgn}\,](y_i - y_j)$. Since the graph is undirected we can make the following assumption

Assumption 1: Each pair of neighbor agents i and j takes the choice of ϕ_{ij} and ϕ_{ji} when $y_i = y_j$ such that

$$\phi_{ij} = -\phi_{ji} \tag{31}$$

and take the initialization as

$$A_i y_i(0) = b_i. \tag{32}$$

We assume that $k(t) \in \mathbb{R}$ is measurable and locally bounded almost everywhere for $t \in [0, \infty)$, and

$$\lim_{t\to\infty} k(t) = \delta \tag{33}$$

$$\int_0^\infty k(t)dt = \infty \tag{34}$$

where δ is a sufficiently small nonnegative number depending on the connection of the network and A, note that 0 is always a feasible choice of δ. One example of a choice of $k(t)$ is $k(t) = \frac{\bar{\delta}}{t+1} + \delta$. One simple case is choosing $\bar{\delta} = 1$, $\delta = 0$, and resulting in $k(t) = \frac{1}{t+1}$, obtained by taking δ to be zero. This choice obviates the need to decide how small one has to be to meet a "sufficiently small" condition, but may result in rather slow convergence. Now from $A_i y_i(0) = b_i$ and the fact that P_i is the projection to the kernel of A_i (which ensures $\dot{y}_i \in \ker A_i$), one has

$$A_i y_i(t) = b_i, \quad t \in [0, \infty) \tag{35}$$

Let $y = \text{col}\,\{y_1, y_2, y_3, ..., y_m\}$, $\bar{P} = \text{diag}\,\{P_1, P_2, ..., P_m\}$, and $\bar{H} = H \otimes I_n$ with H the incidence matrix of $\mathbb{G}$. From the updates in (30) and Assumption 1, one has

$$\dot{y} \in -k(t)\bar{P}F[\text{sgn}\,](y) - \bar{P}\bar{H}F[\text{sgn}\,](\bar{H}'y) \tag{36}$$

Note that $\text{sgn}\,(y)$, $k(t)$ and $\bar{P}\bar{H}F[\text{sgn}\,](\bar{H}'y)$ are measurable and locally bounded almost everywhere. Then by Lemma 2 there exists a Filippov solution to system (36) for any given $y(0)$ satisfying (32), which we denote by $y(t) = \text{col}\,\{y_1(t), y_2(t), y_3(t), ..., y_m(t)\}$. By Lemma 3, there exists a set $\mathcal{I} = [0, \infty)\setminus\mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that $\frac{d\|y(t)\|_1}{dt}$ exists for all $t \in \mathcal{I}$. Then one has the following theorem:

Theorem 2: Under Assumption 1 and the update (30) and with A of full row rank, all $y_i(t)$, $i = 1, 2, ..., m$ converge in finite time to the same value which is a minimum l_1-norm solution to $Ax = b$.

V. SIMULATION RESULT

In this section, we will report several simulations of the proposed algorithms for solving an underdetermined linear equation $Ax = b$ in a four-agent undirected and connected network as in Figure 1.

Here, A and b are partitioned as $A = \begin{bmatrix} A_1' & A_2' & A_3' & A_4' \end{bmatrix}'$ and $b = \begin{bmatrix} b_1' & b_2' & b_3' & b_4' \end{bmatrix}'$, respectively. Each agent i knows

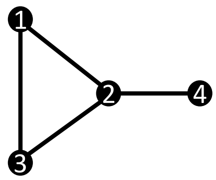

Fig. 1. A four agent network

A_i and b_i with

$$A'_1 = \begin{bmatrix} 0.63 & 0.04 \\ 0.58 & 0.60 \\ 0.65 & 0.50 \\ 0.33 & 0.81 \\ 0.68 & 0.01 \\ 0.22 & 0.51 \\ 0.49 & 0.23 \\ 0.21 & 0.29 \\ 0.62 & 0.25 \\ 0.57 & 0.21 \\ 0.71 & 0.66 \\ 0.28 & 0.90 \end{bmatrix}, \quad A'_2 = \begin{bmatrix} 0.80 & 0.99 \\ 0.25 & 0.65 \\ 0.53 & 0.38 \\ 0.79 & 0.12 \\ 0.13 & 0.76 \\ 0.79 & 0.52 \\ 0.34 & 0.55 \\ 0.45 & 0.24 \\ 0.10 & 0.55 \\ 0.94 & 0.51 \\ 0.78 & 0.58 \\ 0.70 & 0.85 \end{bmatrix} \tag{37}$$

$$A'_3 = \begin{bmatrix} 0.44 & 0.34 \\ 0.06 & 0.94 \\ 0.77 & 0.28 \\ 0.16 & 0.41 \\ 0.84 & 0.75 \\ 0.62 & 0.56 \\ 0.74 & 0.41 \\ 0.26 & 0.89 \\ 0.44 & 0.69 \\ 0.28 & 0.23 \\ 0.50 & 0.88 \\ 0.38 & 0.63 \end{bmatrix}, \quad A'_4 = \begin{bmatrix} 0.05 & 0.23 \\ 0.09 & 0.33 \\ 0.65 & 0.92 \\ 0.69 & 0.66 \\ 0.94 & 0.92 \\ 0.73 & 0.06 \\ 0.51 & 0.13 \\ 0.59 & 0.94 \\ 0.76 & 0.40 \\ 0.95 & 0.69 \\ 0.39 & 0.24 \\ 0.03 & 0.92 \end{bmatrix}$$

$$\begin{aligned} b'_1 &= \begin{bmatrix} 0.47 & 0.52 \end{bmatrix}, \quad b'_2 = \begin{bmatrix} 0.77 & 0.34 \end{bmatrix} \\ b'_3 &= \begin{bmatrix} 0.63 & 0.33 \end{bmatrix}, \quad b'_4 = \begin{bmatrix} 0.31 & 0.65 \end{bmatrix} \end{aligned} \tag{38}$$

Example 1: We employ the centralized update (29) with state vector $y(t)$ to achieve $\bar{x}^*$ which denotes a minimum l_1-norm solution to $Ax = b$. As shown in Figure 2, $\|y(t) - \bar{x}^*\|_1$ reaches 0 in finite time and maintains to be 0 afterwards. This indicates that a minimum l_1-norm solution $\bar{x}^*$ is achieved in finite time corresponding to Theorem 1. It is worth noting that one could observe multiple phases of convergence in Figure 2. This is because $F[\mathrm{sgn}\,](y(t))$ in the update (29) takes different values piece-wisely, and results in different convergence rates.

Example 2: Finally, we utilize the distributed update (30) to achieve a minimum l_1 solution to $Ax = b$ denoted by $\bar{x}^*$ in finite time. Here $k(t)$ is chosen to take the form $\frac{\bar{\delta}}{1+t} + \delta$ with $\bar{\delta}$ and δ constants. We still let $y = \begin{bmatrix} y'_1 & y'_2 & y'_3 & y'_4 \end{bmatrix}'$ where $y_i(t)$ denote the state of agent i that is the estimate of agent i to $\bar{x}^*$. Then $\|y(t) - \mathbf{1}_m \otimes \bar{x}^*\|_1$ measures the difference between all agents' estimations and $\bar{x}^*$. As shown in Figure 3 and Figure 4, all $y_i(t)$ reach the same minimum l_1-norm solution in finite time regardless of different choices of $\bar{\delta}$ and δ. Moreover, by fixing $\bar{\delta}$ and increasing the value of δ in $k(t)$, one achieves a significantly faster convergence as shown in Figure 3. Similarly, increasing $\bar{\delta}$ with a fixed δ also

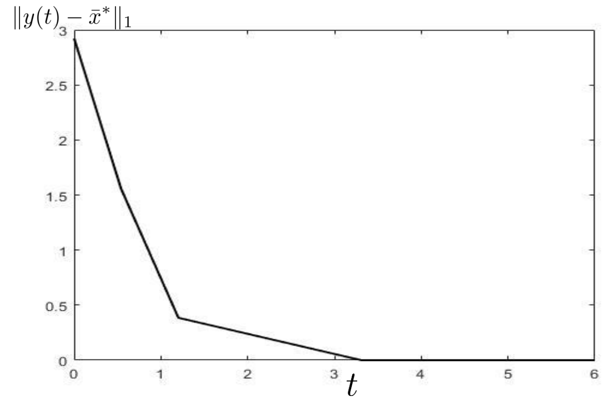

Fig. 2. Centralized solver for achieving a minimum l_1 norm solution under the update (29)

leads to a faster convergence, although not that dramatically, as shown in Figure 4.

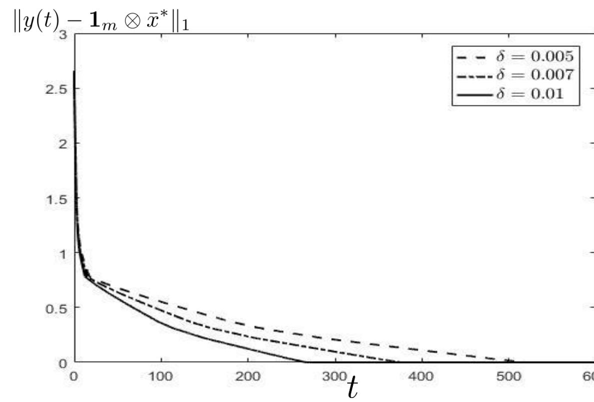

Fig. 3. Distributed solver for achieving minimum l_1 norm solution under the update (30), where $k(t) = \frac{\bar{\delta}}{t+1} + \delta$ with fixed $\bar{\delta} = 0.1$ and different values of δ.

We also note from Figure 3 and Figure 4 that the convergence time required in this distributed way for minimum l_1-norm solutions is dramatically longer, roughly speaking, $\frac{1}{\delta}$ times longer, than that in the centralized case in Figure 2. The major reason for this is that the centralized update appearing in the distributed update (30) is scaled by $k(t)$, which is smaller than 1.

VI. Conclusion

We have developed continuous-time distributed algorithms for achieving minimum l_1-norm solutions to linear equations $Ax = b$ in finite time. The algorithms result from combination of the projection-consensus flow proposed in [13] and the finite-time gradient flow for consensus devised in [34], and work for fixed undirected multi-agent networks. Future work includes the generalization of the proposed update to networks that are directed and time-varying.

References

[1] S. Mou, J. Liu, and A. S. Morse. A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions on Automatic Control*, 60(11):2863–2878, 2015.

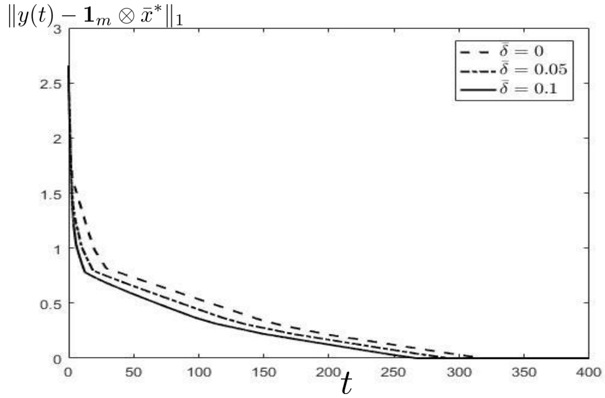

Fig. 4. Distributed solver for achieving minimum l_1 norm solution under update (30) where $k(t) = \frac{\bar{\delta}}{t+1} + \delta$ with fixed $\delta = 0.01$ and different values of $\bar{\delta}$.

[2] B. D. O. Anderson, S. Mou, U. R. Helmke, and A. S. Morse. Decentralized gradient algorithm for solution of a linear equation. *Numerical Algebra, Control and Optimization*, 6(3):319–328, 2016.

[3] J. Lu and C. Y. Tang. A distributed algorithm for solving positive definite linear equations over networks with membership dynamics. *IEEE Transactions on Control of Network Systems*, 2016.

[4] J. Wang and N. Elia. Distributed solution of linear equations over unreliable networks. *Proceedings of American Control Conference*, pages 6471–6476, 2016.

[5] S. Mou and A. S. Morse. A fixed-neighbor, distributed algorithm for solving a linear algebraic equation. *European Control Conference*, pages 2269–2273, 2013.

[6] X. Wang, S. Mou, and D. Sun. Improvement of a distributed algorithm for solving linear equations. *IEEE Transactions on Industrial Electronics*, 64(4):3113–3117, 2017.

[7] P. Wang, W. Ren, and Z. Duan. Distributed minimum weighted norm solution to linear equations associated with weighted inner product. *Proceedings of the 55th Conference on Decision and Control*, pages 5220–5225, 2016.

[8] L. Wang, D. Fullmer, and A. S. Morse. A distributed algorithm with an arbitary initialization for solving a linear algebraic equation. *Proceedings of American Control Conference*, pages 1078–1081, 2016.

[9] S. Mou, Z. Lin, L. Wang, D. Fullmer, and A. S. Morse. A distributed algorithm for efficiently solving linear equations and its applications (special issue jcw). *Systems & Control Letters*, 91:21–27, 2016.

[10] J. Wang and N. Elia. Distributed least square with intermittent communications. In *2012 American Control Conference (ACC)*, pages 6479–6484, June 2012.

[11] J. Wang and N. Elia. A control perspective for centralized and distributed convex optimization. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 3800–3805, Dec 2011.

[12] B. Gharesifard and J. Cortés. Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Transactions on Automatic Control*, 59(3):781–786, March 2014.

[13] G. Shi and B.D. O.Anderson. Distributed network flows solving linear algebraic equations. *Proceedings of American Control Conference*, pages 2864–2869, 2016.

[14] Y. Liu, C. Lageman, B. D. O. Anderson, and G. Shi. Exponential least squares solvers for linear equations over networks. *IFAC World Congress, Toulouse*, pages 2598–2603, 2017.

[15] Y. Liu, Y. Lou, B. D. O. Anderson, and G. Shi. Network flows as least squares solvers for linear equations. *56th IEEE Conference on Decision and Control, Melbourne*, 2017. accepted.

[16] P. M. Shearer. Improving local earthquake locations using the l1 norm and waveform cross correlation: Application to the whittier narrows, california, aftershock sequence. *Journal of Geophysical Research: Solid Earth*, 102, 1997.

[17] Y. Dodge. *Statistical data analysis based on the L1-norm and related methods*. Birkhäuser, 2012.

[18] R. Beucker and H. A. Schlitt. On minimal lp-norm solutions of the biomagnetic inverse problem. Technical report, Zentralinstitut für Angewandte Mathematik, 1996.

[19] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk. Distributed compressive sensing. 2009. arXiv:0901.3403.

[20] Y. C. Eldar and G. Kutyniok. Compressed sensing: Theory and applications. *Cambridge University Press*, 2012.

[21] E. J. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, (12):4203–4215, 2005.

[22] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, (2):489–509, 2006.

[23] A. Y. Yang, A. Genesh, Z. H. Zhou, S. S. Sastry, and Y. Ma. A review of fast l (1)-minimization algorithms for robust face recognition. Technical report, CALIFORNIA UNIV BERKELEY DEPT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, 2010.

[24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[25] K. Frisch. The logarithmic potential method of convex programming. *Memorandum, University Institute of Economics, Oslo*, 5(6), 1955.

[26] M. Kojima, N. Megiddo, and S. Mizuno. Theoretical convergence of large-step primaldual interior point algorithms for linear programming. *Mathematical Programming*, 59:1–21, 1993.

[27] M. Figueiredo, R.D. Nowak, and J.S. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal processing*, 1:586–597, 2007.

[28] M. R. Osborne, B. Presnell, and B. A. Turlach. A new approach to variable selection in least squares problems. *IMA journal of numerical analysis*, 2000.

[29] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on pure and applied mathematics*, pages 1413–1457, 2004.

[30] S. Becker, J. Bobin, and E. J. Candès. Nesta: A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, pages 1–39, 2011.

[31] M. Cao, A. S. Morse, and B.D.O. Anderson. Agreeing asynchronously. *IEEE Transactions on Automatic Control*, 53(8):1826–1838, 2008.

[32] M. Cao, A. S. Morse, and B. D. O. Anderson. Reaching a consensus in a dynamically changing environment: A graphical approach. *SIAM Journal on Control and Optimization*, 47(2):575–600, 2008.

[33] J. Qin and C. Yu. Exponential consensus of general linear multi-agent systems under directed dynamic topology. *Automatica*, 50(9):2327–2333, 2014.

[34] J. Cortés. Finite-time convergent gradient flows with applications to network consensus. *Automatica*, 42(11):1993–2000, 2006.

[35] J. Cortes. Discontinuous dynamical system: A tutorial on solutions, nonsmooth analysis and stability. *IEEE Control System Magazine*, pages 36–73, 2008.

[36] A. F. Filippov. *Differential equations with discontinuous righthand sides: control systems*, volume 18. Springer Science & Business Media.

[37] F. H. Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990.

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

[39] F. Chung. *Spectral graph theory*. American Mathematical Soc, 1997.

[40] P. A. Absil and K. Kurdyka. On the stable equilibrium points of gradient systems. *Systems & control letters*, 55(7):573–577, 2006.