# Solving Linear Equations with Separable Problem Data over Directed Networks

Priyank Srivastava Jorge Cortés

*Abstract*—**This paper deals with linear algebraic equations where the global coefficient matrix and constant vector are given respectively, by the summation of the coefficient matrices and constant vectors of the individual agents. Our approach is based on reformulating the original problem as an unconstrained optimization. Based on this exact reformulation, we first provide a gradient-based, centralized algorithm which serves as a reference for the ensuing design of distributed algorithms. We propose two sets of exponentially stable continuous-time distributed algorithms that do not require the individual agent matrices to be invertible, and are based on estimating non-distributed terms in the centralized algorithm using dynamic average consensus. The first algorithm works for time-varying weight-balanced directed networks, and the second algorithm works for general directed networks for which the communication graphs might not be balanced. Numerical simulations illustrate our results.**

## I. Introduction

The importance of solving linear algebraic equations is paramount. They appear frequently in core mathematics as well as in applications, in physics and engineering. Nonlinear systems can often be well understood by their linear approximation. Due to the recent development of large-scale networks coupled with parallel processing power and fast communication capabilities, there is a growing effort aimed at developing distributed algorithms to solve systems of linear equations. Distributed algorithms preserve the privacy of the agents, are robust against single point of failures, and scale well with the network size. Keeping these considerations in mind, this paper is a contribution to the growing body of distributed algorithms to solve linear algebraic equations.

*Literature Review:* Justifying the ubiquity of linear equations, there is a vast and expanding literature to solve them efficiently, cf. [1]–[3] and references therein. However, most of the works consider the information structure where each agent knows some rows of the coefficient matrix and the constant vector. In those cases, the collective problem has a solution if and only if the individual equations are solvable. Instead, the problem structure considered here is different, and assumes that each agent has a full coefficient matrix and constant vector of its own. This setting appears frequently in distributed sensor fusion, where sensors are spatially distributed and they seek to build a global state estimate (e.g., about the location of a source or the position of a target) from local measurements, cf. [4], [5]. To the best of our knowledge, all the works in this category rely on the communication graph being *undirected*. The work [4] relies on the positive definiteness of the individual matrices to compute the updates and prove stability. [5] uses element-wise average consensus for the coefficient matrix as well as the constant vector, which does not scale with either the problem dimension or the network size, and is not desirable from a privacy standpoint. The work [6] also exploits the positive definite property of the individual matrices and requires the agents to know the state as well as the matrices of the neighbors. The work [7] proposes a distributed algorithm without any positive definiteness condition, but the agents are allowed to converge to different solutions. Our approach here is based on using dynamic average consensus [8], [9] to estimate certain non-distributed terms in a gradient-based algorithm for the reformulated optimization problem. We also draw inspiration from [10], [11] on distributed optimization to extend our treatment to deal with unbalanced networks. However, unlike the aforementioned works where the desired solution is not an equilibrium of the dynamics, requiring a diminishing time-varying stepsize-like parameter to ensure convergence, here we make sure that any solution of the linear equation is indeed an equilibrium of the proposed dynamics. This enables us to employ Lyapunov stability analysis to establish algorithm convergence and offers a framework to study robustness against disturbances and errors. Our work [12] requires bidirectional 2-hop communication. In contrast, the distributed algorithms here require information exchange only with immediate neighbors and work for arbitrary directed graphs.

*Statement of Contributions:* We consider linear algebraic equations where the coefficient matrices and constant vector for the overall problem are given, respectively, by the summation of the individual agents' coefficient matrices and constant vectors. Our starting point is the exact reformulation of this problem as a constrained optimization problem. Using the observation that the optimal value of this optimization is zero, we reformulate it as optimization of an unconstrained function,

This work was partially supported by NSF Awards 1917177 and 1947050.
The authors are with the Department of Mechanical and Aerospace Engineering, UC San Diego, {psrivast,cortes}@ucsd.edu

[0]We employ the following notation. $\mathbb{R}$, $\mathbb{R}_{>0}$ and $\mathbb{Z}$ denote the set of real numbers, positive real numbers, and integers, resp. $|\mathcal{X}|$ denotes the cardinality of a set $\mathcal{X}$. $\mathbf{1}$, $\mathbf{0}$ and $I$ denote a vector or matrix of all ones and zeros, and an identity matrix of appropriate dimension, resp. We let lowercase letters to denote vectors and uppercase letters to denote matrices. $\|x\|$ and $\|A\|$ denote the 2-norm of a vector $x$ and the induced 2-norm of a matrix $A$, resp. $\mathrm{diag}(x)$ denotes the diagonal matrix obtained after arranging the entries of the vector $x$ along the principal diagonal. $A_{ij}$ denotes the $ij$th element of a matrix $A$, $A^\top$ its transpose, $A^{-1}$ its inverse (if it exists) and $\mathrm{null}(A)$ its null space. $A \otimes B$ denotes the Kronecker product between two matrices $A$ and $B$. Unless otherwise stated, $\mathbf{x} \in \mathbb{R}^{mn}$ denotes the concatenated vector obtained after stacking the vectors $\{x_i\}_{i=1}^n \in \mathbb{R}^m$. $A \succ \mathbf{0}$ and $A \succeq \mathbf{0}$ imply that a matrix $A$ is positive definite and semidefinite, resp. For a symmetric matrix $A$, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote its maximum and minimum eigenvalue, resp. Regardless of the multiplicity of eigenvalue 0, $\lambda_2(A)$ denotes the minimum non-zero eigenvalue of a positive semidefinite matrix $A$. For two vectors $x, y \in \mathbb{R}^n$, $[x; y]$ denotes the concatenated vector containing the entries of $x$ and $y$, in that order, and $x > y$ means that the inequality holds elementwise.

and propose a centralized algorithm which works for weight-balanced networks and serves as a reference for the design of distributed algorithms. Using dynamic average consensus, we then propose a distributed algorithm that does not require the agent matrices to be positive definite, works for time-varying weight-balanced networks and is guaranteed to converge to a solution of the original problem exponentially fast. Building on the insights gained in establishing these results, we propose a distributed algorithm that is not limited to weight-balanced networks and is also guaranteed to converge to a solution of the linear equation exponentially fast.

## II. Preliminaries

Here we review basic notions from graph theory [10], [13], [14] and dynamic average consensus [8], [9].

*Graph Theory:* Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$ denote a weighted directed graph (or digraph), with $\mathcal{V}$ as the set of vertices (or nodes) and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ as the set of edges: $(v_i, v_j) \in \mathcal{E}$ iff there is an edge from node $v_i$ to node $v_j$. With $|\mathcal{V}| = n$, the *adjacency matrix* $\mathsf{A} \in \mathbb{R}^{n\times n}$ of $\mathcal{G}$ is such that $\mathsf{A}_{ij} > 0$ if $(v_i, v_j) \in \mathcal{E}$ and $\mathsf{A}_{ij} = 0$, otherwise. A directed path is an ordered sequence of vertices such that any pair of consecutive vertices is an edge. A digraph is strongly connected if there is a directed path between any two distinct vertices. The out- and in-degree of a node are, resp., the number of outgoing edges from and incoming edges to it. The weighted out-degree and weighted in-degree of a node $v_i$ are $d^{\text{out}}(v_i) = \sum_{j=1}^n \mathsf{A}_{ij}$ and $d^{\text{in}}(v_i) = \sum_{j=1}^n \mathsf{A}_{ji}$, resp. The *out-degree matrix* $\mathsf{D}^{\text{out}} \in \mathbb{R}^{n\times n}$ and the *in-degree matrix* $\mathsf{D}^{\text{in}} \in \mathbb{R}^{n\times n}$ are the diagonal matrices defined as $\mathsf{D}^{\text{out}}_{ii} = d^{\text{out}}(v_i)$ and $\mathsf{D}^{\text{in}}_{ii} = d^{\text{in}}(v_i)$, resp. A graph is weight-balanced if $\mathsf{D}^{\text{out}} = \mathsf{D}^{\text{in}}$. The *Laplacian matrix* $\mathsf{L} \in \mathbb{R}^{n\times n}$ is $\mathsf{L} = \mathsf{D}^{\text{in}} - \mathsf{A}$. All eigenvalues of $\mathsf{L}$ have nonnegative real parts, 0 is a simple eigenvalue with left eigenvector $\mathbf{1}$ iff $\mathcal{G}$ is strongly connected, and $\mathsf{L}\,\mathbf{1} = \mathbf{0}$ iff $\mathcal{G}$ is weight-balanced iff $\mathsf{L} + \mathsf{L}^\top$ is positive semidefinite, cf. [13, Theorem 1.37]. If $\mathcal{G}$ is strongly connected, it follows from [14, Lemma 3] that there exists a positive right eigenvector $\bar{v} \in \mathbb{R}^n$ associated to the eigenvalue 0 of $\mathsf{L}$.

*Dynamic Average Consensus:* Consider a group of $n \in \mathbb{Z}_{>1}$ agents communicating over a weight-balanced digraph $\mathcal{G}$ whose Laplacian is denoted by $\mathsf{L}$. Each agent $i \in \{1, \dots, n\}$ has a state $x_i \in \mathbb{R}$ and an input $z_i \in \mathbb{R}$. The dynamic average consensus algorithm aims at making all the agents track the average $\frac{1}{n}\sum_{i=1}^n z_i$ asymptotically. Here we present the algorithm following [8], where it was introduced for undirected graphs. The algorithm is given by

$$\dot{\mathbf{x}} = -\mathsf{L}\,\mathbf{x} + \dot{\mathbf{z}},$$

If $\sum_{i=1}^n x_i(0) = \sum_{i=1}^n z_i(0)$ and the input $\mathbf{z}$ is bounded, then $x_i(t) \to \frac{1}{n}\sum_{i=1}^n z_i(t)$ asymptotically for all $i \in \{1, \dots, n\}$, cf. [8].

## III. Problem Formulation

Consider a group of $n$ agents interacting over a digraph that seek to solve in a distributed way the linear algebraic equation

$$\underbrace{\Big(\sum_{i=1}^n A_i\Big)}_{A} x = \underbrace{\Big(\sum_{i=1}^n b_i\Big)}_{b}, \tag{1}$$

where $x \in \mathbb{R}^m$ is the unknown solution vector, and $A_i \in \mathbb{R}^{m\times m}$ and $b_i \in \mathbb{R}^m$ are the coefficient matrix and constant vector corresponding to each agent $i \in \{1, \dots, n\}$. We assume that (1) has at least one solution. Interestingly, the formulation (1) includes, as a particular case, scenarios where each agent $i$ knows only some rows of the coefficient matrix $A$ and constant vector $b$. Our approach consists of first formulating (1) as a system involving $n$ unknown solution vectors, one per agent, and then reformulating it as a convex optimization problem. Based on this reformulation, we propose two sets of (out-)distributed algorithms (where each agent only needs information from its out-neighbors) to find the solutions of (1).

We start by endowing each agent with its own version $x_i \in \mathbb{R}^m$ of $x$. Then (1) can be equivalently written as

$$\sum_{i=1}^n A_i x_i = \sum_{i=1}^n b_i, \tag{2a}$$

$$x_i = x_j \quad \forall i, j. \tag{2b}$$

Equation (2b) ensures that $x_i = x$ for all the agents. Clearly the set of equations (2) and the original problem (1) are equivalent. Next we formulate (2) as a convex optimization problem. Consider the quadratic function $f : \mathbb{R}^{mn} \to \mathbb{R}$

$$f(\mathbf{x}) = \Big(\sum_{i=1}^n (A_i x_i - b_i)\Big)^\top \Big(\sum_{i=1}^n (A_i x_i - b_i)\Big),$$

which is convex and attains its minimum over the solution set of (2a). For convenience, we use $\mathbf{L} = \mathsf{L} \otimes I$ and $f(\mathbf{x}) = (\mathbf{A}\,\mathbf{x} - \mathbf{b})^\top \mathbb{1}\,\mathbb{1}^\top (\mathbf{A}\,\mathbf{x} - \mathbf{b})$, where $\mathbb{1} = \mathbf{1} \otimes I$, $\mathbf{A} \in \mathbb{R}^{mn\times mn}$ denotes the block-diagonal matrix obtained after putting the matrices $\{A_i\}_{i=1}^n$ along the principal diagonal, and $\mathbf{b} = [b_1; \dots; b_n] \in \mathbb{R}^{mn}$. If $\mathcal{G}$ is strongly connected, the solutions of (2) are the same as the optimizers of

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{L}^\top \mathbf{x} = \mathbf{0}\,. \end{aligned} \tag{3}$$

*Remark* 1. *(Distributed algorithmic solutions to optimization problem):* The problem (3) can be solved over an undirected graph by reformulating it using the techniques in [15] and employing the saddle-point dynamics, cf. [16], [17]. These dynamics involve terms of the form $\mathbf{L}^\top$ and, to be implemented over a digraph, would need information from in- as well as out-neighbors and hence are not suitable for our setup. It is worth mentioning that works that deal with distributed optimization under consensus constraints over digraphs, see e.g. [10], [18] and references therein, require the objective function to be separable, and therefore are not applicable here. •

## IV. Distributed Algorithms Over Weight-Balanced Networks

Here, we present distributed algorithms to solve problem (1) over weight-balanced networks.

### A. Centralized Algorithm

Here, we present a centralized algorithm making use of the observation that the objective function $f$ vanishes at the optimizers of (3). Building on this insight, consider

$$\min_{\mathbf{x}} \quad \frac{1}{2}\alpha\,\mathbf{x}^\top (\mathbf{L} + \mathbf{L}^\top)\,\mathbf{x} + \beta f(\mathbf{x}), \tag{4}$$

where $\alpha, \beta > 0$. Clearly, (3) and (4) have the same set of solutions if $\mathcal{G}$ is strongly connected and weight-balanced. Since problem (4) is unconstrained, one can use gradient descent to find its optimizers. However, the gradient $-\alpha(\mathbf{L}+\mathbf{L}^\top)\mathbf{x} - \beta\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\mathbf{x}-\mathbf{b})$ of the objective function in (4) involves terms with $\mathbf{L}^\top$, whose computation would require information from in-neighbors. Instead, we consider the following gradient-based dynamics

$$\dot{\mathbf{x}} = -\alpha\mathbf{L}\mathbf{x} - \beta\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\mathbf{x}-\mathbf{b}). \tag{5}$$

Whenever convenient, we refer to (5) as $\psi_{\text{grad}}$. Note that the first term in the dynamics (5) is distributed, meaning that each agent can implement it with information from its out-neighbors. The second term, however, requires collective information from all the agents because of the summation across the network. Nevertheless, this algorithm serves as the basis for our distributed algorithm design in the next section.

The next result formally characterizes the equivalence between the equilibria of (5) and the solutions of (1).

**Lemma IV.1.** (Equivalence between (5) and (1)): *Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph. Then for all $\alpha, \beta \in \mathbb{R}_{>0}$, $\mathbf{x}^*$ is an equilibrium of (5) if and only if $\mathbf{x}^* = \mathbf{1}\otimes x^*$, where $x^* \in \mathbb{R}^m$ solves (1).*

*Proof:* The implication from right to left is immediate. To prove the implication in the other direction, let $\bar{x} \in \mathbb{R}^m$ be a solution of (1) and consider $\bar{\mathbf{x}} = \mathbf{1}\otimes\bar{x}$. Since $\mathbf{x}^*$ and $\bar{\mathbf{x}}$ are equilibria of (5),

$$\alpha\mathbf{L}(\mathbf{x}^* - \bar{\mathbf{x}}) + \beta\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}(\mathbf{x}^* - \bar{\mathbf{x}}) = \mathbf{0}. \tag{6}$$

Let $\mathbf{Q}_{11} = \frac{1}{2}\alpha(\mathbf{L}+\mathbf{L}^\top) + \beta\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}$. Then (6) implies

$$(\mathbf{x}^* - \bar{\mathbf{x}})^\top \mathbf{Q}_{11}(\mathbf{x}^* - \bar{\mathbf{x}}) = 0.$$

Since $\mathcal{G}$ is weight-balanced, $(\mathbf{L}+\mathbf{L}^\top) \succeq \mathbf{0}$. This along with the fact that $\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A} \succeq \mathbf{0}$ implies $\mathbf{L}^\top(\mathbf{x}^* - \bar{\mathbf{x}}) = \mathbf{0}$ and $\mathbb{1}^\top \mathbf{A}(\mathbf{x}^* - \bar{\mathbf{x}}) = \mathbf{0}$. Therefore, $\mathbf{x}^* = \mathbf{1}\otimes x^*$, for some $x^* \in \mathbb{R}^m$ which satisfies $Ax^* = A\bar{x} = b$, as claimed. ■

The next result characterizes the convergence of (5).

**Proposition IV.2.** (Exponential stability of (5)): *Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph. Then for all $\alpha, \beta \in \mathbb{R}_{>0}$, any trajectory of (5) converges exponentially to a point of the form $\mathbf{x}^* = \mathbf{1}\otimes x^*$, where $x^* \in \mathbb{R}^m$ solves (1).*

*Proof:* Consider a vector $\mathbf{w} \in \mathbb{R}^{mn}$ in the null space of $\mathbf{Q}_{11}$. Using the same line of arguments as in the proof of Lemma IV.1, this implies that $\mathbf{L}^\top\mathbf{w} = \mathbf{0}$ and $\mathbb{1}^\top \mathbf{A}\mathbf{w} = \mathbf{0}$. Therefore, along (5),

$$\dot{\mathbf{x}}^\top\mathbf{w} = -(\alpha\mathbf{x}^\top\mathbf{L}^\top + \beta(\mathbf{A}\mathbf{x}-\mathbf{b})^\top \mathbb{1}\mathbb{1}^\top \mathbf{A})\mathbf{w} = 0.$$

This means that the dynamics (5) are orthogonal to the null space of $\mathbf{Q}_{11}$ and hence the component of $\mathbf{x}$ in the null space of $\mathbf{Q}_{11}$, say $\mathbf{x}_{\text{null}}$, remains constant. Given the initial condition $\mathbf{x}(0)$, consider the particular equilibrium $\mathbf{x}^*$ of (5) satisfying $\mathbf{x}^*_{\text{null}} = \mathbf{x}(0)_{\text{null}}$. Since different equilibria differ only in their null space component, $\mathbf{x}^*$ defined this way is unique. Consider the Lyapunov function candidate $V : \mathbb{R}^{mn} \to \mathbb{R}$

$$V(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^\top(\mathbf{x}-\mathbf{x}^*).$$

The Lie derivative of $V$ along the dynamics (5) is given by

$$\begin{aligned}\mathcal{L}_{\psi_{\text{grad}}} V &= -(\mathbf{x}-\mathbf{x}^*)^\top(\alpha\mathbf{L}\mathbf{x} + \beta\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\mathbf{x}-\mathbf{b})) \\ &= -(\mathbf{x}-\mathbf{x}^*)^\top\mathbf{Q}_{11}(\mathbf{x}-\mathbf{x}^*) \le -2\lambda_2(\mathbf{Q}_{11})V.\end{aligned}$$

The last inequality follows from applying the Courant-Fischer theorem [19, Theorem 4.2.11] together with the fact that $(\mathbf{x}-\mathbf{x}^*)^\top\mathbf{w} = 0$ as $\mathbf{x}_{\text{null}}$ is constant. Using the monotonicity theorem [19, Corollary 4.3.3], we further have

$$\mathcal{L}_{\psi_{\text{grad}}} V \le -2\min\left\{\frac{1}{2}\alpha\lambda_2(\mathsf{L}+\mathsf{L}^\top), \beta\lambda_2(\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A})\right\}V.$$

Hence, the dynamics (5) is exponentially stable with a rate depending on $\alpha, \beta, \mathsf{L}$ and $\{A_i\}_{i=1}^n$. ■

### *B. Distributed Algorithm*

We present a distributed algorithm to find a solution of (1), which is based on the centralized algorithm (5) and involves employing dynamic average consensus (cf. Section II) to estimate the aggregate $\mathbb{1}^\top(\mathbf{A}\mathbf{x}-\mathbf{b})$. Formally,

$$\dot{\mathbf{x}} = -\alpha\mathbf{L}\mathbf{x} - n\beta\mathbf{A}^\top\mathbf{y}, \tag{7a}$$

$$\dot{\mathbf{y}} = -\alpha\mathbf{A}\mathbf{L}\mathbf{x} - n\beta\mathbf{A}\mathbf{A}^\top\mathbf{y} - \gamma\mathbf{L}\mathbf{y}, \tag{7b}$$

with design parameter $\gamma > 0$. Here, each agent $i \in \{1, \dots, n\}$ updates $y_i \in \mathbb{R}^m$ which estimates the average mismatch $\frac{1}{n}\mathbb{1}^\top(\mathbf{A}\mathbf{x}-\mathbf{b})$. The dynamics (7) is distributed as each agent just needs to know its state and that of its out-neighbors. Whenever convenient, we refer to it as $\psi_{\text{gdac}}$. The following result characterizes the equilibria of (7) and shows that the total deviation from the average mismatch is conserved.

**Lemma IV.3.** (Equilibria of (7) and invariance of total deviation): *Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph. Then, if $(\mathbf{x}^*, \mathbf{0})$ is an equilibrium of (7) then $\mathbf{x}^* = \mathbf{1}\otimes x^*$, where $x^* \in \mathbb{R}^m$. Moreover, for all $\alpha, \beta, \gamma \in \mathbb{R}_{>0}$, $\mathbb{1}^\top(\mathbf{y} - \mathbf{A}\mathbf{x})$ remains constant along the evolution of (7).*

*Proof:* Let $(\mathbf{x}^*, \mathbf{0})$ be an equilibrium of (7). From (7a), it follows that $\mathbf{L}\mathbf{x} = \mathbf{0}$, and hence $\mathbf{x}^* = \mathbf{1}\otimes x^*$ for some $x^* \in \mathbb{R}^m$, establishing the first statement. Now, consider the derivative $\mathbb{1}^\top(\dot{\mathbf{y}} - \mathbf{A}\dot{\mathbf{x}}) = -\gamma\mathbb{1}^\top\mathbf{L}\mathbf{y} = \mathbf{0}$. Hence, $\mathbb{1}^\top(\mathbf{y} - \mathbf{A}\mathbf{x})$ is conserved along the evolution of (7). ■

*Remark 2. (Distributed initialization of the $\psi_{\text{gdac}}$ algorithm):* From Lemma IV.3, we observe that in order for a trajectory of (7) to converge to an equilibrium of the form $(\mathbf{x}^*, \mathbf{y}^*) = (\mathbf{1}\otimes x^*, \mathbf{0})$, where $x^* \in \mathbb{R}^m$ solves (1), its initial condition must satisfy $\mathbb{1}^\top\mathbf{y}(0) = \mathbb{1}^\top(\mathbf{A}\mathbf{x}(0) - \mathbf{b})$. This could be implemented in a distributed way if each agent $i \in \{1, \dots, n\}$ chooses its initial states satisfying $y_i(0) = A_i x_i(0) - b_i$. One trivial selection, for example, is $\mathbf{x}(0) = \mathbf{0}$ and $\mathbf{y}(0) = -\mathbf{b}$. •

The next result characterizes the convergence of (7).

**Theorem IV.4.** (Exponential stability of (7) over balanced networks): *Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph and assume $\text{null}(A) \subseteq \text{null}(A_i)$, for all $i \in \{1, \dots, n\}$. Let $\alpha, \beta \in \mathbb{R}_{>0}$ and define*

$$\bar{\gamma} = \max\left\{\frac{2}{\lambda_2(\mathsf{L}+\mathsf{L}^\top)}\lambda_{\max}\left(\frac{\mathbf{Q}_{12}^\top\mathbf{Q}_{12}}{\lambda_2(\mathbf{Q}_{11})} - n\beta\mathbf{A}\mathbf{A}^\top\right), 0\right\},$$

*where* $\mathbf{Q}_{11}= \frac{1}{2}\alpha(\mathbf{L}+\mathbf{L}^\top) + \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}$ *and* $\mathbf{Q}_{12} = \frac{1}{2}(n\beta\,\mathbf{A}^\top + \alpha\,\mathbf{L}^\top \mathbf{A}^\top + \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}\,\mathbf{A}^\top)$*. Then, for all* $\gamma \in (\bar{\gamma}, \infty)$*, any trajectory of* (7) *with initial condition satisfying* $\mathbb{1}^\top \mathbf{y}(0) = \mathbb{1}^\top(\mathbf{A}\,\mathbf{x}(0) - \mathbf{b})$ *converges exponentially to* $(\mathbf{x}^*, \mathbf{0})$*, where* $\mathbf{x}^* = \mathbf{1} \otimes x^*$ *and* $x^* \in \mathbb{R}^m$ *solves* (1)*.*

*Proof:* Define the error variable

$$\mathbf{e} = \mathbf{y} - \frac{1}{n}\mathbb{1}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x} - \mathbf{b}), \tag{8}$$

measuring the difference between the agents' estimates and the actual value of average mismatch. Note that

$$\begin{aligned}\dot{\mathbf{e}} &= \dot{\mathbf{y}} - \frac{1}{n}\mathbb{1}\mathbb{1}^\top \mathbf{A}\,\dot{\mathbf{x}}, \\ &= -\alpha\mathbf{\Pi}\,\mathbf{A}\,\mathbf{L}\,\mathbf{x} - n\beta\mathbf{\Pi}\,\mathbf{A}\,\mathbf{A}^\top \mathbf{y} - \gamma\,\mathbf{L}\,\mathbf{y},\end{aligned}$$

where $\mathbf{\Pi} = I - \frac{1}{n}\mathbb{1}\mathbb{1}^\top$. Rewriting (7) in terms of $\mathbf{x}$ and $\mathbf{e}$,

$$\dot{\mathbf{x}} = -\alpha\,\mathbf{L}\,\mathbf{x} - \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x} - \mathbf{b}) - n\beta\,\mathbf{A}^\top \mathbf{e}, \tag{9a}$$

$$\begin{aligned}\dot{\mathbf{e}} = &-\alpha\mathbf{\Pi}\,\mathbf{A}\,\mathbf{L}\,\mathbf{x} - \beta\mathbf{\Pi}\,\mathbf{A}\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x} - \mathbf{b}) \\ &- n\beta\mathbf{\Pi}\,\mathbf{A}\,\mathbf{A}^\top \mathbf{e} - \gamma\,\mathbf{L}\,\mathbf{e}.\end{aligned} \tag{9b}$$

From the proof of Proposition IV.2, we know that if $\mathbf{w} \in \mathbb{R}^{mn}$ is in the null space of $\mathbf{Q}_{11}$, then $\mathbf{L}^\top \mathbf{w} = \mathbf{0}$ and $\mathbb{1}^\top \mathbf{A}\,\mathbf{w} = \mathbf{0}$. Therefore, $\mathbf{w} = \mathbf{1} \otimes w$, where $w \in \mathbb{R}^m$ belongs to $w \in \text{null}(A)$. By hypothesis, $A_i w = \mathbf{0}$ for all $i \in \{1, \dots, n\}$. Therefore, from (9a), $\dot{\mathbf{x}}^\top \mathbf{w} = \mathbf{0}$, and the $\mathbf{x}$ component of the equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ of (7) satisfies $\mathbf{x}^*_{\text{null}} = \mathbf{x}(0)_{\text{null}}$ and is unique. With the initialization of the statement, it follows from Lemma IV.3 that $\mathbf{y}^* = \mathbf{1} \otimes \frac{1}{n}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x}^* - \mathbf{b})$. Substituting this value of $\mathbf{y}^*$ in (7a) and following the proof of Lemma IV.1, one can establish that the corresponding equilibrium is of the form $(\mathbf{1} \otimes x^*, \mathbf{0})$, where $x^* \in \mathbb{R}^m$ is a solution of (1). Consider the Lyapunov function candidate $V_2 : \mathbb{R}^{2mn} \to \mathbb{R}$

$$V_2(\mathbf{x}, \mathbf{e}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}\mathbf{e}^\top\mathbf{e}.$$

The Lie derivative of $V_2$ along (9) is given by

$$\begin{aligned}\mathcal{L}_{\psi_{\text{gdac}}} V_2 = &-(\mathbf{x} - \mathbf{x}^*)^\top(\alpha\,\mathbf{L}\,\mathbf{x} + \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x} - \mathbf{b})) \\ &- n\beta(\mathbf{x} - \mathbf{x}^*)^\top\mathbf{A}^\top\mathbf{e} - \mathbf{e}^\top\mathbf{\Pi}\,\mathbf{A}(\alpha\,\mathbf{L}\,\mathbf{x} + n\beta\,\mathbf{A}^\top\mathbf{e}) \\ &- \mathbf{e}^\top(\beta\mathbf{\Pi}\,\mathbf{A}\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top(\mathbf{A}\,\mathbf{x} - \mathbf{b}) + \gamma\,\mathbf{L}\,\mathbf{e}) \\ = &-\begin{bmatrix}\mathbf{x} - \mathbf{x}^* \\ \mathbf{e}\end{bmatrix}^\top \begin{bmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \mathbf{Q}_{22}\end{bmatrix} \begin{bmatrix}\mathbf{x} - \mathbf{x}^* \\ \mathbf{e}\end{bmatrix},\end{aligned}$$

where $\mathbf{Q}_{22} = \frac{1}{2}\gamma(\mathbf{L} + \mathbf{L}^\top) + n\beta\,\mathbf{A}\,\mathbf{A}^\top$ and we have used the fact that due to the mentioned initialization, $\mathbb{1}^\top \mathbf{e} = \mathbf{0}$ from Lemma IV.3. Since $\mathbf{x}_{\text{null}}$ is constant, $(\mathbf{x} - \mathbf{x}^*)^\top\mathbf{w} = \mathbf{0}$ and from the Courant-Fischer theorem [19, Theorem 4.2.11],

$$-(\mathbf{x} - \mathbf{x}^*)^\top\mathbf{Q}_{11}(\mathbf{x} - \mathbf{x}^*) \leq -\lambda_2(\mathbf{Q}_{11})(\mathbf{x} - \mathbf{x}^*)^\top(\mathbf{x} - \mathbf{x}^*).$$

Also, since $\mathbb{1}^\top \mathbf{e} = \mathbf{0}$ and $\mathcal{G}$ is weight-balanced, it again follows from the Courant-Fischer theorem that

$$-\mathbf{e}^\top\mathbf{Q}_{22}\mathbf{e} \leq -\frac{1}{2}\gamma\lambda_2(\mathsf{L} + \mathsf{L}^\top)\mathbf{e}^\top\mathbf{e} - n\beta\mathbf{e}^\top \mathbf{A}\,\mathbf{A}^\top \mathbf{e}.$$

Therefore, we can upper bound the Lie derivative as

$$\mathcal{L}_{\psi_{\text{gdac}}} V_2 \leq -\begin{bmatrix}\mathbf{x} - \mathbf{x}^* \\ \mathbf{e}\end{bmatrix}^\top \underbrace{\begin{bmatrix}\lambda_2(\mathbf{Q}_{11})I & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \bar{\mathbf{Q}}_{22}\end{bmatrix}}_{\bar{\mathbf{Q}}} \begin{bmatrix}\mathbf{x} - \mathbf{x}^* \\ \mathbf{e}\end{bmatrix},$$

where $\bar{\mathbf{Q}}_{22} = \frac{1}{2}\gamma\lambda_2(\mathsf{L} + \mathsf{L}^\top)I + n\beta\,\mathbf{A}\,\mathbf{A}^\top$. Next, we examine the positive definiteness of $\bar{\mathbf{Q}}$. Using the Schur complement [20], $\bar{\mathbf{Q}} \succ \mathbf{0}$ iff

$$\frac{1}{2}\gamma\lambda_2(\mathsf{L} + \mathsf{L}^\top)I + n\beta\,\mathbf{A}\,\mathbf{A}^\top - \frac{1}{\lambda_2(\mathbf{Q}_{11})}\mathbf{Q}_{12}^\top\mathbf{Q}_{12} \succ \mathbf{0}.$$

Hence, if $\gamma > \bar{\gamma}$, then $\bar{\mathbf{Q}} \succ \mathbf{0}$, and $\mathcal{L}_{\psi_{\text{gdac}}} V_2 \leq -2\lambda_{\min}(\bar{\mathbf{Q}})V_2$, concluding the proof. ■

The null space condition in Theorem IV.4 makes sure that $\mathbf{x}^*_{\text{null}}$ remains invariant along the evolution of (7) and all the agents approach the solution of (1) closest to $\mathbf{x}(0)$. This condition is automatically satisfied if the matrix $A$ is full rank, or in other words, equation (1) has a unique solution. We believe (and simulations also suggest) that if this condition is not satisfied, then the $\mathbf{x}$ component of the dynamics still converges to a solution of (1).

*Remark* 3. *(Lower bound on* $\gamma$*):* The lower bound $\bar{\gamma}$ in Theorem IV.4 is conservative in general. In fact, the algorithm may converge even if this condition is not satisfied, something that we have observed in simulation. Note also that although $\alpha$ and $\beta$ are free parameters, they should still be carefully chosen as $\bar{\gamma}$ depends on them. •

The result above can be extended to time-varying networks. In case $\mathcal{G}(t)$ is time-varying, the algorithm in (7) reads as

$$\dot{\mathbf{x}} = -\alpha\,\mathbf{L}(t)\,\mathbf{x} - n\beta\,\mathbf{A}^\top \mathbf{y}, \tag{10a}$$

$$\dot{\mathbf{y}} = -\alpha\,\mathbf{A}\,\mathbf{L}(t)\,\mathbf{x} - n\beta\,\mathbf{A}\,\mathbf{A}^\top \mathbf{y} - \gamma\,\mathbf{L}(t)\mathbf{y}. \tag{10b}$$

The next result formally characterizes the convergence of (10). Its proof is similar to that of Theorem IV.4 and hence omitted.

**Theorem IV.5.** (Exponential stability of (10) over time-varying balanced networks): *Let* $\{\mathcal{G}(t)\}_{t=0}^\infty$ *be a sequence of strongly connected and weight-balanced digraphs with uniformly bounded edge weights (i.e., there exists* $a \in (0, \infty)$ *such that* $\mathsf{A}_{ij}(t) < a$ *for all* $(i, j)$ *and* $t \geq 0$*), and assume* $\text{null}(A) \subseteq \text{null}(A_i)$*, for all* $i \in \{1, \dots, n\}$*. Let* $\alpha, \beta \in \mathbb{R}_{>0}$ *and define* $\bar{\gamma}(t)$ *as*

$$\max\Bigg\{\frac{2}{\lambda_2(\mathsf{L}(t) + \mathsf{L}(t)^\top)}\lambda_{\max}\Bigg(\frac{\mathbf{Q}_{12}(t)^\top\mathbf{Q}_{12}(t)}{\lambda_2(\mathbf{Q}_{11}(t))} - n\beta\,\mathbf{A}\,\mathbf{A}^\top\Bigg), 0\Bigg\},$$

*where* $\mathbf{Q}_{11}(t) = \frac{1}{2}\alpha(\mathbf{L}(t) + \mathbf{L}(t)^\top) + \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}$ *and* $\mathbf{Q}_{12}(t) = \frac{1}{2}(n\beta\,\mathbf{A}^\top + \alpha\,\mathbf{L}(t)^\top \mathbf{A}^\top + \beta\,\mathbf{A}^\top \mathbb{1}\mathbb{1}^\top \mathbf{A}\,\mathbf{A}^\top)$*. Then for all* $\gamma \in (\hat{\gamma}, \infty)$*, where* $\hat{\gamma} = \sup_{t \geq 0} \bar{\gamma}(t)$*, any trajectory of* (10) *with initial conditions* $\mathbb{1}^\top \mathbf{y}(0) = \mathbb{1}^\top(\mathbf{A}\,\mathbf{x}(0) - \mathbf{b})$ *converges exponentially to* $(\mathbf{x}^*, \mathbf{0})$*, where* $\mathbf{x}^* = \mathbf{1} \otimes x^*$ *and* $x^* \in \mathbb{R}^m$ *solves* (1)*.*

## V. Distributed Algorithm Over Unbalanced Networks

In this section, we extend our approach to solve problem (1) over graphs that are not necessarily balanced. In those scenarios, since $\mathsf{L}\,\mathbf{1} \neq \mathbf{0}$, the one-to-one correspondence between the desired equilibria of (5) or (7) and the solutions of (1) does not hold anymore. To overcome this, we propose

$$\dot{\mathbf{x}} = -\alpha\,\mathbf{L}\,\bar{\mathbf{V}}\,\mathbf{x} - n\beta\,\mathbf{A}^\top \mathbf{y}, \tag{11a}$$

$$\dot{\mathbf{y}} = -\alpha\,\mathbf{A}\,\mathbf{L}\,\bar{\mathbf{V}}\,\mathbf{x} - n\beta\,\mathbf{A}\,\mathbf{A}^\top \mathbf{y} - \gamma\,\mathbf{L}\,\bar{\mathbf{V}}\mathbf{y}, \tag{11b}$$

where $\bar{\mathbf{V}} = \mathrm{diag}(\bar{\mathbf{v}})$, and $\bar{\mathbf{v}}$ is a positive right eigenvector with eigenvalue 0 of $\mathbf{L}$. Note that $\bar{\mathbf{v}} = \mathbf{1} \otimes \bar{v}$, where $\bar{v}$ is a positive right eigenvector with eigenvalue 0 of $\mathsf{L}$. Exponential stability of (11) can be established by interpreting $\mathsf{L} \cdot \mathrm{diag}(\bar{v})$ as the Laplacian of a weight-balanced graph and then following the same steps as in the proof of Theorem IV.4, but we omit it here for reasons of space. Although (11) is distributed, it assumes that agents have a priori knowledge of the corresponding entries of $\bar{v}$ which might be limiting in practice. To deal with this limitation, we propose an algorithm that does not require such knowledge by augmenting (11) with an additional dynamics converging to $\bar{\mathbf{v}}$,

$$\dot{\mathbf{x}} = -\alpha \mathbf{L} \mathbf{V} \mathbf{x} - n\beta \mathbf{A}^\top \mathbf{y}, \tag{12a}$$

$$\dot{\mathbf{y}} = -\alpha \mathbf{A} \mathbf{L} \mathbf{V} \mathbf{x} - n\beta \mathbf{A} \mathbf{A}^\top \mathbf{y} - \gamma \mathbf{L} \mathbf{V} \mathbf{y}, \tag{12b}$$

$$\dot{\mathbf{v}} = -\mathbf{L} \mathbf{v}, \tag{12c}$$

where $\mathbf{V} = \mathrm{diag}(\mathbf{v})$. Whenever convenient, we refer to dynamics (12) as $\psi_{\text{dist}}$. Note that, unlike all the dynamics discussed so far, $\psi_{\text{dist}}$ is nonlinear.

*Remark* 4. *(Distributed nature of* (12)*):* The dynamics (12) is out-distributed, but requires each agent $i \in \{1, \dots, n\}$ to have knowledge of its in-degree because $\mathsf{L} = \mathsf{D}^{\text{in}} - \mathsf{A}$ and the graph is not weight-balanced. If we use instead the out-Laplacian $\mathsf{L} = \mathsf{D}^{\text{out}} - \mathsf{A}$, then one could still define an equivalent algorithm for (11) with $\mathbf{L} \bar{\mathbf{V}}$ replaced by $\bar{\mathbf{V}} \mathbf{L}$, but (12c) would look like $\dot{\mathbf{v}} = -\mathbf{L}^\top \mathbf{v}$, which would require state information from in-neighbors too. •

The next result characterizes the convergence of (12).

**Theorem V.1.** (Exponential stability of (12) over unbalanced networks): *Let $\mathcal{G}$ be a strongly connected digraph and assume* $\mathrm{null}(A) \subseteq \mathrm{null}(A_i)$*, for all $i \in \{1, \dots, n\}$. Let $\alpha, \beta \in \mathbb{R}_{>0}$ and define*

$$\bar{\gamma} = \max\left\{ \frac{2}{\lambda_2(\mathsf{L} \bar{V} + \bar{V} \mathsf{L}^\top)} \lambda_{\max}\left( \frac{\mathbf{Q}_{12}^\top \mathbf{Q}_{12}}{\lambda_2(\mathbf{Q}_{11})} - n\beta \mathbf{A} \mathbf{A}^\top \right), 0 \right\},$$

*where* $\mathbf{Q}_{11} = \frac{1}{2}(\alpha \mathbf{L} \bar{\mathbf{V}} + \bar{\mathbf{V}} \mathbf{L}^\top) + \beta \mathbf{A}^\top \mathbb{1} \mathbb{1}^\top \mathbf{A}$, $\mathbf{Q}_{12} = \frac{1}{2}(n\beta \mathbf{A}^\top + \alpha \bar{\mathbf{V}} \mathbf{L}^\top \mathbf{A}^\top + \beta \mathbf{A}^\top \mathbb{1} \mathbb{1}^\top \mathbf{A} \mathbf{A}^\top)$*, $\bar{v}$ is the positive eigenvector with eigenvalue* 0 *of* $\mathsf{L}$ *satisfying* $\mathbf{1}^\top \bar{v} = 1$*, and* $\bar{V} = \mathrm{diag}(\bar{v})$*. Then, for all $\gamma \in (\bar{\gamma}, \infty)$, any trajectory of* (12) *with initial condition satisfying* $\mathbb{1}^\top \mathbf{y}(0) = \mathbb{1}^\top (\mathbf{A} \mathbf{x}(0) - \mathbf{b})$ *and* $\mathbf{v}(0) = \frac{1}{n} \mathbf{1}$*, converges exponentially to* $(\mathbf{x}^*, \mathbf{0}, \bar{\mathbf{v}})$*, where* $\mathbf{x}^* = \mathbf{1} \otimes x^*$ *and* $x^* \in \mathbb{R}^m$ *solves* (1)*, and* $\bar{\mathbf{v}} = \mathbf{1} \otimes \bar{v}$*.*

*Proof:* From [11, Proposition 2.2], we have that $\mathbf{v}(t) > \mathbf{0}$ for all $t \geq 0$. Also, since $\mathbf{1}^\top \mathsf{L} = \mathbf{0}$, $\mathbf{1}^\top \mathbf{v}$ is conserved along the evolution of (12c). Hence $\mathbf{v}(t) \to \bar{\mathbf{v}}$ exponentially fast with a rate determined by the non-zero eigenvalue of $\mathsf{L}$ with the smallest real part. Let us interpret the dynamics (12a)-(12b) as the dynamics (11) with some disturbance $\mathbf{d}(t)$ defined by

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}^{\mathbf{x}} \\ \mathbf{d}^{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} -\alpha \mathbf{L}(\mathbf{V} - \bar{\mathbf{V}}) \mathbf{x} \\ -\alpha \mathbf{A} \mathbf{L}(\mathbf{V} - \bar{\mathbf{V}}) \mathbf{x} - \gamma \mathbf{L}(\mathbf{V} - \bar{\mathbf{V}}) \mathbf{y} \end{bmatrix},$$

which goes to $\mathbf{0}$ as $t \to \infty$. Consider a vector $\mathbf{w} \in \mathrm{null}(\mathbf{Q}_{11})$. Then as in the proof of Theorem IV.4, $\mathbf{w} = \mathbf{1} \otimes w$, where $w \in \mathrm{null}(A)$ and by hypothesis, $A_i w = \mathbf{0}$ for all $i \in \{1, \dots, n\}$. Since $\mathbb{1}^\top \mathbf{L} = \mathbf{0}$, therefore, $\mathbf{w}^\top \mathbf{d}^{\mathbf{x}} = 0$ and we still have $\mathbf{w}^\top \dot{\mathbf{x}} = 0$, and the $\mathbf{x}$ component of the equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \bar{\mathbf{v}})$ of (12) satisfies $\mathbf{x}^*_{\text{null}} = \mathbf{x}(0)_{\text{null}}$ and is unique. With the initialization of the statement and following the same steps as in the proof of Lemma IV.3, one can establish that $\mathbf{y}^* = \mathbf{1} \otimes \frac{1}{n} \mathbb{1}^\top (\mathbf{A} \mathbf{x}^* - \mathbf{b})$. Substituting this value of $\mathbf{y}^*$ in (12a) and following the proof of Lemma IV.1, one can establish that the corresponding equilibrium is of the form $(\mathbf{1} \otimes x^*, \mathbf{0}, \bar{\mathbf{v}})$, where $x^* \in \mathbb{R}^m$ is a solution of (1). Consider now the Lyapunov function candidate $V_3 : \mathbb{R}^{3mn} \to \mathbb{R}$

$$V_3(\mathbf{x}, \mathbf{e}, \mathbf{v}) = V_2(\mathbf{x}, \mathbf{e}) + \frac{\delta}{2} (\mathbf{v} - \bar{\mathbf{v}})^\top \mathbf{P} (\mathbf{v} - \bar{\mathbf{v}}),$$

where $\delta > 0$, $\mathbf{P} = \bar{\mathbf{V}}^{-1}$, $\mathbf{e}$ is defined as in (8), and $V_2$ is the same function as in the proof of Theorem IV.4. The Lie derivative of $V_3$ along (12) is given by

$$\mathcal{L}_{\psi_{\text{dist}}} V_3 = -\begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{e} \end{bmatrix}^\top \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{e} \end{bmatrix} + (\mathbf{x} - \mathbf{x}^*)^\top \mathbf{d}^{\mathbf{x}} + \mathbf{e}^\top \mathbf{d}^{\mathbf{e}} - \delta (\mathbf{v} - \bar{\mathbf{v}})^\top (\mathbf{L}^\top \mathbf{P} + \mathbf{P} \mathbf{L}) (\mathbf{v} - \bar{\mathbf{v}}),$$

where $\mathbf{d}^{\mathbf{e}} = -\alpha \mathbf{\Pi} \mathbf{A} \mathbf{L}(\mathbf{V} - \bar{\mathbf{V}}) \mathbf{x} - \gamma \mathbf{L}(\mathbf{V} - \bar{\mathbf{V}}) \mathbf{e}$, and $\mathbf{Q}_{22} = \frac{1}{2} \gamma (\mathbf{L} \bar{\mathbf{V}} + \bar{\mathbf{V}} \mathbf{L}^\top) + n\beta \mathbf{A} \mathbf{A}^\top$. Interestingly, $\mathsf{L} \bar{V}$ can be interpreted as the Laplacian of a weight-balanced graph and as a result, $\mathbf{L} \bar{\mathbf{V}} + \bar{\mathbf{V}} \mathbf{L}^\top \succeq \mathbf{0}$ implying that $\mathbf{L}^\top \mathbf{P} + \mathbf{P} \mathbf{L} \succeq \mathbf{0}$. Once again, following Lemma IV.3, one can establish that with the initialization of the statement, $\mathbb{1}^\top \mathbf{e} = \mathbf{0}$ and therefore using the Courant-Fischer theorem [19, Theorem 4.2.11] together with the fact that $(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{w} = 0$ due to invariance of $\mathbf{x}_{\text{null}}$, we can upper bound the Lie derivative as

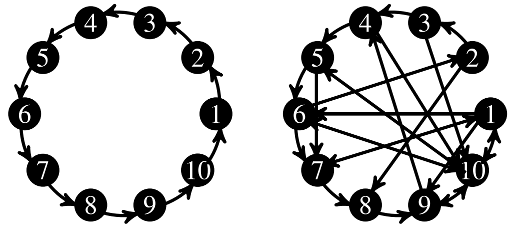
$$\begin{aligned} \mathcal{L}_{\psi_{\text{dist}}} V_3 \leq & -\begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{e} \end{bmatrix}^\top \underbrace{\begin{bmatrix} \lambda_2(\mathbf{Q}_{11}) I & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^\top & \bar{\mathbf{Q}}_{22} \end{bmatrix}}_{\bar{\mathbf{Q}}} \begin{bmatrix} \mathbf{x} - \mathbf{x}^* \\ \mathbf{e} \end{bmatrix} \\ & + \alpha \| \mathbf{x} - \mathbf{x}^* \| \| \mathbf{L} \| \| \mathbf{v} - \bar{\mathbf{v}} \| (\| \mathbf{x} - \mathbf{x}^* \| + \| \mathbf{x}^* \|) \\ & + \alpha \| \mathbf{e} \| \| \mathbf{\Pi} \mathbf{A} \mathbf{L} \| \| \mathbf{v} - \bar{\mathbf{v}} \| (\| \mathbf{x} - \mathbf{x}^* \| + \| \mathbf{x}^* \|) \\ & + \gamma \| \mathbf{e} \| \| \mathbf{L} \| \| \mathbf{v} - \bar{\mathbf{v}} \| \| \mathbf{e} \| - \delta \lambda_2 (\mathbf{L}^\top \mathbf{P} + \mathbf{P} \mathbf{L}) \| \mathbf{v} - \bar{\mathbf{v}} \|^2, \end{aligned}$$

where $\bar{\mathbf{Q}}_{22} = \frac{1}{2} \gamma \lambda_2 (\mathsf{L} \bar{V} + \bar{V} \mathsf{L}^\top) I + n\beta \mathbf{A} \mathbf{A}^\top$. Define $\mathbf{z} = [\| \mathbf{x} - \mathbf{x}^* \|; \| \mathbf{e} \|; \| \mathbf{v} - \bar{\mathbf{v}} \|]$. If $\gamma > \bar{\gamma}$, then $\bar{\mathbf{Q}} \succ \mathbf{0}$ and from the Courant-Fischer theorem, we have

$$\mathcal{L}_{\psi_{\text{dist}}} V_3 \leq -\mathbf{z}^\top \underbrace{\begin{bmatrix} \lambda_{\min}(\bar{\mathbf{Q}}) & 0 & \hat{\mathbf{Q}}_{13}(\mathbf{z}) \\ 0 & \lambda_{\min}(\bar{\mathbf{Q}}) & \hat{\mathbf{Q}}_{23}(\mathbf{z}) \\ \hat{\mathbf{Q}}_{13}(\mathbf{z}) & \hat{\mathbf{Q}}_{23}(\mathbf{z}) & \delta \lambda_2 (\mathbf{L}^\top \mathbf{P} + \mathbf{P} \mathbf{L}) \end{bmatrix}}_{\hat{\mathbf{Q}}(\mathbf{z})} \mathbf{z},$$

where $\hat{\mathbf{Q}}_{23}(\mathbf{z}) = -\frac{1}{2} \alpha \| \mathbf{\Pi} \mathbf{A} \mathbf{L} \| (\mathbf{z} + \| \mathbf{x}^* \|) - \frac{1}{2} \gamma \| \mathbf{L} \| \mathbf{z}$ and $\hat{\mathbf{Q}}_{13}(\mathbf{z}) = -\frac{1}{2} \alpha \| \mathbf{L} \| (\mathbf{z} + \| \mathbf{x}^* \|)$. Using the Schur complement, one can verify that for a given value of $\mathbf{z}$, $\hat{\mathbf{Q}}(\mathbf{z}) \succ \mathbf{0}$ iff $\delta > \bar{\delta}(\mathbf{z}) = \frac{1}{\lambda_{\min}(\bar{\mathbf{Q}}) \lambda_2 (\mathbf{L}^\top \mathbf{P} + \mathbf{P} \mathbf{L})} (\hat{\mathbf{Q}}_{13}(\mathbf{z})^2 + \hat{\mathbf{Q}}_{23}(\mathbf{z})^2)$. Hence, if $\delta > \bar{\delta}(\mathbf{z}(0))$, then $\mathcal{L}_{\psi_{\text{dist}}} V_3 \leq -\lambda_{\min}(\hat{\mathbf{Q}}(\mathbf{z}(0))) \mathbf{z}^\top \mathbf{z}$. This along with the fact that $\frac{1}{2} \min\{1, \delta \lambda_{\min}(\mathbf{P})\} \| \mathbf{z} \|^2 \leq V_3 \leq \frac{1}{2} \max\{1, \delta \lambda_{\max}(\mathbf{P})\} \| \mathbf{z} \|^2$, implies that $V_3$ satisfies the hypotheses of [21, Theorem 4.10] for exponential stability, completing the proof. ■

The exponential convergence of algorithms (5) and (7) for weight-balanced graphs, and (11) for unbalanced graphs follows from their linear nature. For algorithm (12), exponential convergence could be attributed to the fact that the dynamics (12c) converge exponentially and hence, after some time, (12a)-(12b) and (11) are essentially the same.



(a) $\mathcal{G}_1$ (b) $\mathcal{G}_2$

Fig. 1: Communication topologies among the agents. The edge weights are adjusted to make the graphs either weight-balanced or unbalanced, as needed.

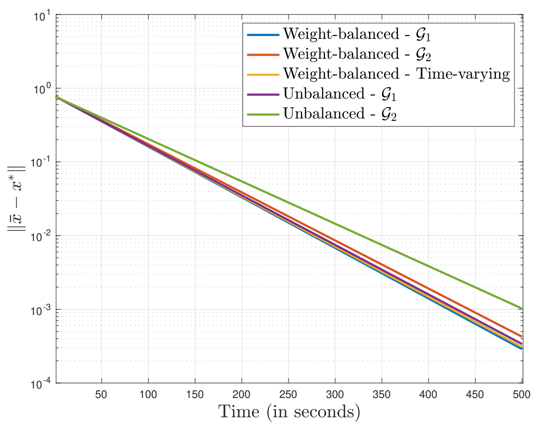



Fig. 2: Evolution of the error between the actual solution and the average state using the proposed algorithms from initial condition $\mathbf{x}(0) = \mathbf{0}$, $\mathbf{y}(0) = -\mathbf{b}$, over the graphs shown in Fig. 1. The algorithms are implemented in discrete time with a stepsize of $2.5 \times 10^{-3}$, and the values of $\alpha = 2$, $\beta = 0.1$ and $\gamma = 20$. Straight lines correspond to exponential convergence.

## VI. Simulations

We consider 10 agents communicating over the digraphs shown in Fig. 1, seeking to solve problem (1) with $\{A_i\}_{i=1}^{10} \in \mathbb{R}^{5\times5}$ and $\{b_i\}_{i=1}^{10} \in \mathbb{R}^5$. Since the proposed dynamics are in continuous time, we use a first-order Euler discretization with stepsize $2.5 \times 10^{-3}$ for the MATLAB implementation. The edge weights for various cases are adjusted to make the graphs weight-balanced and unbalanced, resp. For the time-varying case, at every iteration, the communication graph is switched randomly between $\mathcal{G}_1$ and $\mathcal{G}_2$. In Fig. 2, we plot the evolution of the error between the actual solution of (1) and the average state $\bar{x} = \frac{1}{n}\mathbb{1}^\top \mathbf{x}$ using (7), (10) and (12). The initial conditions for all the algorithms are chosen according to Remark 2. Even though $\mathcal{G}_2$ (with 4.6 as the minimum of the real parts of non-zero eigenvalues of $\mathsf{L}$ and $\lambda_2(\mathsf{L}+\mathsf{L}^\top) = 7.6$, for the weight-balanced case) is more connected than $\mathcal{G}_1$ (with 1.9 as the minimum of the real parts of non-zero eigenvalues of $\mathsf{L}$ and $\lambda_2(\mathsf{L}+\mathsf{L}^\top) = 3.8$, for the weight-balanced case), convergence is slower. The error in the time-varying case is lower and upper bounded by the error for $\mathcal{G}_1$ and $\mathcal{G}_2$, resp.

## VII. Conclusions and Future Work

We have presented continuous-time algorithms to solve linear algebraic equations whose problem data is represented as the summation of the data of individual agents. The proposed algorithms are distributed over general directed networks, do not require the individual agent matrices to be positive definite, and are guaranteed to converge to a solution of the linear equation exponentially fast. Future work will involve formally characterizing the convergence when the null space condition is not satisfied, and explore the design of distributed algorithms for finding least-square solutions when exact ones do not exist, extension to cases where the problem data is time-varying, and the communication graph is unbalanced and time-varying.

## References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

[2] S. Mou, J. Liu, and A. S. Morse, "A distributed algorithm for solving a linear algebraic equation," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2863–2878, 2015.

[3] B. D. O. Anderson, S. Mou, A. S. Morse, and U. Helmke, "Decentralized gradient algorithm for solution of a linear equation," *Numerical Algebra, Control and Optimization*, vol. 6, no. 3, pp. 319–328, 2016.

[4] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Distributed sensor fusion using dynamic consensus," in *IFAC World Congress*, Prague, CZ, July 2005, electronic proceedings.

[5] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Symposium on Information Processing of Sensor Networks*, Los Angeles, CA, Apr. 2005, pp. 63–70.

[6] 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*, vol. 5, no. 1, pp. 215–227, 2018.

[7] X. Wang and S. Mou, "A distributed algorithm for achieving the conservation principle," in *American Control Conference*, Milwaukee, WI, June 2018, pp. 5863–5867.

[8] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Dynamic consensus on mobile networks," in *IFAC World Congress*, Prague, Czech Republic, July 2005.

[9] S. S. Kia, B. V. Scoy, J. Cortés, R. A. Freeman, K. M. Lynch, and S. Martinez, "Tutorial on dynamic average consensus: The problem, its applications, and the algorithms," *IEEE Control Systems*, vol. 39, no. 3, pp. 40–72, 2019.

[10] B. Touri and B. Gharesifard, "Continuous-time distributed convex optimization on time-varying directed networks," in *IEEE Conf. on Decision and Control*, Osaka, Japan, 2015, pp. 724–729.

[11] ——, "A modified saddle-point dynamics for distributed convex optimization on general directed graphs," *IEEE Transactions on Automatic Control*, vol. 65, no. 7, pp. 3098–3103, 2020.

[12] P. Srivastava and J. Cortés, "Network optimization via smooth exact penalty functions enabled by distributed gradient computation," *IEEE Transactions on Control of Network Systems*, 2021, to appear.

[13] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.

[14] Z. Li and Z. Duan, *Cooperative Control of Multi-Agent Systems: A Consensus Region Approach*. CRC Press, 2015.

[15] A. Cherukuri and J. Cortés, "Distributed algorithms for convex network optimization under non-sparse equality constraints," in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, Sept. 2016, pp. 452–459.

[16] K. Arrow, L. Hurwitz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford, CA: Stanford University Press, 1958.

[17] A. Cherukuri, B. Gharesifard, and J. Cortés, "Saddle-point dynamics: conditions for asymptotic stability of saddle points," *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 486–511, 2017.

[18] B. Gharesifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2014.

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[21] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.