Consensus-Based Distributed Optimization with Malicious Nodes

Shreyas Sundaram Bahman Gharesifard

***Abstract*— We investigate the vulnerabilities of consensus-based distributed optimization protocols to nodes that deviate from the prescribed update rule (e.g., due to failures or adversarial attacks). After characterizing certain fundamental limitations on the performance of any distributed optimization algorithm in the presence of adversaries, we propose a robust consensus-based distributed optimization algorithm that is guaranteed to converge to the convex hull of the set of minimizers of the non-adversarial nodes' functions. We also study the distance-to-optimality properties of our proposed robust algorithm in terms of F-local sets of the graph. We show that finding the largest size of such sets is NP-hard.**

I. Introduction

The distributed optimization problem consists of a group of agents that are each equipped with an individual objective function, and the objective is for the agents to agree on a state that optimizes the sum of these functions. As in the consensus problem, the agents only use local information obtained from their neighboring agents, described by a communication network. There is a vast literature devoted to designing distributed algorithms, both in discrete and continuous-time, that guarantee convergence to an optimizer of the sum of the objective functions, under reasonable convexity and continuity assumptions [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12].

As outlined above, the predominant assumption in distributed optimization is that all agents cooperate to calculate the global optimizer. In particular, in most distributed optimization protocols, the individuals update their state using a combination of an agreement term and an appropriately scaled gradient flow of their individual functions. It is hence reasonable to ask how resilient such algorithms are with respect to failure or malicious behavior by certain nodes. Such resilience issues have been recently studied for consensus dynamics (e.g., see [13], [14], [15]). As we argue in this paper, current consensus-based distributed optimization algorithms are vulnerable to adversarial behavior. We subsequently characterize fundamental limitations on the performance of distributed optimization algorithms in the presence of adversaries, and show that there is an inherent tradeoff between performance and resilience. We then design a consensus-based algorithm that provides certain safety guarantees against adversarial behavior, and characterize graph properties that affect the performance of this algorithm.[1] The proofs of our main results are omitted in this paper in the interests of space, and will appear elsewhere.

Shreyas Sundaram is with the School of Electrical and Computer Engineering at Purdue University, W. Lafayette, IN, USA. Email: sundara2@purdue.edu. Bahman Gharesifard is with the Department of Mathematics and Statistics at Queen's University, Kingston, ON, Canada. Email: bahman@queensu.ca. The work of the second author was partially supported by the Natural Sciences and Engineering Research Council of Canada.

II. Mathematical Notation and Terminology

Let $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ denote the real and nonnegative real numbers, respectively, $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^n$, $\mathbf{1} = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}'$, $\mathbf{0} = \begin{bmatrix}0 & 0 & \cdots & 0\end{bmatrix}'$, and I_n the identity matrix in $\mathbb{R}^{n\times n}$. A matrix $A \in \mathbb{R}^{n\times n}$ with nonnegative entries is called (row) stochastic if $A\mathbf{1} = \mathbf{1}$. Throughout this paper, we are concerned with stochastic matrices whose diagonal entries are bounded away from zero.

A graph $\mathcal{G} = (V, \mathcal{E})$ consists of a set of *vertices* (or *nodes*) $V = \{v_1, v_2, \ldots, v_n\}$, and a set of *edges* $\mathcal{E} \subset V \times V$. The graph is said to be *undirected* if $(v_i, v_j) \in \mathcal{E} \Leftrightarrow (v_j, v_i) \in \mathcal{E}$, and *directed* otherwise. The *in-neighbors* and *out-neighbors* of vertex $v_i \in V$ are given by the sets $\mathcal{N}_i^- \triangleq \{v_j \in V \mid (v_j, v_i) \in \mathcal{E}\}$ and $\mathcal{N}_i^+ \triangleq \{v_j \in V \mid (v_i, v_j) \in \mathcal{E}\}$, respectively. The *in-degree* and *out-degree* of vertex $v_i \in V$ are given by $d_i^- \triangleq |\mathcal{N}_i^-|$ and $d_i^+ \triangleq |\mathcal{N}_i^+|$, respectively. For undirected graphs, we denote $\mathcal{N}_i = \mathcal{N}_i^- = \mathcal{N}_i^+$ as the *neighbors* of vertex $v_i \in V$, and $d_i = d_i^- = d_i^+$ as the *degree*.

A *path* from vertex $v_i \in V$ to vertex $v_j \in V$ is a sequence of vertices $v_{k_1}, v_{k_2}, \ldots, v_{k_l}$ such that $v_{k_1} = v_i$, $v_{k_l} = v_j$ and $(v_{k_r}, v_{k_{r+1}}) \in \mathcal{E}$ for $1 \leq r \leq l-1$. A graph $\mathcal{G} = (V, \mathcal{E})$ is said to be *rooted at vertex* $v_i \in V$ if for all vertices $v_j \in V \setminus \{v_i\}$, there a path from v_i to v_j. A graph is said to be *rooted* if it is rooted at some vertex $v_i \in V$. A graph is *strongly connected* if there is a path from every vertex to every other vertex in the graph.

For any $r \in \mathbb{N}$, a subset $S \subset V$ of vertices is said to be *r-local* if $|\mathcal{N}_i^- \cap S| \leq r$ for all $v_i \in V \setminus S$. In other words, if S is r-local, there are at most r vertices from S in the in-neighborhood of any vertex from $V \setminus S$. A *maximum r-local* set is an r-local set of largest cardinality (i.e., there are no r-local sets of larger size). A subset $S \subset V$ of vertices is said to be *r-reachable* if there exists a vertex $v_i \in S$ such that $|\mathcal{N}_i^- \setminus S| \geq r$. In other words, S is r-reachable if it contains a vertex that has at least r in-neighbors from outside S. A graph $\mathcal{G}$ is said to be *r-robust* if for all pairs

[1] The recent work [16], developed independently and in parallel, also considers the problem of distributed optimization with adversaries under different assumptions on the graph topology, faulty behavior and classes of functions than the ones that we consider here.

of disjoint nonempty subsets $S_1, S_2 \subset V$, at least one of S_1 or S_2 is r-reachable.

The following result (from Lemma 6 and Lemma 7 in [13]) will be useful for our analysis.

Lemma 2.1: Suppose a graph $\mathcal{G}$ is r-robust. Let $\mathcal{G}'$ be a graph obtained by removing $r-1$ or fewer incoming edges from each node in $\mathcal{G}$. Then $\mathcal{G}'$ is rooted.

III. Review of Consensus-Based Distributed Optimization

Consider a network of n agents $V = \{v_1, \dots, v_n\}$ whose communication topology is, for now, a time-invariant strongly connected graph $\mathcal{G} = (V, \mathcal{E})$. An edge $(v_i, v_j) \in \mathcal{E}$ indicates that v_j can receive information from v_i. For each $i \in \{1, \dots, n\}$, let $f_i : \mathbb{R}^d \to \mathbb{R}$ be locally Lipschitz and convex, and only available to agent v_i. The objective is for the agents to solve, in a distributed way (i.e., by exchanging information only with their immediate neighbors), the global optimization problem

$$\text{minimize} \quad f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x). \tag{1}$$

A common approach to solve this problem is to use a synchronous iterative consensus-based protocol in which agents use a combination of consensus dynamics and gradient flow to find a minimizer of f [4], [6], [17]. Specifically, at each time-step $t \in \mathbb{N}$, each agent $v_i \in V$ has an estimate $x_i(t) \in \mathbb{R}^d$ of the solution to the problem (1). Each agent $v_i \in V$ sends its estimate to its out-neighbors, receives the estimates of its in-neighbors, and updates its estimate as [4]

$$x_i(t+1) = a_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{N}_i^-} a_{ij}(t)x_j(t) - \alpha_t d_i(t). \tag{2}$$

In the above update rule, the quantities $a_{ij}(t)$, $v_j \in \{v_i\} \cup \mathcal{N}_i^-$ are a set of nonnegative real numbers satisfying $a_{ii}(t) + \sum_{v_j \in \mathcal{N}_i^-} a_{ij}(t) = 1$. In other words, the first portion of the right hand side is a *consensus step*, representing a weighted average of the estimates in node v_i's neighborhood. The quantity $d_i(t)$ is a subgradient of $f_i(x)$, evaluated at $a_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{N}_i^-} a_{ij}(t)x_j(t)$. The quantities α_t, $t \in \mathbb{N}$ are a sequence of *step-sizes*, indicating the influence of the subgradient on the update rule at each time-step. Thus, the last term in the above expression represents a *gradient step*.

There are some typical assumptions that are made on the weights in the update rule (2), which we encapsulate below.

Assumption 3.1 (Lower Bounded Weights): There exists a constant $\eta > 0$ such that $\forall t \in \mathbb{N}$, $\forall v_i \in V$, if $v_j \in \{v_i\} \cup \mathcal{N}_i^-$, then $a_{ij}(t) \geq \eta$.

Assumption 3.2 (Double Stochasticity): For all $t \in \mathbb{N}$, $v_i \in V$, the weights satisfy $a_{ii}(t) + \sum_{v_j \in \mathcal{N}_i^+} a_{ji}(t) = 1$.

The following result was established in [6].[2]

Proposition 3.3: Suppose the network $\mathcal{G}$ is time-invariant and strongly connected. Suppose the subgradients of each of the local functions $f_i(x)$ are bounded, i.e., $\exists L \in \mathbb{R}_{>0}$ such that $\|d\| \leq L$ for all $d \in \partial f_i(x)$ and $x \in \mathbb{R}^d$. Consider the update rule (2), and suppose the weights satisfy Assumption 3.1 and Assumption 3.2. Let the step-sizes satisfy $\sum_{t\in\mathbb{N}} \alpha_t = \infty$ and $\sum_{t\in\mathbb{N}} \alpha_t^2 < \infty$. Then there is a minimizer x^* of (1) such that

$$\lim_{t\to\infty} \|x_i(t) - x^*\| = 0$$

for all $v_i \in V$.

The above result shows that the update rule (2) allows the nodes in the network to distributively solve the global optimization problem (1), under appropriate assumptions on the weights, step-sizes and subgradients. Our main objective in this paper is to investigate the vulnerabilities of such protocols to nodes that *deviate* from the prescribed update rule (e.g., due to failures or adversarial attacks), and to develop a resilient distributed optimization algorithm that has provable safety guarantees in the presence of such deviations. To do this, we will first need to generalize the above analysis to handle cases where the weights are not doubly-stochastic. We restrict attention to scalar optimization problems (i.e., $f_i : \mathbb{R} \to \mathbb{R}$) throughout the rest of the paper.

A. Scenarios with Non-Doubly-Stochastic Weights

Here, we will establish convergence of the node states under the dynamics (2) under certain classes of non-doubly-stochastic consensus weights. At each time-step $t \in \mathbb{N}$, let $A(t) \in \mathbb{R}_{\geq 0}^{n\times n}$ be the matrix containing the weights $a_{ij}(t)$. Note that $a_{ij}(t) = 0$ if $v_j \notin \mathcal{N}_i^-$. Suppose there exists some constant $\beta > 0$ such that at each time-step $t \in \mathbb{N}$, $A(t)$ has a rooted subgraph that has edge-weights lower-bounded by β, and diagonal elements lower-bounded by β. Let $\Phi(t, s) \triangleq A(t)A(t-1)\cdots A(s)$ for $t \geq s \geq 0$. Using the fact that $A(t)$ has a rooted subgraph, and with an argument similar to the one in [18] which we omit here, for each $s \in \mathbb{N}$, there exists a stochastic vector $\mathbf{q}_s$ such that

$$\lim_{t\to\infty} \Phi(t, s) = \mathbf{1}\mathbf{q}_s'. \tag{3}$$

Noting that $\Phi(t, s) = \Phi(t, s+1)A(s)$, we have that

$$\mathbf{q}_s' = \mathbf{q}_{s+1}'A(s) \tag{4}$$

for all $s \in \mathbb{N}$.

For each $t \in \mathbb{N}$, let $x(t) \in \mathbb{R}^n$ be the state vector for the network, and define the quantity

$$y(t) \triangleq \mathbf{q}_t'x(t) \tag{5}$$

[2]The result in [6] is actually for a more general setting involving constrained optimization problems and time-varying graphs. For the purposes of this paper, it suffices to consider the version of the result given in Proposition 3.3.

(i.e., $y(t)$ is a convex combination of the states of the nodes at time-step t). We have

$$\begin{aligned} y(t+1) = \mathbf{q}'_{t+1}x(t+1) &= \mathbf{q}'_{t+1}A(t)x(t) - \alpha_t \mathbf{q}'_{t+1}d(t) \\ &= \mathbf{q}'_t x(t) - \alpha_t \mathbf{q}'_{t+1}d(t) \\ &= y(t) - \alpha_t \mathbf{q}'_{t+1}d(t), \end{aligned} \tag{6}$$

where $d(t)$ is the vector of subgradients. Unrolling the iteration, we obtain

$$y(t) = y(0) - \sum_{s=0}^{t-1} \alpha_s \mathbf{q}'_{s+1} d(s) \tag{7}$$

for $t \in \mathbb{N}$. Using the above definition, we can now provide the following convergence result. The proof of this result closely follows the proof for doubly-stochastic weights provided in [6], with the main difference being in the use of the vector $\mathbf{q}_t$ at appropriate points.

Lemma 3.4: Consider the network $\mathcal{G} = (V, \mathcal{E})$. Suppose that the functions $f_i(x)$, $v_i \in V$ have subgradients bounded by some constant L, and that the nodes run the dynamics (2). Assume that there exists a constant $\beta > 0$ such that at each time-step $t \in \mathbb{N}$, the weight matrix $A(t)$ has diagonal elements lower bounded by β and contains a rooted subgraph whose edge weights are lower bounded by β. Let $y(t)$ be the corresponding sequence defined in (5).

(i) If $\alpha_t \to 0$ as $t \to \infty$, then

$$\limsup_{t\to\infty} \|x(t) - \mathbf{1}y(t)\| = 0.$$

(ii) If $\sum_{t=1}^{\infty} \alpha_t^2 < \infty$, then

$$\sum_{t=1}^{\infty} \alpha_t \|x(t) - \mathbf{1}y(t)\| < \infty.$$

The above results lead to the following theorem, showing convergence of the individual node trajectories under dynamics (2) to a minimizer of a certain convex combination of the individual functions.

Theorem 3.5: Consider the network $\mathcal{G} = (V, \mathcal{E})$. Suppose that the functions $f_i(x)$, $v_i \in V$ have subgradients bounded by some constant L, and that the nodes run the dynamics (2). Assume that there exists a constant $\beta > 0$ such that at each time-step $t \in \mathbb{N}$, the weight matrix $A(t)$ has diagonal elements lower bounded by β and contains a rooted subgraph whose edge weights are lower bounded by β. Furthermore, suppose each matrix $A(t)$, $t \in \mathbb{N}$ has a common left-eigenvector $\mathbf{q}'$ corresponding to eigenvalue 1. Let the step sizes satisfy $\sum \alpha_t = \infty$ and $\sum \alpha_t^2 < \infty$. Then

$$\lim_{t\to\infty} \|x_i(t) - x^*\| = 0$$

for all $v_i \in V$, where x^* is a minimizer of $\sum_{i=1}^n q_i f_i(x)$ (q_i is the i-th element of $\mathbf{q}'$).

Note that if all matrices $A(t)$ do not have a common left-eigenvector, convergence to a constant value is not guaranteed under the dynamics (2). To see this, consider two row-stochastic matrices A_1 and A_2, each with rooted subgraphs and nonzero diagonal elements, with different left eigenvectors $\mathbf{q}_1$ and $\mathbf{q}_2$, respectively. Select the functions for the nodes such that $\sum q_{1i} f_i(x)$ and $\sum q_{2i} f_i(x)$ have different minimizers, where q_{ij} is the j-th component of $\mathbf{q}_i$. Then, if the dynamics evolve according to matrix A_1 for a sufficiently large period of time, all nodes will approach the minimizer of $\sum q_{1i} f_i(x)$, regardless of the initial conditions. Similarly, if the dynamics evolve according to the matrix A_2 for a sufficiently large period of time, all nodes will approach the minimizer of $\sum q_{2i} f_i(x)$, again regardless of the initial conditions. Thus, by appropriately switching between the matrices A_1 and A_2, the nodes will continually oscillate between the two different minimizers.

IV. Adversary Model and Vulnerabilities of Distributed Optimization Algorithms

With the results on distributed optimization with non-doubly-stochastic weights in hand, we now turn our attention to the effect of adversaries on the optimization dynamics.

A. Adversary Model

We partition the set of nodes V into two subsets: $\mathcal{M}$ containing a set of *malicious (or adversarial) nodes*, and $\mathcal{R} = V \setminus \mathcal{M}$ consisting of *regular nodes*. The regular nodes will exactly follow any algorithm that is prescribed, while the adversarial nodes are allowed to update their states in a completely arbitrary (potentially worst-case and coordinated) manner. We will allow the adversarial nodes to know the entire network topology and the optimization functions of all other nodes, in keeping with the goal of providing resilience to worst-case behavior. Clearly there is no hope of achieving any optimization objective if all nodes in the network are adversarial. We will later place restrictions on the number of adversarial nodes in the neighborhood of any regular node, but for now, we will investigate the ability of adversarial nodes to disrupt distributed optimization algorithms of the form (2).

B. Attacking Consensus-Based Distributed Optimization Algorithms

We start with the following result showing that it is extremely simple for even a single malicious node to disrupt dynamics of the form (2). The proof is a consequence of Theorem 3.5.

Proposition 4.1: Consider the network $\mathcal{G} = (V, \mathcal{E})$, and let there be a single adversarial node $\mathcal{M} = \{v_n\}$. Suppose the network is rooted at v_n. Then if v_n keeps its value fixed at some constant $\bar{x} \in \mathbb{R}$ and the stepsizes satisfy $\alpha_t \to 0$, all regular nodes will asymptotically converge to $\bar{x}$ when following the distributed optimization dynamics (2).

C. Fundamental Limitations on the Performance of Any Distributed Optimization Algorithm

Theorem 4.2: Suppose the local objective functions at each node are convex, but otherwise completely arbitrary.

Suppose Γ is a distributed algorithm that guarantees that all nodes calculate the global optimizer of problem (1) when there are no malicious nodes. Then a single adversary can cause all nodes to converge to any arbitrary value when they run algorithm Γ, and furthermore, will remain undetected.

The sketch of the proof of the above result is as follows. Suppose the adversarial node is v_n (without loss of generality), and wishes all nodes to converge to some value $\bar{x}$. Then the adversarial node simply chooses a function $\bar{f}_n(x)$ such that the minimizer of $\sum_{i=1}^{n-1} f_i(x) + \bar{f}_n(x)$ is $\bar{x}$. It then participates in the algorithm pretending that its function is $\bar{f}_n(x)$. Since the functions are arbitrary and known only to the nodes themselves, this deception cannot be detected.

The above theorem applies to *any* algorithm that is guaranteed to output the globally optimum value in the absence of adversaries. The takeaway point is that **the price paid for resilience is a loss in optimality**: it is *impossible* to develop an algorithm that always finds optimal solutions in the absence of adversaries and that is also resilient to carefully crafted attacks.

V. Robust Consensus-Based Distributed Optimization Protocols

In this section, we describe a modification of the traditional distributed optimization dynamics (2) that will allow the regular nodes in the network to mitigate the impact of the adversarial nodes. Specifically, suppose that the adversarial nodes are restricted to form an F-local set, where F is a nonnegative integer. The regular nodes do not know which (if any) of their neighbors are adversarial. At each time-step $t \in \mathbb{N}$, each regular node $v_i \in \mathcal{R}$ performs the following actions in parallel with the other regular nodes:

(i) Node v_i gathers the states $\{x_j(t), v_j \in \mathcal{N}_i^-\}$ of its in-neighbors.

(ii) Node v_i removes the highest F and lowest F states from the set of gathered states, breaking ties arbitrarily. Let $\mathcal{J}_i(t) \subset \mathcal{N}_i^-$ be the set of in-neighbors of v_i whose states were retained by v_i at time-step t.

(iii) Node v_i updates its state as

$$x_i(t+1) = a_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{J}_i(t)} a_{ij}(t)x_j(t) - \alpha_t d_i(t), \tag{8}$$

where $d_i(t)$ is a subgradient of f_i evaluated at $a_{ii}x_i(t) + \sum_{v_j \in \mathcal{J}_i(t)} a_{ij}(t)x_j(t)$, and α_t is a sequence of nonnegative stepsizes. At each time-step t and for each $i \in \mathcal{R}$, the weights $a_{ij}(t)$, $v_j \in \{v_i\} \cup \mathcal{J}_i(t)$ are lower-bounded by some strictly positive real number η and sum to 1 (i.e., they specify a convex combination).

The adversarial nodes are allowed to update their states however they wish. Note that the above dynamics are *purely-local* in the sense that they do not require the regular nodes to know anything about the network topology (other than their own in-neighbors). Also note that even when the underlying network $\mathcal{G}$ is time-invariant, the filtering operation induces *state-dependent switching* (i.e., the effective in-neighbor set $\mathcal{J}_i(t)$ is a function of the states of the in-neighbors of v_i at time-step t). Local filtering operations of the above form have been previously studied in the context of resilient consensus dynamics (i.e., outside of distributed optimization) in [13], [14], [19]. We will refer to the above dynamics as **Local Filtering (LF) Dynamics** with parameter F. We will analyze these dynamics in the remainder of the paper, and show that they are resilient to adversarial behavior under certain conditions on the network topology.

A. A Mathematically Equivalent Representation of Local Filtering Dynamics

Note that in the LF dynamics given by (8), a regular node v_i may use the value of one or more adversarial in-neighbors during its update at some time-step t, as long as the states of those adversarial nodes are not among the most extreme values in v_i's in-neighborhood. As we will only be concerned with understanding the evolution of the states of the regular nodes, it will be useful to consider a *mathematically equivalent* representation of the dynamics (8) that only involves the states of the regular nodes. The key idea of the proof of the following proposition is from [20].

Proposition 5.1: Consider the network $\mathcal{G} = (V, \mathcal{E})$, with a set of regular nodes $\mathcal{R}$ and a set of adversarial nodes $\mathcal{M}$. Suppose that $\mathcal{M}$ is an F-local set, and that each regular node has at least $2F+1$ in-neighbors. Then the update rule (8) for each node $v_i \in \mathcal{R}$ is mathematically equivalent to

$$x_i(t+1) = \bar{a}_{ii}(t)x_i(t) + \sum_{v_j \in \mathcal{N}_i^- \cap \mathcal{R}} \bar{a}_{ij}(t)x_j(t) - \alpha_t d_i(t), \tag{9}$$

where the nonnegative weights $\bar{a}_{ij}(t)$ satisfy the following properties at each time-step t:

(i) $\bar{a}_{ij}(t) = a_{ij}(t)$ if $v_j \in \{v_i\} \cup (\mathcal{J}_i(t) \cap \mathcal{R})$.

(ii) $\bar{a}_{ii}(t) + \sum_{v_j \in \mathcal{N}_i^- \cap \mathcal{R}} \bar{a}_{ij}(t) = 1$.

(iii) $\bar{a}_{ii}(t) \geq \eta$ and at least $|\mathcal{N}_i^-| - 2F$ of the other weights are lower bounded by $\frac{\eta}{2}$.

We emphasize again that the regular nodes run the dynamics (8) (which does not require them to know which of their neighbors is adversarial); the dynamics (9) are *mathematically equivalent* to the dynamics (8) due to the nature of the local filtering that is done by each regular node, and will be more convenient for us to analyze.

Henceforth, we assume without loss of generality that the regular nodes are arranged first in the ordering of the nodes, and define the vectors

$$\boldsymbol{x}_{\mathcal{R}}(t) \triangleq \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_{|\mathcal{R}|}(t) \end{bmatrix}'$$
$$\boldsymbol{d}_{\mathcal{R}}(t) \triangleq \begin{bmatrix} d_1(t) & d_2(t) & \cdots & d_{|\mathcal{R}|}(t) \end{bmatrix}'$$

to be the vectors of states and subgradients of the regular nodes, respectively. Based on Proposition 5.1, the network-wide dynamics of the regular nodes under the Local Filtering dynamics can be written as

$$\boldsymbol{x}_{\mathcal{R}}(t+1) = \bar{A}(t)\boldsymbol{x}_{\mathcal{R}}(t) - \alpha_t \boldsymbol{d}_{\mathcal{R}}(t), \tag{10}$$

where $\bar{A}(t) \in \mathbb{R}_{\geq 0}^{|\mathcal{R}|\times|\mathcal{R}|}$ contains the weights $\bar{a}_{ij}(t)$ from (9).

B. Convergence to Consensus

Theorem 5.2: Consider the network $\mathcal{G} = (V, \mathcal{E})$, with regular nodes $\mathcal{R}$ and an F-local set of adversarial nodes $\mathcal{M}$. Suppose the network is $(2F+1)$-robust, that the functions $f_i(x)$, $v_i \in \mathcal{R}$ have subgradients bounded by some constant L, and that the regular nodes run the Local Filtering dynamics (8) with parameter F. Further suppose that $\alpha_t \to 0$ as $t \to \infty$. Then, there exists a sequence of stochastic vectors $\mathbf{q}_t$, $t \in \mathbb{N}$, such that

$$\limsup_{t\to\infty} \|\boldsymbol{x}_{\mathcal{R}}(t) - \mathbf{1}y(t)\| = 0,$$

where $y(t) = \mathbf{q}_t' \boldsymbol{x}_{\mathcal{R}}(t)$.

The above result shows that the network being $(2F+1)$-robust is sufficient for the nodes to reach consensus under the LF-based distributed optimization dynamics (8) when each node discards the highest F and lowest F values in its neighborhood at each time-step, and $\alpha_t \to 0$. Note that this holds true *regardless* of the actions of the adversaries, as long as those adversaries form an F-local set.

C. A Safety Condition: Convergence to the Convex Hull of the Local Minimizers

The following result shows that the LF dynamics provide a *safety guarantee* for distributed optimization. Specifically, under the conditions of the theorem, all regular nodes will asymptotically converge to the convex hull of the minimizers of the regular nodes' functions.

Theorem 5.3: Suppose that the set of malicious nodes forms an F-local set and that the underlying graph is $(2F+1)$-robust. Suppose that all regular nodes follow the LF dynamics (8) with parameter F. For each node $v_i \in \mathcal{R}$, let the local function $f_i(\cdot)$ have subgradients bounded by L, and have minimizer $m_i \in \mathbb{R}$. Define $\overline{M} = \max\{m_i \mid v_i \in \mathcal{R}\}$ and $\underline{M} = \min\{m_i \mid v_i \in \mathcal{R}\}$. If the stepsizes satisfy $\sum \alpha_t = \infty$ and $\alpha_t \to 0$, then $\limsup_{t\to\infty} x_i(t) \leq \overline{M}$ and $\liminf_{t\to\infty} x_i(t) \geq \underline{M}$ for all $v_i \in \mathcal{R}$, regardless of the actions of the malicious nodes and the initial values.

D. Lack of Convergence to a Constant Value Under Malicious Behavior

Although the LF dynamics prevent adversarial nodes from driving the states of regular nodes to arbitrarily large values (as shown in Theorem 5.3), a single malicious node can behave in such a way that the regular nodes do not converge to a constant value when the stepsizes satisfy $\sum \alpha_t = \infty$ and $\sum \alpha_t^2 < \infty$, as we show below.

Consider a complete network $\mathcal{G} = (V, \mathcal{E})$ with n nodes. Consider two functions $f_a(x)$ and $f_b(x)$, with minimizers $a \in \mathbb{R}$ and $b \in \mathbb{R}$, respectively.

Suppose that up to $F = 1$ node in the network can be malicious. Let the first $n-1$ nodes in the network be regular (i.e., $\mathcal{R} = \{v_1, v_2, \ldots, v_{n-1}\}$), and let the last node be malicious (i.e., $\mathcal{M} = \{v_n\}$). Note that the regular nodes do not know that the last node is malicious.

Further partition the set of regular nodes into two sets $\mathcal{R}_1 = \{v_1, v_2, \ldots, v_{n-2}\}$ and $\mathcal{R}_2 = \{v_{n-1}\}$. For each $v_i \in \mathcal{R}_1$, let $f_i(x) = f_a(x)$, and let $f_{n-2}(x) = f_b(x)$. In other words, there are $n-2$ regular nodes that have the local function $f_a(x)$, and one regular node that has the local function $f_b(x)$. The following results assume that all nodes run the LF dynamics (8) with parameter $F = 1$, and with stepsizes satisfying $\sum \alpha_t = \infty$ and $\sum \alpha_t^2 < \infty$.

Lemma 5.4: Suppose the states of all nodes in $\mathcal{R}_1$ agree at some time-step t_0. Then the states of all nodes in $\mathcal{R}_1$ will agree for all subsequent time-steps, regardless of the actions of the malicious node.

Lemma 5.5: Suppose the states of all nodes in $\mathcal{R}_1$ agree at some time-step t_0. Suppose that starting from t_0, the malicious node v_n updates its state so that $x_n(t) = x_1(t)$ for all $t \geq t_0$ (i.e., the malicious node keeps its value the same as the nodes in $\mathcal{R}_1$). Then the states of all regular nodes will asymptotically converge to a (the minimizer of the function $f_a(x)$).

Lemma 5.6: Suppose the states of all nodes in $\mathcal{R}_1$ agree at some time-step t_0. Suppose that for all $t \geq t_0$, the malicious node v_n updates its state as follows: if $x_{n-1}(t) \geq x_1(t)$, then $x_n(t) > x_{n-1}(t)$, and if $x_{n-1}(t) < x_1(t)$ then $x_n(t) < x_{n-1}(t)$ (i.e., the malicious node always makes its value more extreme than that of v_{n-1}). Then the states of all regular nodes will asymptotically converge to the minimizer of the function $\frac{1}{n-2}\left((n-3)f_a(x) + f_b(x)\right)$.

The above results lead to the following negative result about the ability of malicious nodes to prevent convergence of the regular nodes to a constant value (although they will still reach consensus and asymptotically approach the convex hull of the minimizers of the regular nodes' functions).

Proposition 5.7: Consider a complete graph $\mathcal{G} = (V, \mathcal{E})$. Consider two different functions $f_a(x)$ and $f_b(x)$ with the property that the minimizer of $(n-3)f_a(x) + f_b(x)$ is different from the minimizer of $f_a(x)$. Suppose all regular nodes run the LF dynamics (8) with parameter $F = 1$, and stepsizes satisfying $\sum \alpha_t = \infty$ and $\sum \alpha_t^2 < \infty$. Then there is an allocation of $f_a(x)$ and $f_b(x)$ to regular nodes such that a single malicious node can prevent convergence of the regular nodes to a constant value.

VI. Bounds on Performance

Recall from Theorem 4.2 that an algorithm that provides resilience to adversaries cannot guarantee that the global optimum can be calculated in the absence of adversaries. Here, we show that under the LF dynamics (8), the nature of the individual optimization functions (together with the network topology) plays a role in determining how far away the convergence point is from the minimizer of the average of the regular nodes' functions.

Proposition 6.1: Consider a network $\mathcal{G} = (V, \mathcal{E})$ with n nodes and let $F \in \mathbb{N}$. Suppose that the network is $(2F+1)$-robust, and let $\mathcal{T} \subset V$ be a maximum F-local set. Suppose all nodes are regular. Pick any $a \in \mathbb{R}$, and let the nodes in set $\mathcal{T}$ have local function $f_a(x) = (x-a)^2$, and let the nodes in set $V \setminus \mathcal{T}$ have local function $f_0(x) = x^2$ (both functions can be modified to have their gradients capped at sufficiently large values, so as to not affect the minimizer of any convex combination of the functions). Let $f(x)$ be the average of all of the functions, with minimizer $x^* = \frac{|\mathcal{T}|}{n}a$. Then, under the local filtering dynamics with parameter F, all nodes converge to the value $\bar{x} = 0$ and thus $|\bar{x} - x^*| = \frac{|\mathcal{T}|}{n}|a|$, $f(\bar{x}) - f(x^*) = \frac{|\mathcal{T}|^2}{n^2}a^2$.

The above result shows that loss in performance under LF dynamics will depend on both the network topology and the nature of the local objective functions. Specifically, if the network contains a large F-local set (in relation to the total number of nodes) or the local functions have minimizers that are very different (corresponding to a large $|a|$ in the above result), then the value computed by the LF dynamics will have a greater divergence from the globally optimal solution. Note that the maximum F-local set in a graph will have size at least equal to F (since any set of size F is F-local).

An alternate interpretation of the above result is as follows. Recall from Theorem 5.3 that under the conditions in that theorem, all regular nodes asymptotically converge to the convex hull of the minimizers of the regular nodes' functions. Any point in this convex hull corresponds to the minimizer of some convex combination of the functions of the regular nodes. Suppose that $\mathcal{T}$ is an F-local set of maximum size in the graph. Proposition 6.1 thus indicates that under the LF dynamics, one cannot guarantee that more than $n - |\mathcal{T}|$ of the regular nodes' functions will play a role in the set of minimizers that the regular nodes converge to. It is worth noting that under the F-local model of adversarial behavior, *no resilient algorithm* can guarantee that more than $n - |\mathcal{T}|$ nodes will play a role in the final solution (as all nodes in $\mathcal{T}$ can potentially be malicious).

We conclude the paper by stating the complexity of finding the size of maximum r-local sets in graphs.

Definition 6.2: Let r, k be positive integers. The **r-Local Set Problem** is to determine whether a given graph has an r-local set of size at least k.

Theorem 6.3: The r-Local Set Problem is NP-complete.

The proof of the above theorem is via a reduction from the NP-complete Set Packing problem.

VII. Directions for Future Research

In this paper, we proposed a consensus-based distributed optimization algorithm that is resilient to adversarial behavior, in the sense that the regular nodes will always asymptotically converge to the convex hull of the minimizers of the regular nodes' functions, despite the actions of any F-local set of adversaries. We also identified topological properties (in the form of maximum F-local sets) that affect the performance of the algorithm. There are many interesting directions for future research, including a more explicit characterization of the distance-to-optimality of such algorithms, with corresponding conditions on the network topology and distribution of functions that lead to near-optimal solutions.

References

[1] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.

[2] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, Berkeley, CA, Apr. 2004, pp. 20–27.

[3] L. Xiao and S. Boyd, "Optimal scaling of a gradient method for distributed resource allocation," *Journal of Optimization Theory & Applications*, vol. 129, no. 3, pp. 469–488, 2006.

[4] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[5] P. Wan and M. D. Lemmon, "Event-triggered distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, San Francisco, CA, 2009, pp. 49–60.

[6] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[7] B. Johansson, M. Rabi, and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM Journal on Control and Optimization*, vol. 20, no. 3, pp. 1157–1170, 2009.

[8] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151–164, 2012.

[9] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *IEEE Conf. on Decision and Control*, Orlando, Florida, 2011, pp. 3800–3805.

[10] A. H. Sayed, "Adaptive networks," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 460–497, 2014.

[11] 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.

[12] A. Nedic and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2015.

[13] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 766–781, April 2013.

[14] N. H. Vaidya, L. Tseng, and G. Liang, "Iterative approximate Byzantine consensus in arbitrary directed graphs," in *ACM Symposium on Principles of Distributed Computing*, 2012, pp. 365–374.

[15] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1495–1508, 2011.

[16] L. Su and N. Vaidya, "Byzantine multi-agent optimization," *arXiv preprint arXiv:1506.04681*, 2015.

[17] A. Nedić and A. Ozdaglar, "Cooperative distributed multi-agent optimization," in *Convex optimization in signal processing and communications*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge University Press, 2010, pp. 340–386.

[18] M. Cao, S. A. Morse, and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: a graphical approach," *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 575–600, 2008.

[19] D. Dolev, N. A. Lynch, S. S. Pinter, E. W. Stark, and W. E. Weihl, "Reaching approximate agreement in the presence of faults," *Journal of the ACM*, vol. 33, pp. 499–516, May 1986.

[20] N. H. Vaidya, "Matrix representation of iterative approximate Byzantine consensus in directed graphs," *arXiv preprint arXiv:1203.1888*, 2012.