Stability and Resilience of Distributed Information Spreading in Aggregate Computing

Yuanqiu Mo, Member IEEE, Soura Dasgupta, Fellow IEEE and Jacob Beal, Senior Member IEEE

*Abstract*—Spreading information through a network of devices is a core activity for most distributed systems. Self-stabilizing algorithms for information spreading are one of the key building blocks enabling aggregate computing to provide resilient coordination in open complex distributed systems. This paper improves a general spreading block in the aggregate computing literature by making it resilient to network perturbations, establishes its global uniform asymptotic stability and proves that it is ultimately bounded under persistent disturbances. The ultimate bounds depend only on the magnitude of the largest perturbation and the network diameter, and three design parameters trade off competing aspects of performance. For example, as in many dynamical systems, values leading to greater resilience to network perturbations slow convergence and vice versa.

Index Terms—aggregate computing, multi-agent systems, distributed graph algorithms, nonlinear stability, ultimate bounds.

### I. INTRODUCTION

O of Things (IoT) require seamless distributed device coordination as individual devices must collaborate with their neighbors to perform multiple tasks, often under persistent perturbations caused by mobility, noise and device and link losses. Distributed algorithms using ordinary programming are ill-equipped to perform such coordination and delivering predictable performance as they overly focus on message passing and construction. Device coordination often involves compositions of these algorithms, sometimes in feedback. The analyses of compositions of ordinary programming algorithms are typically limited to self-stabilization. None exist for transient performance like time to converge or bounds under persistent perturbations, needed to guarantee safe operations.

Aggregate Computing [1] avoids these issues by making details of message passing implicit and enforcing abstraction barriers to information access, allowing each block in distributed coordination to be analyzed separately. Crucially,

Mo is with the Institute of Advanced Technology, Westlake Institute for Advanced Study, Westlake University, Hangzhou 310024, China (moyuanqiu@@westlake.edu.cn). Dasgupta is with the University of Iowa, Iowa City, Iowa 52242 USA (soura-dasgupta@uiowa.edu). S. Dasgupta is also a Visiting Professor at Shandong Computer Science Center, Shandong Provincial Key Laboratory of Computer Networks, China. Beal is with Raytheon BBN Technologies, Cambridge, MA, USA 02138 USA (jakebeal@ieee.org) compositions of three basis blocks realize most coordination tasks: G block spreads information from sources across a network according to some optimality criterion, C sends information from devices to sources, and T performs timing operations. Each is a distributed graph algorithm [1].

1

We consider a general G block that spreads such information across a network as, possibly non-Euclidean, distances of devices from a source set, *broadcasts* information from a set of sources to nearest devices, generates pathways away from hazardous areas or creates spanning trees used by C-blocks to apprise leaders of the net resources in a network, [2]. Unlike games networks these algorithms are cooperative.

The stability analysis of this block is limited to selfstablization [1], that too under the assumption that all states lie in Noetherian rings and are thus *a priori bounded*. Unlike global uniform asymptotic stability (GUAS) [3], selfstabilization has no notion of robustness to persistent perturbations *although these are inevitable under feedback*. Such feedback compositions are natural in device coordination. For example, [4] describes a setting where a controller provides a resource allocation plan to distributed services, by accepting load information as input in a feedback loop with G and C blocks. Mobility makes persistent perturbations ubiquitous in open systems: Algorithms that estimate distances from sources require that devices know their distance from neighbors. Mobility and localization errors manifest as persistent noise.

Thus we improve this general G-block to allow removal of the Noetherian assumption, proof of GUAS, and (under an additional Lipschitz condition) ultimate boundedness in face of persistent perturbations. Ultimate bounds permit the use of variants of the small gain theorem [5], or equivalent passivity theory as in [6] for closed loop analysis. This goes well beyond our analysis in [4], of the simplest G-block, the Adaptive Bellman-Ford (ABF) algorithm, which estimates Euclidean distances of nodes from a source set in a distributed manner and, unlike the classical Bellman-Ford algorithm [7], accommodates underestimates. In [8] we have analyzed *without proof* another special case, which generalizes ABF by allowing broadcast and other features given in Section II.

A key problem with both ABF and indeed the generalization in [8] is the *rising value problem*. All *G*-block algorithms generate estimates  $\hat{x}_i(t)$  that must converge to a value  $x_i$ . The rising value problem is when underestimates (i.e.  $\hat{x}_i(t) < x_i$ ) may rise very slowly. The more general *G*-block given in [1] and studied here removes this problem. Though proved to be self-stabilizing under a Noetherian assumption we explain in Section II-B that it is not robust to perturbations. We improve this *G*-block so that it can cope with non-negative numbers

This work has been supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001117C0049. The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This document does not contain technology or technical data controlled under either U.S. International Traffic in Arms Regulation or U.S. Export Administration Regulations. Approved for public release, distribution unlimited (DARPA DISTAR case 33519, 10/14/20).

without assuming a priori that they are in a bounded set and by introducing three design parameters that make the algorithm robust to persistent perturbations. As with most dynamical systems, we argue that these design choices compete between robustness and convergence speed. Greater amelioration of the rising value problem yields less robustness.

The vast literature of multi-agent systems, [9], [10], [11], [12], [13], [14], [15], [16], [17], comprising e.g., distributed filtering, consensus and formation control, also considers distributed algorithms. Distributed filtering, e.g., [9] of course considers performance under persistent noise as does [11]. Further, [15] and [17] consider persistent link losses and show that provided connectivity is maintained over every fixed interval, convergence still occurs. To our knowledge, the only paper, apart from our work in [8], [18], that proves ultimate boundedness is [10]. However, [10] does not provide the ultimate bound. Unlike [15] and [17], perturbations considered by us are on edge weights and we do quantify the ultimate bounds. Most proximate to the problem considered here is the path finding literature, e.g. [12], where the underlying assumption is that *perturbations vanish in the relevant part* of the network, before the objective is achieved. This is in contrast to this paper where the perturbations are persistent.

Section II has the algorithm, assumptions and motivating applications. Section III examines stationary points. Section IV proves GUAS and gives ultimate bounds proportional to bound on the perturbations. Section V discusses design choices. Section VI gives simulations. Section VII concludes.

#### II. ALGORITHM

In this section, we present a general G-block that spreads information through a network in a distributed fashion. Section II-A describes a special case shown to be GUAS in [8], with proofs omitted, plus examples and a shortcoming. Section II-B then presents a more general algorithm that removes this deficiency, and Section II-C provides assumptions and definitions that will be used for proofs in subsequent sections.

#### A. The Spreading block of [8]

Consider an undirected graph  $\mathcal{G} = (V, E)$  with nodes in  $V = \{1, 2, \dots, N\}$  and edge set E. Nodes i and k are *neighbors* if they share an edge. Denote  $\mathcal{N}(i)$  as the set of neighbors of i. The goal of the algorithm is to spread the information  $x_i \ge 0$  to node *i*. The simplest example of  $x_i = d_i$ is the distance of node i from a set of sources. This can be done with ABF, [4], that updates the distance estimate  $\hat{d}_i(t)$ of  $d_i$ . In this case, suppose  $e_{ik} > 0$  is the edge length between i and k and the signature of node i, obeys  $s_i = 0$  if i is a source and  $s_i = \infty$  otherwise. Then as for a source  $j, d_i = 0$ , ABF proceeds as

$$\hat{d}_{i}(t+1) = \begin{cases} 0 & s_{i} = 0\\ \min_{k \in \mathcal{N}(i)} \left\{ \hat{d}_{k}(t) + e_{ik} \right\} & s_{i} = \infty \end{cases}$$
(1)

Observe, with  $\hat{x}_i(t) = d_i(t)$ , (1) is a special case of the Gblock in [8] where:

$$\hat{x}_{i}(t+1) = \min\left\{\min_{k \in \mathcal{N}(i)} \left\{ f\left(\hat{x}_{k}(t), \bar{e}_{ik}(t)\right) \right\}, s_{i} \right\}, \forall t \ge t_{0}.$$
(2)

with initial estimates  $\hat{x}_i(t_0) \ge 0$ . For ABF,

$$f(\hat{x}_k(t), e_{ik}) = \hat{x}_k(t) + e_{ik}.$$
(3)

As in ABF source nodes have  $s_i = 0$ , under  $\hat{x}_i(t_0) > 0$ , the outer minimum ensures that all sources have  $d_i(t) = \hat{x}_i(t) =$ 0. Likewise, as non-source nodes in ABF have infinite  $s_i$ , (2) reduces to the second bullet of (1) for such nodes. As noted in the introduction, among other things this generalization captures potential non-Euclidean distance metrics.

The  $\bar{e}_{ik}(t)$  define such structural aspects as edge lengths between neighbors;  $s_i \ge 0$ , may be either finite or infinite. Due to the outer minimum in (2),  $s_i$  is the maximum value that  $\hat{x}_i(t)$  can acquire after the initial time. The argument t in  $\bar{e}_{ik}(t)$ anticipates analysis when there are persistent perturbations in them. In particular, we assume that perturbations are around the nominal values  $e_{ij}$  i.e.

$$0 < e_{\min} \le \bar{e}_{ij}(t) = e_{ij} + \epsilon_{ij}(t) \tag{4}$$

$$|\epsilon_{ij}(t)| \le \epsilon < e_{\min}.$$
(5)

We permit

$$\epsilon_{ij}(t) \neq \epsilon_{ji}(t),$$
 (6)

though we require

$$e_{\min} < e_{ij} = e_{ji}.\tag{7}$$

Such perturbations could reflect noise, localization error, or (if coherent) movement of devices. The asymmetry in (6) recognizes that  $\bar{e}_{ik}(t)$  is the noisy estimate of  $e_{ik}$  seen by *i* rather than k. The function  $f(\cdot, \cdot)$  is progressive i.e.

$$f(a,b) > a + \sigma, \ \sigma > 0, \tag{8}$$

$$f(a_1, b) \ge f(a_2, b), \text{ if } a_1 \ge a_2.$$
 (9)

and is finite for finite a and b. Note (3) obeys both. The initialization in (2) ensures that  $\hat{x}_i(t) \geq 0$ , for all  $t \geq t_0$ . Define  $S^*$  as the set of nodes with *finite maximum values*  $s_i$ :  $\mathcal{S}^* = \{ i \in V | s_i < \infty \}.$ (10)

We will assume that this set is nonempty. The state  $\hat{x}_i(t)$  is the estimate of the  $x_i$  to be spread to node i, in the perturbation free case of  $\epsilon = 0$  and are the *stationary* values of (2):

$$x_{i} = \min\left\{\min_{k \in \mathcal{N}(i)} \left\{f\left(x_{k}, e_{ik}\right)\right\}, s_{i}\right\}, \forall i \in V.$$
(11)

We show in Section III) that the  $x_i$  are unique, finite, with  $x_i = s_i$  for at least one *i*.

In (1) all nodes with finite signatures have distance estimates equaling  $s_i$ , in particular zero. Figure 1 presents another example of (3) where this need not be the case. Here, node A (red) is a high-speed gateway with  $s_A = 1$ , node D (purple) is a low-speed gateway with  $s_D = 5$ , and the others are nongateways with  $s_i = \infty$ . Through (2), all nodes try to route to external networks through the shortest effective path. After 3 rounds, all nodes, including the low-speed link, converge to route through the high-speed link. In this case the stationary state of the low-speed link is  $x_D = 4$  and does not equal  $s_i = 5$  even though  $s_i$  is finite. Should the high-speed link represented by node A disappear, then the state estimate of D does converge to its maximum value 5, while those of nodes B and C converge to 7 and 6, respectively, i.e. nodes reroute through the still available low-speed link.



Fig. 1: Non-zero  $s_i$  representing external gateways of a tactical wireless network: red and purple nodes are high and low speed external links, respectively, while blue are nodes without external links. Black numbers represent edge weights  $e_{ik}$ , green numbers represent state estimates  $\hat{x}_i$ . After 3 rounds, all nodes, including the low-speed link, have converged to route through the high-speed link.

In the previous examples f(a, b) is linear and increasing in b. A specialization of (2) violating both these properties finds the most probable path (MPP) from each node in a network to a source. In this case  $e_{ik}$  represents the probability of successful traversal or delivery between neighbors i and j. The stationary value  $x_i$  is the *smallest* probability of failure in delivery from node i to the source. In this case  $x_i = 0$  for sources. Assuming independence, for all other nodes

$$x_i = \min_{k \in \mathcal{N}(i)} \{1 - (1 - x_k)e_{ik}\}.$$
 (12)

The sequence of minimizing nodes k then indicates the MPP from node i to any source and can be computed using (2) with

$$f(\hat{x}_k(t), e_{ik}) = 1 - (1 - \hat{x}_k(t))e_{ik}.$$
(13)

If  $0 < \sigma \leq e_{ik} < 1 - \sigma$ , this is progressive and increasing in  $\hat{x}_k(t)$ , though decreasing in  $e_{ik}$ .

A key shortcoming of (2), however, is that underestimates can rise very slowly in the presence of small  $e_{ik}$ . Consider for example (1) with nodes 1 and 2 having the smallest estimates and sharing a short edge e. At successive instants  $d_1(t+1) =$  $d_2(t) + e$  and  $d_2(t+1) = d_1(t) + e$ , i.e. each rises in small increments of e (and as shown in [4]) converge slowly. The generalization below accelerates this slow convergence.

### B. A more general spreading block

A more general G-block, given in [1], ameliorates this rising value problem by speeding up the rise in  $\hat{x}_i(t)$ . Specifically, it introduces an auxiliary state  $\tilde{x}_i$  which is updated in a manner similar to (2), though on the basis of the immediate past value of neighboring  $\hat{x}_k(t)$  rather than  $\tilde{x}_k(t)$ . Thus,

$$\tilde{x}_{i}(t+1) = \min\left\{\min_{k \in \mathcal{N}(i)} \left\{ f\left(\hat{x}_{k}(t), \bar{e}_{ik}(t)\right) \right\}, s_{i} \right\}, \forall t \ge t_{0}.$$
(14)

The first argument in  $f(\cdot, \cdot)$  is  $\hat{x}_k$  rather than  $\tilde{x}_k$ . On the other hand  $\hat{x}_i(t)$  is itself raised by a fixed amount, independent of  $e_{ii}$ , unless it exceeds a value M or if it is within a dead zone parameter D of the updated  $\tilde{x}_i(t+1)$ . Specifically,

$$\hat{x}_i(t+1) = F\left(\tilde{x}_i(t+1), \hat{x}_i(t)\right)$$
 (15)

where  $f(\cdot, \cdot)$  remains progressive and monotonic. The function  $F(\ell_1, \ell_2)$  is raising, i.e. for finite  $M \ge 0$ ,  $\delta > 0$  and  $D \ge 0$ ,

$$F(\ell_1, \ell_2) = \begin{cases} \ell_1 & \ell_2 \ge M \text{ or } |\ell_2 - \ell_1| \le D\\ g(\ell_2) & \text{otherwise} \end{cases}, \quad (16)$$

where g(x) is finite for finite x and obeys

$$g(x) \ge x + \delta, \ \delta > 0. \tag{17}$$

The second bullet of (16), speeds the ascent of  $\hat{x}_i(t)$ , and because of (8) ameliorating the problem of the slow rise in underestimates experienced by (2). The first bullet renders (14)identical to (2). As the second bullet of (16) must change  $\hat{x}_i(t)$ , the stationary point of (14-17) is identical to that of (11). Thus this algorithm spreads the same information as (2), while accelerating the rise of underestimates. Observe also that  $D = \infty$  and/or M = 0, reduces (14-17) to (2).

The version of (14-17) in [1] sets the *dead zone* variable as D = 0. In face of persistent structural perturbations in  $e_{ik}$ ,  $l_2 = l_1$  cannot be sustained. Consequently, regardless of the size of perturbations, with D = 0,  $\hat{x}_i$  will regularly rise to the limit of the *modulation threshold* M, then fall, and then rise again. On the other hand we will show that if D is sufficiently greater than  $\epsilon$ , the bound on the perturbation, then (14-17) will have ultimate bounds proportional to  $\epsilon$ . This raises an essential trade-off. Too large a D slows convergence though imparts greater robustness to perturbations. Such a compromise is inherent to most dynamic systems. Slower convergence improves noise performance. Another key difference is that [1] assumes that  $\hat{x}_i$  belong to a Noetherian ring with M its maximal element. This implicitly assumes that the algorithm is a priori bounded. For distance estimation this means a prior assumption on the diameter, which is unappealing in the context of open systems. The generalized Adaptive Bellman-Ford algorithm (GABF), presented and analyzed without proofs in [18], is a specific example of (15), with  $f(\cdot, \cdot)$  as in (3) with  $e_{ik}$ the edge length between i and k,  $\hat{x}_k(t)$  the distance estimate of k at time t,  $s_i = 0$  if i is a source while  $s_i = \infty$  otherwise.

# C. Definitions and Assumptions

We merge the analysis without perturbations, i.e. when  $\epsilon = 0$  in (4), (5) with the perturbation analysis when  $\epsilon > 0$ . The latter requires some additional conditions: First that the function  $f(\cdot, \cdot)$  is monotonically increasing with respect to its second argument as well, i.e. f(a, b) obeys

$$f(a, b_1) \ge f(a, b_2), \text{ if } b_1 \ge b_2.$$
 (18)

Second, that there exist  $L_i > 0$ , such that

$$|f(a,b_1) - f(a,b_2)| \le L_1 |b_1 - b_2|$$
(19)

$$|f(a_1, b) - f(a_2, b)| \le L_2 |a_1 - a_2|.$$
 (20)

Most perturbation analyses require such Lipschitz conditions. Formally, the following assumption holds in this paper.

**Assumption 1.** Graph  $\mathcal{G}$  is connected, and  $\bar{e}_{ii}(t)$  obey (4-7). The set  $S^*$  defined in (10) in nonempty and

$$s_{\min} = \min_{j \in \mathcal{S}^*} \{ s_j \} \ge 0, \ \forall i \in V.$$
(21)

Further

$$\mathcal{S}_{\min} = \{ i \in V | s_i = s_{\min} \}.$$

$$(22)$$

If  $\epsilon = 0$  then  $f(\cdot, \cdot)$  only obeys (8,9). If  $\epsilon > 0$  then  $f(\cdot, \cdot)$ additionally obeys (18-20).

As in any given iteration the estimated state of a node is obtained by one of the bullets in (16), at each t, we partition V into two sets defined below.

<sup>3</sup> 0018-9286 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications\_standards/publications/rights/index.html for more information. Authorized licensed use limited to: Raytheon Technologies. Downloaded on September 30,2022 at 13:51:55 UTC from IEEE Xplore. Restrictions apply.

**Definition 1.** The set  $\mathcal{A}(t)$  (ABF type nodes) comprises all nodes that use the first case in (16) to obtain  $\hat{x}_i(t)$ , i.e. in (14),  $\hat{x}_i(t) = \tilde{x}_i(t)$ . Define the set of extraordinary nodes  $\mathcal{E}(t) =$  $V \setminus \mathcal{A}(t)$  to be those that use the second case in (16).

The next definition defines a (current) constraining node.

**Definition 2.** For  $i \in A(t)$ , if  $\hat{x}_i(t) = s_i = x_i$  then *i* is its own current constraining, or constraining node at t. Otherwise the minimizing  $k \neq i$  in (14) used to find  $\hat{x}_i(t)$ , is i's constraining node at t. If  $i \in \mathcal{E}(t)$ , then i is its own constraining node at t. The constraining node of i at t is said to constrain i at t.

We define S(t) as a subset of  $S^*$  comprising nodes that acquire their maximum values at time t,

$$\mathcal{S}(t) = \{ i \in \mathcal{S}^* \mid \hat{x}_i(t) = s_i \},\tag{23}$$

and we say i is a source at time t if  $i \in \mathcal{S}(t)$ . We also need one other type of partitioning of V.

**Definition 3.** The set of nodes rooted to sources is  $\mathcal{R}(t+1) =$  $S(t+1) \bigcup P(t+1)$  with S(t+1) as in (23) and P(t+1)comprising those whose constraining node at t+1 is in  $\mathcal{R}(t)$ . Further  $\mathcal{R}(t_0) = \mathcal{S}(t_0)$ . The unrooted set is  $\mathcal{U}(t) = V \setminus \mathcal{S}(t)$ .

As  $\mathcal{U}(t+1) \bigcap \mathcal{S}(t+1) = \emptyset$ , every node has a constraining node at every t, and  $i \in \mathcal{R}(t+1)$  is either in  $\mathcal{S}(t+1)$  or are constrained by some  $j \in \mathcal{R}(t)$ , each  $k \in \mathcal{U}(t+1)$  must be constrained by an  $\ell \in \mathcal{U}(t)$ . Thus

> $\mathcal{U}(t) = \emptyset \implies \mathcal{U}(t+1) = \emptyset.$ (24)

# **III. PROPERTIES OF STATIONARY POINTS**

As explained in Section II, (11) are also the stationary values of (14-17) in the perturbation free case of  $\epsilon = 0$ . Observe that stationary points if they exist comprise two sorts of values. Those where  $x_i = s_i$ . Those where  $x_i < s_i$ . We call the former sources and their set is defined as

$$\mathcal{S}_{\infty} = \{i | x_i = s_i\}.$$
(25)

Though  $\mathcal{S}_{\infty} \subset \mathcal{S}^*$ , as the example in Figure 1 shows,  $\mathcal{S}_{\infty}$ need not equal  $S^*$ . Evidently

$$x_{i} = \begin{cases} s_{i} & i \in \mathcal{S}_{\infty} \\ \min_{k \in \mathcal{N}(i)} \{ f(x_{k}, e_{ik}) \} & i \notin \mathcal{S}_{\infty} \end{cases}$$
(26)

As shown by example in Section II-A, not all members of  $S^*$ are sources. Like the definition of constraining nodes we now define true constraining nodes.

**Definition 4.** In (26), if  $x_i = s_i$ , then we say that i is its own true constraining node. Otherwise, any minimizing k in the second bullet of (26) is a true constraining node of *i*. As *i* may have more than one true constraining node, its set of true constraining nodes is designated as C(i).

As  $f(\cdot, \cdot)$  is progressive we have that,

$$> x_k, \ \forall k \in \mathcal{C}(i) \text{ and } i \notin \mathcal{S}_{\infty}.$$
 (27)

 $x_i$ We have the following lemma.

**Lemma 1.** The following hold under Assumption 1 with  $\epsilon = 0$ . (A) Consider any sequence of nodes, without loss of generality  $\{1, 2, \dots, l\}$  such that  $i + 1 \in C(i)$  as defined in Definition 4. Then this sequence is finite and its last element is in  $S_{\infty}$ , defined in (25). (B) The set  $S_{\infty}$  is nonempty. (C) The set  $S_{\min} \subset S_{\infty}$ . (D) All  $x_i$  in (11) are finite.

*Proof.* Due to (27) the chain in (A) cannot have cycles. As there are only N nodes it must end, and the last element l must be its own true constraining node i.e.  $l \in S_{\infty}$ . This proves (A), and also (B). Without loss of generality suppose  $s_1 = s_{\min}$ . To establish a contradiction, suppose  $1 \notin S_{\infty}$ . Then from (A) there is a sequence of nodes starting from 1 and terminating in  $j \in S_{\infty}$ , such that each is the true constraining node of its predecessor. Thus from (27)  $x_j = s_j < s_1 = s_{\min}$ , violating the definition of  $s_{\min}$ , proving (C). To prove (D) consider  $i \neq 1$ . As the graph is connected there is a path from 1 to *i*, comprising nodes  $\{1 = l_1 \rightarrow l_2 \rightarrow \cdots , l_k = i\}$  Then from (26) for each  $n \in \{2, \dots, k\}$  there holds

$$x_{l_n} \le f(x_{l_{n-1}}, e_{l_n, l_{n-1}}).$$

Due to the fact that f(a, b) is finite for finite  $a, b, x_{l_n}$  is finite if  $x_{l_{n-1}}$  is finite. The result follows as  $x_1$  is finite.

We make another definition for the stability analysis.



Fig. 2: Illustration of graph where  $S_{\infty}$  is not a subset of  $\mathcal{F}_0$ . Here  $s_1 = 0, s_2 = 1$  and  $s_3 = \infty$ . All edge lengths are 1 and f(a, b) =a+b. In this case  $x_1 = 0, x_2 = 1$  and  $x_3 = 1$ . Here  $2 \in S_{\infty}$  as  $x_2 = s_2$ . However, as  $x_2 = x_1 + 1, 2 \in \mathcal{F}_1$ .

**Definition 5.** We call a path from a node i to  $j \in S_{\infty}$  a shortest path, if it starts at i, ends with  $j \in S_{\infty}$ , and each node in the path is a true constraining node of its predecessor. We call a shortest path from i the longest shortest path if it has the most nodes among all shortest paths of i. The set  $\mathcal{F}_i$ is the set of nodes whose longest shortest paths to the source set have i+1 nodes. We call  $\mathcal{D}(\mathcal{G})$  the effective diameter of  $\mathcal{G}$ if the longest shortest path among all  $i \in V$  has  $\mathcal{D}(\mathcal{G})$  nodes.

From Lemma 1, the effective diameter is always finite. If a node *i* has two shortest paths, one with two and the other with three nodes, then  $i \notin \mathcal{F}_1$  but  $i \in \mathcal{F}_2$ . It is tempting to believe that  $\mathcal{F}_0 = \mathcal{S}_\infty$ . However, the scenario of Figure 2 provides a counterexample. In this case  $s_1 = 0$ ,  $s_2 = 1$  and  $s_3 = \infty$ . All edge lengths are 1 and f(a, b) = a + b. In this case  $x_1 = 0, x_2 = 1$  and  $x_3 = 1$ . Here  $2 \in \mathcal{S}_{\infty}$  as  $x_2 = s_2$ . However, as  $x_2 = x_1 + 1, 2 \in \mathcal{F}_1$ .

**Lemma 2.** Under the conditions of Lemma 1, consider  $\mathcal{F}_i$ given in Definition 5. If for some  $k \in \{1, \dots, \mathcal{D}(\mathcal{G}) - 1\}, \mathcal{F}_k$ is nonempty then every node in  $\mathcal{F}_k$  has a true constraining node in  $\mathcal{F}_{k-1}$ . Further  $\mathcal{S}_{\min} \subset \mathcal{F}_0 \subset \mathcal{S}_{\infty}$ .

*Proof.* Consider any  $i \in \mathcal{F}_k$ . From Definition 5, starting from *i* there is a sequence containing k + 1 nodes to a  $j \in S_{\infty}$  in which each node is the true constraining node of its predecessor. Suppose the second node in this sequence is l. By definition l is a true constraining node of i. Also by definition  $l \in \mathcal{F}_m$ , where  $m \ge k - 1$ . If m > k - 1, then for some  $M > k, i \in \mathcal{F}_M$ . This contradicts the assumption that  $i \in \mathcal{F}_k$ . Thus  $l \in \mathcal{F}_{k-1}$ . By definition, every node in  $\mathcal{F}_0$  is its own true constraining node as otherwise it will belong to some  $\mathcal{F}_i$ , i > 0. Thus from Definition 4,  $\mathcal{F}_0 \subset \mathcal{S}_{\infty}$ .

Finally consider  $j \in S_{\min}$ . By definition  $s_j = s_{\min} \leq s_i$ for all i. If  $j \in \mathcal{F}_k$ , with k > 0, then there is a sequence starting from j to an  $l \in \mathcal{S}_{\infty}$ , such that each node is the true constraining of its predecessor. Thus from the progressive property of  $f(\cdot, \cdot)$ ,  $s_{\min} = s_j \ge x_j > s_l$ , establishing a contradiction. Thus  $j \in \mathcal{F}_0$  and  $\mathcal{S}_{\min} \subset \mathcal{F}_0$ .

These lemmas lead to another critical to our analysis.

**Lemma 3.** Under the conditions of Lemma 1, with  $\mathcal{F}_i$  defined in Definition 5,  $\mathcal{F}_i \neq \emptyset$ ,  $\forall i \in \{0, 1, \cdots, \mathcal{D}(\mathcal{G}) - 1\}$ , and for all  $i \in \{0, 1, \dots, \mathcal{D}(\mathcal{G}) - 2\}$  each node in  $\mathcal{F}_{i+1}$  has a true constraining node in  $\mathcal{F}_i$ .

*Proof.* Use induction to show that  $\forall k \in \{1, \dots, \mathcal{D}(\mathcal{G}) - 1\},\$  $\mathcal{F}_k \neq \emptyset$ .  $\mathcal{F}_{\mathcal{D}(\mathcal{G})-1} \neq \emptyset$  by Definition 5. Suppose for some  $L \in \{1, \cdots, \mathcal{D}(\mathcal{G}) - 1\}, \mathcal{F}_L \neq \emptyset$ . From Lemma 2 every  $i \in \mathcal{F}_L$ has a true constraining node in  $\mathcal{F}_{L-1}$ , i.e.  $\mathcal{F}_{L-1} \neq \emptyset$ . From Lemma 2 every  $i \in \mathcal{F}_1$  has a true constraining node in  $\mathcal{F}_0$ making the latter nonempty. Lemma 2 proves the result. 

We conclude by proving the uniqueness of the  $x_i$ .

**Theorem 1.** Under the conditions of Lemma 1, there is a unique stationary point obeying (11).

*Proof.* The existence of stationary values is guaranteed by the fact that (26) is a Bellman's equation.

As the graph is connected, there is a path from every node to every other node. Define  $\mathcal{P}_{ji}$  to be the set of all paths from j to i, including j = i. Denote such a path  $\mathcal{P} \in \mathcal{P}_{ji}$ , e.g.  $l_0 \rightarrow l_1 \rightarrow \cdots, \rightarrow l_L = i$ , by the ordered set  $\mathcal{P} = \{j = i\}$  $l_0, l_1, \dots, l_L = i$ . In particular the path from *i* to *i*, will be  $\mathcal{P}_{ii} = \{\{i\}\}\}$ . Consider the recursion,

$$x_{l_k}^*(\mathcal{P}) = \begin{cases} s_{l_k} & k = 0\\ f(x_{l_{k-1}}^*(\mathcal{P}), e_{l_{k-1}l_k}) & k \in \{1, \cdots, L\} \end{cases}$$
(28)

For  $x_{l_k}^*(\mathcal{P})$  to be finite  $\mathcal{P} \in \mathcal{P}_{ji}$  with  $j \in \mathcal{S}^*$ . Then, from (9) and the principle of optimality,

$$x_{i} = \min\left\{\min_{j \in V}\left\{\min_{\mathcal{P} \in \mathcal{P}_{ji}}\left\{x_{i}^{*}(\mathcal{P})\right\}\right\}, s_{i}\right\}.$$

Because  $S^*$  is nonempty, V is finite, and  $f(\cdot, \cdot)$  is well defined, this value is also well defined and unique.

# **IV. STABILITY ANALYSIS**

We now prove GUAS with  $\epsilon = 0$  and ultimate boundedness with  $\epsilon > 0$ . The latter result specializes to the former. Even though it uses (18-20) we show in the proofs in the appendix, why (18-20) are not needed for  $\epsilon = 0$ . We first provide three key insights into the convergence of  $\hat{x}_i(t)$  to  $x_i$  when  $\epsilon = 0$ .

The second bullet of (16), increases  $\hat{x}_i(t)$ . Also, because of (15), the use of the first bullet of (16) forces  $\hat{x}_i(t+1) =$  $\tilde{x}_i(t+1)$ , which due to (14) and (17), must exceed the state of at least one neighbor. Together these cause all underestimates to eventually disappear, i.e., for some T and all i and t > T,  $\hat{x}_i(t) \ge x_i$ , i.e., from (9),  $f(\hat{x}_k(t), e_{ik}) \ge f(x_k, e_{ik})$ .

Note,  $\epsilon = 0$  implies  $\bar{e}_{ij}(t) = e_{ij}$ . Also,  $\forall i \in \mathcal{F}_0$ ,  $x_i = s_i$  and in (11),  $\min\{\min_{k \in \mathcal{N}(i)} \{f(x_k, e_{ik})\}, s_i\} = s_i$ . Thus, whenever the first bullet of (16) is invoked at  $i \in$  $\mathcal{F}_0$  and  $t \geq T$ , from (15),  $\hat{x}_i(t+1) = \tilde{x}_i(t+1) =$  $\min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(\hat{x}_{k}(t),e_{ik}\right)\right\},s_{i}\right\} = s_{i} = x_{i}.$  Hereafter, the first bullet will be always invoked by *i* because  $\hat{x}_i(t+1) =$  $\tilde{x}_i(t+1)$  and  $\hat{x}_i(t)$  converges to  $x_i$ . Similarly, as long as all  $j \in \mathcal{F}_L$  have converged, if  $i \in \mathcal{F}_{L+1}$  invokes the first case of (16) then  $\hat{x}_i(t)$  converges to  $x_i$ . Further, as from (17), the second bullet of (16) keeps increasing  $\hat{x}_i(t)$ ,  $\hat{x}_i(t)$  must eventually exceed M, making it invoke the first bullet.

When  $\epsilon > 0$ , the ultimate bound on the overestimates are similarly acquired. The bound on underestimates is obtained using a *shrunken* graph  $\mathcal{G}^-$  with

$$\bar{e}_{ij}(t) = e_{ij} - \epsilon = e_{ij}^{-}.$$
(29)

The estimates  $\hat{X}_i(t)$  given by (14-16) when run on  $\mathcal{G}^-$ , lower bound  $\hat{x}_i(t)$ . As  $\mathcal{G}^-$  is static,  $\hat{X}_i(t)$  converges to the desired values  $X_i$ . Underestimates are bounded by relating  $X_i$  to  $x_i$ 

**Definition 6.** Define  $\mathcal{G}^-$  as  $\mathcal{G}$ 's shrunken version:  $\forall i \in V$ and  $j \in \mathcal{N}(i)$  in  $\mathcal{G}$ , (29) holds in  $\mathcal{G}^-$ . Also consider (15) implemented on this shrunken graph, i.e.

$$\hat{X}_{i}(t+1) = F(\tilde{X}_{i}(t+1), \hat{X}_{i}(t)), \ \hat{X}_{i}(0) \le \hat{x}_{i}(0), \quad (30)$$

$$\tilde{X}_{i}(t+1) = \min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(\hat{X}_{k}(t), e_{ik}^{-}\right)\right\}, s_{i}\right\}$$
(31)

As  $\mathcal{G}^-$  is perturbation free, we define  $X_i$  as the unique stationary values in  $\mathcal{G}^-$  to which (30) converges, i.e.

$$X_{i} = \min\left\{\min_{k \in \mathcal{N}(i)} f\left(X_{k}, e_{ik}^{-}\right), s_{i}\right\}, \forall i \in V.$$
(32)

Also,  $\mathcal{G}^-$  has the effective diameter  $\mathcal{D}(\mathcal{G}^-)$  and the source set  $\mathcal{S}_{\infty}^{-} = \{i | X_i = s_i\},\$ (33)

Evidently, the following holds in  $\mathcal{G}^-$ :

$$s_{\min} \le X_i \le s_i, \ \forall i \in V.$$
(34)

Define

$$X_{\max} = \max_{k \in V} \{X_k\},\tag{35}$$

and

$$T = \left\lceil \frac{X_{\max} - \hat{x}_{\min}(t_0)}{\min\{\delta, \sigma\}} \right\rceil$$
(36)

We prove the elimination of underestimates in  $\mathcal{U}(t_0 + T)$ .

**Lemma 4.** Consider (14)-(16), under Assumption 1 and  $\mathcal{U}(t)$ as in Definition 3. Define  $\hat{x}_{\min}(t) = \min_{\substack{i \in \mathcal{U}(t)}} \{\hat{x}_j(t)\}$  if  $\mathcal{U}(t) \neq i$  $j \in \mathcal{U}(t)$  $\emptyset$ . Then as long as  $\mathcal{U}(t) \neq \emptyset$ , there holds:

$$\hat{x}_i(t) \ge \hat{x}_{\min}(t_0) + \min\{\sigma, \delta\}(t - t_0), \ \forall i \in \mathcal{U}(t)$$
(37)

$$\hat{x}_i(t) \ge X_{\max} \ge X_i, \ \forall i \in \mathcal{U}(t) \ and \ \forall t \ge t_0 + T$$
 (38)

We next prove that all  $\hat{x}_i(t) \geq x_i$  for all in  $i \in \mathcal{R}(t)$ .

**Lemma 5.** Under the conditions of Lemma 4, consider  $\mathcal{A}(t)$ ,  $\mathcal{E}(t)$ ,  $\mathcal{R}(t)$  and  $X_i$  as in definitions 1, 3 and 6, respectively. There holds:

$$\hat{x}_i(t) \ge X_i, \ \forall i \in \mathcal{R}(t).$$
 (39)

Consequently, with Lemma 4 and Lemma 5, there holds:

$$\dot{x}_i(t) \ge X_i, \ \forall i \in V, \ \forall t \ge t_0 + T.$$
(40)

We now quantify the relation between  $X_i$  and  $x_i$  by using the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140253, IEEE Transactions on Automatic Control

the following function for  $P \ge 1$ :

$$W(L,P) = \sum_{i=0}^{P-1} L^{i}.$$
 (41)

**Lemma 6.** Under Assumption 1, consider  $\epsilon$ ,  $W(\cdot,\cdot)$  and  $\mathcal{D}(\mathcal{G}^{-})$  defined in (5), (41) and Definition 6, respectively. Then for all  $i \in V$  and  $t \geq T + t_0$ , there holds:

$$x_i \le X_i + W(L_2, \mathcal{D}(\mathcal{G}^-) - 1)L_1\epsilon.$$
(42)

With  $\mathcal{F}_i$  defined in Definition 5, define  $X_{i\min}$  as

$$X_{i\min} = \min_{j \in \mathcal{F}_i} \{X_j\}$$
(43)

As  $S_{\min} \subset \mathcal{F}_0 \subset S_{\infty}$ ,  $X_{0\min} = s_{\min}$ . Define the sequence alluded to at the begin of this section:

$$T_{i} = \max\left\{0, \left\lceil\frac{M - X_{i\min}}{\delta}\right\rceil\right\} + 2.$$
(44)

Lemma 7 shows that acceptable behavior under perturbations requires the dead zone D in (16) to be sufficiently large.

**Lemma 7.** Consider (14)-(16) under Assumption 1, with  $\epsilon \geq$ 0,  $W(\cdot, \cdot)$  as in (5), (41), respectively. Suppose D in (16) obeys (45)

 $D \ge (W(L_2, \mathcal{D}(\mathcal{G}^-) - 1) + W(L_2, \mathcal{D}(\mathcal{G}) - 1))L_1\epsilon$ and at a time  $t' \ge t_0 + T$  defined in (36)

 $\hat{x}_i(t) \leq x_i + W(L_2, L)L_1\epsilon, \ \forall i \in \mathcal{F}_L, \ \forall t \geq t',$ (46)for some  $L \in \{0, 1, \dots, \mathcal{D}(\mathcal{G}) - 2\}$ . Then with  $T_i$  as in (44):  $\hat{x}_i(t) \le x_i + W(L_2, L+1)L_1\epsilon, \ \forall \ i \in \mathcal{F}_{L+1}, \ t \ge t' + T_{L+1}.$ 

The main result proves that the algorithm is ultimately bounded under bounded persistent perturbations and provides an upper bound on the time to attain the ultimate bound.

**Theorem 2.** Under the conditions of Lemma 7, for all 
$$i \in V$$
,  
 $t \ge t_0 + T + \sum_{i=0}^{\mathcal{D}(\mathcal{G})-1} T_i$  and  $D$  as in (45),  
 $|\hat{x}_i(t) - x_i| \le \epsilon L_1 \max \{ W(L_2, \mathcal{D}(\mathcal{G}) - 1), W(L_2, \mathcal{D}(\mathcal{G}^-) - 1) \} \}$ 

This is a classical ultimate bound proportional to  $\epsilon$ , the maximum disturbance level. For  $\epsilon = 0$  this reduces to  $\hat{x}_i(t) = x_i$  for all  $t \ge t_0 + T + \sum_{i=0}^{\mathcal{D}(\mathcal{G})-1} T_i$  and  $D \ge 0$ . As  $T + \sum_{i=0}^{\mathcal{D}(\mathcal{G})-1} T_i$  is independent of  $t_0$  this proves GUAS in the perturbation free case of  $\epsilon = 0$ . As explained in the proofs in the appendix, (18-20) are not needed when  $\epsilon = 0$ .

#### V. DESIGN CHOICES AND DISCUSSION

Theorem 2 verifies the intuitively clear requirement that the dead zone D should grow proportionally to the disturbance bound  $\epsilon$ . However, as this is a worst-case analysis, it masks the full effects of parameters M,  $\delta$ , and D. Looking beyond worst-case analysis, however, we can find that choosing these parameters involves tradeoffs between the convergence speed of underestimates and overestimates.

The convergence of underestimates is upper bounded by T in (36), which is in turn determined by (37), and thus conservatively by the smaller of  $\sigma$  and  $\delta$ . In practice, if  $\sigma$ is small and the first bullet of raising is invoked too often then underestimates rise slowly, i.e. the rising value problem will persist. If the second bullet of (16) is invoked at most times and  $\delta \gg \sigma$  then underestimates decline fast. Large D or small M makes this less likely and slows convergence, while a large  $\delta \geq M$  speeds convergence by reducing T and  $T_i$ .

For the convergence of overestimates,  $T_i$  gives the worst case time to invoke the first case of (16), whereupon all elements in  $\mathcal{F}_i$  converge forthwith. The worst case analysis quantifies  $T_i$  by how long it takes for  $\hat{x}_i(t)$  to exceed M and assumes that the second clause of (16) is invoked until this happens. With a large D, however, this time shortens as the first bullet is likely to be invoked more quickly.

In most cases, the need to alleviate the rising value problem is more compelling as overestimates in algorithms like plain ABF converge in at most  $\mathcal{D}(\mathcal{G}) - 1$  steps. Accordingly, the desirability of a smaller dead zone D competes with the requirement of resilience to persistent perturbations as quantified by (45). This of course is common to most dynamical systems where faster convergence generally comes at the price of reduced resilience. We note, however, the following appealing fact: both the ultimate bound and the required D are determined exclusively by the perturbation magnitude  $\epsilon$  and the effective diameters of the original and shrunken graph.

Complementarily, note that in the special case of the algorithm in [8], we effectively have M = 0 and  $D = \infty$ . In this case the second bullet of (16) is never invoked. Accordingly, a small  $\sigma$  leads to large T and T and the rising value problem. In particular, the algorithm remains GUAS with the same ultimate bound as (45) is automatically satisfied. Overestimates however, converge quickly as  $T_i = 2$ .

# **VI. SIMULATIONS**

Because of space constraints simulations under perturbations are in https://arxiv.org/pdf/2102.10319.pdf. We first consider GABF. We use 500 nodes, randomly distributed in a  $4 \times 1 \text{ km}^2$  area, and communicating within a 0.25 km }radius. One node is designated as a source and initial distance estimates are in  $\mathcal{U}[0,\sqrt{17}]$ km. Define the greatest overestimate  $\Delta^+(t) = \max\{0, \max_i \{\Delta_i(t)\}\}\$  and the least underestimate  $\Delta^{-}(t) = \max\{0, -\min_{i} \{\Delta_{i}(t)\}\}, \text{ with } \Delta_{i}(t) = \hat{d}_{i}(t) - d_{i}$ the estimation error of *i*. We average over 100 runs.

The bottom plots in Figure 3(a) and (b) use M = 5, D = 0, and  $\delta$  varying from 0.2M to M. For a fixed M, a smaller  $\delta$ slows convergence, as  $\hat{x}_i(t)$  reaches M slower, while D has the opposing effect. The middle plots concerns  $\delta = M = 5$ and D varying from 0 to 4K, using  $K = (\mathcal{D}(\mathcal{G}) + \mathcal{D}(\mathcal{G}^{-}) - 2)\epsilon$ , where the average values of  $\epsilon$ ,  $\mathcal{D}(\mathcal{G})$  and  $\mathcal{D}(\mathcal{G}^{-})$  are  $2.3 \times 10^{-3}$ km and 18.5 and 26.6, respectively. A small D accelerates convergence of  $\Delta^+$  as in this case GABF acts more like ABF where  $\Delta^+(t)$  converges fast. A large D has a negative impact on the convergence of  $\Delta^-$  as GABF approaches ABF.

The top plots in Figure 3(a) and (b) illustrate the impact of M when D = 0 and  $\delta = M$  with M increasing from 4.108 to 4.124 in steps of 0.004. Convergence improves with higher M. Though  $T_i$  in (44) satisfies  $T_i = 3$  for all M, as  $\delta = M$ , overestimates disappear more quickly with larger M because T in (36) become smaller;  $\Delta^-$  converges in two steps when M is exceeds both the largest true distance and  $d_i(t_0)$ .

Figure 4 concerns a nonlinear  $f(\cdot, \cdot)$  in (2), for nodes to compute paths minimizing exposure to a hazard: 400 nodes are randomly distributed in a  $4 \times 4 \text{ km}^2$  field except the upleft corner, communicating over a 0.5 km radius. A source is



Fig. 3: Convergence time for (a)  $\Delta^+(t)$  and (b)  $\Delta^-(t)$  without perturbations. The bottom plots in (a) and (b) correspond to M = 5, D = 0 and  $\delta$  varying from 0.2M to M. The middle plots correspond to  $\delta = M = 5$  and D varying from 0 to 4K. The top plots correspond to D = 0,  $\delta = M$  and M varying from 4.108 to 4.124.



(a) The spreading block (2) (b) The general spreading block (15)-(16)

Fig. 4: In this example, 400 nodes are randomly located in a  $4 \times$  $4 \text{ km}^2$  field. There is a source marked by red asterisk located at (0.3, 0.3), and a  $2.5 \times 2.5 \text{ km}^2$  radiation zone in the middle. Color represents degree of contamination, with a logarithmic scale. While both spreading block and the general spreading block can achieve the shortest available path, the fast convergence of the general spreading block greatly reduces total contamination.

located at (0.3, 0.3), and there exists a  $2.5 \times 2.5$  km<sup>2</sup> radiation zone centered at (1.95, 1.95). Define  $\mathcal{M}$  as the set of radioactive nodes. All nodes ever constrained by a  $i \in \mathcal{M}$  are deemed radioactive. We use  $f(\hat{x}_k(t), e_{ik}) = \hat{x}_k(t) + e_{ik}, k \notin \mathcal{M},$ with  $e_{ik}$  the edge length between i and k. When  $k \in \mathcal{M}$ ,  $f(\hat{x}_k(t), e_{ik}) = h(\hat{x}_k(t) + 1000e_{ik})$ , where  $h(a) = a^{1.5}$  when a > 1 and a otherwise;  $s_i = 0$  if i is a source and  $\infty$  otherwise. This severely penalizes passage through hazards. In particular, the factor 1000 amplifies the distance. The exponent 1.5 does so even more when  $\hat{x}_k(t) + 1000e_{ik} > 1$ . In each round a node i will receive  $100 \sim 120$  units of radiation dose if it is radioactive and  $0 \sim 1$  unit otherwise. Figure 4(a) uses (2), while Figure 4(b) uses (15)-(16) with  $D = 0, \delta = M > x_{max}$ . Nodes outside the radiation zone never cross over, and nodes inside take the shortest path of exit. Contamination is greatly reduced when using (15)-(16), due to faster convergence.

# VII. CONCLUSION

We have improved a general algorithm for spreading information across a network of devices by making it resilient

to perturbations and by removing a prior boundedness assumption. This algorithm, a key building block for aggregate computing and applicable to a wide range of distributed systems, has parameters that remove the rising value problem that appears in some of its special cases, such as ABF. Unlike ABF, the general algorithm covers a much wider class of applications with non-Euclidean metrics. We have proved GUAS and provided ultimate bounds in the face of persistent network disturbances using an additional Lipschitz condition. These bounds depend only on the largest perturbation and structural network properties. Finally, we provide design guidelines demonstrating how algorithm parameters have competing effects on performance. These results are a crucial stepping stone in our long term goal of determining stability conditions for feedback interconnections of aggregate computing blocks, using possibly new small gain theorems, [5]. Progress in this program has broad applicability for the engineering of resilient distributed systems.

#### REFERENCES

- [1] M. Viroli, G. Audrito, J. Beal, F. Damiani, and D. Pianini, "Engineering resilient collective adaptive systems by self-stabilisation," ACM Transactions on Modeling and Computer Simulation, pp. 16:1–16:28, 2018.
- H. Zainab, G. Audrito, S. Dasgupta, and J. Beal, "Improving collection [2] dynamics by monotonic filtering," 2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C), 2020.
- [3] H. K. Khalil, "Nonlinear systems," Upper Saddle River, 2002.
- [4] Y. Mo, S. Dasgupta, and J. Beal, "Robustness of the Adaptive Bellman-Ford Algorithm: Global stability and ultimate bounds," IEEE Transactions on Automatic Control, pp. 4121-4136, 2019.
- [5] Z.-P. Jiang, I. M. Mareels, and Y. Wang, "A Lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems," Automatica, vol. 32, no. 8, pp. 1211-1215, 1996.
- [6] S. Dasgupta, B. D. O. Anderson, and R. Kaye, "Output-error identification methods for partially known systems," Int. J. of Control, 1985.
- R. Bellman, "On a routing problem," Quarterly of applied mathematics, [7] vol. 16, no. 1, pp. 87-90, 1958.
- Y. Mo, S. Dasgupta, and J. Beal, "Robust stability of spreading blocks [8] in aggregate computing," in 2018 IEEE Conference on Decision and Control (CDC), 2018, pp. 6007-6012.
- [9] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in 2007 46th IEEE Conference on Decision and Control, 2007.
- [10] Z. Li, W. Ren, X. Liu, and L. Xie, "Distributed consensus of linear multiagent systems with adaptive dynamic protocols," Automatica, vol. 49, no. 7, pp. 1986-1995, 2013.
- [11] J. Wang, A. Lanzon, and I. R. Petersen, "Robust output feedback consensus for networked negative-imaginary systems," IEEE Transactions on Automatic Control, vol. 60, no. 9, pp. 2547-2552, 2015.
- [12] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," The International Journal of Robotics Research, vol. 30, no. 7, pp. 846-894, 2011.
- [13] J. Baillieul and A. Suri, "Information patterns and hedging brockett's theorem in controlling vehicle formations," in 42nd IEEE International Conference on Decision and Control, 2003.
- [14] T. H. Summers, C. Yu, S. Dasgupta, and B. D. Anderson, "Control of minimally persistent leader-remote-follower and coleader formations in the plane," IEEE Transactions on Automatic Control, vol. 56, no. 12, pp. 2778-2792, 2011.
- [15] L. Moreau, "Stability of multiagent systems with time-dependent communication links," IEEE Transactions on Automatic Control, 2005.
- [16] T. H. Summers, C. Yu, B. D. Anderson, and S. Dasgupta, "Formation shape control: Global asymptotic stability of a four-agent formation," in Proceedings of the 48h IEEE CDC, 2009.
- [17] M. Arcak, "Passivity as a design tool for group coordination," IEEE Transactions on Automatic Control, vol. 52, no. 8, pp. 1380-1390, 2007.
- [18] Y. Mo, S. Dasgupta, and J. Beal, "Global uniform asymptotic stability of a generalized adaptive Bellman-Ford algorithm," in 2019 IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 1868–1873.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140253, IEEE Transactions on Automatic Control

#### APPENDIX

**Proof of Lemma 4:** Because of (24)  $\mathcal{U}(t)$  can be nonempty only on a single contiguous time interval commencing at  $t_0$ .

We prove (37) by induction. As it holds for  $t = t_0$ , suppose (37) holds for some  $t \ge t_0$ . Consider  $i \in \mathcal{U}(t+1)$  with  $\hat{x}_i(t+1) = \hat{x}_{\min}(t+1)$ . Then  $\mathcal{U}(t) \neq \emptyset$  and j the constraining node of i is in  $\mathcal{U}(t)$ . If  $i \in \mathcal{E}(t+1)$  in Definition 1 then from Definition 2, j = i. From the induction hypothesis and (17)

 $\hat{x}_i(t)$ 

$$\begin{aligned} +1) &= \hat{x}_{\min}(t+1) \\ &\geq \hat{x}_j(t) + \delta \ge \hat{x}_{\min}(t) + \min\{\sigma, \delta\} \\ &\geq \hat{x}_{\min}(t_0) + \min\{\sigma, \delta\}(t+1-t_0). \end{aligned}$$

If  $i \in \mathcal{A}(t+1) \cap \mathcal{U}(t+1)$ , then,  $i \notin \mathcal{S}(t+1)$ , i.e.  $\hat{x}_i(t+1) \neq s_i$ . From the induction hypothesis and (14) we have

$$\hat{x}_i(t+1) = \hat{x}_{\min}(t+1) = f(\hat{x}_j(t), \bar{e}_{ij}(t))$$

$$\geq \hat{x}_j(t) + \sigma \geq \min\{\sigma, \delta\}(t+1-t_0).$$
Further with (35) and (36), (38) follows.

**Proof of Lemma 5:** If  $\mathcal{R}(t) \neq \emptyset$ , then  $\exists t_5, t_6$  such that  $\forall t_0 \leq t_5 \leq t \leq t_6$  and  $\mathcal{R}(t_5) = \mathcal{S}(t_5)$ . As  $\hat{x}_i(t_5) = s_i$  for all  $i \in \mathcal{S}(t_5) = \mathcal{R}(t_5)$ , from (34) the result holds for  $t = t_5$ .

Suppose  $\hat{x}_i(t) \geq X_i$  for some  $t_5 \leq t < t_6$  and all  $i \in$  $\mathcal{R}(t)$ . Consider any  $i \in \mathcal{R}(t+1)$ . From Definition 3, either  $i \in \mathcal{S}(t+1)$  in which case the result holds, or i is constrained by some  $j \in \mathcal{R}(t)$ . If  $i \in \mathcal{A}(t+1)$ , then by the induction hypothesis,  $\hat{x}_i(t) \geq X_i$ . As  $i \notin \mathcal{S}(t+1)$ , there follows:

$$\hat{x}_i(t+1) = f(\hat{x}_j(t), \bar{e}_{ij}(t)) \ge f(X_j, \bar{e}_{ij})$$
 (47)  
>  $X_i$  (48)

Here if  $\epsilon > 0$ , then (47) uses  $\bar{e}_{ij}(t) \ge e_{ij}^{-}$  for all t and the fact that  $f(\cdot, \cdot)$  is increasing in each argument. If  $\epsilon = 0$  then only (9) is needed. Further, (48) uses (32) and the fact that  $x_i = X_i$ when  $\epsilon = 0$ . If  $i \in \mathcal{E}(t+1)$ , then *i* is its own constraining node and  $i \in \mathcal{R}(t)$ . Thus by our induction hypothesis,  $\hat{x}_i(t) \geq X_i$ . From (17),  $\hat{x}_i(t+1) \ge \hat{x}_i(t) + \delta > X_i$ .

**Proof of Lemma 6:** The result holds when  $\epsilon = 0$  as  $X_i = x_i$ and  $\mathcal{G} = \mathcal{G}^-$ . For  $\epsilon > 0$ , consider nodes  $n_0, n_1, \cdots, n_Q$  such that  $n_0 \in S_{\infty}^-$ , and for all  $i \in \{0, \ldots, Q-1\}$ ,  $n_i$  is a true constraining node of  $n_{i+1}$  in  $\mathcal{G}^-$ . Each node in  $\mathcal{G}^-$  is in one such sequence. As from Definition 6,  $Q \leq \mathcal{D}(\mathcal{G}^{-}) - 1$ , the result holds if

$$x_{n_i} - X_{n_i} \le W(L_2, i) L_1 \epsilon, \ \forall i \in \{0, \cdots, Q\}.$$
 (49)

Evidently,  $x_{n_0} \leq s_{n_0} = X_{n_0}$ . Suppose (49) holds for some  $i \in \{0, \dots, T-1\}$ . As  $n_i$  and  $n_{i+1}$  are neighbors in both  $\mathcal{G}$  and  $\mathcal{G}^-$ ,  $n_i$  is a true constraining node of  $n_{i+1}$  in  $\mathcal{G}^-$  and  $x_{n_i} \leq X_{n_i} + W(L_2, i)L_1\epsilon$  by our induction hypothesis,

$$x_{n_{i+1}} \le f(x_{n_i}, e_{n_i n_{i+1}}) \le f(X_{n_i} + W(L_2, i)L_1\epsilon, e_{n_i n_{i+1}}) \le f(X_{n_i}, e_{n_i n_{i+1}}) + L_2 W(L_2, i)L_1\epsilon$$
(50)

$$= f(X_{n_i}, e^- + \epsilon) + L_2 W(L_2, i) L_1 \epsilon$$
(50)
$$= f(X_{n_i}, e^- + \epsilon) + L_2 W(L_2, i) L_1 \epsilon$$
(51)

$$< f(X e^{-}) + L_{1}\epsilon + L_{2}W(L_{2},i)L_{1}\epsilon$$
(61)

$$\leq \int (A_{n_i}, e_{n_i n_{i+1}}) + L_1 e + L_2 w (L_2, i) L_1 e \tag{52}$$

$$= X_{n_{i+1}} + W(L_2, i+1)L_1\epsilon$$
(53)

where (50) uses (20), (51) uses (29), (52) uses (19), and (53) uses the fact that  $n_i$  is a true constraining node of  $n_{i+1}$  in  $\mathcal{G}^-$ . **Proof of Lemma 7:** Consider  $i \in \mathcal{F}_{L+1}$ . From (16) and (17), there is a  $t' < t \leq t' + T_{L+1}$  such that  $i \in \mathcal{A}(t)$ . This is so as  $i \in \mathcal{E}(t)$  implies  $\hat{x}_i(t+1) \ge \hat{x}_i(t) + \delta$  and at some time in the interval  $(t', t' + T_{L+1}], \hat{x}_i(\cdot) > M$ . From Lemma 3, there

is a  $j \in \mathcal{F}_L$  that is a true constraining node of i in  $\mathcal{G}$ . Thus  $\hat{x}_j(t-1) \le x_j + W(L_2, L)L_1\epsilon$  by (46). When  $\epsilon > 0$ ,

$$\hat{x}_{i}(t) = \min\left\{\min_{k \in \mathcal{N}(i)} \left\{f\left(\hat{x}_{k}(t-1), \bar{e}_{ik}(t-1)\right)\right\}, s_{i}\right\} \\
\leq f(\hat{x}_{j}(t-1), \bar{e}_{ij}(t-1)) \\
\leq f(x_{j} + W(L_{2}, L)L_{1}\epsilon, e_{ij} + \epsilon) \\
\leq f(x_{j}, e_{ij}) + L_{2}W(L_{2}, L)L_{1}\epsilon + L_{1}\epsilon \tag{55} \\
= x_{i} + W(L_{2}, L+1)L_{1}\epsilon \tag{56}$$

where (54) uses (4), (5), (18) and (46), (55) uses (19) and (20). When  $\epsilon = 0$ , (56) holds without (18-20) as  $f(x_i +$  $W(L_2, L)L_1\epsilon, e_{ij} + \epsilon) = f(x_j, e_{ij}) = x_i.$ 

Similarly, as (46) holds for all 
$$t \ge T + t_0$$
 for all  $j \in \mathcal{F}_L$ ,  
 $\tilde{x}_i(t+1) \le x_i + W(L_2, L+1)L_1\epsilon$ . (57)

As  $t > t_0 + T$ , (40) implies that  $\hat{x}_k(t) \ge X_k$  for all  $k \in V$ . From (9), (18) and the fact that  $X_i \leq s_i$ , (32) implies that

$$\tilde{x}_{i}(t+1) = \min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(\hat{x}_{k}(t),\bar{e}_{ik}(t)\right)\right\},s_{i}\right\}\right\}$$
$$\geq \min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(X_{k},e_{ik}^{-}\right)\right\},s_{i}\right\} = X_{i},$$

i.e.  $[X_i, x_i + W(L_2, L+1)L_1\epsilon]$  contains both  $\hat{x}_i(t)$  and  $\tilde{x}_i(t+1)$ . This also holds with  $\epsilon = 0$  as then  $\bar{e}_{ik}(t) = e_{ik} = e_{ik}^$ and  $X_i = x_i$ . Thus, from (45) and Lemma 6

$$\begin{aligned} |\tilde{x}_i(t+1) - \hat{x}_i(t)| &\leq |x_i + W(L_2, L+1)L_1\epsilon - X_i| \\ &\leq W(L_2, \mathcal{D}(\mathcal{G}^-) - 1)L_1\epsilon + \\ &W(L_2, L+1)L_1\epsilon \leq D, \end{aligned}$$

i.e,  $\hat{x}_i(t+1) = \tilde{x}_i(t+1)$ . An induction proves the result. Proof of Theorem 2: From Lemma 6 and (40)

 $\hat{x}_i(t) - x_i \ge -W(L_2, \mathcal{D}(\mathcal{G}^-) - 1)L_1 \epsilon \ \forall \ t \ge t_0 + T,$  (58) proving the lower bound on  $\hat{x}_i(t) - x_i$  implicit in the theorem statement. To prove the upper bound we will first show that  $\hat{x}_i(t) \le x_i = x_i + W(L_2, 0)\epsilon, \ \forall i \in \mathcal{F}_0, \ t \ge t_0 + T + T_0.$  (59) Then the repeated application of Lemma 7 will prove that  $\mathcal{D}(\mathcal{G}) = 1$ 

$$\hat{x}_i(t) - x_i \le W(L_2, \mathcal{D}(\mathcal{G}) - 1)L_1 \epsilon \ \forall \ t \ge t_0 + T + \sum_{i=0}^{\infty} T_i$$

and thus the theorem.

Consider  $i \in \mathcal{F}_0$ . As  $i \in \mathcal{E}(t)$  implies  $\hat{x}_i(t+1) \ge \hat{x}_i(t) + \delta$ from (16), (17) and (44), there is a  $t_0 + T < t \le t_0 + T + T_0$ such that  $i \in \mathcal{A}(t)$ . As  $\mathcal{F}_0 \subset \mathcal{S}_\infty$ , from (26)

$$\hat{x}_i(t) = \tilde{x}_i(t) \le s_i = x_i.$$
(60)

As  $t > t_0 + T$ , it follows from (40) that  $\hat{x}_k(t) \geq X_k$  for all  $k \in V$ . As  $f(\cdot, \cdot)$  is monotonically increasing in both its arguments and  $X_i \leq s_i$ , we obtain

$$\tilde{x}_{i}(t+1) = \min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(\hat{x}_{k}(t),\bar{e}_{ik}(t)\right)\right\},s_{i}\right\}$$
$$\geq \min\left\{\min_{k\in\mathcal{N}(i)}\left\{f\left(X_{k},e_{ik}^{-}\right)\right\},s_{i}\right\} = X_{i} \qquad (61)$$

where (61) uses (18), (29) and (32) when  $\epsilon > 0$ . When  $\epsilon = 0$ (61) holds without (18) as  $\bar{e}_{ik}(t) = e_{ik} = e_{ik}^{-}$  and  $X_k = x_k$ . Therefore,  $[X_i, x_i]$  contains both  $\hat{x}_i(t)$  and  $\tilde{x}_i(t+1)$ . Then (45) and Lemma 6 yield

$$|\tilde{x}_i(t+1) - \hat{x}_i(t)| \le |x_i - X_i| = W(L_2, \mathcal{D}(\mathcal{G}^-) - 1)L_1\epsilon \le D$$
  
From (15-16),  $\hat{x}_i(t+1) = \tilde{x}_i(t+1)$ . An induction proves (59).