Distributed Asynchronous Algorithms for Solving Positive Definite Linear Equations over Dynamic Networks

Jie Lu
ACCESS Linnaeus Center
School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm 10044, Sweden
jielu@kth.se

Choon Yik Tang
School of Electrical and Computer Engineering
University of Oklahoma
Norman, OK 73019, USA
cytang@ou.edu

May 7, 2014

Abstract

This paper develops Subset Equalizing (SE), a distributed algorithm for solving a symmetric positive definite system of linear equations over a network of agents with arbitrary asynchronous interactions and membership dynamics, where each agent may join and leave the network at any time, for infinitely many times, and may lose all its memory upon leaving. To design and analyze SE, we introduce a time-varying Lyapunov-like function, defined on a state space with changing dimension, and a generalized concept of network connectivity, capable of handling such interactions and membership dynamics. Based on them, we establish the boundedness, asymptotic convergence, and exponential convergence of SE, along with a bound on its convergence rate. Finally, through extensive simulation, we demonstrate the effectiveness of SE in a volatile agent network and show that a special case of SE, termed Groupwise Equalizing, is significantly more bandwidth/energy efficient than two existing algorithms in multi-hop wireless networks.

1 Introduction

Solving a system of linear equations \( Pz = q \) is a fundamental problem with numerous applications spanning various areas of science and engineering. In this paper, we address the problem of decentralizedly solving such equations over a network of \( N \) agents, whereby each agent \( i \) observes a symmetric positive definite matrix \( P_i \in \mathbb{R}^{n \times n} \) and a vector \( q_i \in \mathbb{R}^n \), and all of them wish to find the unique solution \( z \in \mathbb{R}^n \) to

\[
\left( \sum_{i=1}^{N} P_i \right) z = \sum_{i=1}^{N} q_i. \tag{1}
\]

*This work was supported by the National Science Foundation under grant CMMI-0900806.
Since each agent $i$ knows only its own $P_i$ and $q_i$, none of them is able to individually solve (1). As a result, they must cooperatively do so, preferably without any centralized coordination.

The need to solve (1) arises in many applications of multi-agent systems, mobile ad hoc networks, and wireless sensor networks. For instance, suppose each wireless sensor $i$ in a sensor network takes a noisy measurement $y_i = A_i \theta + v_i$, where $A_1, A_2, \ldots, A_N$ are full column-rank matrices, $v_1, v_2, \ldots, v_N$ are jointly Gaussian random vectors with zero mean and covariance $V_1 \oplus V_2 \oplus \cdots \oplus V_N$ ($\oplus$ denotes the direct sum), and $A_i$ and $V_i$ are known only to sensor $i$. Then, finding the maximum-likelihood estimate of the unknown parameter $\theta$ is equivalent to solving (1) for $z$ with $P_i = A_i^T V_i^{-1} A_i$ and $q_i = A_i^T V_i^{-1} y_i$ for every $i$. As another example, suppose each mobile robot $i$ in a robotic network uses a quadratic function $(x - c_i)^T R_i (x - c_i)$ to model the risk it perceives (or energy it consumes) to meet everyone else at location $x$, where $R_1, R_2, \ldots, R_N$ are positive definite matrices, $c_1, c_2, \ldots, c_N$ are vectors, and $R_i$ and $c_i$ are known only to robot $i$. Then, the optimal rendezvous location, which minimizes the total perceived risk, can be determined by solving (1) for $z$ with $P_i = R_i$ and $q_i = R_i c_i$ for every $i$. Finally, the widely studied distributed averaging problem [3–23] that finds many applications is a notable special case of (1) with $n = 1$ and $P_i = 1$ for all $i$.

Given its broad applications, problem (1) has received considerable attention in the literature. Most of the studies, however, focus on the special case of distributed averaging, as is evident by the rich collection of continuous-time [5,11,16], discrete-time synchronous [4–7,12,14,16–22], and discrete-time asynchronous [3,8–10,13,15,18,23] algorithms that are available to date. Nonetheless, a few distributed algorithms devoted to the regular case of (1) with arbitrary $n$ and $P_i$’s have been proposed, including the continuous-time algorithm from [24], which computes the solution $z$ to (1) by exploiting the positive definiteness of the $P_i$’s, and the two discrete-time synchronous, average-consensus-based algorithms from [1,2], which do so by element-wise averaging the $P_i$’s and $q_i$’s.

In this paper, we consider the regular case of (1) and present a series of developments—from modeling to results—that differ substantially from those in [1,2,24], and that generalize some of those in [3,23] for distributed averaging. More specifically, we first introduce, in Section 2, a novel agent network model that can handle arbitrary asynchronous interactions and membership dynamics, so that agents may freely interact with one another, or spontaneously join and leave the network, at any time, for infinitely many times. Unlike existing models in [1,2,24] that require fixed agent memberships (i.e., graphs with fixed vertices), this model can handle dynamic ones, making it more general and allowing it to cope with practical situations, where agents must join or leave the network during runtime, either temporarily or permanently, voluntarily or involuntarily. We also formulate, in Section 2, a problem of designing a distributed asynchronous iterative algorithm that enables the agents to cooperatively and asymptotically solve (1), despite possibly having no control over their actions, essentially no knowledge about the network, very limited physical memories, and having to lose all their memories upon leaving.

We next borrow ideas from Lyapunov stability theory and optimization to construct, in Section 3, an algorithm named Subset Equalizing (SE) that exhibits the aforementioned properties.
More precisely, by defining a time-varying Lyapunov-like function that quantifies how far away the agents are from solving (1), and by repeatedly minimizing this function in hope of incrementally dropping its value to zero, we obtain SE, which is essentially a networked dynamical system that evolves by repeatedly equalizing different subsets of its state variables. Unlike the algorithms from [1,2,24] that require network-wide clock synchronization, SE is fully asynchronous. It also is quite flexible and general: for example, SE can be tailored to multi-hop wireless networks, leading to a gossip version called Pairwise Equalizing (PE) and a local broadcast version called Groupwise Equalizing (GE), which happen to generalize three existing distributed averaging schemes known as Pairwise Averaging [3], Randomized Gossip Algorithm [8], and Distributed Random Grouping [10].

To analyze SE, we subsequently develop, in Section 4, a few notions of network connectivity—including instantaneous connectivity, connectivity, and uniform connectivity—which are applicable to the agent network model (and which, together with the model, might be of interest in their own right). We also clarify these notions via examples and show that one of them reduces to a classic definition of connectivity [3] when the agent memberships are fixed. Building upon these notions, we then derive, in Section 5, sufficient conditions for establishing the boundedness, asymptotic convergence, and exponential convergence of SE, as well as a bound on its convergence rate. As a highlight of the results, we show that connectivity leads to asymptotic convergence, while uniform connectivity leads to exponential convergence.

As additional contributions of this paper, we demonstrate, through simulation in Section 6, that SE is effective in a volatile agent network, while GE is several times more bandwidth/energy efficient than PE and the two algorithms from [1,2] in multi-hop wireless networks. Finally, we state in Section 7 the conclusion of the paper and include in Appendix A the proofs of all assertions. Throughout the paper, we let \( N, \mathbb{N}, S^+_n, \) and \(|\cdot|\) denote, respectively, the sets of nonnegative integers, positive integers, \( n \times n \) symmetric positive definite matrices over \( \mathbb{R} \), and the cardinality of a set.

2 Network Modeling and Problem Formulation

Consider a nonempty, finite set of \( M \geq 2 \) agents, taking actions at each time \( k \in \mathbb{N} \) according to the following model:

A1. At time \( k = 0 \), a nonempty subset \( \mathcal{F} \) of the \( M \) agents form a network and become members of the network.

A2. Upon forming, each member \( i \in \mathcal{F} \) observes a matrix \( P_i \in S^+_n \) and a vector \( q_i \in \mathbb{R}^n \).

A3. The rest of the \( M \) agents become non-members of the network and make no observations.

A4. At each time \( k \in \mathbb{P} \), three disjoint subsets of the \( M \) agents—namely, a possibly empty subset \( \mathcal{J}(k) \) of the non-members, a nonempty subset \( \mathcal{I}(k) \) of the members, and a possibly empty, proper subset \( \mathcal{L}(k) \) of the members—take actions \( \text{A5\text{-}A7} \) below.

A5. The set \( \mathcal{J}(k) \) of non-members join the network and become members.
A6. Upon joining, the set \( J(k) \cup I(k) \cup L(k) \) of members interact, sharing information with one another and acknowledging their joining (i.e., \( J(k) \)), staying (i.e., \( I(k) \)), and leaving (i.e., \( L(k) \)).

A7. Upon interacting, the set \( L(k) \) of members leave the network and become non-members.

A8. The rest of the \( M \) agents (i.e., the complement of \( J(k) \cup I(k) \cup L(k) \)) take no actions.

Actions A1–A8 above define a general agent network model, where: (i) initially, an arbitrary subset of the agents form the network (i.e., A1) and make one-time observations (A2), but the rest of them do not (A3); (ii) at each subsequent time, arbitrary subsets of the agents (A4) spontaneously join the network (A5), interact with one another (A6), and leave the network (A7); and (iii) the agents take actions asynchronously (A8). With this model, \( M \) represents the maximum number of members the network may have, and each agent at any time is either a member or a non-member, but may change membership infinitely often. Labeling the \( M \) agents as \( 1, 2, \ldots, M \) and letting \( M(k) \subset \{1, 2, \ldots, M\} \) denote the set of members upon completing the actions at each time \( k \in \mathbb{N} \), the membership dynamics may be expressed as

\[
M(0) = F, \quad M(k) = (M(k-1) \cup J(k)) - L(k), \quad \forall k \in \mathbb{P},
\]

where, since \( F \neq \emptyset \) and \( L(k) \subsetneq M(k-1) \forall k \in \mathbb{P} \), the network always has at least one member, i.e., \( M(k) \neq \emptyset \forall k \in \mathbb{N} \). Moreover, since \( J(k) \) and \( L(k) \) may be empty for some \( k \in \mathbb{P} \) but \( I(k) \neq \emptyset \forall k \in \mathbb{P} \), while there may not always be membership changes, there are always member interactions, among the agents in

\[
J(k) \cup I(k) \cup L(k), \quad \forall k \in \mathbb{P}.
\]

Since the membership dynamics (2) and the member interactions (3) are completely characterized by the sets \( F, J(k), I(k), \) and \( L(k) \) \( \forall k \in \mathbb{P} \), the network is driven by a sequence \( A \) of agent actions given by

\[
A = (F, J(1), I(1), L(1), J(2), I(2), L(2), \ldots).
\]

Remark 1. Although it is common to model networks using graphs, we use the sets \( F, J(k), I(k), \) and \( L(k) \) \( \forall k \in \mathbb{P} \) to model the above agent network because they enable convenient handling of the membership dynamics. We note that in the absence of membership changes (i.e., \( J(k) = L(k) = \emptyset \forall k \in \mathbb{P} \))—which is the de facto assumption in the literature—specifying \( F \) and \( I(k) \) \( \forall k \in \mathbb{P} \) or \( A \) in (1) is the same as specifying an interaction graph.

Given the agent network modeled by A1–A8 the objective of this paper is to design and analyze a distributed asynchronous algorithm of iterative nature, which allows the ever-changing members of the network to cooperatively and asymptotically compute the constant solution \( z \in \mathbb{R}^n \).
of the following symmetric positive definite system of linear equations, defined by the one-time observations $P_i$ and $q_i \forall i \in \mathcal{M}(0)$ of the initial members:

$$
\left( \sum_{i \in \mathcal{M}(0)} P_i \right) z = \sum_{i \in \mathcal{M}(0)} q_i. \tag{5}
$$

The algorithm should also exhibit the following desirable properties:

P1. It should allow the sequence $\mathcal{A}$ of agent actions to be dictated by an exogenous source, for which the agents have no control over, since, for example, in a sensor network, $J(k)$, $I(k)$, and $L(k)$ may be governed by sensor redeployment, reseeding, mobility, failures, and recoveries, all of which may be forced exogenously.

P2. It should allow the agents to not know the values of $M$, $k$, $F$, $J(k)$, $I(k)$, and $L(k)$ $\forall k \in \mathcal{P}$, since in many practical situations they are not available, or at least not known ahead of time.

P3. It should not impose large memory requirements on the agents, and should allow them to lose all their memories upon leaving the network, since the departure may be caused by, for instance, software or hardware failures.

Note that solving (5) is equivalent to solving an unconstrained quadratic optimization problem of the form $\min_{y \in \mathbb{R}^n} \frac{1}{2} y^T (\sum_{i \in \mathcal{M}(0)} P_i) y - y^T \sum_{i \in \mathcal{M}(0)} q_i$. Moreover, due to property P1 and the fact that $\mathcal{A}$ in (4) dictates all but how members share and process information whenever they interact in $\mathcal{A}6$ designing such an algorithm amounts to defining what information to share, and how to process it, during each iteration. Furthermore, due to property P3 sharing of information via simple but memory-intensive flooding of the $P_i$’s and $q_i$’s is prohibited (flooding, however, will serve as a benchmark for performance comparison in Section 6.2).

3 Subset Equalizing

In this section, using ideas from Lyapunov stability theory and optimization, we construct an algorithm that possesses properties P1–P3 and strives to solve (5).

Consider a networked dynamical system formed by the $M$ agents, in which each agent $i \in \{1, 2, \ldots, M\}$ maintains in its memory two state variables $z_i \in \mathbb{R}^n \cup \{\#\}$ and $Q_i \in \mathbb{S}^+_n \cup \{\#\}$, where $z_i$ represents its estimate of the unknown solution $z$ of (5), $Q_i$ plays the part of helping $z_i$ approach $z$, and the symbol $\#$ means undefined. To describe the system dynamics, let $z_i(k)$ and $Q_i(k)$ be the values of $z_i$ and $Q_i$ upon completing the actions at each time $k \in \mathbb{N}$. In addition, let $z_i(k) \in \mathbb{R}^n$ and $Q_i(k) \in \mathbb{S}^+_n$ if $i \in \mathcal{M}(k)$, and $z_i(k)$ and $Q_i(k)$ be undefined otherwise, i.e.,

$$
z_i(k) = \#, \quad \forall k \in \mathbb{N}, \forall i \in \{1, 2, \ldots, M\} - \mathcal{M}(k), \tag{6}
$$

$$
Q_i(k) = \#, \quad \forall k \in \mathbb{N}, \forall i \in \{1, 2, \ldots, M\} - \mathcal{M}(k). \tag{7}
$$
It follows that for each \( i \in \{1, 2, \ldots, M\} \) and \( k \in \mathbb{P} \), the value of \( z_i \) (and, likewise, that of \( Q_i \)) changes from \( z_i(k-1) = \# \) to \( z_i(k) \in \mathbb{R}^n \) if agent \( i \) joins the network at time \( k \), changes from \( z_i(k-1) \in \mathbb{R}^n \) to \( z_i(k) = \# \) if it leaves, remains in \( \mathbb{R}^n \) if it remains a member, and remains equal to \( \# \) if it remains a non-member. Since being in \( \mathbb{R} \) is, to fully specify the system dynamics, it suffices to only specify what the initial \( Q \) \( \forall i \in \mathcal{M}(0) \) are, and how the new \( z_i(k) \) and \( Q_i(k) \forall i \in \mathcal{M}(k) \) depend on the old \( z_i(k-1) \) and \( Q_i(k-1) \forall i \in \mathcal{M}(k-1) \) for every \( k \in \mathbb{P} \).

To this end, consider a time-varying Lyapunov-like function \( V \) of the state variables \( z_i(k)'s \) and \( Q_i(k)'s \), defined for each \( k \in \mathbb{N} \) as

\[
V(k, z_1(k), z_2(k), \ldots, z_M(k), Q_1(k), Q_2(k), \ldots, Q_M(k)) = \sum_{i \in \mathcal{M}(k)} (z_i(k) - z)^T Q_i(k)(z_i(k) - z). \tag{8}
\]

Note that, as the left-hand side of (S) is lengthy, we will write it as \( V(k) \) in the sequel, omitting all but the first of its arguments for brevity. Also, as the right-hand side of (S) is a sum over \( i \in \mathcal{M}(k) \), which excludes all the non-members \( i \in \{1, 2, \ldots, M\} - \mathcal{M}(k) \) and their \( z_i(k) = \# \) and \( Q_i(k) = \# \) (see (10) and (11)), \( V(k) \) is always in \( \mathbb{R} \) and well-defined. Furthermore, as the sum involves a time-varying subset of the \( z_i(k)'s \) and \( Q_i(k)'s \), \( V(k) \) is akin to a function defined on a state space with growing and shrinking dimension. Finally, although not a standard Lyapunov function candidate, \( V(k) \) exhibits some similar features that make it useful for the problem at hand: as \( Q_i(k) \in \mathbb{S}^n_+ \) \( \forall k \in \mathbb{N} \) \( \forall i \in \mathcal{M}(k) \), we have \( V(k) \geq 0 \) \( \forall k \in \mathbb{N} \), with \( V(k) = 0 \) if and only if \( z_i(k) = z \forall i \in \mathcal{M}(k) \), i.e., (S) is exactly solved. However, there is a caveat, in that \( \lim_{k \to \infty} V(k) = 0 \) does not necessarily imply \( \lim_{k \to \infty} z_i(k) = z \forall i \in \mathcal{M}(k) \) because the \( Q_i(k)'s \) may be “losing” their positive definiteness as \( k \to \infty \). In fact, one should be careful with the statement “\( \lim_{k \to \infty} z_i(k) = z \forall i \in \mathcal{M}(k) \)” because \( z_i(k) \) may be alternating between \( z_i(k) \in \mathbb{R}^n \) and \( z_i(k) = \# \). Nevertheless, if there exists a constant \( \alpha > 0 \) such that \( Q_i(k) - \alpha I \in \mathbb{S}^n_+ \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k) \), then \( \lim_{k \to \infty} V(k) = 0 \) does imply that the \( z_i(k)'s \) of the members asymptotically converge to \( z \) as \( k \to \infty \). Hence, \( V(k) \) in (S) possesses certain key features of a standard Lyapunov function candidate, explaining why we define it as such and call it a Lyapunov-like function.

Having introduced \( V(k) \), we now use it to devise the system dynamics, expressing the new \( z_i(k) \) and \( Q_i(k) \forall i \in \mathcal{M}(k) \) in terms of the old \( z_i(k-1) \) and \( Q_i(k-1) \forall i \in \mathcal{M}(k-1) \) for every \( k \in \mathbb{P} \). To begin, observe from (14) and (15) that for each \( k \in \mathbb{P} \), the members in \( \mathcal{M}(k) \) can be partitioned into those in \( \mathcal{M}(k) - (\mathcal{J}(k) \cup \mathcal{I}(k)) \) who take no actions at time \( k \), and those in \( \mathcal{J}(k) \cup \mathcal{I}(k) \) who interact with the leaving members in \( \mathcal{L}(k) \). For those in \( \mathcal{M}(k) - (\mathcal{J}(k) \cup \mathcal{I}(k)) \), because none of them gain any new information, it is reasonable to let their \( z_i \)'s and \( Q_i \)'s be unchanged, i.e.,

\[
\begin{align*}
z_i(k) &= z_i(k-1), \quad \forall i \in \mathcal{M}(k) - (\mathcal{J}(k) \cup \mathcal{I}(k)), \\
Q_i(k) &= Q_i(k-1), \quad \forall i \in \mathcal{M}(k) - (\mathcal{J}(k) \cup \mathcal{I}(k)).
\end{align*} \tag{9}
\]

As for those in \( \mathcal{J}(k) \cup \mathcal{I}(k) \), because they get to share information among themselves and with
those in $\mathcal{L}(k)$, they get to jointly determine
\begin{equation}
    z_i(k) \text{ and } Q_i(k), \quad \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k),
\end{equation}
based on
\begin{equation}
    z_i(k-1) \text{ and } Q_i(k-1), \quad \forall i \in \mathcal{I}(k) \cup \mathcal{L}(k).
\end{equation}
To enable such determination, notice from (8), (2), (9), and (10) that the change in the value of $V$ can be written as
\begin{equation}
    V(k) - V(k-1) = \left[ \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} z_i(k)^T Q_i(k) z_i(k) \right] - \left[ \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} z_i(k-1)^T Q_i(k-1) z_i(k-1) \right] \\
    - 2z^T \left[ \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} Q_i(k) z_i(k) - \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} Q_i(k-1) z_i(k-1) \right] + z^T \left[ \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} Q_i(k) - \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} Q_i(k-1) \right] z.
\end{equation}
Also note that the first bracket in (13) contains only the to-be-determined variables in (11), the second bracket contains only the given variables in (12), and the unknown $z$ appears only right by the third and fourth brackets, each of which contains both the to-be-determined and given variables in the form of a subtraction. Therefore, by having the members in $\mathcal{J}(k) \cup \mathcal{I}(k)$ choose the to-be-determined variables so that the third and fourth brackets disappear, i.e.,
\begin{equation}
    \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} Q_i(k-1) z_i(k-1) = \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} Q_i(k) z_i(k),
\end{equation}
\begin{equation}
    \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} Q_i(k-1) = \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} Q_i(k),
\end{equation}
the change $V(k) - V(k-1)$ in (13) would be unaffected by the unknown $z$ and, thus, would be known to those members. In addition, as the second bracket is fixed, by having those members use the remaining freedom in the to-be-determined variables to jointly minimize the first bracket, i.e.,
\begin{equation}
    \minimize_{(z_i(k), Q_i(k)) \in \mathcal{J}(k) \cup \mathcal{I}(k)} \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} z_i(k)^T Q_i(k) z_i(k)
\begin{array}{c}
\text{subject to} \\
\text{(14) and (15),}
\end{array}
\end{equation}
the change $V(k) - V(k-1)$ in (13) would be minimized, perhaps even made negative, i.e., $V(k) < V(k-1)$. Since $V(k-1)$ is also fixed, solving (16) is equivalent to robustly minimizing $V(k)$ without letting the uncertain $z$ influence the outcome, and without requiring any agent to ever know the value of $V(k)$. Lastly, since (16) is solved at each time $k \in \mathbb{P}$ by a generally different subset of the $M$ agents with the rest of them being idle, this paragraph describes a Lyapunov-based optimization approach to solving (14), whereby the agents repeatedly and asynchronously minimize
the Lyapunov-like function $V(k)$, hoping that such minimization would incrementally decrease its value to zero and ultimately drive the $z_i(k)$'s of the members to the unknown solution $z$ of (5)

The following lemma shows that the optimization problem (16) admits a nonempty, convex set of solutions with appealing properties:

**Lemma 1.** For any $\mathcal{A}$ and any $k \in \mathbb{P}$, $(z_i(k), Q_i(k))_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)}$ is an optimal solution to problem (16) if and only if $Q_i(k)$ for all $i \in \mathcal{J}(k) \cup \mathcal{I}(k)$ satisfy (15) and

$$z_i(k) = \left( \sum_{j \in \mathcal{J}(k) \cup \mathcal{L}(k)} Q_j(k-1) \right)^{-1} \sum_{j \in \mathcal{I}(k) \cup \mathcal{L}(k)} Q_j(k-1) z_j(k-1), \quad \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k).$$

(17)

Moreover, if (9), (10), (14), and (15) hold, then $V(k) \leq V(k-1)$, where the equality holds if and only if $z_i(k-1) \forall i \in \mathcal{I}(k) \cup \mathcal{L}(k)$ are equal.

**Proof.** See Appendix A.1.

Lemma 1 says that the optimal action, which minimizes $V(k)$ under conditions (14) and (15), is an equalizing action, whereby the $z_i(k)$'s of the members in $\mathcal{J}(k) \cup \mathcal{I}(k)$ are set equal to the same value given by the right-hand side of (17). Indeed, this equalizing action (17), along with (15), enables the agents in $\mathcal{J}(k) \cup \mathcal{I}(k) \cup \mathcal{L}(k)$ to jointly make the value of $V$ decrease, unless the $z_i(k-1)$'s of those in $\mathcal{I}(k) \cup \mathcal{L}(k)$ are identical, in which case the value of $V$ is unchanged. Moreover, although $z_i(k) \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k)$ are uniquely determined by (17), there are infinitely many ways for $Q_i(k)$ for all $i \in \mathcal{J}(k) \cup \mathcal{I}(k)$ to satisfy (15). For simplicity, we adopt the following way to determine $Q_i(k)$ for all $i \in \mathcal{J}(k) \cup \mathcal{I}(k)$ so that (15) holds: when there are no membership changes, i.e., $\mathcal{J}(k) = \mathcal{L}(k) = \emptyset$, the members in $\mathcal{J}(k) \cup \mathcal{I}(k)$ do not update their $Q_i(k)$'s, i.e.,

$$Q_i(k) = Q_i(k-1), \quad \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k),$$

(18)

whereas when there are membership changes, i.e., $\mathcal{J}(k) \cup \mathcal{L}(k) \neq \emptyset$, their $Q_i(k)$'s are set equal to the same value while satisfying (15), i.e.,

$$Q_i(k) = \frac{1}{|\mathcal{J}(k) \cup \mathcal{I}(k)|} \sum_{j \in \mathcal{J}(k) \cup \mathcal{L}(k)} Q_j(k-1), \quad \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k).$$

(19)

Having specified the evolution of the state variables $z_i(k)$'s and $Q_i(k)$'s, we next define the initial states $z_i(0)$ and $Q_i(0)$ for all $i \in \mathcal{M}(0)$. Notice that problem (5) is solved only if $z_i(k) \forall i \in \mathcal{M}(k)$ asymptotically reach a consensus. Also note from (9), (10), (14), and (15) that

$$\sum_{i \in \mathcal{M}(k)} Q_i(k) z_i(0) = \sum_{i \in \mathcal{M}(k)} Q_i(0) z_i(0), \quad \forall k \in \mathbb{N},$$

(20)

$$\sum_{i \in \mathcal{M}(k)} Q_i(k) = \sum_{i \in \mathcal{M}(k)} Q_i(0), \quad \forall k \in \mathbb{N}.$$
Hence, the consensus, if achieved, is the solution \( z \) of problem (5) if

\[
\sum_{i \in \mathcal{M}(0)} Q_i(0) z_i(0) = \sum_{i \in \mathcal{M}(0)} q_i, \quad (22)
\]

\[
\sum_{i \in \mathcal{M}(0)} Q_i(0) = \sum_{i \in \mathcal{M}(0)} P_i. \quad (23)
\]

To satisfy (22) and (23), it suffices to set the initial states as follows:

\[
z_i(0) = P_i^{-1} q_i, \quad \forall i \in \mathcal{M}(0), \quad (24)
\]

\[
Q_i(0) = P_i, \quad \forall i \in \mathcal{M}(0). \quad (25)
\]

Observe that (24) and (25) are not the only way to guarantee (22) and (23), but they allow each initial member \( i \in \mathcal{M}(0) \) to initialize its \( z_i(0) \) and \( Q_i(0) \) using only its own observations \( P_i \) and \( q_i \).

Expressions (24), (25), (6), (7), (9), (10), (17), (18), and (19) collectively define a distributed asynchronous iterative algorithm. Since at each time \( k \in \mathbb{P} \), this algorithm involves an equalizing action taken by a subset \( \mathcal{J}(k) \cup \mathcal{I}(k) \cup \mathcal{L}(k) \) of the agents, we refer to the algorithm as Subset Equalizing (SE). A complete description of SE is as follows:

**Algorithm 1** (Subset Equalizing).

**Initialization:** At time \( k = 0 \):

1. Each agent \( i \in \{1, 2, \ldots, M\} \) creates variables \( z_i \in \mathbb{R}^n \cup \{\#\} \) and \( Q_i \in \mathbb{S}_+^n \cup \{\#\} \) and initializes them according to

\[
z_i(0) = \begin{cases} P_i^{-1} q_i, & \text{if } i \in \mathcal{M}(0), \\ \# & \text{otherwise}, \end{cases} \quad (26)
\]

\[
Q_i(0) = \begin{cases} P_i, & \text{if } i \in \mathcal{M}(0), \\ \# & \text{otherwise}. \end{cases} \quad (27)
\]

**Operation:** At each time \( k \in \mathbb{P} \):

2. Each agent \( i \in \{1, 2, \ldots, M\} \) updates \( z_i(k) \) according to

\[
z_i(k) = \begin{cases} \left( \sum_{j \in \mathcal{I}(k)} Q_j(k-1) \right)^{-1} \sum_{j \in \mathcal{I}(k)} Q_j(k-1) z_j(k-1), & \text{if } i \in \mathcal{J}(k) \cup \mathcal{I}(k), \\ \# & \text{if } i \in \mathcal{L}(k), \\ z_i(k-1), & \text{otherwise}. \end{cases} \quad (28)
\]

3. If \( \mathcal{J}(k) = \mathcal{L}(k) = \emptyset \), then each agent \( i \in \{1, 2, \ldots, M\} \) updates \( Q_i(k) \) according to

\[
Q_i(k) = Q_i(k-1), \quad \forall i \in \{1, 2, \ldots, M\}. \quad (29)
\]
Otherwise, each agent $i \in \{1, 2, \ldots, M\}$ updates $Q_i(k)$ according to
\[
Q_i(k) = \begin{cases} 
\frac{1}{|J(k) \cup L(k)|} \sum_{j \in J(k) \cup L(k)} Q_j(k-1), & \text{if } i \in J(k) \cup I(k), \\
\# , & \text{if } i \in L(k), \\
Q_i(k-1), & \text{otherwise.}
\end{cases}
\] (30)

Thus far in the paper, we focus mostly on the mathematical aspects of the agent network model and the algorithm SE, saying little about their communication and computation aspects. In the following two examples, we discuss such aspects in the context of multi-hop wireless networks and point out the relationship between SE and a few distributed averaging algorithms:

**Example 1.** Consider a multi-hop wireless network modeled as a connected, undirected graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \ldots, N\}$ represents the set of $N \geq 2$ nodes forming the network, and $\mathcal{E} \subseteq \{(i, j) : i, j \in \mathcal{V}, i \neq j\}$ represents the set of bidirectional links connecting the nodes. Suppose each node $i \in \mathcal{V}$ observes a matrix $P_i \in \mathbb{S}_+^n$ and a vector $q_i \in \mathbb{R}^n$, and the goal is for all of them to solve (1) for $z \in \mathbb{R}^n$ through gossiping, i.e., through every node $i \in \mathcal{V}$ gossiping with a neighbor $j \in N_i = \{j \in \mathcal{V} : \{i, j\} \in \mathcal{E}\}$ from time to time. To accomplish this goal, note that such a network is a special case of the agent network [8] obtained by letting $M = N$, $F = \mathcal{V}$, $J(k) \equiv \emptyset$, $L(k) \equiv \emptyset$, and $I(k) \in \mathcal{E}$ be the gossiping pair of nodes. Thus, the goal can be accomplished as follows using SE, which in this special case may be referred to as *Pairwise Equalizing* (PE):

**Algorithm 2 (Pairwise Equalizing).**

**Initialization:**
1. Each node $i \in \mathcal{V}$ transmits $P_i$ to every node $j \in N_i$.
2. Each node $i \in \mathcal{V}$ creates a variable $z_i \in \mathbb{R}^n$ and initializes it: $z_i \leftarrow P_i^{-1}q_i$.

**Operation:** At each iteration:
3. A node, say, node $i \in \mathcal{V}$, initiates the iteration and selects a neighbor, say, node $j \in N_i$, to gossip.
4. Node $i$ transmits $z_i$ to node $j$.
5. Node $j$ updates $z_j$: $z_j \leftarrow (P_i + P_j)^{-1}(P_i z_i + P_j z_j)$.
6. Node $j$ transmits $z_j$ to node $i$.
7. Node $i$ updates $z_i$: $z_i \leftarrow z_j$.

Notice that: (i) Step 1 of PE enables each node $j \in \mathcal{V}$ to learn, once and for all, the $P_i$ of every neighbor $i \in N_j$ so that it can readily perform Step 5; (ii) since the $P_i$’s are symmetric and the communications are wireless, $\frac{n(n+1)}{2}N$ real-number transmissions are needed to execute Step 1; (iii) in Step 2, the nodes need not create the variables $Q_i$’s because they are constant and equal to the $P_i$’s; (iv) Step 3 can be implemented by having each node periodically initiate an iteration and cyclically select a neighbor to gossip, or randomly do so; (v) Steps 4–7 carry out the pairwise
equalizing action, through which $z_i$ and $z_j$ are equalized; (vi) PE requires one-time inversion of positive definite matrices, in Steps 2 and 5; (vii) PE requires $2n$ real-number transmissions per iteration, in Steps 4 and 6; and (viii) just like SE is a generalization of PE, PE is a generalization of Pairwise Averaging [8] and Randomized Gossip Algorithm [9], reducing to these distributed averaging algorithms when $n = 1$ and $P_i = 1 \forall i \in \mathcal{V}$.

Example 2. Although simple, PE may have slow convergence because at each iteration, only two of the $N$ $z_i$’s are updated. Conceivably, allowing more $z_i$’s to be updated at once may speed up convergence. It turns out that SE is flexible enough to allow that. To see this, reconsider the wireless network in Example 1 and suppose the goal is instead to solve (1) via groupwise interactions, i.e., via every node $i \in \mathcal{V}$ interacting with all its neighbors in $\mathcal{N}_i$ as a group from time to time. As in Example 1, by letting $I(k) \in \{\{i\} \cup \mathcal{N}_i : i \in \mathcal{V}\}$ be the interacting group of nodes, this goal can be achieved as follows using SE, which in this case may be termed Groupwise Equalizing (GE):

Algorithm 3 (Groupwise Equalizing).

Initialization:
1. Each node $i \in \mathcal{V}$ transmits $P_i$ to every node $j \in \mathcal{N}_i$.
2. Each node $i \in \mathcal{V}$ creates a variable $z_i \in \mathbb{R}^n$ and initializes it: $z_i \leftarrow P_i^{-1} q_i$.

Operation: At each iteration:
3. A node, say, node $i \in \mathcal{V}$, initiates the iteration and transmits a message to every node $j \in \mathcal{N}_i$, requesting their $z_j$’s.
4. Each node $j \in \mathcal{N}_i$ transmits $z_j$ to node $i$.
5. Node $i$ updates $z_i$: $z_i \leftarrow (\sum_{j \in \{i\} \cup \mathcal{N}_i} P_j)^{-1} \sum_{j \in \{i\} \cup \mathcal{N}_i} P_j z_j$.
6. Node $i$ transmits $z_i$ to every node $j \in \mathcal{N}_i$.
7. Each node $j \in \mathcal{N}_i$ updates $z_j$: $z_j \leftarrow z_i$.

Observe that: (i) Steps 1 and 2 of GE are identical to those of PE; (ii) similar to PE, Step 3 of GE can be implemented deterministically or randomly; (iii) Steps 4–7 perform the groupwise equalizing action, through which $z_i$ and $z_j \forall j \in \mathcal{N}_i$ are equalized; (iv) like PE, GE also involves matrix inversion, in Steps 2 and 5; (v) unlike PE, GE requires $n(|\mathcal{N}_i| + 1)$ real-number transmissions per iteration initiated by node $i$, in Steps 4 and 6; and (vi) when $n = 1$ and $P_i = 1 \forall i \in \mathcal{V}$, GE simplifies to Distributed Random Grouping [10], so that SE also includes this distributed averaging algorithm from [10] as a special case.

4 Network Connectivity

With SE, every time a subset of the $M$ agents interact, they update their state variables $z_i(k)$’s and $Q_i(k)$’s in such a manner that the value of the Lyapunov-like function $V(k)$ is non-increasing. While this ensures that $V(k)$ must converge, it does not guarantee that $V(k)$ would converge to zero, which is desired. In fact, it is not difficult to think of a sequence $\mathcal{A}$ of agent actions, with
which $V(k)$ is bounded away from zero by a positive constant—just imagine an agent network with no membership changes, in which two groups of members never interact with each other, i.e., are never “connected.” This suggests that, for SE to drive $V(k)$ to zero, the agent network modeled by $\text{A1 A2 A3}$ must be connected in some sense. In this section, we develop a few notions of connectivity, which—unlike those in basic graph theory—are applicable to such a network.

To begin, consider a hypothetical scenario, in which $k \in \mathbb{N}$ denotes the initial time and $\ell \geq k$ the actual time. Before the initial time $\ell = k$, each agent $i \in \{1, 2, \ldots, M\}$ has an empty memory. At time $\ell = k$, each member $i \in \mathcal{M}(\ell)$ creates a message called message $i$. At each subsequent time $\ell \geq k + 1$, besides action $\text{A5}$, each joining member $i \in \mathcal{J}(\ell)$ creates a message called message $i$ (if it has never been created) or recreates message $i$ (if it has been destroyed). Upon joining, through action $\text{A6}$, all interacting members in $\mathcal{J}(\ell) \cup \mathcal{I}(\ell) \cup \mathcal{L}(\ell)$ learn from one another the messages they have gathered so far. Upon interacting, besides action $\text{A7}$, each leaving member $i \in \mathcal{L}(\ell)$ empties its memory and asks all the staying members in $\mathcal{M}(\ell)$ to erase message $i$ from their memories, effectively destroying message $i$. This process is then repeated indefinitely for every $\ell \geq k + 1$.

For the hypothetical scenario stated above, an intriguing question that arises is: with messages being created, shared, and destroyed as agents join the network, interact, and leave, what would be an appropriate definition of connectivity? To answer this question, recall that an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is connected if every pair of vertices in $\mathcal{V}$ is connected by a path of edges in $\mathcal{E}$. In other words, it is not possible to partition $\mathcal{V}$ into two nonempty subsets $\mathcal{V}_1$ and $\mathcal{V}_2$ and have no paths connecting the vertices in $\mathcal{V}_1$ with those in $\mathcal{V}_2$. Motivated by this simple observation, we say that the agent network is disconnected under $\mathcal{A}$ at time $k \in \mathbb{N}$ if $\mathcal{A}$ in (11) is such that for every $\ell \geq k$, $\mathcal{M}(\ell)$ can be partitioned into two nonempty subsets $\mathcal{M}_1(\ell)$ and $\mathcal{M}_2(\ell)$, such that all the members in $\mathcal{M}_1(\ell)$ are unaware of any messages created by those in $\mathcal{M}_2(\ell)$, and vice versa. In this definition, the phrase “under $\mathcal{A}$ at time $k \in \mathbb{N}$” is needed because the statement may be true for some $\mathcal{A}$ and $k$, and false for others. Likewise, the quantifier “for every $\ell \geq k$” is added so that being disconnected means there are always two groups of messages, which are separable.

Although it is mathematically precise, the above definition may not be readily useful in analysis because checking whether $\mathcal{M}(\ell)$ can be so partitioned for infinitely many $\ell$’s may be cumbersome. Also, if the network is not disconnected (i.e., is connected), the definition says nothing about how well-connected it is. To overcome these two limitations, let us associate with each initial time $k \in \mathbb{N}$, each subsequent time $\ell \geq k$, and each agent $i \in \{1, 2, \ldots, M\}$ a set $C_i(k, \ell) \subset \mathcal{M}(\ell)$ which, roughly speaking, keeps track of the subset of members that cannot be partitioned without message crossovers. More precisely, for each $k \in \mathbb{N}$, let $C_i(k, \ell) \forall i \in \{1, 2, \ldots, M\}$ be initialized at $\ell = k$ to

$$C_i(k, k) = \begin{cases} \{i\}, & \text{if } i \in \mathcal{M}(k), \\ \emptyset, & \text{otherwise}, \end{cases} \quad (31)$$
and defined recursively for each \( \ell \geq k + 1 \) as

\[
C_i(k, \ell) = \begin{cases} 
(\bigcup_{j \in I(\ell)} C_j(k, \ell - 1) \cup J(\ell)) - \mathcal{L}(\ell), & \text{if } i \in (\bigcup_{j \in I(\ell)} C_j(k, \ell - 1) \cup J(\ell)) - \mathcal{L}(\ell), \\
\emptyset, & \text{if } i \in \mathcal{L}(\ell), \\
C_i(k, \ell - 1), & \text{otherwise.}
\end{cases}
\]

Then, by induction on \( \ell \) using (31) and (32), we see that: (i) \( \forall k \in \mathbb{N}, \forall \ell \geq k, \text{ and } \forall i \in \{1, 2, \ldots, M\} \), if \( i \in \mathcal{M}(\ell) \) then \( i \in C_i(k, \ell) \subset \mathcal{M}(\ell) \), otherwise \( C_i(k, \ell) = \emptyset \); (ii) \( \forall k \in \mathbb{N}, \forall \ell \geq k, \text{ and } \forall i, j \in \mathcal{M}(\ell) \), either \( C_i(k, \ell) \cap C_j(k, \ell) = \emptyset \) or \( C_i(k, \ell) = C_j(k, \ell) \); and (iii) \( \forall k \in \mathbb{N}, \forall \ell \geq k, \text{ and } \forall i \in \mathcal{M}(\ell) \), \( C_i(k, \ell) \) is the largest subset of \( \mathcal{M}(\ell) \) containing agent \( i \) that cannot be partitioned into two nonempty subsets, such that all the members in one are unaware of any messages from those in the other. It follows from (i)–(iii) that the agent network is connected under \( \mathcal{A} \) at time \( k \in \mathbb{N} \) if and only if \( \mathcal{A} \) in (4) is such that there exists \( \ell' \geq k \) with \( \ell' < \infty \), such that \( C_i(k, \ell') = \mathcal{M}(\ell') \forall i \in \mathcal{M}(\ell') \) (note that if such an \( \ell' \) exists, then \( C_i(k, \ell) = \mathcal{M}(\ell) \forall i \in \mathcal{M}(\ell) \forall \ell > \ell' \)). This necessary and sufficient condition is more useful in analysis than the original definition (i.e., checking whether \( \mathcal{M}(\ell) \) can be partitioned) because it leverages (31) and (32) and eliminates the need to record what messages are known to which agents at what times, which is rather cumbersome. Additionally, if the network is connected at time \( k \), the smallest such \( \ell' \), denoted as \( \ell^* \), is a measure of how well-connected it is because \( \ell^* - k \) represents the number of time instants required for the messages to become inseparable. Thus, this condition bypasses the two aforementioned limitations.

Observe that for a given \( \mathcal{A} \), the network may be disconnected at certain times, and connected at others, during which it may require different number of time instants (i.e., \( \ell^* - k \)) for the messages to become inseparable (note that \( \ell^* \) depends on \( k \)). To reflect these different levels of connectedness, let us introduce a function \( h : \mathbb{N} \rightarrow \mathbb{N} \cup \{\infty\} \) and a constant \( h^* \in \mathbb{N} \cup \{\infty\} \), defined as

\[
h(k) = \inf_{k \in \mathbb{N}} D_k - k, \quad \forall k \in \mathbb{N}, \quad (33)
\]

\[
h^* = \sup_{k \in \mathbb{N}} h(k), \quad (34)
\]

where the set \( D_k \subset \{k, k + 1, \ldots\} \) is given by

\[
D_k = \{\ell \geq k : C_i(k, \ell) = \mathcal{M}(\ell) \forall i \in \mathcal{M}(\ell)\}, \quad \forall k \in \mathbb{N}. \quad (35)
\]

With (33)–(35), we have \( h(k) = \ell^* - k \) if the network is connected at time \( k \) (due to definition of \( \ell^* \)), \( h(k) = \infty \) otherwise (due to \( D_k = \emptyset \) and \( \inf \emptyset = \infty \)), and \( h^* < \infty \) if and only if \( h \) is bounded. Hence, the smaller \( h(k) \) and \( h^* \), the better the “instantaneous” and “worst-case” connectedness, respectively. Putting all of the above together, we arrive at the following formal definition:

**Definition 1.** The agent network modeled by [A1]–[A8] is said to be **connected under \( \mathcal{A} \) at time \( k \in \mathbb{N} \)** if \( h(k) < \infty \). It is said to be **connected under \( \mathcal{A} \)** if \( h(k) < \infty \forall k \in \mathbb{N} \), and **uniformly connected under \( \mathcal{A} \)** if \( h^* < \infty \).
To illustrate the above ideas, consider Figure 1, which shows a 6-agent network at some time $k$ and its evolution until time $k + 4$. In this figure, an agent $i$ is a member at time $\ell$ if and only if it is enclosed by a black dashed curve (e.g., agent 6 is not a member at time $k$). Moreover, if an agent $i$ at time $\ell$ is enclosed by a gray solid curve, then $C_i(k, \ell)$ is the set of agents enclosed by the same curve (e.g., $C_1(k, k + 1) = \{1\}$, $C_2(k, k + 1) = \{2\}$, $C_3(k, k + 1) = C_5(k, k + 1) = \{3, 5\}$, and $C_4(k, k + 1) = \{4\}$). Otherwise, $C_i(k, \ell)$ is empty (e.g., $C_6(k, k + 1) = \emptyset$). Note that the black dashed curve at time $k$ is arbitrarily selected, whereas those at subsequent times are consequences of (2). Similarly, the gray solid curves at time $k$ are due to (31), whereas those at subsequent times are due to (32). Examining these curves along with (35), we deduce that the set $D_k$ does not contain $k$, $k + 1$, $k + 2$, and $k + 3$ but contains $k + 4$. From (33) and Definition 1, we conclude that $h(k) = 4$ and, hence, the network is connected under $\mathcal{A}$ at time $k$.

To further illustrate Definition 1 consider the following examples:

**Example 3.** Consider the agent network $\mathcal{A}1 \mathcal{A}8$ and suppose $M = 3$. Let $\mathcal{F} = \{1, 2\}$ and $(\mathcal{J}(k), \mathcal{I}(k), \mathcal{L}(k))$ be equal to $(\{3\}, \{1\}, \emptyset)$ if $(k - 1)/6 \in \mathbb{N}$, to $(\emptyset, \{3\}, \{1\})$ if $(k - 2)/6 \in \mathbb{N}$, to $(\{1\}, \{2\}, \emptyset)$ if $(k - 3)/6 \in \mathbb{N}$, to $(\emptyset, \{1\}, \{2\})$ if $(k - 4)/6 \in \mathbb{N}$, to $(\{2\}, \{3\}, \emptyset)$ if $(k - 5)/6 \in \mathbb{N}$, and to $(\emptyset, \{2\}, \{3\})$ if $(k - 6)/6 \in \mathbb{N}$, thereby defining $\mathcal{A}$ in (4). Investigating $\mathcal{A}$, we see two groups of messages being passed around the agents, but never getting a chance to “mix.” Thus, we expect the network to be disconnected under $\mathcal{A}$ at all times. Indeed, applying (31), (32), (33), (35), and Definition 1 yields $h(k) = \infty \forall k \in \mathbb{N}$, confirming the expectation.

**Example 4.** Reconsider the agent network in Example 3 but let $(\mathcal{J}(k), \mathcal{I}(k), \mathcal{L}(k))$ be equal to $(\{3\}, \{1\}, \emptyset)$ if $(k - 1)/6 \in \mathbb{N}$, to $(\emptyset, \{2\}, \{1\})$ if $(k - 2)/6 \in \mathbb{N}$, to $(\{1\}, \{2\}, \emptyset)$ if $(k - 3)/6 \in \mathbb{N}$, to $(\emptyset, \{3\}, \{2\})$ if $(k - 4)/6 \in \mathbb{N}$, to $(\{2\}, \{3\}, \emptyset)$ if $(k - 5)/6 \in \mathbb{N}$, and to $(\emptyset, \{1\}, \{3\})$ if $(k - 6)/6 \in \mathbb{N}$. Observe that unlike the $\mathcal{A}$ in Example 3, the $\mathcal{A}$ here causes the messages to quickly become inseparable no matter the initial time. Hence, the network is expected to not only be connected, but uniformly so, under $\mathcal{A}$. It follows from (33) that $h(k) = 2$ if $k$ is even and $h(k) = 3$ if $k$ is odd, from (34) that $h^* = 3$, and from Definition 1 that the network is indeed uniformly connected.

**Example 5.** Reconsider the agent network in Example 3 but let $\mathcal{F} = \{1, 2, 3\}$ and $(\mathcal{J}(k), \mathcal{I}(k), \mathcal{L}(k))$ be equal to $(\emptyset, \{1, 2\}, \emptyset)$ if $k \in \ell(\ell + 1)/2 : \ell \in \mathbb{P}$ and to $(\emptyset, \{2, 3\}, \emptyset)$ otherwise. Notice that although agent 2 takes turn to interact with agent 1 and with agent 3, its interaction with agent 1 becomes less and less frequent as time elapses, as if the network is gradually losing its connectivity.
Therefore, the network is expected to be connected, but not uniformly so, under $A$. Indeed, it is connected because $h(0) = 2$ and $h(k) \leq \ell + 1 < \infty \forall \ell \in \mathbb{P} \forall k \in [\ell(\ell + 1)/2, (\ell + 1)(\ell + 2)/2 - 1]$. It is not uniformly connected because $h(\ell(\ell + 1)/2) = \ell + 1 \forall \ell \in \mathbb{P}$ so that $h^* = \infty$. □

Finally, it might be of interest to see how Definition 1 is related to existing definitions of connectivity in the literature. The following proposition sheds light on this question, showing that when there are no membership changes, the connectivity of the agent network under $A$ is equivalent to the connectivity of an infinite interaction graph first introduced in [3], so that the former may be viewed as a generalization of the latter:

**Proposition 1.** If $\mathcal{J}(k) = \mathcal{L}(k) = \emptyset \forall k \in \mathbb{P}$, then the agent network $A_1$–$A_8$ is connected under $A$ if and only if the graph $(\mathcal{F}, \mathcal{E}_\infty)$ is connected, where

$$\mathcal{E}_\infty = \{\{i, j\} \subset \mathcal{F} : \{i, j\} \subset \mathcal{I}(k) \text{ for infinitely many } k \in \mathbb{P}\}. \quad (36)$$

*Proof.* See Appendix A.2. □

## 5 Boundedness and Convergence

In this section, we analyze the boundedness, asymptotic convergence, and exponential convergence of SE and derive a bound on its convergence rate. To streamline the presentation of the results, we defer their proofs to Appendix A. Moreover, we let $\beta > 0$ denote the spectral radius of $\sum_{i \in M(0)} P_i$ and introduce the following definition:

**Definition 2.** The sequence $\{Q_i(k)\}_{k \in \mathbb{N}, i \in M(k)}$ produced by SE is said to be uniformly positive definite under $A$ if $\exists \alpha > 0$ such that $\forall k \in \mathbb{N}, \forall i \in M(k), Q_i(k) - \alpha I \in S^n_+$.\]

Although the initial values $Q_i(0)$'s depend on the observations $P_i$'s via (27), the proposition below says that the uniform positive definiteness of $\{Q_i(k)\}_{k \in \mathbb{N}, i \in M(k)}$ depends only on the agent actions $A$ and not on the $P_i$'s nor the $q_i$'s:

**Proposition 2.** Whether or not the sequence $\{Q_i(k)\}_{k \in \mathbb{N}, i \in M(k)}$ produced by SE is uniformly positive definite under $A$ is independent of the observations $P_i \in S^n_+$ and $q_i \in \mathbb{R}^n \forall i \in \mathcal{F}$.

*Proof.* See Appendix A.3. □

Our first result is a sufficient condition on the boundedness of SE:

**Theorem 1.** Consider the agent network modeled by $A_1$–$A_8$ and the use of SE described in Algorithm 1. Let $A$ be given. Then, $Q_i(k)$ is bounded as follows:

$$Q_i(k) \leq \beta I, \quad \forall k \in \mathbb{N}, \forall i \in M(k). \quad (37)$$

15
If, in addition, the sequence \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) is uniformly positive definite under \( \mathcal{A} \), then \( z_i(k) \) is bounded as follows:

\[
\|z_i(k) - \beta\|^2 \leq \frac{V(k)}{\alpha} \leq \frac{V(0)}{\alpha}, \quad \forall k \in \mathbb{N}, \forall i \in \mathcal{M}(k),
\]

where \( \alpha \) is any positive number satisfying \( Q_i(k) - \alpha I \in \mathbb{S}_+^n \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k) \).

**Proof.** See Appendix A.4.

Theorem 1 implies that all the \( Q_i(k) \)'s, whenever not equal to \#\, are unconditionally bounded from above by \( \beta \), irrespective of the agent actions \( \mathcal{A} \). In addition, if they turn out to be bounded from below by some \( \alpha > 0 \), then all the \( z_i(k) \)'s, whenever not equal to \#, are guaranteed to stay within a ball centered at the solution \( z \), whose radius \( \sqrt{V(k)/\alpha} \) decreases over time. We note, however, that such an \( \alpha \) does not always exist, as the following example reveals:

**Example 6.** Consider the agent network \( \mathcal{A}_1 \mathcal{A}_8 \) and suppose SE is used. Let \( M = 3 \), \( \mathcal{F} = \{1, 2\} \), and \((\mathcal{J}(k), \mathcal{I}(k), \mathcal{L}(k))\) be equal to \((\{3\}, \{1\}, \emptyset)\) if \( k \) is odd and to \((\emptyset, \{2\}, \{3\})\) if \( k \) is even, thus defining \( \mathcal{A} \) in [4]. Also, let \( P_1 = P_2 = 1 \), \( q_1 = 1 \), and \( q_2 = 2 \), so that \( z = 1.5 \) from (5). With this \( \mathcal{A} \), agent 3 repeatedly does the following: joins the network, interacts with agent 1 upon joining, leaves the network subsequently, and interacts with agent 2 prior to leaving. Hence, the network is connected under \( \mathcal{A} \). Moreover, it is straightforward to show that \( \forall k \in \mathbb{N}, Q_1(k) = (\frac{1}{2})^\lceil \frac{k}{2} \rceil, Q_2(k) = 2 - (\frac{1}{2})^\lceil \frac{k}{2} \rceil, Q_3(k) = (\frac{1}{2})^\lceil \frac{k}{2} \rceil \) if \( k \) is odd, \( Q_3(k) = \# \) if \( k \) is even, \( z_1(k) = 1, z_2(k) = (3 - (\frac{1}{2})^\lceil \frac{k}{2} \rceil)/(2 - (\frac{1}{2})^\lceil \frac{k}{2} \rceil), z_3(k) = 1 \) if \( k \) is odd, and \( z_3(k) = \# \) if \( k \) is even. It follows that \( \lim_{k \to \infty} Q_1(k) = 0, \lim_{k \to \infty} Q_2(k) = 2, \lim_{k \to \infty} z_1(k) = 1, \) and \( \lim_{k \to \infty} z_2(k) = 1.5 \).

In Example 6, the required \( \alpha > 0 \), with which \( Q_i(k) - \alpha I \in \mathbb{S}_+^n \forall k \in \mathbb{N} \), does not exist, implying that \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) is not uniformly positive definite under \( \mathcal{A} \). Yet, all the \( z_i(k) \)'s are bounded. Therefore, \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) being uniformly positive definite under \( \mathcal{A} \) is a sufficient, but not necessary, condition for the boundedness of SE.

In general, given \( \mathcal{A} \), it is not easy to check whether the resulting sequence \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) is uniformly positive definite under \( \mathcal{A} \). However, if \( \mathcal{A} \) happens to be such that every agent joins and leaves the network arbitrarily but finitely many times—a rather mild condition that is often satisfied in practice—then the uniform positive definiteness of \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) can be immediately verified. The following definition and corollary to Theorem 1 formalize this claim:

**Definition 3.** The membership dynamics (2) of the agent network \( \mathcal{A}_1 \mathcal{A}_8 \) are said to be **ultimately static under** \( \mathcal{A} \) if \( \exists k \in \mathbb{N} \) such that \( \forall \ell > k, \mathcal{M}(\ell) = \mathcal{M}(k) \), i.e., \( \mathcal{J}(\ell) = \mathcal{L}(\ell) = \emptyset \).

**Corollary 1.** If the membership dynamics (2) are ultimately static under \( \mathcal{A} \), then \( Q_i(k) \) and \( z_i(k) \) are bounded as in (37) and (38) for some \( \alpha > 0 \).

**Proof.** See Appendix A.5.
In Theorem 1 and Corollary 1, the network is not assumed to be connected since such an assumption is not needed for the boundedness of SE. For convergence, however, this assumption is crucial. The following lemma, which makes use of this assumption, is a key step toward establishing both the asymptotic and exponential convergence of SE:

**Lemma 2.** Consider the agent network modeled by $A_1–A_8$ and the use of SE described in Algorithm 1. Let $A$ be given. Suppose the agent network is connected under $A$ at some time $k \in \mathbb{N}$, so that $h(k) < \infty$ by Definition 1. Then,

$$V(k + h(k)) \leq \left(\frac{4^2}{\alpha}\right)^{M-1} \cdot M \cdot M! \cdot V(k),$$

where $\alpha$ is any positive number satisfying $Q_i(\ell) - \alpha I \in S^+ \forall i \in [k, k + h(k)] \forall i \in M(\ell)$.

**Proof.** See Appendix A.6.

Lemma 2 asserts that regardless of how the agents interact, and how they join and leave the network, as long as it is connected at some time $k$, the value of $V$ must strictly decrease from $V(k)$ at time $k$ to $V(k + h(k))$ at time $k + h(k)$, by a factor that can be explicitly calculated, in (39). This result suggests that the better the “instantaneous” connectedness (i.e., the smaller $h(k)$), the faster the value of $V$ drops, which makes intuitive sense.

As was mentioned in Section 3, even if $V(k)$ decreases asymptotically to zero as $k \to \infty$, it does not necessarily imply that SE is asymptotically convergent: in Example 6, $\lim_{k \to \infty} V(k) = 0$. Also, $h(k) = 2$ if $k$ is even and $h(k) = 3$ if $k$ is odd, so that the network is connected under $A$ by Definition 1. Yet, $z_1(k)$ fails to converge to $z$, due to the fact that $Q_1(k)$ keeps “losing” its positive definiteness as $k \to \infty$. This phenomenon suggests that network connectivity and the uniform positive definiteness of $\{Q_i(k)\}_{k \in \mathbb{N}, i \in M(k)}$ together might be all that are needed to establish the asymptotic convergence of SE. The following theorem shows that this is indeed the case:

**Theorem 2.** Consider the agent network modeled by $A_1–A_8$ and the use of SE described in Algorithm 1. Let $A$ be given. Suppose the agent network is connected under $A$ and the sequence $\{Q_i(k)\}_{k \in \mathbb{N}, i \in M(k)}$ is uniformly positive definite under $A$. Then, $z_i(k)$ asymptotically converges to the solution $z$, i.e.,

$$\forall \varepsilon > 0, \exists k \in \mathbb{N} \text{ such that } \forall \ell \geq k, \forall i \in M(\ell), \|z_i(\ell) - z\| < \varepsilon.$$  \hspace{1cm} (40)

**Proof.** See Appendix A.7.

Theorem 2 has the following corollary:

**Corollary 2.** If the agent network is connected under $A$ and the membership dynamics (2) are ultimately static under $A$, then (40) holds.
The proof is an immediate consequence of Theorem 2 and the proof of Corollary 1.

Note that the conclusion of Theorem 2 is written as \( \lim_{k \to \infty} z_i(k) = z \) instead of \( \lim_{k \to \infty} z_i(k) = \# \), because the former excludes cases where \( z_i(k) = \# \), while the latter does not and, thus, is not well-defined. More important, with Theorem 2 and Corollary 2, we achieve the paper’s objective of developing a distributed asynchronous algorithm SE that asymptotically solves (5) over the agent network \( \mathbb{A}_1 - \mathbb{A}_8 \) while possessing properties \( \mathbb{P}_1 - \mathbb{P}_3 \) stated in Section 2.

Finally, we provide a sufficient condition on the exponential convergence of SE and derive a bound on its convergence rate, in terms of \( h^* \). Since \( h^* = 0 \) is a trivial case (that corresponds to \( \mathcal{M}(k) \) containing exactly one and the same agent \( i \in \{1, 2, \ldots, M\} \) with \( z_i(k) = z \ \forall k \in \mathbb{N} \)), below it is assumed that \( h^* > 0 \):

**Theorem 3.** Consider the agent network modeled by \( \mathbb{A}_1 - \mathbb{A}_8 \) and the use of SE described in Algorithm 4. Let \( \mathcal{A} \) be given. Suppose the agent network is uniformly connected under \( \mathcal{A} \) with \( h^* > 0 \) and the sequence \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) is uniformly positive definite under \( \mathcal{A} \). Then,

\[
V(k) \leq V(0) \left( \frac{(4\beta^\alpha)^{M-1} \cdot M \cdot M!}{(4\alpha^\alpha)^{M-1} \cdot M \cdot M! + 1} \right)^{|k| h^*}, \quad \forall k \in \mathbb{N}, \quad (41)
\]

\[
\|z_i(k) - z\|^2 \leq \frac{V(0)}{\alpha} \left( \frac{(4\beta^\alpha)^{M-1} \cdot M \cdot M!}{(4\alpha^\alpha)^{M-1} \cdot M \cdot M! + 1} \right)^{|k| h^*}, \quad \forall k \in \mathbb{N}, \ \forall i \in \mathcal{M}(k), \quad (42)
\]

where \( \alpha \) is any positive number satisfying \( Q_i(k) - \alpha I \in \mathbb{S}^n_+ \ \forall k \in \mathbb{N} \ \forall i \in \mathcal{M}(k) \).

**Proof.** See Appendix A.8.

Similar to Theorems 1 and 2, we have the following corollary to Theorem 3:

**Corollary 3.** If the agent network is uniformly connected under \( \mathcal{A} \) with \( h^* > 0 \) and the membership dynamics (2) are ultimately static under \( \mathcal{A} \), then (41) and (42) hold for some \( \alpha > 0 \).

**Proof.** The proof follows immediately from Theorem 3 and the proof of Corollary 1.

Comparing Theorems 2 and 3 (or Corollaries 2 and 3), we see that connectivity helps ensure asymptotic convergence, while uniform connectivity helps ensure exponential convergence.

### 6 Simulation Studies

In this section, we complement the above theoretical analysis with simulation studies. Section 6.1 illustrates through simulation the behavior of SE in a volatile agent network. Section 6.2 compares through simulation the performance of PE and GE from Examples 1 and 2, the two algorithms from [1,2], and flooding in multi-hop wireless networks.
6.1 Illustration of SE in an Agent Network

In this subsection, we simulate SE in an agent network described by $A1$–$A8$. The simulation settings are as follows: $M = 100$; $F = \{1, 2, \ldots, 50\}$; $n = 4$; for each $i \in F$, the matrix $P_i \in \mathbb{S}^n_+$ and the vector $q_i \in \mathbb{R}^n$ are randomly generated; and for each $k \in P$, the sets $J(k)$, $I(k)$, and $L(k)$ are random subsets of the sets $\{1, 2, \ldots, M\} - \mathcal{M}(k - 1)$, $\mathcal{M}(k - 1)$, and $\mathcal{M}(k - 1)$, respectively, such that $I(k) \cap L(k) = \emptyset$, $I(k) \neq \emptyset$, and $L(k) \subseteq \mathcal{M}(k - 1)$ according to $A4$. Note that with these settings, the agent network is volatile with random, unpredictable member interactions and membership dynamics. Thus, the behavior of SE in such a network is indicative of its effectiveness.

Figure 2 depicts the simulation results. The top subplot of Figure 2 shows the number of members $|\mathcal{M}(k)|$ as a function of time $k$. The middle subplot shows the actions taken by two selected agents, agent 1 and agent 51, at each time $k$, where a total of five actions are possible as labeled on the vertical axis, and only the actions of two agents are shown to avoid clogging the plot. Also, the actions labeled “$i \in \mathcal{M}(k)$ but idle” and “$i \notin \mathcal{M}(k)$ but idle” are abbreviations for $i \in \mathcal{M}(k - 1) - (I(k) \cup L(k))$ and $i \in \{1, 2, \ldots, M\} - (\mathcal{M}(k - 1) \cup J(k))$, respectively. Lastly, the bottom subplot shows, on a logarithmic scale and as functions of time $k$, the maximum estimation error $\max_{i \in \mathcal{M}(k)} \|z_i(k) - z\|$ among the members in $\mathcal{M}(k)$, the minimum such error $\min_{i \in \mathcal{M}(k)} \|z_i(k) - z\|$, the estimation error $\|z_1(k) - z\|$ of agent 1 whenever it is a member, and the estimation error $\|z_{51}(k) - z\|$ of agent 51 whenever it is a member. Observe from the figure that, despite the rapidly fluctuating number of members, and despite the randomly generated actions of agents that include numerous membership changes, all the estimates $z_i(k)$’s gradually approach the unknown solution $z$, demonstrating the effectiveness of SE.

6.2 Comparison of PE and GE with Existing Algorithms in Wireless Networks

In this subsection, we compare through simulation the performance of five algorithms—namely, PE and GE from Examples 1 and 2, the two average-consensus-based algorithms from $[1, 2]$ called Maximum-Degree Weights (MDW) and Metropolis Weights (MW), and flooding—in solving problem $[1]$ of different sizes, over multi-hop wireless networks modeled by random geometric graphs of different sizes and densities. The simulation settings are as follows: to methodically evaluate the algorithm performance, we let the simulation be governed by three parameters—the number of nodes $N$ that represents network sizes, the average number of neighbors $\frac{2L}{N}$ that represents network densities (the meaning of $\frac{2L}{N}$ will be clear shortly), and the number of dimensions $n$ that represents problem sizes—and write them as a 3-tuple $(N, \frac{2L}{N}, n)$. To understand their individual impact, we vary these parameters one at a time, choosing the values of $(N, \frac{2L}{N}, n)$ as:

- S1. $(50, 20, 4), (100, 20, 4), \ldots, (500, 20, 4)$;
- S2. $(200, 10, 4), (200, 20, 4), \ldots, (200, 100, 4)$; and
- S3. $(200, 20, 2), (200, 20, 4), \ldots, (200, 20, 20)$. 

19
Figure 2: Behavior of SE in a volatile agent network with random, unpredictable member interactions and membership dynamics.

For each value of \((N, \frac{2L}{N}, n)\) in \(S_1\)–\(S_3\) we consider 50 random scenarios. For each scenario, we generate a wireless network with \(N\) nodes and \(L\) links by randomly and equiprobably placing \(N\) nodes on a unit square in \(\mathbb{R}^2\) and gradually increasing the one-hop radius until the number of links is \(L\) or, equivalently, the average number of neighbors is \(\frac{2L}{N}\) (this explains the meaning of \(\frac{2L}{N}\)). If the resulting network is not connected, it is discarded and the preceding process is repeated. We also generate an instance of problem \(\Pi\) with \(n\) dimensions by factoring each \(P_i \in S_+^n\) as \(P_i = X_i^T X_i\) and letting both \(X_i \in \mathbb{R}^{n \times n}\) and \(q_i \in \mathbb{R}^n\) have random entries that are independent and standard normally distributed. Subsequently, we simulate PE, GE, MDW, and MW and let the gossiping pair in Step 3 of PE, as well as the interacting group in Step 3 of GE, be randomly
and equiprobably chosen. We then count the number of real-number transmissions needed for each algorithm to converge (including initialization overhead, if any), where the convergence criterion is $\max_{i \in V} \| z_i(k) - z \| < 0.005$. To count such numbers, we use the fact that PE, GE, MDW, and MW require, respectively, $\frac{n(n+1)}{2}N$, $\frac{n(n+1)}{2}N$, 0, and 0 real-number transmissions to initialize and $2n$, $n(|N_i| + 1)$, $\left(\frac{n(n+1)}{2} + n\right)N$, and $\left(\frac{n(n+1)}{2} + n\right)N$ real-number transmissions per iteration (in the case of GE, per iteration initiated by node $i$). Finally, for each value of $(N, \frac{2L}{N}, n)$ in $S_1$–$S_3$ and for each algorithm, we average over the 50 scenarios and record the resulting number needed to converge. As a benchmark, we also record the number needed by flooding to exactly solve (1) (i.e., $\left(\frac{n(n+1)}{2} + n\right)N^2$).

Figure 3 displays the simulation results, showing in its subplots (a), (b), and (c) the number of real-number transmissions needed as a function of the values of $(N, \frac{2L}{N}, n)$ in $S_1$–$S_3$ respectively. Notice from the figure that:

- Generally, the larger the network size, or the lower the network density, or the larger the problem size, the higher the number needed. One exception to this trend is flooding in subplot (b), which is expected since its number depends only on $N$ and $n$ and not on $L$.
- Among the five algorithms, MDW has, on average, the worst bandwidth/energy efficiency, requiring by far the most real-number transmissions to converge. Nonetheless, MDW does outperform flooding when the network is sufficiently dense.
- PE is not as efficient as MW in subplots (a) and (b). However, it becomes more efficient than MW when the problem size is sufficiently large, in subplot (c). This is likely due to PE being $O(n)$ and MW being $O(n^2)$ in the number of real-number transmissions per iteration.
- Among the five algorithms, GE has the best bandwidth/energy efficiency and scalability with respect to $N$ and $n$. Indeed, GE is at least 2.5 times and up to 8 times more efficient than the next best algorithm—be it MW or PE—in all the values of $(N, \frac{2L}{N}, n)$ considered.

7 Conclusion

In this paper, we have developed SE, a distributed asynchronous algorithm for solving symmetric positive definite systems of linear equations over agent networks with arbitrary member interactions and membership dynamics. To facilitate the development, we have introduced a time-varying Lyapunov-like function and a generalized concept of network connectivity. Based on these entities, we have derived sufficient conditions for ensuring the boundedness, asymptotic convergence, and exponential convergence of SE, as well as a bound on its convergence rate. We have also shown that SE reduces to known algorithms in very special cases. Finally, we have demonstrated through extensive simulation the effectiveness and efficiency of SE in a variety of settings.
A Appendix

Throughout the Appendix, for any $x \in \mathbb{R}^n$ and any $P \in \mathbb{S}_+^n$, we write $x^T x$ and $x^T P x$ as $\|x\|^2$ and $\|x\|^2_P$, respectively. Moreover, for any $k \in \mathbb{N}$ and any nonempty $X \subset \mathcal{M}(k)$, we let

$$z_X^k = \left( \sum_{i \in X} Q_i(k) \right)^{-1} \sum_{i \in X} Q_i(k) z_i(k),$$

(43)

so that from (17),

$$z_i(k) = z_X^k - \frac{1}{2} \mathbb{I}(k)_{j \cup l(k)}, \quad \forall k \in \mathbb{P}, \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k).$$

(44)

A.1 Proof of Lemma 1

Let $A$ and $k \in \mathbb{P}$ be given. To prove the first statement, pick any $Q_i(k) \in \mathbb{S}_+^n \forall i \in \mathcal{J}(k) \cup \mathcal{I}(k)$ satisfying (15) and consider the following equality-constrained, convex optimization problem:

$$\begin{align*}
\text{minimize}_{(z_i(k))_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)}} & \quad \sum_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)} z_i(k)^T Q_i(k) z_i(k) \\
\text{subject to} & \quad (17). 
\end{align*}$$

(45)

By forming the Lagrangian of problem (45) and setting its gradient to zero, we see that problem (45) has a unique solution $(z_i(k))_{i \in \mathcal{J}(k) \cup \mathcal{I}(k)}$ given by (17). Moreover, by substituting (17) into the
objective function and using (15), we see that the optimal value of problem (45) depends only on \((z_i(k-1), Q_i(k-1))_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)}\) and not on the arbitrary \((Q_i(k))_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)}\). Hence, problem (46) has a nonempty, convex set of solutions given by \(\{(z_i(k), Q_i(k))_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} : (17) \text{ and } (15) \text{ hold}\}\), i.e., the first statement is true. For the second statement, note from (13), (14), (15), and (44) that

\[
V(k) - V(k-1) = \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} z_i(k)^T Q_i(k) z_i(k) - \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} z_i(k-1)^T Q_i(k-1) z_i(k-1)

= -\left( \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} z_i(k-1)^T Q_i(k-1) z_i(k-1) + \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} (z_i(k-1)^{k-1}_{\mathcal{I}(k) \cup \mathcal{L}(k)})^T Q_i(k-1) z_i(k-1)^{k-1}_{\mathcal{I}(k) \cup \mathcal{L}(k)} - 2 \sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} z_i(k-1)^T Q_i(k-1) z_i(k-1)^{k-1}_{\mathcal{I}(k) \cup \mathcal{L}(k)} \right)

= -\sum_{i \in \mathcal{I}(k) \cup \mathcal{L}(k)} (z_i(k-1) - z_i^{k-1}_{\mathcal{I}(k) \cup \mathcal{L}(k)})^T Q_i(k-1) (z_i(k-1) - z_i^{k-1}_{\mathcal{I}(k) \cup \mathcal{L}(k)}).
\]

(46)

Since the right-hand side of (46) is nonpositive, \(V(k) \leq V(k-1)\). Moreover, from (43) and (46), \(V(k) = V(k-1)\) if and only if \(z_i(k-1) \forall i \in \mathcal{I}(k) \cup \mathcal{L}(k)\) are equal.

### A.2 Proof of Proposition [1]

First, suppose the graph \((\mathcal{F}, \mathcal{E}_{\infty})\) is connected. Pick any \(k \in \mathbb{N}\) and let \(k' = \inf \{\tilde{k} \geq k + 1 : \forall \{i,j\} \in \mathcal{E}_{\infty}, \exists \tilde{k} \in [k+1, \tilde{k}] \text{ such that } \{i,j\} \subset \mathcal{I}(\tilde{k})\}\). Then, from (36), \(k' < \infty\). Due to (31), (32), and \((\mathcal{F}, \mathcal{E}_{\infty})\) being connected, we have \(C_i(k, k') = \mathcal{F} \forall i \in \mathcal{F}\). It follows from (35) and (33) that \(k' \in D_k\) and, thus, \(h(k) \leq k' - k < \infty\). From Definition [1], the network is connected under \(\mathcal{A}\). Conversely, suppose the network is connected under \(\mathcal{A}\), i.e., \(h(k) < \infty \forall k \in \mathbb{N}\). For each \(k \in \mathbb{N}\), let \(\tilde{\mathcal{E}}(k) = \bigcup_{k'=k+1}^{k+h(k)} \{i,j\} \subset \mathcal{F} : \{i,j\} \subset \mathcal{I}(k')\}\). Then, due to (31), (32), (35), and (33), the graph \((\mathcal{F}, \tilde{\mathcal{E}}(k))\) is connected. Let \(\mathcal{E}\) denote the collection of all nonempty edge sets associated with the vertex set \(\mathcal{F}\). Clearly, \(\mathcal{E}\) contains \(2^{\lvert \mathcal{F} \rvert}((\lvert \mathcal{F} \rvert)-1)/2 - 1\) sets and \(\tilde{\mathcal{E}}(k) \in \mathcal{E} \forall k \in \mathbb{N}\). Then, \(\exists \tilde{\mathcal{E}} \in \mathcal{E}\) such that \(\tilde{\mathcal{E}}(k) = \tilde{\mathcal{E}}\) for infinitely many \(k \in \mathbb{N}\). From (36), we see that \(\tilde{\mathcal{E}} \subset \mathcal{E}_{\infty}\). Therefore, the graph \((\mathcal{F}, \mathcal{E}_{\infty})\) is connected.

### A.3 Proof of Proposition [2]

Let \(\mathcal{A}\) be given. In addition, let \(\hat{P}_i \in \mathbb{S}_+^n\) and \(\hat{q}_i \in \mathbb{R}^n \forall i \in \mathcal{F}\), and \(\hat{P}_i \in \mathbb{S}_+^n\) and \(\hat{q}_i \in \mathbb{R}^n \forall i \in \mathcal{F}\), be two distinct sets of observations. Let \(\{\hat{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}\) and \(\{\tilde{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}\) be state variables associated with these two sets of observations, respectively, and suppose they are determined by the same \(\mathcal{A}\). To prove the proposition, it suffices to show that if the sequence \(\{\hat{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}\) is...
uniformly positive definite under the given $A$, then so is the sequence $\{\tilde{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}$. Note from (27), (29), and (30) that $\forall k \in \mathbb{N}, \forall i \in \mathcal{M}(k)$, $\tilde{Q}_i(k) = \sum_{j \in \mathcal{F}} a_{ij}(k)\tilde{P}_j$ and $\tilde{Q}_i(k) = \sum_{j \in \mathcal{F}} a_{ij}(k)\tilde{P}_j$, where each $a_{ij}(k) \geq 0$ is completely determined by $A$. For each $j \in \mathcal{F}$, let $\tilde{\lambda}_j > 0$ be the largest eigenvalue of $\tilde{P}_j$ and $\tilde{\lambda}_j > 0$ be the smallest eigenvalue of $\tilde{P}_j$. Also, let $\tilde{\lambda} = \max_{j \in \mathcal{F}} \tilde{\lambda}_j$ and $\tilde{\lambda} = \min_{j \in \mathcal{F}} \tilde{\lambda}_j$. Since $\{\tilde{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}$ is uniformly positive definite under $A$, $\exists \tilde{\alpha} > 0$ such that $\forall k \in \mathbb{N}, \forall i \in \mathcal{M}(k)$, $\tilde{Q}_i(k) > \tilde{\alpha} I$. Hence, $\sum_{j \in \mathcal{F}} a_{ij}(k) > \tilde{\alpha} I \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. It follows that $\tilde{Q}_i(k) \geq \tilde{\lambda} \sum_{j \in \mathcal{F}} a_{ij}(k) I > \frac{\tilde{\alpha}}{\tilde{\lambda}} I \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. Since $\frac{\tilde{\alpha}}{\tilde{\lambda}} > 0$, $\{\tilde{Q}_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}$ is uniformly positive definite under $A$.

### A.4 Proof of Theorem 1

Let $A$ be given. From (21) and (23), $Q_i(k) \leq \sum_{i \in \mathcal{M}(0)} P_i \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. Thus, (37) holds, i.e., each $Q_i(k)$ is bounded. To derive (38), suppose $\{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}$ is uniformly positive definite under $A$ and let $\alpha > 0$ be such that $\tilde{Q}_i(k) - \alpha I \in \mathbb{S}_+^n \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. Then, from (3) and Lemma 1, $\alpha \sum_{i \in \mathcal{M}(k)} \|z_i(k) - z\|^2 \leq V(k) \leq V(0) \forall k \in \mathbb{N}$. Therefore, (38) is satisfied, i.e., each $z_i(k)$ is bounded.

### A.5 Proof of Corollary 1

Suppose the membership dynamics (2) are ultimately static under $A$. Then, by Definition 3, $\exists k \in \mathbb{N}$ such that $\forall 0 \leq k, \mathcal{M}(\ell) = \mathcal{M}(k)$. Due to (27), (29), and (30), $\exists \alpha > 0$ such that $Q_i(\ell) > \alpha I \forall \ell \leq k \forall i \in \mathcal{M}(k)$. Due again to (29), $Q_i(\ell) = Q_i(k) \forall \ell \geq k + 1 \forall i \in \mathcal{M}(k)$. Hence, $\{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)}$ is uniformly positive definite under $A$. It follows from Theorem 1 that (37) and (38) hold.

### A.6 Proof of Lemma 2

Let $A$ be given. Suppose the agent network is connected under $A$ at some time $k \in \mathbb{N}$, i.e., $h(k) < \infty$. If $h(k) = 0$, then from (33), (35), and (34), $\mathcal{M}(k) = \{i\}$ for some $i \in \{1, 2, \ldots, M\}$. Also, from (20), (21), (22), (23), and (5), $z_i(k) = z$. It follows that $V(k) = 0$ and, thus, (39) holds. Now suppose $h(k) \in \mathbb{P}$ and consider the following:

**Lemma 3.** For any $\ell \in \mathbb{N}$, any nonempty $X \subset \mathcal{M}(\ell)$, and any $\eta \in \mathbb{R}^n$, $\sum_{i \in X} \|z_X^\ell - \eta\|^2_{Q_i(\ell)} \leq \sum_{i \in X} \|z_i(\ell) - \eta\|^2_{Q_i(\ell)}$.

**Proof.** Due to (43), we have $\sum_{i \in X} Q_i(\ell)z_X^\ell = \sum_{i \in X} Q_i(\ell)z_i(\ell)$, implying that $\sum_{i \in X} (z_X^\ell)^T Q_i(\ell)\eta = \sum_{i \in X} z_i(\ell)^T Q_i(\ell)\eta$. Also, $\sum_{i \in X} (z_X^\ell)^T Q_i(\ell)z_X^\ell = \sum_{i \in X} z_i(\ell)^T Q_i(\ell)z_X^\ell$. Hence, $\sum_{i \in X} \|z_X^\ell - \eta\|^2_{Q_i(\ell)} - \sum_{i \in X} \|z_i(\ell) - \eta\|^2_{Q_i(\ell)} = - \sum_{i \in X} \|z_i(\ell) - z_X^\ell\|^2_{Q_i(\ell)} \leq 0$.  

**Lemma 4.** For any $\ell \in \mathbb{N}$, any nonempty $X \subset \mathcal{M}(\ell)$, and any $\eta \in \mathbb{R}^n$, $\sum_{i \in X} \|z_i(\ell) - z_X^\ell\|^2_{Q_i(\ell)} \leq \sum_{i \in X} \|z_i(\ell) - \eta\|^2_{Q_i(\ell)}$.
Proof. Using the two properties in the proof of Lemma 3, \( \sum_{i \in X} \|z_i(\ell) - z^\ell_X\|^2_{Q_i(\ell)} = -\sum_{i \in X} \|z^\ell_X - \eta\|^2_{Q_i(\ell)} \leq 0. \)

Let \( \alpha > 0 \) be such that \( Q_i(\ell) - \alpha I \in S^+_{\ell} \) \( \forall \ell \in [k, k + h(k)] \) \( \forall i \in M(\ell) \). This and (37) imply that
\[
\alpha I < Q_i(\ell) \leq \beta I, \quad \forall \ell \in [k, k + h(k)], \quad \forall i \in M(\ell).
\]
Assume, to the contrary, that (39) does not hold, i.e.,
\[
\epsilon = \frac{V(k)}{(4\beta^2)^{M-1-M!} M \cdot M! + 1}.
\]
Then, \( V(k) - V(k + h(k)) \leq \epsilon \). It follows from Lemma 1 that
\[
V(\ell - 1) - V(\ell) \leq \epsilon, \quad \forall \ell \in [k + 1, k + h(k)].
\]

Due to (49), (46), and (47),
\[
\|z_i(\ell - 1) - z^\ell_{(\ell-1),\ell}\|^2 \leq \frac{\epsilon}{\alpha}, \quad \forall \ell \in [k + 1, k + h(k)], \quad \forall i \in I(\ell) \cup L(\ell).
\]
Next, let \( d_i(\ell) = \sum_{j \in C_i(\ell,k)} \|z_j(\ell) - z^\ell_{C_i(\ell,k)}\|^2_{Q_j(\ell)} \) \( \forall \ell \geq k \) \( \forall i \in M(\ell) \). In addition, let \( m(\ell) \) be the number of distinct sets in the collection \( \{C_i(\ell,k)\}_{i \in M(\ell)} \) \( \forall \ell \geq k \). Notice from (51) and (52) that \( 1 \leq m(\ell) \leq |M(\ell)| \leq M \forall \ell \geq k \) and \( m(\ell) \leq m(\ell - 1) \) \( \forall \ell \geq k + 1 \). Moreover, let \( B(\ell) = \{k\} \cup \{k' \in [k + 1, \ell] : m(k') < m(k' - 1)\} \) \( \forall \ell \geq k + 1 \). Consider the following lemma:

**Lemma 5.** For each \( \ell \in [k, k + h(k)] \),
\[
d_i(\ell) \leq \left(\frac{4\beta}{\alpha}\right)^{M-m(\ell)} (M + 1 - m(\ell)) \left( \prod_{k' \in B(\ell)} (M + 1 - m(k')) \right) \epsilon, \quad \forall i \in M(\ell).
\]

**Proof.** By induction over \( \ell \in [k, k + h(k)] \). Let \( \ell = k \). For any \( i \in M(\ell) \), from (51), \( C_i(\ell,k) = \{i\} \), which, together with (43), implies that \( z_i(\ell) = z^\ell_{C_i(\ell,k)} \). Hence, \( d_i(\ell) = 0 \) \( \forall i \in M(\ell) \). Since the right-hand side of (51) is positive, (51) holds for \( \ell = k \). Next, let \( \ell \in [k + 1, k + h(k)] \) and suppose
\[
d_i(\ell - 1) \leq \left(\frac{4\beta}{\alpha}\right)^{M-m(\ell-1)} (M + 1 - m(\ell - 1)) \left( \prod_{k' \in B(\ell-1)} (M + 1 - m(k')) \right) \epsilon, \quad \forall i \in M(\ell - 1).
\]
Below, we show that (52) implies (51). To do so, consider the following two mutually exclusive and exhaustive cases:

*Case (I):* \( I(\ell) \cup L(\ell) \subset C_{i^*}(k, \ell - 1) \) for some \( i^* \in M(\ell - 1) \). Due to (52),
\[
m(\ell) = m(\ell - 1),
\]

25
so that

\[ B(\ell) = B(\ell - 1). \tag{54} \]

Let \( i \in \mathcal{M}(\ell) \). Suppose \( i \in \mathcal{M}(\ell) - (C_{i^*}(k, \ell - 1) \cup \mathcal{J}(\ell)) \). Then, due to (32), (9), and (10), \( C_{i}(k, \ell) = C_{i}(k, \ell - 1) \), \( z_{j}(\ell) = z_{j}(\ell - 1) \forall j \in C_{i}(k, \ell) \), and \( Q_{j}(\ell) = Q_{j}(\ell - 1) \forall j \in C_{i}(k, \ell) \), implying that \( d_{i}(\ell) = d_{i}(\ell - 1) \). Now suppose \( i \in (C_{i^*}(k, \ell - 1) \cup \mathcal{J}(\ell)) - \mathcal{L}(\ell) \). From (32), \( C_{i}(k, \ell) = (C_{i^*}(k, \ell - 1) \cup \mathcal{J}(\ell)) - \mathcal{L}(\ell) \). Thus, from (9), (10), (14), and (15), we have \( \sum_{j \in C_{j}(k, \ell - 1)} Q_{j}(\ell - 1) = \sum_{j \in C_{j}(k, \ell)} Q_{j}(\ell) \). These and (43) indicate that \( z_{C_{i}(k, \ell)}^{\ell} = z_{C_{i^*}(k, \ell - 1)}^{\ell} \). It follows from (44), (15), (9), and Lemma 3 that

\[
d_{i}(\ell) = \sum_{j \in \mathcal{J}(\ell) \cup \mathcal{I}(\ell)} \| z_{j}(\ell) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell)}^{2} + \sum_{j \in C_{i}(k, \ell - 1) - (\mathcal{I}(\ell) \cup \mathcal{L}(\ell))} \| z_{j}(\ell) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2}
\]

\[
= \sum_{j \in \mathcal{I}(\ell) \cup \mathcal{L}(\ell)} \| z_{j}(\ell) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2} + \sum_{j \in C_{i}(k, \ell - 1) - (\mathcal{I}(\ell) \cup \mathcal{L}(\ell))} \| z_{j}(\ell - 1) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2}
\]

\[
\leq \sum_{j \in \mathcal{I}(\ell) \cup \mathcal{L}(\ell)} \| z_{j}(\ell - 1) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2} + \sum_{j \in C_{i}(k, \ell - 1) - (\mathcal{I}(\ell) \cup \mathcal{L}(\ell))} \| z_{j}(\ell - 1) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2}.
\]

\[
= d_{i}(\ell - 1).
\]

It follows from (32), (55), and (56) that (51) holds for Case (I).

Case (II): \( \mathcal{I}(\ell) \cup \mathcal{L}(\ell) \not\subset C_{i}(k, \ell - 1) \forall i \in \mathcal{M}(\ell - 1) \). Due to (32),

\[ m(\ell - 1) - m(\ell) \geq 1, \tag{55} \]

\[ B(\ell) = B(\ell - 1) \cup \{ \ell \}. \tag{56} \]

Let \( i \in \mathcal{M}(\ell) \). Suppose \( i \in \mathcal{M}(\ell) - \left( \bigcup_{j \in \mathcal{I}(\ell) \cup \mathcal{L}(\ell)} C_{j}(k, \ell - 1) \cup \mathcal{J}(\ell) \right) \). Then, observe from (52), (9), and (10) that \( C_{i}(k, \ell) = C_{i}(k, \ell - 1) \), \( z_{j}(\ell) = z_{j}(\ell - 1) \forall j \in C_{i}(k, \ell) \), and \( Q_{j}(\ell) = Q_{j}(\ell - 1) \forall j \in C_{i}(k, \ell) \). Hence, \( d_{i}(\ell) = d_{i}(\ell - 1) \). Because of this, (52), (55), and (56), and because \( \frac{4\beta}{\alpha} > 1 \), we have \( d_{i}(\ell) \leq \left( \frac{4\beta}{\alpha} \right)^{M - m(\ell)}(M + 1 - m(\ell)) \prod_{k' \in B(\ell)}(M + 1 - m(k')) \epsilon \). Now suppose \( i \in \left( \bigcup_{j \in \mathcal{I}(\ell) \cup \mathcal{L}(\ell)} C_{j}(k, \ell - 1) \cup \mathcal{J}(\ell) \right) - \mathcal{L}(\ell) \). Also, write \( \{ C_{j}(k, \ell - 1) \}_{j \in \mathcal{I}(\ell) \cup \mathcal{L}(\ell)} \) as \( \{ C_{j_{1}}(k, \ell - 1), C_{j_{2}}(k, \ell - 1), \ldots, C_{j_{p}}(k, \ell - 1) \} \), where \( 2 \leq p \leq m(\ell - 1) \). Then, from (32),

\[ C_{i}(k, \ell) = \left( \bigcup_{q=1}^{p} C_{j_{q}}(k, \ell - 1) \cup \mathcal{J}(\ell) \right) - \mathcal{L}(\ell). \tag{57} \]

Let \( s_{q} \in C_{j_{q}}(k, \ell - 1) \cap (\mathcal{I}(\ell) \cup \mathcal{L}(\ell)) \forall q \in \{ 1, 2, \ldots, p \} \). Then, because of Lemma 4 (57), (47), (9), (44), the triangle inequality, (50), (52), (55), and (56), we have

\[
d_{i}(\ell) \leq \beta \sum_{j \in \left( \bigcup_{q=1}^{p} C_{j_{q}}(k, \ell - 1) \cup \mathcal{J}(\ell) \right) - \mathcal{L}(\ell)} \| z_{j}(\ell) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell)}^{2} = \beta \sum_{q=1}^{p} \sum_{j \in C_{j_{q}}(k, \ell - 1) - (\mathcal{I}(\ell) \cup \mathcal{L}(\ell))} \| z_{j}(\ell - 1) - z_{C_{i^*}(k, \ell - 1)}^{\ell - 1} \|_{Q_{j}(\ell - 1)}^{2}.
\]
\[
\leq \beta \sum_{q=1}^{p} \sum_{j \in C_{pq}(k, \ell-1)} \left( \|z_j(\ell - 1) - z_{s_q}(\ell - 1)\| + \|z_{s_q}(\ell - 1) - z_{C_{pq}(k, \ell-1)}^{\ell-1}\| \right)^2 \\
\leq \beta \sum_{q=1}^{p} \sum_{j \in C_{pq}(k, \ell-1)} 2\left( \|z_j(\ell - 1) - z_{C_{pq}(k, \ell-1)}^{\ell-1}\| + \|z_{s_q}(\ell - 1) - z_{C_{pq}(k, \ell-1)}^{\ell-1}\| \right)^2 \\
+ \|z_{s_q}(\ell - 1) - z_{C_{pq}(k, \ell-1)}^{\ell-1}\|^2 \\
\leq \beta \sum_{q=1}^{p} \sum_{j \in C_{pq}(k, \ell-1)} 2\left( \frac{d_{j_q}(\ell - 1)}{\alpha} + \frac{\epsilon}{\alpha} \right) \\
\leq |C_i(k, \ell)| \left( \frac{4\beta}{\alpha} M - m(\ell - 1) + 1 \right) (M + 1 - m(\ell - 1)) \left( \prod_{k' \in B(\ell - 1)} (M + 1 - m(k')) \right) \epsilon + \frac{2\beta}{\alpha} \epsilon \\
\leq |C_i(k, \ell)| \left( \frac{4\beta}{\alpha} M - m(\ell - 1) \right) (M - 1 - m(\ell)) \left( \prod_{k' \in B(\ell - 1)} (M + 1 - m(k')) \right) \epsilon \\
= |C_i(k, \ell)| \left( \frac{4\beta}{\alpha} M - m(\ell) \right) \left( \prod_{k' \in B(\ell)} (M + 1 - m(k')) \right) \epsilon.
\]

This, along with the fact that \( |C_i(k, \ell)| \leq M + 1 - m(\ell) \), implies that \( d_i(\ell) \leq \left( \frac{4\beta}{\alpha} \right) M - m(\ell) (M + 1 - m(\ell)) \left( \prod_{k' \in B(\ell)} (M + 1 - m(k')) \right) \epsilon \). Therefore, (51) holds for Case (II). \( \square \)

Since \( C_i(k, k + h(k)) = \mathcal{M}(k + h(k)) \forall i \in \mathcal{M}(k + h(k)) \), we have \( m(k + h(k)) = 1 \). Also, notice that \( \Pi_{k' \in B(k+h(k))} (M + 1 - m(k')) \leq M! \). Furthermore, note from (45), (20), (21), (22), and (23) that \( z = z_{C_{pq}(k, \ell)}^{\ell} \forall \ell \in \mathbb{N} \), implying that \( d_i(k + h(k)) = V(k + h(k)) \forall i \in \mathcal{M}(k + h(k)) \). It follows from Lemma 5 and (48) that \( V(k + h(k)) \leq \left( \frac{4\beta}{\alpha} \right) M - 1 \cdot M! \cdot \epsilon \leq \left( \frac{4\beta}{\alpha} \right) M - 1 \cdot M! V(k) \), which contradicts the assumption that (39) is violated. Consequently, (39) holds.

A.7 Proof of Theorem 2

Let \( \mathcal{A} \) be given. Suppose the agent network is connected under \( \mathcal{A} \), i.e., \( h(k) < \infty \forall k \in \mathbb{N} \), and \( \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} \) is uniformly positive definite under \( \mathcal{A} \). Let \( \alpha > 0 \) be such that \( Q_i(k) - \alpha I \in \mathbb{S}^+ \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k) \). Then, (41) holds if and only if \( \lim_{k \to \infty} V(k) = 0 \). To show that \( \lim_{k \to \infty} V(k) = 0 \), note from (8) and Lemma 11 that \( V(k)^{\infty}_{k=0} \) is nonnegative and non-increasing. Thus, \( \exists \epsilon \geq 0 \) such that \( \lim_{k \to \infty} V(k) = c \). To show that \( c = 0 \), assume, to the contrary, that \( c > 0 \). Let \( \epsilon = \frac{c}{\left( \frac{4\beta}{\alpha} \right) M - 1 \cdot M!} \). Then, \( \exists k \in \mathbb{N} \) such that \( c \leq V(\ell) < c + \epsilon \forall \ell \geq k \). However, by Lemma 2
we have $V(k + h(k)) < \frac{(\frac{4\beta}{\alpha})^{M-1}M!}{(\frac{4\beta}{\alpha})^{M-1}M! + 1} (c + \epsilon) = c$, which contradicts the inequality $c \leq V(\ell)$ for $\ell = k + h(k)$. Therefore, $c = 0$, i.e., $\lim_{k \to \infty} V(k) = 0$, so that (40) holds.

A.8 Proof of Theorem 3

Let $\mathcal{A}$ be given. Suppose the agent network is uniformly connected under $\mathcal{A}$, i.e., $h^* < \infty$, and \{Q_i(k)\}_{k \in \mathbb{N}, i \in \mathcal{M}(k)} is uniformly positive definite under $\mathcal{A}$. Let $\alpha > 0$ be such that $Q_i(k) - \alpha I \in S^n_+$ $\forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. Then, it follows from (34), Lemma 1 and Lemma 2 that $\forall \ell \in \mathbb{N}, V((\ell + 1)h^*) \leq V(\ell h^* + h(\ell h^*)) \leq \frac{(\frac{4\beta}{\alpha})^{M-1}M!}{(\frac{4\beta}{\alpha})^{M-1}M! + 1} V(\ell h^*)$, which implies that $V(\ell h^*) \leq \left(\frac{(\frac{4\beta}{\alpha})^{M-1}M!}{(\frac{4\beta}{\alpha})^{M-1}M! + 1}\right)^\ell V(0)$. Due again to Lemma 1, (41) holds. In addition, from (38), $\alpha \|z_i(k) - z\|^2 \leq V(k) \forall k \in \mathbb{N} \forall i \in \mathcal{M}(k)$. Therefore, (42) is satisfied.

References

[1] L. Xiao, S. Boyd, and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” in Proc. International Symposium on Information Processing in Sensor Networks, Los Angeles, CA, 2005, pp. 63–70.

[2] ———, “A space-time diffusion scheme for peer-to-peer least-squares estimation,” in Proc. International Conference on Information Processing in Sensor Networks, Nashville, TN, 2006, pp. 168–176.

[3] J. N. Tsitsiklis, “Problems in decentralized decision making and computation,” Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1984.

[4] D. Kempe, A. Dobra, and J. Gehrke, “Gossip-based computation of aggregate information,” in Proc. IEEE Symposium on Foundations of Computer Science, Cambridge, MA, 2003, pp. 482–491.

[5] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” IEEE Transactions on Automatic Control, vol. 49, no. 9, pp. 1520–1533, 2004.

[6] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” Systems & Control Letters, vol. 53, no. 1, pp. 65–78, 2004.

[7] D. S. Scherber and H. C. Papadopoulos, “Distributed computation of averages over ad hoc networks,” IEEE Journal on Selected Areas in Communications, vol. 23, no. 4, pp. 776–787, 2005.
[8] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[9] M. Cao, D. A. Spielman, and E. M. Yeh, “Accelerated gossip algorithms for distributed computation,” in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2006, pp. 952–959.

[10] J.-Y. Chen, G. Pandurangan, and D. Xu, “Robust computation of aggregates in wireless sensor networks: Distributed randomized algorithms and analysis,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 9, pp. 987–1000, 2006.

[11] J. Cortés, “Finite-time convergent gradient flows with applications to network consensus,” *Automatica*, vol. 42, no. 11, pp. 1993–2000, 2006.

[12] D. B. Kingston and R. W. Beard, “Discrete-time average-consensus under switching network topologies,” in *Proc. American Control Conference*, Minneapolis, MN, 2006, pp. 3551–3556.

[13] C. C. Moallemi and B. Van Roy, “Consensus propagation,” *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4753–4766, 2006.

[14] A. Olshevsky and J. N. Tsitsiklis, “Convergence rates in distributed consensus and averaging,” in *Proc. IEEE Conference on Decision and Control*, San Diego, CA, 2006, pp. 3387–3392.

[15] M. Mehyar, D. Spanos, J. Pongsajapan, S. H. Low, and R. M. Murray, “Asynchronous distributed averaging on communication networks,” *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 512–520, 2007.

[16] A. Tahbaz-Salehi and A. Jadbabaie, “Small world phenomenon, rapidly mixing Markov chains, and average consensus algorithms,” in *Proc. IEEE Conference on Decision and Control*, New Orleans, LA, 2007, pp. 276–281.

[17] L. Xiao, S. Boyd, and S.-J. Kim, “Distributed average consensus with least-mean-square deviation,” *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.

[18] F. Fagnani and S. Zampieri, “Randomized consensus algorithms over large scale networks,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 634–649, 2008.

[19] M. Zhu and S. Martínez, “Dynamic average consensus on synchronous communication networks,” in *Proc. American Control Conference*, Seattle, WA, 2008, pp. 4382–4387.

[20] A. Olshevsky and J. N. Tsitsiklis, “Convergence speed in distributed consensus and averaging,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.
[21] B. N. Oreshkin, M. J. Coates, and M. G. Rabbat, “Optimization and analysis of distributed averaging with short node memory,” *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2850–2865, 2010.

[22] A. Tahbaz-Salehi and A. Jadbabaie, “Consensus over ergodic stationary graph processes,” *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 225–230, 2010.

[23] J. Lu and C. Y. Tang, “Controlled hopwise averaging and its convergence rate,” *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 1005–1012, 2012.

[24] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, “Distributed sensor fusion using dynamic consensus,” in *Proc. IFAC World Congress*, Prague, Czech Republic, 2005.