Finite-Time Distributed Linear Equation Solver for Solutions with Minimum l_1-Norm

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

***Abstract*—This paper proposes a continuous-time distributed algorithm for multi-agent networks to achieve a solution with the minimum l_1-norm to under-determined linear equations. The proposed algorithm comes from a combination of the Filippov set-valued map with the projection-consensus flow. Given the underlying network is undirected and fixed, it is shown that the proposed algorithm drives all agents' individual states to converge in finite time to a common value, which is the minimum l_1-norm solution.**

***Index Terms*—Distributed Algorithms; Minimum l_1-Norm Solutions; Finite-Time Convergence.**

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 an overall linear equation and controls a state vector that can be looked at as an estimate to the solution of the overall equation [1]–[5]. One key idea of these algorithms is the so-called agreement principle [1], in which, each agent updates its estimate subject to its own equation and tries to reach a consensus with its nearby neighbors' estimates; the continuous-time version is also referred as to the projection-consensus flow [6], [7]. These distributed algorithms enable all agents in strongly-connected networks to asymptotically reach the same solution to the overall equation as long as solutions exist. When it comes to under-determined linear equations, l_1-*min solutions* (i.e. solutions with the minimum l_1 norm) are often of particular interest especially in image and signal processing [8]–[10], but unfortunately these recently developed discrete-time distributed algorithms [1]–[5] or continuous-time ones [6], [7] are not applicable. While classical methods based on the idea of LASSO, such as the Alternating Directions Method [11], Iterative Shrinkage-thresholding Algorithm [12], Proximal Gradient Method [13], and so on, could achieve l_1-min solutions in a centralized way, there has been little progress in the development of distributed algorithms for l_1-min solutions. Recognition of this gap has motivated the objective of this paper to achieve a distributed linear equation solver with the capability of achieving l_1-min solutions. More specifically, we will perform further improvement to the projection-consensus flow [6], [7] through the concept of Filippov set-valued maps, which has played a significant role in achieving finite-time consensus in [14], with the goal of developing a continuous-time distributed algorithm for l_1-min solutions.

This work was supported by a funding from Northrop Grumman Cooperation (NGC-REALM Award). 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, DP-160104500 and DP190100887. Corresponding Author: Shaoshuai Mou.

The organization of the paper is as follows: We formulate the problem of interest in Section II and present a distributed update in Section III. Finite convergence of the proposed distributed update is established in Section IV with conclusion in Section V. Proofs of lemmas and propositions are provided in Section VI.

Notation: For a positive integer r, 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; let $\text{col}\,\{A_1, A_2, \cdots, A_r\}$ 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 $\text{diag}\,\{A_1, A_2, \cdots, A_r\}$ denote a block diagonal matrix with A_i the ith diagonal block entry, $i = 1, 2, \cdots, r$. Let $\ker M$ and $\text{image}\, M$ 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 continuously access 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 these neighbor relations can be described by an m-node-$\bar{m}$-edge undirected graph $\mathbb{G}$ (with $\bar{m}$ the total number of edges), such that there is an undirected edge connecting i and j if and only if i and j are neighbors. In this paper we suppose $\mathbb{G}$ 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, we assume A to have full row-rank, but not necessarily be square as in [8]–[10]. Further assume $Ax = b$ has at least one solution. Let x^* denote a l_1-min solution to $Ax = b$, (i.e. solution with the minimum l_1 norm), that is,

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

The problem of interest in this paper is to develop a distributed algorithm 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 l_1-min solution x^*.

III. The Update

In this section we will develop a continuous-time distributed update for l_1-min solutions with finite-time convergence, for which one needs the concept of Filippov set-valued maps. 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. Note that the Filippov set-valued map has played a significant role in achieving finite-time consensus in [14]. By using the Filippov set-valued map of the sign function in a distributed linear equation solver based on the projection-consensus flow in [7], one expects a distributed linear equation solver with finite-time convergence. Including one additional term for minimizing the solution's l_1-norm leads to the following distributed update for agent i, $i=1,2,...,m$:

$$\dot{y}_i\in-k(t)P_iF[\text{sgn}\,](y_i)-P_i\sum_{j\in\mathcal{N}_i}F[\text{sgn}\,](y_i-y_j)\tag{3}$$

with

$$A_iy_i(0)=b_i.\tag{4}$$

Here, $k(t)\in\mathbb{R}$ is measurable and locally bounded almost everywhere for $t\in[0,\infty)$, and further constraints will be imposed subsequently; P_i denotes the projection matrix to the kernel of A_i; $\text{sgn}\,(x):\mathbb{R}^n\to\mathbb{R}^n$ is the sign function. Thus the kth entry, $k=1,2,...,n$, of $\text{sgn}\,(x)$ and its Filippov set-valued map $F[\text{sgn}\,](x)$ are given as follows.

	$\left(\text{sgn}\,(x)\right)_k$	$\left(F[\text{sgn}\,](x)\right)_k$
$(x)_k>0$	1	1
$(x)_k=0$	0	$[-1,1]$
$(x)_k<0$	-1	-1

IV. Analysis and Main Result

In this section, we will analyze the system consisting of (3) and present the main result of the paper.

Before proceeding on, we will record a compact form for the distributed updates (3). Label all the nodes as $1,2,...,m$ and all the edges as $1,2,...,\bar{m}$ for the m-node-$\bar{m}$-edge graph $\mathbb{G}$. 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 (for undirected edge, any of its vertices can be the head and the other one be the tail)

$$h_{ik}=\begin{cases}1, & i\text{ is the head of the }k\text{th edge;}\\ -1, & i\text{ is the tail of the }k\text{th edge;}\\ 0, & \text{otherwise.}\end{cases}\tag{5}$$

Since $\mathbb{G}$ is connected, then $\ker H'$ is the span of $\mathbf{1}_m$ [15]. It follows that

$$\bar{H}'y=0\quad\text{if and only if}\quad y_1=y_2=\cdots=y_m\tag{6}$$

where

$$\bar{H}=H\otimes I_n\tag{7}$$

and $y=\text{col}\,\{y_1,y_2,y_3,...,y_m\}$. Since the underlying network $\mathbb{G}$ is undirected, one has, for a neighbor pair i and j, that $F[\text{sgn}\,](y_i-y_j)$ appears in the update of $\dot{y}_i$ if and only if $F[\text{sgn}\,](y_j-y_i)$ appears in j's update $\dot{y}_j$. From this, the fact that

$$F[\text{sgn}\,](y_i-y_j)=-F[\text{sgn}\,](y_j-y_i)\tag{8}$$

and the definition of incident matrix, one can write updates in (3) in the following compact form:

$$\dot{y}\in-k(t)\bar{P}F[\text{sgn}\,](y)-\bar{P}\bar{H}F[\text{sgn}\,](\bar{H}'y)\tag{9}$$

with

$$\bar{P}=\text{diag}\,\{P_1,P_2,...,P_m\}.\tag{10}$$

A. Existence of a Filippov Solution

We first show the existence of a solution to distributed updates in (3) or equivalently its compact form (9), for which one needs the concept of Filippov solutions. By *a Filippov solution* for $\dot{x}\in F[f](x)$ is meant a solution $x(t)$ such that $\dot{x}\in F[f](x)$ for almost all t, $x(t)$ is absolutely continuous. Here, since we consider the Filippov map of *sgn* function, and 0 is always in the Filippov set valued map, we claim that it is also a Caratheodory solution, thus it can be written in the form of an indefinite integral [16]. The existence of a Filipov solution can be guaranteed by the following lemma

Lemma 1: (Theorem 8 in page 85 of [16]) 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 $F[f](x)$ be given as in (2) Then for any point $(t_0,x_0)\subset G$, there exists a Filippov solution of $\dot{x}\in F[f](t,x)$ with $x(t_0)=x_0$.

Since $\text{sgn}\,(y)$, $k(t)$ and $\bar{P}\bar{H}F[\text{sgn}\,](\bar{H}'y)$ in (9) are measurable and locally bounded almost everywhere, by Lemma 1 there exists a Filippov solution to system (9) for any given $y(0)$ satisfying (4), which we denote by $y(t)$. Further more, since the right side of (9) has a global bound, the existence interval of $y(t)$ would be $t\in[0,\infty)$.

B. Key Lemmas and Proposition

In order to establish finite-time convergence of the Filipov solution $y(t)$, we introduce some key lemmas.

First, we investigate some sets regarding to Fillipov set-valued maps. For any positive semi-definite matrix $M\in\mathbb{R}^{r\times r}$, one defines

$$\Phi(M)=\{q\in\mathbb{R}^r|\,\exists\phi\in F[\text{sgn}\,](q),M\phi(q)=0\}\tag{11}$$

and its complement

$$\Phi_c(M)=\{q\in\mathbb{R}^r|\,\forall\phi\in F[\text{sgn}\,](q),M\phi(q)\neq0\}\tag{12}$$

and

$$\Lambda(M)=\{\phi|\phi\in F[\text{sgn}\,](q),q\in\Phi_c(M)\}\tag{13}$$

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 $\phi' M\phi$ has a nonzero minimum on $\Lambda(M)$ with respect to ϕ, for which we denote

$$\lambda(M) = \min_{\phi\in\Lambda(M)} \phi' M\phi. \tag{14}$$

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

Lemma 2: For any nonnegative-definite matrix M, let $\Phi_c(M)$, $\Lambda(M)$ and $\lambda(M)$ be defined as above. Suppose that $\Phi_c(M)$ is nonempty. Then $\lambda(M)$ is a positive constant.

Second, in order to study systems involving functions which are not differentiable everywhere such as $||x||_1$, one needs more general concepts of derivatives. For a locally Lipschitz function $w : \mathbb{R}^r \to \mathbb{R}$, the *generalized gradient* [17] of w is defined as

$$\partial w(x) \triangleq \text{co}\{\lim_{i\to\infty} \nabla w(x_i) : x_i \to x, x_i \notin \mathcal{S} \cup \Omega_w\} \tag{15}$$

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 has

$$F[\text{sgn }](x) = \partial\|x\|_1 \tag{16}$$

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

$$\begin{aligned}\tilde{\mathcal{L}}_{\mathcal{F}} w(x) = \{&q \in \mathbb{R} : \text{there exists } \alpha \in \mathcal{F}(x)\\ &\text{such that } \forall\beta \in \partial w(x), q = \beta'\alpha\}\end{aligned} \tag{17}$$

Lemma 3: (Proposition 10 in [18] and Proof of Lemma 1 in [19]) Let $x : [0,\infty] \to \mathbb{R}^r$ be a solution for $\dot{x} \in \mathcal{F}(x(t))$, where $\mathcal{F}$ is a set-valued map. Let $w(x) : \mathbb{R}^r \to \mathbb{R}$ be locally Lipschitz and regular[1]. Then there exists a set $\mathcal{I} = [0,\infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that such that $w(x(t))$ is differentiable for $t \in \mathcal{I}$; The derivative of $w(x(t))$ satisfies $\frac{dw(x(t))}{dt} \in \tilde{\mathcal{L}}_{\mathcal{F}} w(x(t))$ for almost all $t \in \mathcal{I}$; and

$$\frac{dw(x(t)))}{dt} = \beta'\dot{x}(t) \tag{18}$$

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

Third, recall that we restrict our attention to the case when A has full row rank, for which the following lemma holds.

Lemma 4: Suppose A has full row rank and $\mathbb{G}$ is connected. Then

$$\text{image } \bar{H} \cap \ker \bar{P} = 0 \tag{19}$$

and

$$\text{image } \bar{H}' \cap \Phi(\bar{H}'\bar{P}\bar{H}) = 0 \tag{20}$$

The proof of Lemma 4 will be given in the Appendix.

Based on the above lemmas, one can prove in the Appendix the following proposition, which will serve as the foundation to establish finite-time convergence of distributed updates (3).

Proposition 1: Given a system

$$\dot{x} \in -MF[\text{sgn }](x) \tag{21}$$

with any positive semi-definite matrix $M \in \mathbb{R}^{r\times r}$ and any given $x(0) \in \mathbb{R}^r$. Then there exists a Filippov solution to (21) denoted by $x(t)$ for $t \in [0,\infty)$. Moreover,

(a.) there exists a set $\mathcal{I} = [0,\infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that

$$\frac{d\|x(t)\|_1}{dt} \text{ exists for all } t \in \mathcal{I} \tag{22}$$

and

$$\frac{d\|x(t)\|_1}{dt} \leq 0, \quad t \in \mathcal{I}; \tag{23}$$

(b.) there exists a finite time T such that

$$x(T) \in \Phi(M); \tag{24}$$

(c.) if

$$\frac{d\|x(t)\|_1}{dt} = 0, \quad t \in [T,\infty) \setminus \mathcal{T}, \tag{25}$$

one has $x(t) = x(T),\ \forall t \in [T,\infty)$.

(d.) if $Ax(0) = b$ and M is the projection matrix to $\ker A$, then $x(t)$ converges in finite time to be a constant x^*, which is the l_1-min solution to $Ax = b$.

Remark 1: Convergence of the system (21) with M positive-definite has been studied in [21], but the approach is not applicable here, because we require M to be a projection matrix, which is necessarily singular. Proposition 1 considers the case when M in (21) is positive semi-definite. Then if M is a projection matrix to $\ker A$, the right-hand side of (21) is the gradient flow of a potential function $\|x(t)\|_1$ subject to $Ax = b$. It is 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 [22]. 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[2], and one cannot immediately assert that (21) will drive $\|x(t)\|_1$ to its minimum, not to mention the finite-time convergence. When the number of agents $m = 1$, the proposed distributed update (3) naturally becomes a centralized update $\dot{y} \in PF[\text{sgn }](y)$ with P the projection matrix to $\ker A$ and $k(t) = 1$, convergence of which could be directly established by Proposition 1.

C. Main Result

Theorem 1: Given $\mathbb{G}$ is fixed, undirected and connected. Suppose A has full row rank; $Ax = b$ has solutions; and $k(t)$ is such that

$$\lim_{t\to\infty} k(t) = \delta, \quad \int_0^\infty k(t)dt = \infty, \tag{26}$$

with δ a sufficiently small nonnegative scalar. Then under the update (3) with initializations (4), one has there exists a finite time T such that

$$y_1(t) = y_2(t) = \cdots = y_m(t) = x^*, \quad t \in [T,\infty) \tag{27}$$

where x^* is a constant l_1-min solution to $Ax = b$.

Proof of Theorem 1: To prove Theorem 1, it is sufficient to prove the followings:

[1] A function is termed regular if and only if it is analytic and single-valued throughout a region $\mathcal{R}$ [20].

[2] A real analytic function is a real function, the values of which can be obtained locally from a convergent power series.

- **Claim 1.** There exists a finite time T_2 and a vector $\bar{y}(t)$ such that
$$y_1(t) = y_2(t) = \cdots = y_m(t) = \bar{y}(t), \;\; t \in [T_2, \infty). \quad (28)$$
- **Claim 2.** There exists a finite time T such that
$$\bar{y}(t) = x^*, \quad t \in [T, \infty). \quad (29)$$

First, we prove Claim 1, which by (6) is equivalent to showing
$$z(t) = 0, \quad t \in [T_2, \infty) \quad (30)$$
with $z(t) = \bar{H}'y(t)$. Multiplying both sides of (9) by $\bar{H}'$, one has
$$\dot{z} \in -k(t)\bar{H}'\bar{P}F[\text{sgn}\,](y) - \bar{H}'\bar{P}\bar{H}F[\text{sgn}\,](z) \quad (31)$$
By Lemma 3, there exists a set $\mathcal{I} = [0, \infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that
$$\frac{d\|z(t)\|_1}{dt} = \beta'\dot{z}, \quad t \in \mathcal{I}$$
where β can be any vector in $\partial\|z\|_1$. Note that $\partial\|z\|_1 = F[\text{sgn}\,](z)$. Then
$$\frac{d\|z(t)\|_1}{dt} = -k(t)\gamma'\bar{H}'\bar{P}\eta - \gamma'\bar{H}'\bar{P}\bar{H}\gamma, \quad t \in \mathcal{I} \quad (32)$$
where $\eta \in F[\text{sgn}\,](y)$, $\gamma \in F[\text{sgn}\,](z)$ and β is chosen to be equal to γ. Since $z(t) \in \text{image}\, \bar{H}'$, then by Lemma 4, if also $z(t) \in \Phi(\bar{H}'\bar{P}\bar{H})$, one will have $z(t) = 0$. Thus
$$z(t) \in \Phi_c(\bar{H}'\bar{P}\bar{H}) \text{ as long as } \|z(t)\|_1 \neq 0. \quad (33)$$
Now by the definition of $\lambda(\bar{H}'\bar{P}\bar{H})$ and Lemma 2, one has
$$\gamma'\bar{H}'\bar{P}\bar{H}\gamma \geq \lambda \text{ as long as } \|z(t)\|_1 \neq 0 \quad (34)$$
where λ is a positive constant. Let κ denote an upper bound on $|\gamma'\bar{H}'\bar{P}\eta|$. Choose δ sufficiently small such that
$$\delta < \frac{\lambda}{\kappa}. \quad (35)$$
Then there must exist a finite time T_1 such that
$$|-k(t)\gamma\bar{H}'\bar{P}\eta| \leq \rho, \quad t \in [T_1, \infty). \quad (36)$$
where we have $\rho = \kappa\delta < \lambda$. From (32), (34) and (36), one has
$$\frac{d\|z(t)\|_1}{dt} \leq -(\lambda - \rho) < 0 \quad (37)$$
as long as $\|z(t)\|_1 \neq 0$, $t \in [T_1, \infty) \setminus \mathcal{T}$. Thus there must exist a finite time $T_2 \geq T_1$ such that
$$z(T_2) = 0. \quad (38)$$
Next, to complete the proof of Claim 1, we prove $z(t) = 0$, $t \in [T_2, \infty)$ in (30) by contradiction. Suppose (30) is not true. Then there exists a time $\bar{T}_2 > T_2$ such that $z(\bar{T}_2) \neq 0$. Then $\|z(\bar{T}_2)\|_1 > 0$. Since $\|z(t)\|_1$ is continuous, there exists a time T_2^* such that $\|z(T_2^*)\|_1$ takes its maximum value for $t \in [T_2, \bar{T}_2]$. Again, since $\|z(t)\|_1$ is continuous, there exists a sufficiently small but positive ϵ such that $\|z(t)\|_1 > 0$ is differentiable for $t \in [T_2^* - \epsilon, T_2^*]$. Because $\|z(t)\|_1$ is differentiable almost everywhere, we know that
$$\|z(T_2^*)\|_1 = \int_{T_2^*-\epsilon}^{T_2^*} \frac{d\|z(t)\|_1}{dt}dt + \|z(T_2^* - \epsilon)\|_1 \quad (39)$$
Because $\|z(t)\|_1 > 0$ in $[T_2^* - \epsilon, T_2^*]$, by (37) and (39) we have
$$\|z(T_2^*)\|_1 \leq -\epsilon(\rho - \rho_2) + \|z(T_2^* - \epsilon)\|_1 < \|z(T_2^* - \epsilon)\|_1 \quad (40)$$
This contradicts the fact that $\|z(T_2^*)\|_1$ is the maximum value on $[T_2, \bar{T}_2]$. Thus (30) is true. So Claim 1 is true.

Second, we prove Claim 2. From (28) and $A_iy_i(t) = b_i$ for $i = 1, 2, ..., m$, one has $A\bar{y}(T_2) = b$. Let P denote the projection matrix to the kernel of A. Then $PP_i = P$ for $i = 1, 2, ..., m$. Multiplying both sides of (3) by P, one has
$$\dot{y}_i \in -k(t)PF[\text{sgn}\,](y_i) - P\sum_{j\in\mathcal{N}_i} F[\text{sgn}\,](y_i - y_j). \quad (41)$$
Since $\mathbb{G}$ is undirected, then $F[\text{sgn}\,](y_i - y_j)$ appears in the update i if and only if $F[\text{sgn}\,](y_j - y_i)$ appears in its neighbor j's update. By adding the updates in (41) for $i = 1, 2, ..., m$ and recalling $F[\text{sgn}\,](y_i - y_j) = -F[\text{sgn}\,](y_j - y_i)$ for any two neighbors i and j in (8), one has
$$\sum_{i=1}^m \dot{y}_i \in -k(t)P\sum_{i=1}^m F[\text{sgn}\,](y_i) \quad (42)$$
which and (28) lead to
$$\dot{\bar{y}} \in -k(t)PF[\text{sgn}\,](\bar{y}), \quad t \in [T_2, \infty) \quad (43)$$
Let $\tau = \int_{T_2}^t k(s)ds$. From
$$\frac{d\bar{y}}{d\tau} = \frac{d\bar{y}}{dt} \cdot \frac{dt}{d\tau}$$
and (43), one has
$$\frac{d\bar{y}}{d\tau} \in -PF[\text{sgn}\,](\bar{y}), \quad \tau \in [0, \infty) \quad (44)$$
with $A\bar{y}(\tau) = b$ for $\tau = 0$. By Proposition 1. d, one has there exists a finite time Γ such that
$$\bar{y}(\tau) = x^*, \quad \tau \in [\Gamma, \infty),$$
This equation and the relation between τ and t imply Claim 2 is true. This completes our proof. ■

Remark 2: If the matrix A does not have full row rank, simulations suggests the proposed update (3) still enables all agents to achieve the same l_1-min solution. Second, if $Ax = b$ has no feasible solution, there does not exist a vector for all y_i to reach a consensus, for which one may refer to distributed algorithms for least square solution in [23].

Remark 3: Note that $k(t)$ in (26) is not a diminishing step size since the δ is not necessary to be zero. As indicated by (35), δ relies on both properties of the underlying network $\bar{H}$ and properties of the linear equation $\bar{P}$. By maximizing the smallest eigenvalue of $\bar{H}'\bar{P}\bar{H}$, the algorithm may allow choosing a bigger δ, which usually leads to a faster convergence as will be seen later in an numerical example.

If a solution to $Ax = b$ is needed instead of a minimum l_1-norm solution, one could simply chooses $k(t) = 0$ in (3).

Corollary 1: Suppose A has full row rank and $\mathbb{G}$ is fixed, undirected and connected. Under the distributed update

$$\dot{y}_i \in -P_i \sum_{j\in\mathcal{N}_i} F[\operatorname{sgn}](y_i - y_j) \tag{45}$$

with initializations $A_i y_i(0) = b_i$, all $y_i(t)$ converge in finite time to be the same solution to $Ax = b$.
The proof of Corollary 1 is provided in the Appendix.

D. A Numerical Example

We also present a numerical simulation of the proposed distributed update (3) in MATLAB. A widely used Runge-Kutta method with adaptive time step-sizes known as ODE45 is employed to approximate the solution to the continuous-time update (3). Since Erdös-Rényi random graphs are usually utilized in modeling complex networks and testing algorithms in practical networks [24], we perform simulations in a 20-node Erdös-Rényi random graph, for which any two vertices are connected with probability $p = 0.7$. Let each agent i knows $A_i \in \mathbb{R}^{1\times 23}$ and $b_i \in \mathbb{R}$ with entries randomly selected from the interval $[0, 1]$. Let x^* denote a l_1-min solution to linear equations $A_i x = b_i$, $i = 1, 2, \cdots, 20$. Let $y = \operatorname{col}\{y_1(t), y_2(t), \cdots, y_{20}(t)\}$ be the stack vector of all agents' states. Then $\|y(t) - \mathbf{1}_m \otimes x^*\|_1$ measures the closeness between all agents' states to x^*. As shown in Figure 1, the distributed update (3) with $k(t) = \frac{1}{t+1} + \delta$ drives all $y_i(t)$ to reach l_1-min solution x^* in finite time for different values of δ. Moreover, by increasing the value of δ in $k(t)$, which satisfies equation (35), one achieves a faster convergence.

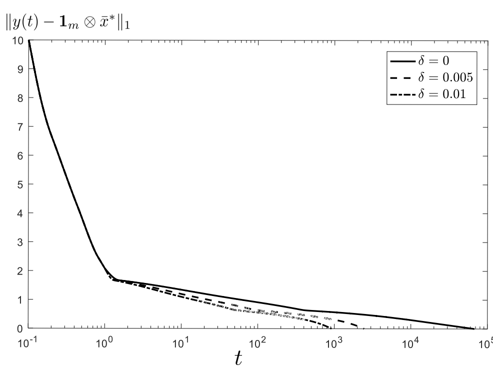

Fig. 1. Finite-time convergence of the distributed update (3) with different values of δ ($\delta \le 0.0143$ by (35)). The time t is in seconds, scaled by log.

V. Conclusion

For a fixed and undirected multi-agent network, we have developed a continuous-time distributed algorithm for achieving minimum l_1-norm solutions to underdetermined linear equations $Ax = b$ in finite time. The algorithm results from combination of the projection-consensus flow proposed in [7] and the finite-time gradient flow for consensus devised in [14]. In future, we will focus on choosing a larger δ based on information locally available to each agent in order to achieve faster convergence of the proposed update.

VI. Appendix

Proof of Lemma 4: For any $v \in \operatorname{image} \bar{H} \cap \ker \bar{P}$, one has $\bar{P}v = 0$ and there exists a vector u such that $v = \bar{H}u$. To prove (19), it is sufficient to show $v = 0$. Let $\bar{A} = \operatorname{diag}\{A_1, A_2, ..., A_m\}$, which is full row rank since A has full row rank. Then $\bar{P} = I - \bar{A}'(\bar{A}\bar{A}')^{-1}\bar{A}$. It follows from $\bar{P}\bar{H}u = 0$ that

$$\bar{H}u = \bar{A}'(\bar{A}\bar{A}')^{-1}\bar{A}\bar{H}u \tag{46}$$

Multiplying both sides of the above equation by $\mathbf{1}'_m \otimes I_n$, one has $0 = (\mathbf{1}_m \otimes I_n)'\bar{A}'(\bar{A}\bar{A}')^{-1}\bar{A}\bar{H}u$. Since $(\mathbf{1}'_m \otimes I'_n)\bar{A}' = A'$, we obtain

$$0 = A'(\bar{A}\bar{A}')^{-1}\bar{A}\bar{H}u \tag{47}$$

Since A' is full column rank, one has

$$0 = (\bar{A}\bar{A}')^{-1}\bar{A}\bar{H}u \tag{48}$$

and then multiplying by $\bar{A}'$ implies the right side of (46) is zero, i.e. $\bar{H}u = 0$. Thus $v = 0$ and (19) is true.

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

$$q = \bar{H}'p \tag{49}$$

and a vector $\phi \in F[\operatorname{sgn}](q)$ such that $\bar{H}'\bar{P}\bar{H}\phi = 0$. Note that $\bar{P}$ is a projection matrix; then $\bar{P}\bar{H}\phi = 0$, which with (19) imply

$$\bar{H}\phi = 0. \tag{50}$$

From the definition of $F[\operatorname{sgn}](x)$, one can verify that

$$\bar{\phi}'x = \|x\|_1, \quad \forall \bar{\phi} \in F[\operatorname{sgn}](x). \tag{51}$$

From $\phi \in F[\operatorname{sgn}](q)$ and (51), one has $||q||_1 = \phi' q$ which together with (49) and (50) implies $||q||_1 = 0$. Then one has $q = 0$ and (20) is true. ■

Proof of Proposition 1: The existence of a Filippov solution to (21) 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 (21). Let $x(t)$ denote such a Filippov solution for any given $x(0)$.

(a.) Note that the function $\|x\|_1$ is Lipschitz and regular [20]. Then by Lemma 3, there exists a set $\mathcal{I} = [0, \infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that (22) holds, and

$$\frac{d||x(t)||_1}{dt} = \beta'\dot{x}, \quad t \in \mathcal{I} \tag{52}$$

holds for all $\beta \in \partial\|x(t)\|_1$. It follows from $\dot{x} \in -MF[\operatorname{sgn}](x)$ in (21) that at each t there exists a $\gamma(t) \in F[\operatorname{sgn}](x)$ such that

$$\dot{x} = -M\gamma(t) \tag{53}$$

Since β can be chosen as any vector in $\partial\|x(t)\|_1$, $\gamma(t) \in F[\operatorname{sgn}](x)$ and $\partial\|x\|_1 = F[\operatorname{sgn}](x)$ from (16), one can choose $\beta = \gamma(t)$ in (52), which together with (53) leads to

$$\frac{d\|x(t)\|_1}{dt} = -\gamma(t)'M\gamma(t), \quad t \in \mathcal{I} \tag{54}$$

Note that M is positive semi-definite. Thus (23) is true.

(b.) We use the method of contradiction to prove that there exists a finite time T such that $x(T) \in \Phi(M),\ T \in \mathcal{I}$. Suppose such a finite time does not exist. One has $x(t) \in \Phi_c(M)$, for $t \in \mathcal{I}$. Then $\frac{d\|x(t)\|_1}{dt} \leq -\lambda(M)$, for $t \in \mathcal{I}$, where $\lambda(M)$ is as defined in (14). It follows that

$$\|x(t)\|_1 \leq \|x(0)\|_1 - \lambda(M)t, \quad t \in [0, \infty).$$

This contradicts the fact that $\|x(t)\|_1 \geq 0, t \in [0, \infty)$ since $\lambda(M)$ is a positive constant by Lemma 2. Thus there exists a finite time T such that $x(T) \in \Phi(M)$.

(c.) From the assumption (25), the fact that M is semipositive definite and (54), one has

$$M\gamma(t) = 0, \quad t \in [T, \infty) \setminus \mathcal{T}. \tag{55}$$

It follows from this and (53) that

$$\dot{x}(t) = 0, \quad t \in [T, \infty) \setminus \mathcal{T}.$$

By integration of $\dot{x}(t)$ from T to ∞, one has $x(t) = x(T), \quad t \in [T, \infty)$. This completes the proof.

(d.) Since a projection matrix is positive semi-definite,then

$$\frac{d\|x(t)\|_1}{dt} \leq 0, t \in \mathcal{I}; \tag{56}$$

and there exists a finite time $T \in \mathcal{I}$ such that $x(T) \in \Phi(P)$. Then there exists a vector $\phi \in F[\text{sgn}\,](x(T))$ such that $P\phi = 0$. This and (51) imply

$$\phi' x(T) = \|x(T)\|_1 \tag{57}$$

Moreover, let $\bar{x}$ denote any solution to $Ax = b$. Then

$$\phi' \bar{x} \leq \|\bar{x}\|_1 \tag{58}$$

Since $\phi \in \ker P$, one has $\phi \in \text{image } A'$. This implies that there exists a vector q such that

$$q'A = \phi' \tag{59}$$

From $Ax(T) = b = A\bar{x}$ and (57)-(59), one has

$$\|x(T)\|_1 = \phi' x(T) = q'Ax(T) = q'A\bar{x} = \phi'\bar{x} \leq \|\bar{x}\|_1 \tag{60}$$

where $\bar{x}$ is any solution to $Ax = b$. Thus $x(T)$ is a minimum l_1 norm solution to $Ax = b$. This and (56) implies $||x(t)||_1$ reaches its minimum value subject to $Ax(t) = b$ for $t \in [T, \infty) \setminus \mathcal{T}$. Thus

$$\frac{d\|x(t)\|_1}{dt} = 0, \quad t \in [T, \infty) \setminus \mathcal{T} \tag{61}$$

which satisfies the assumption (25) in Proposition 1 again. Then $x(t) = x(T),\ t \in [T, \infty)$. ■

Proof of Corollary 1: Since $A_i y_i(0) = b_i$ and $\dot{y}_i$ is in the kernel of A_i, one has $A_i y_i(t) = b_i$ for all $t \in [0, \infty)$. To prove Corollary 1, it is sufficient to show that all y_i reach consensus in finite time, in other words to prove $z(t) = \bar{H}' y(t)$ converges to 0 in finite time. Multiplying both sides of (9) on the left by $\bar{H}'$ with, $k(t) = 0$, one has

$$\dot{z} \in -\bar{H}'\bar{P}\bar{H}F[\text{sgn}\,](z) \tag{62}$$

which by Proposition 1 implies there exists a set $\mathcal{I} = [0, \infty) \setminus \mathcal{T}$ with $\mathcal{T}$ of Lebesgue measure 0 such that

$$\frac{d\|z(t)\|_1}{dt} \leq 0, t \in \mathcal{I}; \tag{63}$$

and there exists a finite time $T \in \mathcal{I}$ such that

$$z(T) \in \Phi(\bar{H}'\bar{P}\bar{H}). \tag{64}$$

From $z(T) \in \text{image } \bar{H}'$, (20) in Lemma 4 and (64), one has $z(T) = 0$. This together with (63) implies $\|z(t)\|_1 = 0,\ t \in [T, \infty)$. That is, $z(t) = 0,\ t \in [T, \infty)$. ■

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.

[2] 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.

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

[4] P. Wang, W. Ren, and Z. Duan. Distributed algorithm to solve a system of linear equations with unique or multiple solutions from arbitrary initializations. *IEEE Transactions on Control of Network Systems*, 2018. DOI:10.1109/TCNS.2018.2797805.

[5] 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*, (1):215–227, 2018.

[6] 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.

[7] G. Shi, B. D. O. Anderson, and U. Helmke. Network flows that solve linear equations. *IEEE Transactions on Automatic Control*, (6):2659–2674, 2017.

[8] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, 2009.

[9] A. Y. Yang, Z. Zhou, A. G. Balasubramanian, S. S. Sastry, and Y. Ma. Fast l_1-minimization algorithms for robust face recognition. *IEEE Transactions on Image Processing*, 22(8):3234–3246.

[10] Y. Zhao. RSP-based analysis for sparsest and least l_1-norm solutions to underdetermined linear systems. *IEEE Transactions on Signal Processing*, 61(22):5777–5788, 2013.

[11] J. Yang and Y.Zhang. Alternating direction algorithms for l_1-problems in compressive sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, 2011.

[12] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[13] 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.

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

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

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

[17] Frank H Clarke. Generalized gradients of lipschitz functionals. *Advances in Mathematics*, 40(1):52–67, 1981.

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

[19] A. Bacciotti and F. Ceragioli. Stability and stabilization of discontinuous systems and nonsmooth lyapunov functions. 4:361–376, 1999.

[20] George B Arfken and Hans J Weber. Mathematical methods for physicists, 1999.

[21] L. V. Ferreira, E. Kaszkurewicz, and A. Bhaya. Solving systems of linear equations via gradient systems with discontinuous righthand sides: Application to ls-svm. *IEEE Transactions on Neural Networks*, 16(2):501–505, 2005.

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

[23] X. Wang, J. Zhou, S. Mou, and M. J. Corless. A distributed algorithm for least square solutions. *IEEE Transactions on Automatic Control*, 2019. DOI: 10.1109/TAC.2019.2894588.

[24] Edgar N Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.