

Noise Resilient Distributed Average Consensus Over Directed Graphs

Vivek Khatana , *Student Member, IEEE*, and Murti V. Salapaka , *Fellow, IEEE*

Abstract—Motivated by the needs of resiliency, scalability, and plug-and-play operation, distributed decision making is becoming increasingly prevalent. The problem of achieving consensus in a multi-agent system is at the core of distributed decision making. In this article, we study the problem of achieving average consensus over a *directed* multi-agent network when the communication links are corrupted with *noise*. We propose an algorithm where each agent updates its estimates based on the local mixture of information and adds its weighted noise-free initial information to its updates during every iteration. We demonstrate that, with appropriately designed weights, the agents achieve consensus despite additive communication noise. We establish that when the communication links are *noiseless*, the proposed algorithm moves towards consensus at a geometric rate. Under communication noise, we prove that the agent estimates reach a consensus value *almost surely*. We present numerical experiments to corroborate the efficacy of the proposed algorithm under different noise realizations and various algorithm parameters.

Index Terms—Almost sure convergence, average consensus, communication noise, directed graphs, multi-agent networks, push sum, ratio consensus.

I. INTRODUCTION

ENABLED by a large number of low-cost sensors, distributed computation, and storage, multi-agent systems (MAS) are being leveraged across diverse applications in modern-day systems, including power systems [1], [2], [3], banking and financial systems [4], control of unmanned aerial and autonomous underwater vehicles [5], and machine learning [6], which involve contain multiple decision-makers. Due to the inherently distributed nature of the available information in MAS, a distributed form of decision-making is required [7], [8]. Moreover, distributed control applications require estimates of global quantities to be locally available for designing control policies of the agents [9], [10]. A consensus algorithm that enables all agents to agree on the final decision (or state value) is an efficient way of obtaining global information/decisions in MAS. The problem of reaching consensus in MAS is a

critical issue in multi-agent signal processing and control. An important special case of agreement among the agents is when the agreed-upon state is the average of all the initial values of the agents; this is known as average consensus (see [11], [12], [13], [14], [15], [16], [17], [18], [19], [20] and references therein).

In recent years, there has been a wide interest in achieving distributed decision making over general directed graph topologies. In the *noiseless* case, the article [11] introduced an average consensus algorithm called the PushSum/Ratio-consensus algorithm, focusing on achieving average consensus in MAS with a directed interconnection structure. An important characteristic of the PushSum algorithm is its ability to be *synthesized locally* by each agent in the MAS without requiring any global information. The PushSum algorithm is extensively utilized in the distributed optimization literature for distributed decision-making over general directed graphs with noiseless communication (refer to [21], [22], [23], [24], [25], [26] and references therein).

Articles [12], [13] have explored the convergence of the PushSum algorithm when the underlying communication topology of the MAS is time-varying. The behavior of the PushSum algorithm under delays in communication channels is investigated in [14], [15]. Authors in [16], [17], [21] have extended the PushSum algorithm to higher dimensions, where the decision variables are vector-valued. Furthermore, finite-time termination criteria for the PushSum algorithm are developed in [13], [15], [16], [17], [21].

A major challenge in applying consensus protocols in MAS arises from the presence of noise in wireless communication channels between the agents. Noisy communication links can obscure the true states of agents and potentially push their estimates far from the original consensus value (the average), or even cause the consensus protocol to diverge. Despite the extensions of the applicability of the PushSum algorithm under more flexible graph and communication scenarios in articles such as [12], [13], [14], [15], [16], [17], [21], the performance guarantees of the PushSum algorithm *in the presence of communication noise remains unaddressed*. In light of this, the current work aims to tackle the following open research questions:

- 1 Can the celebrated PushSum algorithm, which is widely utilized in distributed decision making applications over directed graphs, work under communication noise?
- 2 If the PushSum algorithm does not perform well under noisy communication, then design a consensus algorithm that has guarantees for agents to achieve consensus and

Manuscript received 22 March 2023; revised 26 August 2023; accepted 3 October 2023. Date of publication 16 October 2023; date of current version 7 November 2023. This work was supported by the Advanced Research Projects Agency-Energy (ARPA-E) OPEN under the project “Rapidly Viable Sustained Grid” under Grant DE-AR0001016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gang Wang. (Corresponding author: Vivek Khatana.)

The authors are with the Department of Electrical Engineering, University of Minnesota, Twin Cities, MN 55455 USA (e-mail: khata010@umn.edu; murtis@umn.edu).

Digital Object Identifier 10.1109/TSIPN.2023.3324583

converge towards the average of their initial values, even in the presence of communication noise. This new algorithm should also have all the appealing features of the PushSum algorithm, including distributed synthesis and implementation.

Literature on Consensus Problems under noisy communication: The articles [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60] focus on undirected graphs where the agents are connected via bidirectional edges, and the communication topology is symmetric. Moreover, most of the existing algorithms are *synthesized centrally*, utilizing global information of the interconnection structure. Here, a symmetric double stochastic matrix (with the sum of entries of all rows and columns equal to one) or a balanced Laplacian matrix is utilized to capture the communication interconnection among the agents [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [42], [43], [44], [45], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62]. The consensus protocols in [41], [63], [64] utilize a row stochastic update (with a matrix having the sum of entries in each row equal to one). However, when using row stochastic weight matrices, the estimates incur a bias away from the average of the initial values of the agents.

In most schemes that address communication channel noise, it is assumed that the noise is a zero mean random variable independent across different agents and time instants [27], [28], [29], [30], [31], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [46], [47], [48], [49], [50], [51], [52], [53], [54], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65]. Article [30] considers the channel noise to follow a Poisson distribution, while [32] assumes a zero median communication noise. Additionally, [63] analyzes the algorithm under the assumption of communication noise with bounded p -th moments i.e., $\mathbb{E}|w_t|^p < \infty, p \in (1, 2]$. In [55], a multiplicative noise model is considered, where the noise intensities depend on the relative states of agents and form a martingale. A bounded noise model is assumed in [43], [44], [45], [48], where the communication link noise is allowed to randomly vary within upper and lower bounds.

In terms of convergence guarantees, the existing literature can be categorized into three categories: (i) the articles [27], [28], [29], [30], [31], [33], [37], [39], [40], [41], [42], [47], [50], [51], [52], [53], [54], [55], [57], [58], [59], [60], [63], [64], provide guarantees about weak consensus, where the agreement between the agents is reached in the mean square sense; (ii) the works in [27], [32], [34], [35], [37], [38], [50], [51], [55], [59], [63], [64], [65], [66] establish a strong consensus guarantee, ensuring an almost sure agreement between the agent states. *However, no analysis about the proximity of the asymptotic consensus value to the average of the initial values of the agents is provided;* (iii) under the assumption of bounded unknown noise, the analyses in [43], [44], [45], [48], [56], [62] offer an upper bound on the state residuals (differences between the estimates of different agents) within the MAS; *however the agent estimates not necessarily achieve consensus.*

Contributions of the current work: In this article, we tackle the challenge of designing an algorithm to attain average consensus within directed multi-agent graphs, even when the communication links between agents are affected by noise. Our proposed algorithm entails agents updating their states through local interactions with their neighbors connected via a directed link. We establish guarantees about the agreement between the agents' states under the proposed scheme. Furthermore, we substantiate the effectiveness of our developed algorithm through numerical simulations. The paper's technical contributions are outlined as follows:

- 1) The present article establishes that the PushSum algorithm fails to converge under imperfect communication with additive noise and is not applicable as such in these scenarios.
- 2) An important contribution of the current work is the proposed **Noise Resilient PushSum (NR-PushSum)** algorithm, which enhances and extends the PushSum algorithm by accommodating practical settings with noisy communication channels. Notably, this paper is the *first* to present such a significant extension.
- 3) The developed **NR-PushSum** algorithm possesses the following features:
 - a) In contrast to the existing algorithms in the literature that focus on achieving average consensus under noisy communication, the proposed **NR-PushSum** algorithm achieves average consensus over general directed graph topologies. Importantly, it doesn't necessitate restrictive assumptions regarding the symmetry and/or weight-balance of the communication graph.
 - b) The **NR-PushSum** algorithm can be synthesized in a distributed manner. Notably, **NR-PushSum** can be *designed* (rather than solely implemented distributively) using decentralized information. This noteworthy advantage facilitates the seamless integration of new agents into the network.
 - c) The **NR-PushSum** algorithm attains average consensus at a geometric rate when communication is noiseless. In the presence of communication noise, the agents' estimates under the **NR-PushSum** algorithm achieve consensus (perfect agreement) *almost surely*. Additionally, the algorithm parameters can be adjusted to ensure the consensus value is in close proximity to the average of the agents' initial states.

The structure of the remaining article is as follows. In Section I-A, we present some preliminary definitions and notations employed throughout the article. Section II introduces the problem statement and outlines the performance of the PushSum algorithm in the presence of communication noise. Section III delves into the comprehensive presentation of the proposed **NR-PushSum** algorithm. Convergence analysis of the **NR-PushSum** algorithm is provided in Section IV, encompassing both scenarios: perfect and noisy communication. To demonstrate the efficacy of the **NR-PushSum** algorithm, numerical simulations are presented in Section V. Finally, Section VI offers concluding remarks.

A. Definitions and Notations

First, we present some definitions that will be useful for the rest of the development. A detailed description of most of these notions is available in [67] and [68].

Definition 1: (Directed Graph) A directed graph $\mathcal{G}$ is a pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is a set of vertices (or nodes) and $\mathcal{E}$ is a set of edges, which are ordered subsets of two distinct elements of $\mathcal{V}$. If an edge from $j \in \mathcal{V}$ to $i \in \mathcal{V}$ exists then it is denoted as $(i, j) \in \mathcal{E}$.

Definition 2: (Path) In a directed graph, a directed path from node i to j exists if there is a sequence of distinct directed edges of $\mathcal{G}$ of the form $(k_1, i), (k_2, k_1), \dots, (j, k_m)$.

Definition 3: (Strongly Connected Graph) A directed graph is strongly connected if it has a directed path between each pair of distinct nodes i and j .

Definition 4: (In-neighborhood) The set of in-neighbors of node $i \in \mathcal{V}$ is denoted by $\mathcal{N}_i^I = \{j : (i, j) \in \mathcal{E}\}$. The number of agents $|\mathcal{N}_i^I|$ in the in-neighborhood of an agent i is called the in-degree of agent i .

Definition 5: (Out-neighborhood) The set of out-neighbors of node $i \in \mathcal{V}$ is denoted by $\mathcal{N}_i^O = \{j : (j, i) \in \mathcal{E}\}$. The number of agents $|\mathcal{N}_i^O|$ in the out-neighborhood of an agent i is called the out-degree of agent i .

Definition 6: (Column Stochastic Matrix) A real $n \times n$ matrix $A = [a_{ij}]$ is called a column-stochastic matrix if $0 \leq a_{ij} \leq 1$ and $\sum_{i=1}^n a_{ij} = 1$ for $1 \leq j \leq n$.

Definition 7: (Irreducible Matrix) An $n \times n$ matrix A is reducible if we may partition $\{1, \dots, n\}$ into two non-empty subsets E, F such that $A_{ij} = 0$ if $i \in E, j \in F$. Matrix A is called irreducible if it is not reducible.

Definition 8: (Aperiodic Matrix) Let A be a $n \times n$ matrix. Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a directed graph induced by the zero/non-zero structure of A . The matrix A is aperiodic if $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is strongly connected with at least one node that has a self-loop.

Define $[n] := \{1, 2, \dots, n\}$. We denote by A_{ij} the entry in i th $\in [m]$ row and j th $\in [n]$ column of the matrix $A \in \mathbb{R}^{m \times n}$. $\mathbf{1}$ denotes a vector of appropriate dimension with all entries equal to one. $\mathbf{I}_q \in \mathbb{R}^{q \times q}$ denotes the identity matrix of dimension q . For a vector $b \in \mathbb{R}^q$, we denote $\text{diag}(b)$, to be a $q \times q$ square diagonal matrix with the elements of vector b on the main diagonal.

We denote the Hadamard (or element-wise/entry-wise) product and division of matrices by $\odot$ and $\oslash$ respectively. $O(\cdot)$ and $o(\cdot)$ denote the standard *Big-O* and *Little-o* notations respectively. We use the term *almost surely* to state that an event happens with probability 1. We use the words node and agent interchangeably in the text.

II. CONSENSUS PROBLEM AND PUSHSUM ALGORITHM BEHAVIOR IN THE PRESENCE OF COMMUNICATION NOISE

Consider a set of n agents connected through a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The agents communicate via uni-directional communication links in $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The communication link between any agent i and its neighbor j is corrupted by noise. Each agent $i \in \mathcal{V}$ has an initial state $u_i \in \mathbb{R}^p$. The objective is to design an algorithm that allows each agent to distributively compute the

average,

$$\bar{u} := \frac{1}{n} \sum_{i=1}^n u_i, \quad (1)$$

by interacting only with its neighboring agents in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ under the communication constraints: uni-directional communication links with additive noise.

A. PushSum With Noisy Communication

Under the PushSum algorithm, at any discrete time-instant k each agent i maintains three estimates $x_i(k) \in \mathbb{R}^p, y_i(k) \in \mathbb{R}$ and $z_i(k) \in \mathbb{R}^p$. Due to noise in communication channels in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ the agent i receives a noisy version of the information sent by the agent j . Node i updates its estimates at $(k+1)$ th discrete time-instant, by performing a weighted combination of the (noisy) information received from its in-neighbors and its own information (available without any additive noise at the agent i) as follows:

$$x_i(k+1) = p_{ii}x_i(k) + \sum_{j \in \mathcal{N}_i^I} (p_{ij}x_j(k) + \eta_{x_{ij}}(k)), \quad (2)$$

$$y_i(k+1) = p_{ii}y_i(k) + \sum_{j \in \mathcal{N}_i^I} (p_{ij}y_j(k) + \eta_{y_{ij}}(k)), \quad (3)$$

where, $x_i(0) = u_i$, and $y_i(0) = 1$ for all $i \in [n]$ and the terms $\eta_{x_{ij}}(k)$ and, $\eta_{y_{ij}}(k)$ are the additive noise realizations in the estimates of agent j sent to agent i at any instant k . The variable $z_i(k)$ is the estimate of the average $\bar{u}$ maintained by agent $i \in \mathcal{V}$ at time-instant k which is updated as,

$$z_i(k+1) = \frac{1}{y_i(k+1)} x_i(k+1). \quad (4)$$

Let $\mathbf{P}$ be weight matrix corresponding to (2) and (3) with $\mathbf{P}_{ij} \in [0, 1]$ and,

$$\mathbf{P}_{ij} = \begin{cases} p_{ij} \neq 0, & \text{if } i = j \text{ or } (i, j) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The following assumption is typically made to study the PushSum algorithm:

Assumption 1: The directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is strongly connected, and the associated weight matrix $\mathbf{P}$ is column stochastic.

One choice of weights that satisfy Assumption 1 is the *out-degree based equal neighbor weights rule*: Here,

$$p_{ij} = \begin{cases} \frac{1}{1+|\mathcal{N}_j^O|}, & \text{if } i \in \mathcal{N}_j^O, \forall j, \\ 0, & \text{if } i \notin \mathcal{N}_j^O, \forall j. \end{cases} \quad (6)$$

Define for all $i \in [n]$,

$$\eta_{x_i}(k) := \sum_{j \in \mathcal{N}_i^I} \eta_{x_{ij}}(k), \quad \eta_{y_i}(k) := \sum_{j \in \mathcal{N}_i^I} \eta_{y_{ij}}(k). \quad (7)$$

We stack the agents' estimates in a column vector form and define the following quantities:

$$\mathbf{X}(k) := [x_1(k); \dots; x_n(k)] \in \mathbb{R}^{n \times p}, \quad (8)$$

$$\mathbf{y}(k) := [y_1(k); \dots; y_n(k)] \in \mathbb{R}^n, \quad (9)$$

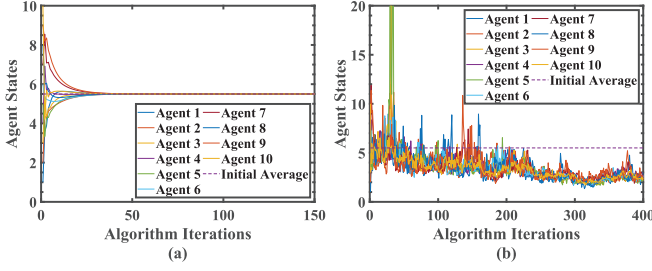


Fig. 1. Performance of the PushSum algorithm under the perfect communication (a) and with communication noise (b).

$$\eta_x(k) := [\eta_{x_1}(k); \dots; \eta_{x_n}(k)] \in \mathbb{R}^{n \times p}, \text{ and} \quad (10)$$

$$\eta_y(k) := [\eta_{y_1}(k); \dots; \eta_{y_n}(k)] \in \mathbb{R}^n. \quad (11)$$

Thus, updates (2) and (3) can be written as:

$$\mathbf{X}(k+1) = \mathbf{P}\mathbf{X}(k) + \eta_x(k), \mathbf{y}(k+1) = \mathbf{P}\mathbf{y}(k) + \eta_y(k),$$

where, $\eta_x(k)$ and $\eta_y(k)$ are the additive noise. Left multiplying by $\mathbf{1}^\top$ on both sides of the above equations, we have

$$\mathbf{1}^\top \mathbf{X}(k+1) = \mathbf{1}^\top \mathbf{X}(k) + \mathbf{1}^\top \eta_x(k), \text{ and}$$

$$\mathbf{1}^\top \mathbf{y}(k+1) = \mathbf{1}^\top \mathbf{y}(k) + \mathbf{1}^\top \eta_y(k).$$

Thus, the sum of estimates $\mathbf{X}(k)$ and $\mathbf{y}(k)$ integrate the communication noise and thus is equivalent to a random walk with unbounded variance. Note that in the case when the communication links are noiseless the sum of the estimates $\mathbf{X}(k)$ and $\mathbf{y}(k)$ are conserved during the iterations [2] that plays a crucial role in the convergence of the PushSum algorithm to the average of the initial values. The random walk behavior of the sums implies that the estimates $\mathbf{X}(k)$ and $\mathbf{y}(k)$ and hence, $\mathbf{Z}(k) = [z_1(k), \dots, z_n(k)] \in \mathbb{R}^{n \times p}$ do not converge to a meaningful value. This is clear from the trajectories of the agent states $\mathbf{Z}_i(k)$ under the PushSum algorithm for a 10 node directed graph with noisy communication shown in Fig. 1. Here, the initial values of the agents are chosen to be [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] (with the average being 5.5). The communication links have an additive noise modeled as a normal random variable, $\mathcal{N}(0, 0.1)$. These observations suggest that the PushSum algorithm *does not* perform adequately under communication noise.

III. PROPOSED APPROACH FOR ACHIEVING AVERAGE CONSENSUS IN NOISY DIRECTED NETWORKS

In this section, we propose a distributed algorithm that allows all the agents to determine the initial average $\bar{u}$ (defined in (1)) under noisy communication. We call it the **NR-PushSum** algorithm as it is based on the PushSum algorithm, similar to which here each agent i , at any discrete time-instant k , maintains three estimates $x_i(k) \in \mathbb{R}^p$, $y_i(k) \in \mathbb{R}$ and $z_i(k) \in \mathbb{R}^p$. We utilize two strategies to design the **NR-PushSum** algorithm that counter the effect of the communication noise:

- 1) The estimates $x_i(k)$, $y_i(k)$ are updated via a local mixing across the neighborhood of the agent $i \in \mathcal{V}$. Every agent i , weighs the information sent by its neighbors (agent $j \in \mathcal{N}_i^I$) and its own information to appropriately control the amount of the noise to be added to its updated estimates.

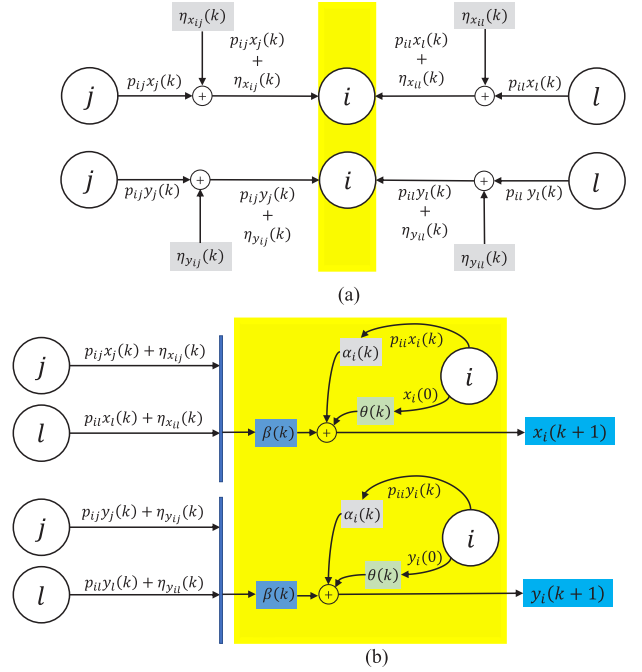


Fig. 2. Estimate updates of the **NR-PushSum** algorithm: (a) agent i receives noisy information from neighbors j and l , (b) agent i updates its estimates using the received information, its own current estimate and its initial values.

- 2) Each agent i , utilizes its initial values $x_i(0)$ and $y_i(0)$ to update its estimates. The initial values $x_i(0)$ and $y_i(0)$ are deterministic and known to the agent i . Utilizing the noise-free initial values provides a tether to the agents' estimates in tracking the initial average in presence of the communication noise.

Incorporating these two strategies together, the estimates in the **NR-PushSum** algorithm are updated as follows, for $i \in \mathcal{V}$:

$$x_i(k+1) = \alpha_i(k) p_{ii} x_i(k) + \beta(k) \sum_{j \in \mathcal{N}_i^I} (p_{ij} x_j(k) + \eta_{x_{ij}}(k)) + \theta(k) x_i(0), \quad (12)$$

$$y_i(k+1) = \alpha_i(k) p_{ii} y_i(k) + \beta(k) \sum_{j \in \mathcal{N}_i^I} (p_{ij} y_j(k) + \eta_{y_{ij}}(k)) + \theta(k) y_i(0), \quad (13)$$

respectively, where, $x_i(0) = u_i$, and $y_i(0) = 1$ for all $i \in [n]$. $\alpha_i(k) \in [0, 1]$, $\forall i \in [n]$, k , and $\beta(k) \in [0, 1]$, $\forall k$ are the relative weights that agent i utilize to control the information mixing, and $\theta(k) \in \mathbb{R}$, $\forall k$, is utilized to control the influence of the noise-free initial values on the estimate updates. The variable $z_i(k)$ maintained by agent $i \in \mathcal{V}$ at time-instant k is updated as in the PushSum algorithm described by

$$z_i(k+1) = \frac{1}{y_i(k+1)} x_i(k+1). \quad (14)$$

Fig. 2 shows the estimate updates for an agent $i \in \mathcal{V}$ at any discrete-time instant k . Fig. 2(a) shows the communication path where agent i receives information from its in-neighbors j and l . Fig. 2(b) shows the computation step where agent i updates its estimates based on the information from the neighbors j and l , its own current state values $(x_i(k), y_i(k))$ and the noise-free initial values $(x_i(0), y_i(0))$. Define $\hat{\mathbf{P}}$ such that

$$\hat{\mathbf{P}}_{ij} = \begin{cases} \mathbf{P}_{ij}, & \text{if } i \neq j \\ 0, & \text{if } i = j. \end{cases} \quad (15)$$

Note that due to Assumption 1 the matrix $\hat{\mathbf{P}}$ defined in (15) is a column sub-stochastic matrix, i.e., $\sum_{i=1}^n \hat{\mathbf{P}}_{ij} < 1$ for all $j \in [n]$. The parameter $\alpha_i(k)$, $i \in [n]$ is set as:

$$\alpha_i(k) = \frac{1 - \beta(k)(1 - p_{ii})}{p_{ii}}. \quad (16)$$

Let $\mathbf{A}(k) := (\text{diag}(\alpha_i(k)) \odot \mathbf{P})$ and $\mathbf{B}(k) := \beta(k)\hat{\mathbf{P}}$. Using the defined notations in (7)–(11), updates (12)–(14) can be written as:

$$\begin{aligned} \mathbf{X}(k+1) &= \mathbf{A}(k)\mathbf{X}(k) + \mathbf{B}(k)\mathbf{X}(k) + \theta(k)\mathbf{X}(0) \\ &\quad + \beta(k)\eta_x(k), \end{aligned} \quad (17)$$

$$\begin{aligned} \mathbf{y}(k+1) &= \mathbf{A}(k)\mathbf{y}(k) + \mathbf{B}(k)\mathbf{y}(k) + \theta(k)\mathbf{y}(0) \\ &\quad + \beta(k)\eta_y(k), \end{aligned} \quad (18)$$

$$\mathbf{Z}(k+1) = \mathbf{X}(k+1) \odot \mathbf{y}(k+1). \quad (19)$$

The initial values of the agents in updates (12)–(14) help preserve the true information of the agents and reduce the effect of the communication noise. Moreover, the agents vary the weights $\beta(k)$ and $\theta(k)$ dynamically which allows the agents' estimates $x_i(k)$ and $y_i(k)$ to reach a steady state. Specifically, the weights $\beta(k)$ and $\theta(k)$ satisfy the following assumptions:

Assumption 2: The sequence $\{\beta(k)\}_{k \geq 0}$ is such that:

- i) $\beta(k) \in [0, 1]$, for all k ,
- ii) there exists a $K_\beta \geq 1$ such that

$$ak^{-q} \leq \beta(k) \leq bk^{-q}, \quad (20)$$

for all $k \geq K_\beta$, where $q > 1$ and $0 \leq a \leq b < \infty$.

Note that the requirement (20) with a suitable K_β , allows more flexibility in choosing $\beta(k)$ as larger values of a and b can be chosen while maintaining $\beta(k) \in [0, 1]$, $k \geq K_\beta$. In the sequel, the parameters a, b and q will be treated as fixed constants.

Assumption 3: The sequence $\{\theta(k)\}_{k \geq 0}$ satisfies:

$$\theta(k) \in [0, \infty), \text{ for all } k, \text{ and } \sum_{k=0}^{\infty} \theta(k) < \infty. \quad (21)$$

Note that one can appropriately choose a $K_\theta > 0$ with large values of $\theta(k)$ chosen early on, i.e. when $k < K_\theta$, to give more importance to the initial values of the agents and then $\theta(k)$ be gradually reduced, i.e. $\theta(k) \in [0, 1]$ for $k > K_\theta$ to satisfy Assumption 3. In Section V, we demonstrate the performance of the proposed **NR-PushSum** algorithm under different choices of the $\beta(k)$ and $\theta(k)$ sequences.

Remark 1: The proposed **NR-PushSum** algorithm inherits nice properties of the PushSum algorithm, i.e. the

NR-PushSum algorithm is applicable on general directed graph topologies and can be both implemented and synthesized/designed distributively. Recall Fig. 2(a), every agent $j \in \mathcal{V}$ decides the weights p_{ij} for sending its information over the out-going links to the neighbor agent i , independent of the other agents. The agent $i \in \mathcal{V}$ upon receiving the information from its in-neighbors adds up its own estimate and the initial value with the received information after weighing them by weights $\alpha_i(k), \theta(k)$ and $\beta(k)$ respectively. Note that the agents do not need to have any centralized information for weights $\beta(k)$ and $\theta(k)$. Each agent can dynamically update the weights $\beta(k)$ and $\theta(k)$ based on Assumptions 2 and 3.

IV. CONVERGENCE ANALYSIS

The proposed **NR-PushSum** algorithm being a new algorithm, we start by examining its convergence properties in the noiseless case. Following that, our analysis expands to investigate the *almost sure* convergence of the proposed **NR-PushSum** algorithm to a consensus value, even in the presence of additive noise affecting communication channels.

We believe that as the proposed **NR-PushSum** algorithm itself is new, its properties in the noiseless case have to be established. This examination helps establish that the proposed **NR-PushSum** algorithm doesn't compromise convergence properties in the absence of noise, all while attaining favorable guarantees in the presence of noise. Furthermore, the exploration of the noise-free scenario provides a solid mathematical foundation that can be extended to establish guarantees for our method in situations where communication is affected by additive noise.

A. Analysis Under Noiseless Communication Channels

Here, we will analyze a special case of the scheme (12)–(14) where the communication channels are assumed to be noiseless. In particular, $\eta_{x_{ij}}(k) = 0 = \eta_{y_{ij}}(k)$ for all $i, j \in [n]$ at all time-instants $k \geq 0$. Let $\mathbf{M}(k) := \mathbf{A}(k) + \mathbf{B}(k)$ for all $k \geq 0$ and define the following operation,

$$\prod_{r=s}^l \mathbf{M}(r) := \begin{cases} \prod_{r=s}^l \mathbf{M}(r), & \text{if } l \geq s \\ \mathbf{I}_n, & \text{otherwise.} \end{cases} \quad (22)$$

Further define,

$$W(l, k) := \prod_{r=l}^k \mathbf{M}(r). \quad (23)$$

We begin by establishing that the following relation holds.

Lemma 1: $\sum_{j=1}^n W_{ij}(l, k) > w_o^{k-l+1}$, for all $k \geq l \geq 1$, where, $w_o > 0$ is a constant.

Proof: For $\mathbf{M}(l)$, $l \geq 1$ note that,

$$\begin{aligned} \sum_{j=1}^n \mathbf{M}_{ij}(l) &= 1 - \beta(l)(1 - p_{ii}) + \beta(l) \sum_{j \neq i}^n p_{ij} \\ &= 1 - \beta(l) + \beta(l) \sum_{j=1}^n p_{ij} \\ &> 1 - \beta(l) > 1 - \max_l \beta(l) = w_o > 0. \end{aligned} \quad (24)$$

We use an argument based on induction on the matrices of the form $W(l, l+t)$. Let $T = k - l$. For $t = 1$, note that $W(l, l+1) = \mathbf{M}(l)\mathbf{M}(l+1)$ and the i - j th entry of the matrix $W(l, l+1)$ is given by $W_{ij}(l, l+1) = [\mathbf{M}(l)\mathbf{M}(l+1)]_{ij} = \sum_{s=1}^n \mathbf{M}_{is}(l)\mathbf{M}_{sj}(l+1)$. Taking the sum over j we have,

$$\begin{aligned} \sum_{j=1}^n W_{ij}(l, l+1) &= \sum_{j=1}^n \sum_{s=1}^n \mathbf{M}_{is}(l)\mathbf{M}_{sj}(l+1) \\ &= \sum_{s=1}^n \left[\mathbf{M}_{is}(l) \sum_{j=1}^n \mathbf{M}_{sj}(l+1) \right] \\ &> \sum_{s=1}^n \mathbf{M}_{is}(l)w_o \\ &> w_o^2, \end{aligned}$$

where, we used (24) in the last two inequalities. Assume, for $t = T = k - l$, that $\sum_{j=1}^n W_{ij}(l, l+T) = \sum_{j=1}^n W_{ij}(l, k) > w_o^{k-l+1}$. For $t = T + 1$,

$$\begin{aligned} \sum_{j=1}^n W_{ij}(l, l+T+1) &= \sum_{j=1}^n \sum_{s=1}^n W_{is}(l, l+T)\mathbf{M}_{sj}(l+T+1) \\ &= \sum_{s=1}^n \left[W_{is}(l, l+T) \sum_{j=1}^n \mathbf{M}_{sj}(l+T+1) \right] \\ &> \sum_{s=1}^n W_{is}(l, l+T)w_o \\ &> w_o^{k-l+1}w_o = w_o^{k-l+2}, \end{aligned}$$

where, we used (24) and the induction hypothesis in the last inequality. Therefore, induction holds. Thus,

$$\sum_{j=1}^n W_{ij}(l, k) > w_o^{k-l+1}, \quad (25)$$

for all $k \geq l \geq 1$. $\square$

Lemma 2: Let Assumption 1 hold. Assume that the communication channels in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are **noiseless**. Let the sequences $\{\beta(k)\}_{k \geq 0}$ and $\{\theta(k)\}_{k \geq 0}$ satisfy Assumptions 2 and 3 respectively. Under the initialization $y_i(0) = 1$ for all $i \in [n]$, there exist $y_o > 0$ such that we have $y_i(k) > y_o$ for all $k \geq 1, i \in [n]$. As a consequence the states $z_i(k)$ are well-defined for all $k \geq 1, i \in [n]$.

Proof: For $\eta_y(k) = 0$, with $\mathbf{M}(k) := \mathbf{A}(k) + \mathbf{B}(k)$ the update $\mathbf{y}(k)$ in the iteration (18) becomes:

$$\mathbf{y}(k+1) = \mathbf{M}(k)\mathbf{y}(k) + \theta(k)\mathbf{y}(0). \quad (26)$$

Recursive application of (26) using the notation in (22) gives,

$$\mathbf{y}(k+1) = \prod_{r=0}^k \mathbf{M}(r)\mathbf{y}(0) + \sum_{s=0}^k \prod_{r=s+1}^k \mathbf{M}(r)\theta(s)\mathbf{y}(0). \quad (27)$$

From the definition of matrices $\mathbf{A}(k)$ and $\mathbf{B}(k)$ under Assumption 1 it is easy to check that the matrix $\mathbf{M}(k)$ is column-stochastic for all $k \geq 0$ (see Appendix A, [69]). Further, note that $W(l, k)$ for all $k, l \geq 0$ is a column-Stochastic, Irreducible, and Aperiodic (cSIA) matrix (see Appendix A, [69]). Column stochasticity follows by induction and the column stochasticity of $\mathbf{M}(k)$. From the expression for $W(l, k)$ in (23), the zero/non-zero structure of $W(l, k)$ corresponds to a graph of n nodes that includes all the edges in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and since the original graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is strongly connected, $W(l, k)$ corresponds to a graph that is strongly connected and hence is irreducible [70]. Since the graph corresponding to $W(l, k)$ is strongly connected and the diagonal entries of $W(l, k)$ are positive, therefore, $W(l, k)$ is aperiodic. We utilize the following result from [71] on the product of column stochastic matrices:

Theorem 1: Let $\mathcal{M} := \{\mathbf{M}(1), \mathbf{M}(2), \dots\}$ be a collection of $n \times n$ column stochastic matrices. Define a word in $\mathcal{M}$ of length ℓ as the product of ℓ matrices from the collection $\mathcal{M}$ (repetitions permitted). Let $\mathcal{M}$ be such that any word in $\mathcal{M}$ is cSIA. Then given $\varepsilon > 0$ there exists an integer ℓ_ε such that any word $\mathbf{N} \in \mathbb{R}^{n \times n}$ in $\mathcal{M}$ of length $\ell \geq \ell_\varepsilon$ satisfies $\max_j \max_{i_1, i_2} |\mathbf{N}_{j, i_1} - \mathbf{N}_{j, i_2}| < \varepsilon$.

Theorem 1 implies that any sufficiently long word in the $\mathcal{M}$ has all its columns such that the difference in the columns element-wise can be made arbitrarily small. Thus, given $\varepsilon > 0$, there exists a $\ell_\varepsilon > 0$ such that $W(l, k)$ (a word in $\mathcal{M}$ that is cSIA) is equal to $\mathbf{v}(k)\mathbf{1}^\top + E(k)$, for all $k > l + \ell_\varepsilon$, where $\mathbf{v}(k)$ is an appropriate vector with positive entries, and $E(k)$ is an error matrix such that the absolute value of its elements is smaller than $\varepsilon/2$ i.e., $|E_{ij}(k)| < \varepsilon/2$ for all $i, j \in [n]$.

Further, recall $\mathbf{M}(k) := \mathbf{A}(k) + \mathbf{B}(k)$. Based on the definition of $\mathbf{A}(k)$ and $\mathbf{B}(k)$, we can re-write $\mathbf{M}(k)$ as

$$\mathbf{M}(k) := \mathbf{I}_n + \beta(k)\tilde{\mathbf{P}}, \quad (28)$$

where, $\tilde{\mathbf{P}} = \text{diag}(p_{ii} - 1) + \hat{\mathbf{P}}$.

Given any $\varepsilon > 0$ for $k > \ell_\varepsilon$, note that $W(0, k) = \mathbf{v}(k)\mathbf{1}^\top + E(k)$, with $|E_{ij}(k)| < \varepsilon/2$ for all i, j .

$$\begin{aligned} W(0, k) &= W(0, \ell_\varepsilon) \prod_{r=\ell_\varepsilon+1}^k \mathbf{M}(r) \\ &= [\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top + E(\ell_\varepsilon)] \prod_{r=\ell_\varepsilon+1}^k \mathbf{M}(r) \\ &= [\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top + E(\ell_\varepsilon)] \prod_{r=\ell_\varepsilon+1}^k (\mathbf{I}_n + \beta(r)\tilde{\mathbf{P}}) \\ &= \mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top \prod_{r=\ell_\varepsilon+1}^k (\mathbf{I}_n + \beta(r)\tilde{\mathbf{P}}) \\ &\quad + E(\ell_\varepsilon) \prod_{r=\ell_\varepsilon+1}^k (\mathbf{I}_n + \beta(r)\tilde{\mathbf{P}}) \\ &= \mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top + E(\ell_\varepsilon)W(\ell_\varepsilon + 1, k), \end{aligned} \quad (29)$$

where, we used the fact that $\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top \tilde{\mathbf{P}} = \mathbf{0}_{n \times n}$.

From (27), (23) (28), (29), and initialization $y_j(0) = 1$ for all j we have for $k > \ell_\varepsilon$,

$$\begin{aligned}
y_i(k+1) &= \sum_{j=1}^n W_{ij}(0, k) + \sum_{j=1}^n \sum_{s=0}^k \theta(s) W_{ij}(s+1, k) \\
&> \sum_{j=1}^n W_{ij}(0, k) \\
&= \sum_{j=1}^n [\mathbf{v}(\ell_\varepsilon) \mathbf{1}^\top + E(\ell_\varepsilon) W(\ell_\varepsilon + 1, k)]_{ij} \\
&> \sum_{j=1}^n \left[\mathbf{v}(\ell_\varepsilon) \mathbf{1}^\top - \frac{\varepsilon}{2} \mathbf{1} \mathbf{1}^\top W(\ell_\varepsilon + 1, k) \right]_{ij} \\
&> \sum_{j=1}^n \left[\mathbf{v}(\ell_\varepsilon) \mathbf{1}^\top - \frac{\varepsilon}{2} \mathbf{1} \mathbf{1}^\top \right]_{ij} \\
&= n \mathbf{v}_i(\ell_\varepsilon) - n \frac{\varepsilon}{2} \\
&> n \min_{1 \leq i \leq n} \mathbf{v}_i(\ell_\varepsilon) - \varepsilon \frac{n}{2} > 0,
\end{aligned}$$

where, we used the fact that $W(l, k)$ is column stochastic for all $k, l \geq 0$ in the third last inequality. For $k \leq \ell_\varepsilon$, from Lemma 1 we have,

$$\begin{aligned}
y_i(k) &> w_o^{k-1} + \sum_{s=0}^{k-1} \theta(s) w_o^{k-s-1} \\
&> w_o^{k-1} + \theta(0) w_o^{k-1} \geq (1 + \theta(0)) w_o^{\ell_\varepsilon} > 0.
\end{aligned}$$

Therefore, $y_i(k) > y_o := \min\{n \min_{1 \leq i \leq n} \mathbf{v}_i(\ell_\varepsilon), (1 + \theta(0)) w_o^{\ell_\varepsilon}\}$, for all $k \geq 1$. Thus, $z_i(k) = \frac{1}{y_i(k)} x_i(k)$ is well defined for all $i \in [n], k \geq 0$. $\square$

Theorem 2: Let Assumption 1 hold. Assume the edges in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are **not** corrupted by additive noise in communication channels. Let $\{z_i(k)\}_{k \geq 0}^{i \in [n]}$ be the sequences generated by the iteration (14). Let the sequence $\{\theta(k)\}_{k \geq 0}$ satisfy Assumption 3. Then, the estimates $z_i(k)$ converges to the average $\bar{u}$, i.e., $\lim_{k \rightarrow \infty} z_i(k) = \bar{u}$, for all $i \in [n]$.

Proof: Similar to (26) the update $\mathbf{X}(k)$ in the iteration (17) when $\eta_x(k) = 0$ becomes:

$$\mathbf{X}(k+1) = \mathbf{M}(k) \mathbf{X}(k) + \theta(k) \mathbf{X}(0). \quad (30)$$

Recursive application of (30) with the notation (23) gives,

$$\mathbf{X}(k+1) = \underbrace{W(0, k) \mathbf{X}(0)}_{\text{first}} + \underbrace{\sum_{s=0}^k \theta(s) W(s+1, k) \mathbf{X}(0)}_{\text{second}}. \quad (31)$$

Note that as $W(l, k)$ is column stochastic for all $k, l \geq 1$, $\|W(l, k)\| \leq \bar{w} < \infty$ for all k, l .

Under Assumption 3 and from Theorem 1, given $\varepsilon > 0$. Define $\hat{\varepsilon} := \frac{\varepsilon / \sum_{s=0}^{\infty} \theta(s)}{\max\{\|\mathbf{X}(0)\|, n\}}$. There exists $\tilde{k}$ and $\mathbf{v}(k) \in \mathbb{R}^n$ such that $\sum_{s=\tilde{k}}^{\infty} \theta(s) \leq \frac{\hat{\varepsilon}}{2\bar{w}}$ and $W(l, k) = \mathbf{v}(k) \mathbf{1}^\top + E(k)$, for all $k > \tilde{k} + l$, where, $|E_{ij}(k)| \leq \frac{\hat{\varepsilon}}{2}$ for all $i, j \in [n]$.

Thus, for all $k > \tilde{k}$, the first term in (31) becomes,

$$W(0, k) \mathbf{X}(0) = [\mathbf{v}(k) \mathbf{1}^\top + E(k)] \mathbf{X}(0) \quad (32)$$

where, $|E_{ij}(k)| < \hat{\varepsilon}/2$ for all $i, j \in [n]$. Consider the second term in (31),

$$\begin{aligned}
\sum_{s=0}^k \theta(s) W(s+1, k) \mathbf{X}(0) &= \underbrace{\sum_{s=0}^{\tilde{k}} \theta(s) W(s+1, k) \mathbf{X}(0)}_{:=\psi_1(k)} \\
&+ \underbrace{\sum_{s=\tilde{k}+1}^k \theta(s) W(s+1, k) \mathbf{X}(0)}_{:=\psi_2(k)}. \quad (33)
\end{aligned}$$

Note, $k > \tilde{k} + s + 1$ for all the terms in $\psi_1(k)$. Substituting $W(s+1, k) := \mathbf{v}(k) \mathbf{1}^\top + \hat{E}(k)$, with $|\hat{E}_{ij}(k)| < \hat{\varepsilon}/2$ for all $i, j \in [n]$ we get,

$$\psi_1(k) \mathbf{X}(0) = \sum_{s=0}^{\tilde{k}} \theta(s) [\mathbf{v}(k) \mathbf{1}^\top + \hat{E}(k)] \mathbf{X}(0). \quad (34)$$

Consider $\psi_2(k)$,

$$\begin{aligned}
\psi_2(k) &= \sum_{s=\tilde{k}+1}^k W(s+1, k) \theta(s) \\
&= \theta(k) + \mathbf{M}(k) \theta(k-1) + W(k-1, k) \theta(k-2) + \dots \\
&\quad + W(\tilde{k}+2, k) \theta(\tilde{k}+1).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\psi_2(k)\| &\leq \theta(k) + \bar{w} \theta(k-1) + \dots + \bar{w} \theta(\tilde{k}+1) \\
&\leq \bar{w} \sum_{s=\tilde{k}}^{\infty} \theta(s) \leq \frac{\hat{\varepsilon}}{2}. \quad (35)
\end{aligned}$$

Using (31)–(35) we have for all $k > \tilde{k}$,

$$\mathbf{X}(k+1) \leq \underbrace{\left(1 + \sum_{s=0}^{\tilde{k}} \theta(s)\right) \mathbf{v}(k) \mathbf{1}^\top \mathbf{X}(0)}_{:=\chi_1(k)} + \varepsilon \mathbf{1}_n \mathbf{1}_p^\top, \quad (36)$$

$$\mathbf{X}(k+1) \geq \left(1 + \sum_{s=0}^{\tilde{k}} \theta(s)\right) \mathbf{v}(k) \mathbf{1}^\top \mathbf{X}(0) - \varepsilon \mathbf{1}_n \mathbf{1}_p^\top, \quad (37)$$

where, $\mathbf{1}_n$ and $\mathbf{1}_p$ are the vectors of dimensions n and p with all entries equal to 1. We drop the subscript when the dimension of the vector is clear from the context. Following the similar procedure for (26) we get, for $k \geq \tilde{k}$,

$$\mathbf{y}(k+1) \leq \chi_1(k) \mathbf{y}(0) + \varepsilon \mathbf{1}, \text{ and}, \quad (38)$$

$$\mathbf{y}(k+1) \geq \chi_1(k) \mathbf{y}(0) - \varepsilon \mathbf{1}. \quad (39)$$

From (19), and (36)–(39), for any $i \in [n]$,

$$\frac{[\chi_1(k) \mathbf{X}(0) - \varepsilon \mathbf{1}_n \mathbf{1}_p^\top]_i}{[\chi_1(k) \mathbf{y}(0) + \varepsilon \mathbf{1}]_i} \leq z_i(k+1) \leq \frac{[\chi_1(k) \mathbf{X}(0) + \varepsilon \mathbf{1}_n \mathbf{1}_p^\top]_i}{[\chi_1(k) \mathbf{y}(0) - \varepsilon \mathbf{1}]_i}, \quad (40)$$

where, $[\cdot]_i$ denote the i th row (i th entry) of the corresponding matrices (vectors). Denote $\bar{\theta} := 1 + \sum_{s=0}^{\tilde{k}} \theta(s)$, and $\mathbf{v}_i(k)$ as the i th entry of $\mathbf{v}(k)$.

For $k > \tilde{k}$, for any $i \in [n]$, from (40)

$$\begin{aligned} z_i(k+1) &\leq \frac{[\chi_1(k)\mathbf{X}(0) + \varepsilon \mathbf{1}_n \mathbf{1}_p^\top]_i}{[\chi_1(k)\mathbf{y}(0) - \varepsilon \mathbf{1}]_i} \\ &= \frac{\bar{\theta} \mathbf{v}_i(k) \mathbf{1}^\top \mathbf{X}(0) + \varepsilon \mathbf{1}^\top}{\bar{\theta} \mathbf{v}_i(k) \mathbf{1}^\top \mathbf{y}(0) - \varepsilon} \\ &= \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}} \mathbf{1}^\top}{1 - \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}} \\ &= \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \left(\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \mathbf{1}^\top \right) \left(\frac{\frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}}{1 - \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}} \right) \end{aligned}$$

Similarly, for $k > \tilde{k}$, for any $i \in [n]$,

$$\begin{aligned} z_i(k+1) &\geq \frac{[\chi_1(k)\mathbf{X}(0) - \varepsilon \mathbf{1}_n \mathbf{1}_p^\top]_i}{[\chi_1(k)\mathbf{y}(0) + \varepsilon \mathbf{1}]_i} \\ &= \frac{\bar{\theta} \mathbf{v}_i(k) \mathbf{1}^\top \mathbf{X}(0) - \varepsilon \mathbf{1}^\top}{\bar{\theta} \mathbf{v}_i(k) \mathbf{1}^\top \mathbf{y}(0) + \varepsilon} \\ &= \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} - \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}} \mathbf{1}^\top}{1 + \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}} \\ &= \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} - \left(\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \mathbf{1}^\top \right) \left(\frac{\frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}}{1 + \frac{\varepsilon}{n \mathbf{v}_i(k) \bar{\theta}}} \right) \end{aligned}$$

Let $\bar{\varepsilon}$ be such that, $\varepsilon \leq \frac{\bar{\varepsilon}}{(1+\bar{\varepsilon})} (n \bar{\theta} \mathbf{v}_i(k))$. Therefore, for all $k > \tilde{k}$, and for all $i \in [n]$,

$$-\left(\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \mathbf{1}^\top \right) \bar{\varepsilon} \leq z_i(k) - \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} \leq \left(\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \mathbf{1}^\top \right) \bar{\varepsilon},$$

where, $\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} = \bar{u}$. Note that $\bar{\varepsilon}$ can be made arbitrarily small. Thus, we have the desired result. $\square$

Theorem 3: Let Assumption 1 hold. Assume the edges in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are **not** corrupted by additive noise in communication channels. Let $\{x_i(k), y_i(k), z_i(k)\}_{k \geq 0}^{i \in [n]}$ be the sequences generated by iterations (12)–(14). Let the sequence $\{\theta(k)\}_{k \geq 0}$ satisfy Assumption 3. Then, for all $j \in [n]$ and $k \geq 0$,

$$\|z_j(k) - \bar{u}\| \leq C_1 \left[\lambda^k + \sum_{s=0}^{k-1} \theta(s) \lambda^{k-s-1} + \frac{1}{C} \theta(k) \right],$$

where, C, C_1 and $\lambda \in (0, 1)$ are some positive constants.

Proof: Recall (27) and (31),

$$\mathbf{y}(k+1) = W(0, k) \mathbf{y}(0) + \sum_{s=0}^k \theta(s) W(s+1, k) \mathbf{y}(0),$$

$$\mathbf{X}(k+1) = W(0, k) \mathbf{X}(0) + \sum_{s=0}^k \theta(s) W(s+1, k) \mathbf{X}(0).$$

Since, $W(l, k), k \geq l \geq 0$ is column stochastic we have,

$$\begin{aligned} \mathbf{1}^\top \mathbf{X}(k+1) &= \mathbf{1}^\top W(0, k) \mathbf{X}(0) \\ &+ \sum_{s=0}^k \theta(s) \mathbf{1}^\top W(s+1, k) \mathbf{X}(0) \\ &= \mathbf{1}^\top \mathbf{X}(0) + \sum_{s=0}^k \theta(s) \mathbf{1}^\top \mathbf{X}(0). \end{aligned} \quad (41)$$

From [25] Corollary 2, there exists stochastic vectors $\mathbf{v}(k) \in \mathbb{R}^n$ and constants $C, \lambda \in (0, 1)$ such that for any given i , for all $j = 1, \dots, n$,

$$|W_{ij}(l, k) - \mathbf{v}_i(k)| \leq C \lambda^{k-l}, \text{ for all } k \geq l \geq 0. \quad (42)$$

Multiplying $\mathbf{M}(k+1)$ on both sides to (31) and $\mathbf{v}(k+1)$ to (41) and subtracting we get,

$$\begin{aligned} &(\mathbf{M}(k+1) - \mathbf{v}(k+1) \mathbf{1}^\top) \mathbf{X}(k+1) \\ &= (W(0, k+1) - \mathbf{v}(k+1) \mathbf{1}^\top) \mathbf{X}(0) \\ &+ \sum_{s=0}^k \theta(s) (W(s+1, k+1) - \mathbf{v}(k+1) \mathbf{1}^\top) \mathbf{X}(0) \\ &\Rightarrow \mathbf{M}(k+1) \mathbf{X}(k+1) = \mathbf{v}(k+1) \mathbf{1}^\top \mathbf{X}(k+1) \\ &+ (W(0, k+1) - \mathbf{v}(k+1) \mathbf{1}^\top) \mathbf{X}(0) \\ &+ \sum_{s=0}^k (W(s+1, k+1) - \mathbf{v}(k+1) \mathbf{1}^\top) \mathbf{X}(0) \theta(s). \end{aligned} \quad (43)$$

Define $\mathbf{D}(l, k) := W(l, k) - \mathbf{v}(k) \mathbf{1}^\top$. Replacing $k+1$ by k in (43) and adding $\theta(k) \mathbf{X}(0)$ on both sides. Comparing with (30) we get,

$$\begin{aligned} \mathbf{X}(k+1) &= \mathbf{M}(k) \mathbf{X}(k) + \theta(k) \mathbf{X}(0) \\ &= \mathbf{v}(k) \mathbf{1}^\top \mathbf{X}(k) + \mathbf{D}(0, k) \mathbf{X}(0) \\ &+ \sum_{s=0}^{k-1} \mathbf{D}(s+1, k) \mathbf{X}(0) \theta(s) + \theta(k) \mathbf{X}(0). \end{aligned} \quad (44)$$

Similarly, for (27)

$$\begin{aligned} \mathbf{y}(k+1) &= \mathbf{M}(k) \mathbf{y}(k) + \theta(k) \mathbf{y}(0) \\ &= \mathbf{v}(k) \mathbf{1}^\top \mathbf{y}(k) + \mathbf{D}(0, k) \mathbf{y}(0) \\ &+ \sum_{s=0}^{k-1} \mathbf{D}(s+1, k) \mathbf{y}(0) \theta(s) + \theta(k) \mathbf{y}(0). \end{aligned} \quad (45)$$

Using (19), (44), and (45) for any agent i and $k \geq 0$, **unnumbered eqn. shown at the bottom of the next page**, where, $[\cdot]_i$ denote the i th row of the corresponding matrix and in the last equality we used the fact $\frac{n \mathbf{v}_i(k) \mathbf{1}^\top \mathbf{X}(k)}{n y_i(k)} - \frac{\mathbf{v}_i(k) \mathbf{1}^\top \mathbf{y}(k) \mathbf{1}^\top \mathbf{X}(0)}{n y_i(k)} = 0$. Note $\|n \mathbf{X}(0) - \mathbf{y}(0) \mathbf{1}^\top \mathbf{X}(0)\| \leq \|n \mathbf{I}_n - \mathbf{1} \mathbf{1}^\top\| \|\mathbf{X}(0)\| \leq 2n^2 \|\mathbf{X}(0)\|$. From Lemma 2, there exists $y_o > 0$ such that $y_i(k) > y_o$ for all i and $k \geq 0$. Therefore,

$$\left\| z_i(k) - \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} \right\|$$

$$\begin{aligned}
&\leq \frac{\|\mathbf{D}_i(0, k)\| \|n\mathbf{X}(0) - \mathbf{y}(0)\mathbf{1}^\top \mathbf{X}(0)\|}{n\|y_i(k)\|} \\
&\quad + \frac{\theta(k)\|n\mathbf{X}_i(0) - \mathbf{y}_i(0)\mathbf{1}^\top \mathbf{X}(0)\|}{n\|y_i(k)\|} \\
&\quad + \frac{\sum_{s=0}^{k-1} \theta(s)\|\mathbf{D}_i(s+1, k)\| \|n\mathbf{X}(0) - \mathbf{y}(0)\mathbf{1}^\top \mathbf{X}(0)\|}{n\|y_i(k)\|}, \\
&\leq \frac{2n^3\|\mathbf{X}(0)\|C\lambda^k}{ny_o} + \frac{2n^3\|\mathbf{X}(0)\|\theta(k)}{ny_o} \\
&\quad + \frac{2n^3\|\mathbf{X}(0)\|C\sum_{s=0}^{k-1} \theta(s)\lambda^{k-s-1}}{ny_o} \\
&= \frac{2n^2\|\mathbf{X}(0)\|C}{y_o} \left[\lambda^k + \sum_{s=0}^{k-1} \theta(s)\lambda^{k-s-1} + \frac{1}{C}\theta(k) \right].
\end{aligned}$$

□

Corollary 3.1: Let Assumption 1 hold. Assume the edges in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are **not** corrupted by additive noise in communication channels. Let $\{x_i(k), y_i(k), z_i(k)\}_{k \geq 0}^{i \in [n]}$ be the sequences generated by the iterations (12)–(14). Let $\theta(k) = \rho^k$, for all $k \geq 1$ such that $\rho \in (0, 1/\alpha)$, where $\alpha > 1$. Then, for all $j \in [n]$ and $k \geq 1$,

$$\|z_j(k) - \bar{u}\| \leq \Upsilon \varrho^k,$$

where, Υ and $\varrho \in (0, 1)$ are positive constants.

Proof: Let $\varrho := \max\{\lambda, \alpha\rho\}$. From Theorem 3,

$$\begin{aligned}
\|z_j(k) - \bar{u}\| &\leq C_1 \left[\lambda^k + \sum_{s=0}^{k-1} \theta(s)\lambda^{k-s-1} + \frac{1}{C}\theta(k) \right] \\
&\leq C_1 \left[\lambda^k + \sum_{s=0}^{k-1} \lambda^{k-s-1} \rho^s + \frac{1}{C}\rho^k \right] \\
&\leq C_1 \left[\lambda^k + \sum_{s=0}^{k-1} \varrho^{k-s-1} \rho^s + \frac{1}{C}\rho^k \right]
\end{aligned}$$

$$\begin{aligned}
&\leq C_1 \left[\varrho^k + \frac{\varrho^k}{\varrho} \sum_{s=0}^{k-1} \left(\frac{\rho}{\varrho} \right)^s + \frac{1}{C}\varrho^k \right] \\
&\leq C_1 \left[1 + \frac{1}{\varrho} \sum_{s=0}^{\infty} \left(\frac{\rho}{\varrho} \right)^s + \frac{1}{C} \right] \varrho^k \\
&= C_1 \left[1 + \frac{1}{\varrho - \rho} + \frac{1}{C} \right] \varrho^k = \Upsilon \varrho^k,
\end{aligned}$$

where, $\Upsilon := C_1 \left[1 + \frac{1}{\varrho - \rho} + \frac{1}{C} \right]$. □

B. Convergence Under Noisy Communication Channels

Here, we analyze the performance of the proposed **NR-PushSum** algorithm under the communication noise. We first introduce conditions on the additive noise in the communication links. In particular, we make the following assumption:

Assumption 4: The noise in communication channels in $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is bounded, i.e. there exists $\delta_{i,j}$ such that $|\eta_{x_{ij}}(k)| \leq \delta_{i,j}$, $|\eta_{y_{ij}}(k)| \leq \delta_{i,j}$ for $i, j \in [n]$ and $k \geq 1$.

Consider a family Θ of sequences with any element $\theta(k) \in \Theta$ such that,

$$\theta(k) = \begin{cases} d_1 n \delta \beta(k) \in \mathbb{R}, & k \leq K_\theta, \\ \frac{d_2 n b \delta}{k^q} \in [0, 1], & k > K_\theta, \end{cases} \quad (46)$$

where, $1 \leq d_1 < \infty$ and $1 \leq d_2 \leq K_\theta^q / n b \delta$ and $K_\theta \geq \max\{K_\beta, (n \delta b)^{1/q}\}$, where K_β is the time index in Assumption 2, such that $a k^{-q} \leq \beta(k) \leq b k^{-q}$ for all $k \geq K_\beta$, with $q > 1$ and $0 \leq a \leq b < \infty$. Different choices of d_1 and d_2 lead to different sequences in Θ . Note that every sequence $\theta(k)$ in the family Θ satisfies Assumption 3. We have the following result for any sequence $\theta(k) \in \Theta$.

Lemma 3: Let $\{\theta(k)\}_{k \geq 0} \in \Theta$, then for all $k \geq 0$, $(\theta(k)\mathbf{y}(0) + \beta(k)\eta_y(k)) > 0$.

Proof: Let $k \leq K_\theta$,

$$\begin{aligned}
\theta(k)\mathbf{y}(0) + \beta(k)\eta_y(k) &= d_1 n \delta \beta(k)\mathbf{y}(0) + \beta(k)\eta_y(k) \\
&\geq d_1 n \delta \beta(k)\mathbf{1}^\top - \delta \beta(k)\mathbf{1}^\top = (d_1 n - 1)\delta \beta(k)\mathbf{1}^\top > 0.
\end{aligned}$$

$$\begin{aligned}
z_i(k+1) - \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} &= \frac{[\mathbf{v}(k)\mathbf{1}^\top \mathbf{X}(k) + \mathbf{D}(0, k)\mathbf{X}(0)]_i + \sum_{s=0}^{k-1} [\mathbf{D}(s+1, k)\mathbf{X}(0)\theta(s)]_i + \theta(k)\mathbf{X}_i(0)}{y_i(k)} - \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} \\
&= \frac{\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{X}(k) + \mathbf{D}_i(0, k)\mathbf{X}(0) + \sum_{s=0}^{k-1} \theta(s)\mathbf{D}_i(s+1, k)\mathbf{X}(0) + \theta(k)\mathbf{X}_i(0)}{y_i(k)} - \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} \\
&= \frac{n(\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{X}(k) + \mathbf{D}_i(0, k)\mathbf{X}(0) + \sum_{s=0}^{k-1} \theta(s)\mathbf{D}_i(s+1, k)\mathbf{X}(0) + \theta(k)\mathbf{X}_i(0))}{ny_i(k)} \\
&\quad - \frac{(\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{y}(k) + \mathbf{D}_i(0, k)\mathbf{y}(0) + \sum_{s=0}^{k-1} \theta(s)\mathbf{D}_i(s+1, k)\mathbf{y}(0))\mathbf{1}^\top \mathbf{X}(0)}{ny_i(k)} - \frac{\theta(k)\mathbf{y}_i(0)\mathbf{1}^\top \mathbf{X}(0)}{ny_i(k)} \\
&= \frac{\mathbf{D}_i(0, k)(n\mathbf{X}(0) - \mathbf{y}(0)\mathbf{1}^\top \mathbf{X}(0))}{ny_i(k)} + \frac{\theta(k)(n\mathbf{X}_i(0) - \mathbf{y}_i(0)\mathbf{1}^\top \mathbf{X}(0))}{ny_i(k)} \\
&\quad + \frac{\sum_{s=0}^{k-1} \theta(s)\mathbf{D}_i(s+1, k)(n\mathbf{X}(0) - \mathbf{y}(0)\mathbf{1}^\top \mathbf{X}(0))}{ny_i(k)},
\end{aligned}$$

For $k > K_\theta$,

$$\begin{aligned}\theta(k)\mathbf{y}(0) + \beta(k)\eta_y(k) &= \frac{d_2 n b \delta}{k^q} \mathbf{y}(0) + \beta(k)\eta_y(k) \\ &\geq \frac{d_2 n b \delta}{k^q} \mathbf{1}^\top - \frac{b \delta}{k^q} \mathbf{1}^\top = (d_2 n - 1) \frac{b \delta}{k^q} \mathbf{1}^\top > 0.\end{aligned}$$

This completes the proof. $\square$

Lemma 4: Let Assumption 1 hold. Let sequence $\{\beta(k)\}_{k \geq 0}$ satisfy Assumption 2. Assume that the channel noise satisfies Assumption 4. Let $\{\theta(k)\}_{k \geq 0} \in \Theta$ (defined in (46)), then there exist $\tilde{y}_o > 0$ such that $y_i(k) > \tilde{y}_o$ for all $k \geq 1$ and $i \in [n]$, i.e., the estimate $y_i(k)$ remain positive at all time-instants and the states $z_i(k)$ are well-defined for all $k, i \in [n]$.

Proof: Let $\delta := \max_{i,j} \delta_{ij}$. Using (18) and proceeding as in (27) we have,

$$\begin{aligned}\mathbf{y}(k+1) &= W(0, k)\mathbf{y}(0) + \sum_{s=0}^k \theta(s)W(s+1, k)\mathbf{y}(0) \\ &\quad + \sum_{s=0}^k \beta(s)W(s+1, k)\eta_y(s) = W(0, k)\mathbf{y}(0) \\ &\quad + \sum_{s=0}^k W(s+1, k)(\theta(s)\mathbf{y}(0) + \beta(s)\eta_y(s)).\end{aligned}\tag{47}$$

Let the sequence $\theta(k) \in \Theta$ as defined in (46). We follow the same steps in the proof of Lemma 2. Given, $\varepsilon > 0$, for any $i \in [n]$ and $k > \ell_\varepsilon$,

$$\begin{aligned}y_i(k+1) &= \sum_{j=1}^n W_{ij}(0, k)y_j(0) \\ &\quad + \sum_{j=1}^n \sum_{s=0}^k W_{ij}(s+1, k)(\theta(s)y_j(0) + \beta(s)\eta_{y_j}(s)) \\ &> \sum_{j=1}^n W_{ij}(0, k) \\ &= \sum_{j=1}^n [\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top + E(\ell_\varepsilon)W(\ell_\varepsilon + 1, k)]_{ij} \\ &> \sum_{j=1}^n \left[\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top - \frac{\varepsilon}{2} \mathbf{1}\mathbf{1}^\top W(\ell_\varepsilon + 1, k) \right]_{ij} \\ &> \sum_{j=1}^n \left[\mathbf{v}(\ell_\varepsilon)\mathbf{1}^\top - \frac{\varepsilon}{2} \mathbf{1}\mathbf{1}^\top \right]_{ij} \\ &= n\mathbf{v}_i(\ell_\varepsilon) - n\frac{\varepsilon}{2} \\ &> n \min_{1 \leq i \leq n} \mathbf{v}_i(\ell_\varepsilon) - \varepsilon \frac{n}{2} > 0,\end{aligned}$$

where, we used (25), the result of Lemma 3 and the fact that $W(l, k)$ is column stochastic for all $k, l \geq 0$.

For $k \leq \ell_\varepsilon$, from (25) and Lemma 3 we have,

$$y_i(k) > w_o^{k-1} > \theta(0)w_o^{\ell_\varepsilon} > 0.$$

Therefore, $y_i(k) > \tilde{y}_o := \min\{n \min_{1 \leq i \leq n} \mathbf{v}_i(\ell_\varepsilon), \theta(0)w_o^{\ell_\varepsilon}\}$ for all $k \geq 1$ with any $\theta(k) \in \Theta$ and $\beta(k)$ satisfying Assumption 2. $\square$

Theorem 4: Let Assumption 1 hold. Let sequence $\{\beta(k)\}_{k \geq 0}$ satisfy Assumption 2. Assume that the channel noise satisfies Assumption 4. Let $\{z_i(k)\}_{k \geq 0}^{i \in [n]}$ be the sequences generated by iteration (14). Given $\mu > 0$, let $\{\theta(k)\}_{k \geq 0} \in \Theta$ (defined in (46)) with $\theta(0) := n\delta \sum_{s=0}^\infty \beta(s)/\mu$, where, $\delta = \max_{i,j} \delta_{ij}$. Then, for all $i \in [n]$, almost surely, $z_i(k)$ converge to the same vector z_∞ . In particular,

$$\frac{\bar{u} - \mu}{1 + \mu} \leq \lim_{k \rightarrow \infty} z_i(k) \leq \frac{\bar{u} + \mu}{1 - \mu}.$$

Proof: Recall (17) and proceed as in (31),

$$\begin{aligned}\mathbf{X}(k+1) &= \underbrace{W(0, k)\mathbf{X}(0)}_{\text{first}} + \underbrace{\sum_{s=0}^k \theta(s)W(s+1, k)\mathbf{X}(0)}_{\text{second}} \\ &\quad + \underbrace{\sum_{s=0}^k \beta(s)W(s+1, k)\eta_x(s)}_{\text{third}}\end{aligned}\tag{48}$$

Note that as $W(l, k)$ is column stochastic for all $k, l \geq 1$, $\|W(l, k)\| \leq \bar{w} < \infty$ for all k, l .

Under Assumption 3 and from Theorem 1, given $\varepsilon > 0$. Define $\hat{\varepsilon} := \frac{\varepsilon / \sum_{s=0}^\infty \theta(s)}{\max\{\|\mathbf{X}(0)\|, n\}}$. There exists $\tilde{k}$ and $\mathbf{v}(k) \in \mathbb{R}^n$ such that $\sum_{s=\tilde{k}}^\infty \theta(s) < \frac{\hat{\varepsilon}}{3\bar{w}}$, $\sum_{s=\tilde{k}}^\infty \beta(s) \leq \frac{\hat{\varepsilon}}{3n\delta\bar{w}}$ and $W(l, k) = \mathbf{v}(k)\mathbf{1}^\top + E(k)$, for all $k > \tilde{k} + l$, where, $|E_{ij}(k)| \leq \frac{\hat{\varepsilon}}{3}$ for all $i, j \in [n]$.

Thus first and second terms in (48) can be re-formulated in the same manner as (32)–(35). Consider, the third term in (48),

$$\begin{aligned}\sum_{s=0}^k \beta(s)W(s+1, k)\eta_x(s) &= \underbrace{\sum_{s=0}^{\tilde{k}} \beta(s)W(s+1, k)\eta_x(s)}_{:=\zeta_1(k)} \\ &\quad + \underbrace{\sum_{s=\tilde{k}+1}^k \beta(s)W(s+1, k)\eta_x(s)}_{:=\zeta_x(k)}\end{aligned}\tag{49}$$

Note, $k > \tilde{k} + s + 1$ for all the terms in $\zeta_1(k)$. Substituting $W(s+1, k) := \mathbf{v}(k)\mathbf{1}^\top + \tilde{E}(k)$, with $|\tilde{E}_{ij}(k)| < \hat{\varepsilon}/3$ for all $i, j \in [n]$ we get,

$$\zeta_1(k) = \sum_{s=0}^{\tilde{k}} \beta(s)[\mathbf{v}(k)\mathbf{1}^\top + \tilde{E}(k)]\eta_x(s).\tag{50}$$

Now consider $\zeta_x(k)$,

$$\begin{aligned}\zeta_x(k) &= \sum_{s=\tilde{k}+1}^k \beta(s)W(s+1, k)\eta_x(s) \\ &= \beta(k)\eta_x(k) + \beta(k-1)\mathbf{M}(k)\eta_x(k-1) + \dots \\ &\quad + \beta(\tilde{k}+1)W(\tilde{k}+2, k)\eta_x(\tilde{k}+1).\end{aligned}$$

Therefore,

$$\|\zeta_x(k)\| \leq n\delta\beta(k) + n\delta\bar{w}\beta(k-1) \dots + n\delta\bar{w}\beta(\tilde{k}+1)$$

$$\leq n\delta\bar{w} \sum_{s=\tilde{k}}^{\infty} \beta(s) \leq \frac{\bar{\varepsilon}}{3}. \quad (51)$$

Using (48), (32)–(35), and (49)–(51), for all $k > \tilde{k}$,

$$\begin{aligned} \mathbf{X}(k+1) &\leq \mathbf{v}(k)\mathbf{1}^\top (\mathbf{X}(0) + \underbrace{\sum_{s=0}^{\tilde{k}} \theta(s)\mathbf{X}(0) + N_x(\tilde{k})}_{:=\chi_x(k)}) + \varepsilon\mathbf{1}_n\mathbf{1}_p^\top \\ &\quad (52) \end{aligned}$$

$$\begin{aligned} \mathbf{X}(k+1) &\geq \mathbf{v}(k)\mathbf{1}^\top (\mathbf{X}(0) + \sum_{s=0}^{\tilde{k}} \theta(s)\mathbf{X}(0) + N_x(\tilde{k})) - \varepsilon\mathbf{1}_n\mathbf{1}_p^\top \quad (53) \end{aligned}$$

where, $N_x(\tilde{k}) := \sum_{s=0}^{\tilde{k}} \beta(s)\eta_x(s)$, $N_y(\tilde{k}) := \sum_{s=0}^{\tilde{k}} \beta(s)\eta_y(s)$ and $\mathbf{1}_n$ and $\mathbf{1}_p$ are the vectors of dimensions n and p with all entries equal to 1. We drop the subscript when the dimension of the vector is clear from the context. Following the similar procedure for (47) we get, for $k \geq 1$,

$$\mathbf{y}(k+1) \leq \mathbf{v}(k)\mathbf{1}^\top \chi_y(k) + \varepsilon\mathbf{1}, \quad (54)$$

$$\mathbf{y}(k+1) \geq \mathbf{v}(k)\mathbf{1}^\top \chi_y(k) - \varepsilon\mathbf{1}. \quad (55)$$

From (19), and (52)–(55),

$$\begin{aligned} \frac{[\mathbf{v}(k)\mathbf{1}^\top \chi_x(k) - \varepsilon\mathbf{1}_n\mathbf{1}_p^\top]_i}{[\mathbf{v}(k)\mathbf{1}^\top \chi_y(k) + \varepsilon\mathbf{1}]_i} &\leq z_i(k+1) \\ &\leq \frac{[\mathbf{v}(k)\mathbf{1}^\top \chi_x(k) + \varepsilon\mathbf{1}_n\mathbf{1}_p^\top]_i}{[\mathbf{v}(k)\mathbf{1}^\top \chi_y(k) - \varepsilon\mathbf{1}]_i}, \quad (56) \end{aligned}$$

where, $[\cdot]_i$ denote the i th row (i th entry) of the corresponding matrices (vectors). Denote $\bar{\theta} := 1 + \sum_{s=0}^{\tilde{k}} \theta(s)$, and $\mathbf{v}_i(k)$ as the i th entry of $\mathbf{v}(k)$.

For $k > \tilde{k}$, for any $i \in [n]$, from (56),

$$\begin{aligned} z_i(k+1) &\leq \frac{[\mathbf{v}(k)\mathbf{1}^\top \chi_x(k) + \varepsilon\mathbf{1}_n\mathbf{1}_p^\top]_i}{[\mathbf{v}(k)\mathbf{1}^\top \chi_y(k) - \varepsilon\mathbf{1}]_i} \\ &\leq \frac{\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{X}(0) + \bar{\theta}\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{X}(0) + \mathbf{v}_i(k)\mathbf{1}^\top N_x(\tilde{k}) + \varepsilon\mathbf{1}^\top}{\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{y}(0) + \bar{\theta}\mathbf{v}_i(k)\mathbf{1}^\top \mathbf{y}(0) + \mathbf{v}_i(k)\mathbf{1}^\top N_y(\tilde{k}) - \varepsilon} \\ &\leq \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\bar{\theta}\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n} + \varepsilon\mathbf{1}^\top}{1 + \bar{\theta} + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n} - \varepsilon} \\ &\leq \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} (1 + \bar{\theta}) + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n} + \varepsilon\mathbf{1}^\top}{1 + \bar{\theta} + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n} - \varepsilon} \\ &\leq \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n(1+\bar{\theta})} + \frac{\varepsilon}{(1+\bar{\theta})}\mathbf{1}^\top}{1 + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n(1+\bar{\theta})} - \frac{\varepsilon}{(1+\bar{\theta})}} \\ &\leq \frac{a}{b} + \left(\frac{a}{b} + \mathbf{1}^\top\right) \left(\frac{\frac{\varepsilon}{b(1+\bar{\theta})}}{1 + \frac{\varepsilon}{b(1+\bar{\theta})}}\right). \end{aligned}$$

where, $a := \frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n(1+\bar{\theta})}$, $b := 1 + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n(1+\bar{\theta})}$. Similarly, for $k > \tilde{k}$, for any $i \in [n]$,

$$z_i(k+1) \geq \frac{a}{b} - \left(\frac{a}{b} + \mathbf{1}^\top\right) \left(\frac{\frac{\varepsilon}{b(1+\bar{\theta})}}{1 - \frac{\varepsilon}{b(1+\bar{\theta})}}\right)$$

Let $\bar{\varepsilon}$ be such that, $\varepsilon \leq \frac{\bar{\varepsilon}}{b(1+\bar{\theta})}(1 + \bar{\theta})$. Therefore, for all $k > \tilde{k}$, and for all $i \in [n]$,

$$-\left(\frac{a}{b} + \mathbf{1}^\top\right) \bar{\varepsilon} \leq z_i(k) - \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n(1+\bar{\theta})}}{1 + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n(1+\bar{\theta})}} \leq \left(\frac{a}{b} + \mathbf{1}^\top\right) \bar{\varepsilon}.$$

Note that $\bar{\varepsilon}$ can be made arbitrarily small. Thus, for all $i \in [n]$,

$$\lim_{k \rightarrow \infty} z_i(k) = \frac{\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} + \frac{\mathbf{1}^\top N_x(\tilde{k})}{n(1+\bar{\theta})}}{1 + \frac{\mathbf{1}^\top N_y(\tilde{k})}{n(1+\bar{\theta})}} := z_\infty.$$

Moreover, under Assumptions 2–4,

$$\frac{\bar{u} - \frac{n\delta \sum_{s=0}^{\infty} \beta(s)}{(1 + \sum_{s=0}^{\infty} \theta(s))}\mathbf{1}^\top}{1 + \frac{\delta \sum_{s=0}^{\infty} \beta(s)}{(1 + \sum_{s=0}^{\infty} \theta(s))}} \leq z_\infty \leq \frac{\bar{u} + \frac{n\delta \sum_{s=0}^{\infty} \beta(s)}{(1 + \sum_{s=0}^{\infty} \theta(s))}\mathbf{1}^\top}{1 - \frac{\delta \sum_{s=0}^{\infty} \beta(s)}{(1 + \sum_{s=0}^{\infty} \theta(s))}},$$

where, we used $\frac{\mathbf{1}^\top \mathbf{X}(0)}{n} = \bar{u}$. Note that if given $\mu > 0$, if $\theta(0) := n\delta \sum_{s=0}^{\infty} \beta(s)/\mu$, we have

$$\frac{\bar{u} - \mu}{1 + \mu} \leq z_\infty \leq \frac{\bar{u} + \mu}{1 - \mu}.$$

□

Remark 2: The upper and lower bounds on z_∞ in Theorem 4 can be made very close to the average $\bar{u}$ by appropriately choosing the sequences $\theta(k)$ and $\beta(k)$. In particular, if $\sum_{s=0}^{\infty} \theta(s) \gg \sum_{s=0}^{\infty} \beta(s)$ then z_∞ is very close to the actual average $\bar{u}$. This can be achieved via suitable choices of K_θ and the starting values of the sequence $\theta(k)$, $k < K_\theta$ in (46). We demonstrate this using numerical simulation studies with different noise realizations in Section V.

Remark 3: Unlike the state-of-the-art in [35], [41], [44], [50], [53], the **NR-PushSum** algorithm remains applicable even for noise distributions that are not necessarily zero-mean and independent. Furthermore, it achieves *almost sure* average consensus as shown in Theorem 4 and substantiated through simulation results in Section V-C.

Moreover, the articles [35], [44], [50], [53] are confined to undirected graphs with agents connected via bidirectional edges, in a symmetric communication topology. Additionally, these articles follow a *centralized synthesis* approach, utilizing global information of the interconnection structure. In contrast, the **NR-PushSum** algorithm operates differently. Notably, the consensus protocol presented in [41] diverges from the average of the initial values of the agents, setting it apart from the **NR-PushSum** algorithm.

The next result establishes the existence of the moments of the estimates $z_i(k)$ generated by the **NR-PushSum** algorithm.

Theorem 5: Let Assumption 1 hold. Let sequence $\{\beta(k)\}_{k \geq 0}$ satisfy Assumption 2. Assume that the channel noise satisfies Assumption 4. Let $\{\theta(k)\}_{k \geq 0} \in \Theta$ (defined in (46)). Let $\{x_i(k), y_i(k), z_i(k)\}_{k \geq 0}^{i \in [n]}$ be the sequences generated by the iterations (12)–(14). Then, for any m non-negative, $\mathbb{E}(|z_{ij}(k)|^m) < \infty$, for all $i \in [n], j = 1, \dots, p$ and $k \geq 0$.

Proof: From (48) and (52), under Assumptions 2–4 there exists $x_{up} < \infty$ such that $|\mathbf{X}_{ij}(k)| \leq x_{up}$ for all $i \in [n], j = 1, \dots, p$ and $k \geq 0$. Thus, for any m non-negative, $\mathbb{E}(|\mathbf{X}_{ij}(k)|^m) = x_{up}^m < \infty$. From Lemma 4, there exists $\tilde{y}_o$ such that we have for all $i \in [n]$ and $k \geq 0$, **Probability** $(|y_i(k)| < \tilde{y}_o) = 0$. Using Hölder's inequality for $(\mathbf{X}_{ij}(k))^m$ and $(y_i(k))^{-m}$. Since, **Probability** $((\tilde{y}_o)^{-m} < (y_i(k))^{-m} < (\tilde{y}_o)^{-m}) = 1$, we have

$$\mathbb{E}(|z_{ij}(k)|^m) \leq \tilde{y}_o^{-m} \mathbb{E}(|\mathbf{X}_{ij}(k)|^m) = \tilde{y}_o^{-m} x_{up}^m < \infty, \quad (57)$$

for all k , for any $i \in [n]$ and $j = 1, \dots, p$. $\square$

Remark 4: As a consequence of Theorem 5, it follows that the variance $\text{Var}(z_i(k))$ of the estimates $z_i(k)$, $\text{Var}(z_i(k)) = \mathbb{E}(|z_i(k)|^2) - (\mathbb{E}(z_i(k)))^2 \leq 2\sqrt{n}\tilde{y}_o^{-2}x_{up}^2 < \infty$ is bounded.

V. SIMULATION STUDY

In this section, we demonstrate the performance of the proposed **NR-PushSum** algorithm. A network of 10 agents generated using the Erdos-Renyi model [72] with a connectivity probability of 0.35 is considered for demonstration. The initial values of the agents chosen for the numerical study are $u = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$ with the average value being $\bar{u} = 5.5$. The examples in the numerical study are implemented in MATLAB and run on a desktop computer with 16 GB RAM and an Intel Core i7 processor running at 1.90 GHz. The noise realizations are obtained by simulating random variables that are drawn independently from a probability distribution at each iteration of the algorithm under study. We present the plots of the evolution of agents' estimates against the algorithm iterations.

First, we present the performance of the proposed **NR-PushSum** algorithm under perfect communication. Then we present the performance under imperfect communication via the following case studies: (i) performance of **NR-PushSum** algorithm under different algorithm parameters, (ii) performance under different magnitudes of the additive noise, and (iii) robustness of the **NR-PushSum** algorithm under the violation of the upper bound in Assumption 4 during the run-time of the algorithm. Lastly, we present a comparison between the proposed **NR-PushSum** algorithm and other existing algorithms in the literature. Consensus error for the **NR-PushSum** algorithm at any iteration k is defined as, $\sum_{i=1}^n (z_i(k) - \bar{u})^2 = \sum_{i=1}^n (z_i(k) - 5.5)^2$.

A. Performance Under Perfect Communication

Fig. 3 presents the agent state trajectories for the **NR-PushSum** algorithm when the communication links are noiseless. The algorithm parameters are: $\theta(k) = 0.7^k$ for all $k \geq 1$ and $K_\beta = 200$ such that $\beta(k) = 0.35, \forall k < 200$ and $\beta(k) =$

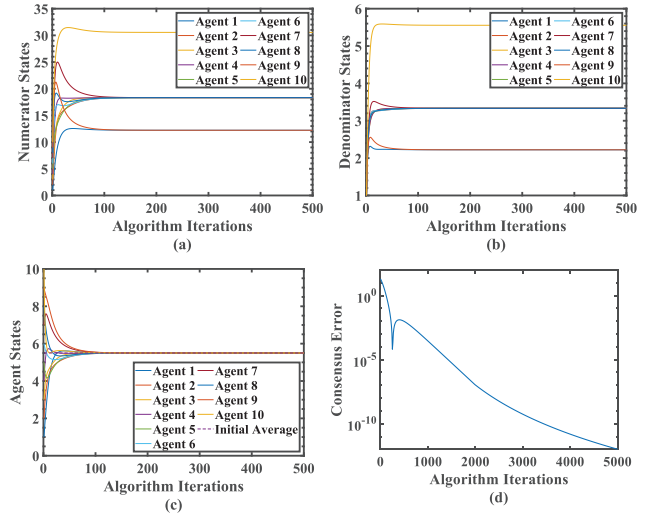


Fig. 3. Performance of the **NR-PushSum** algorithm under the noiseless case: (a) the sequence of states $x_i(k)$, (b) $y_i(k)$ sequences, (c) ratio states of the 10 agents achieve consensus, (d) geometric decay of the consensus error.

TABLE I
EXAMPLE SEQUENCES FOR **NR-PushSum**

(i)	$\beta(k) = 0.2, \forall k < 500,$ $\beta(k) = 1/k^{1.5}, \forall k \geq 500$	$\theta(k) = 100, \forall k < 500,$ $\theta(k) = 10/k^{1.5}, \forall k \geq 500$
(ii)	$\beta(k) = 0.35, \forall k < 200,$ $\beta(k) = 100/k^{1.1}, \forall k \geq 200$	$\theta(k) = 100, \forall k < 200,$ $\theta(k) = 150/k^{1.1}, \forall k \geq 200$

$100/k^{1.1}, \forall k \geq 200$. Fig. 3(a) and (b) show the estimates $x_i(k)$ and $y_i(k)$ (called the numerator and denominator states respectively) for all the agents. The numerator and denominator states reach a steady state. The agent states $z_i(k)$ are plotted in Fig. 3(c); it can be seen that the agents' states $z_i(k)$ achieve consensus to the average value of 5.5. Fig. 3(d) shows the decay of consensus error with respect to the number of iterations of the **NR-PushSum** algorithm.

B. Performance Under Communication Noise

In this section, we present the performance of the proposed **NR-PushSum** algorithm under noisy communication links.

1) *Different Choices of Algorithm Parameters:* Here, we demonstrate that the **NR-PushSum** algorithm achieves average consensus under the additive noise. The noise in communication links is realized from a uniform distribution $\mathcal{U}(-1, 1)$. We employ two different choices of $\beta(k)$ sequences and accordingly choose the $\theta(k)$ sequences from the family of sequences in (46). The choices are summarized in Table I. Fig. 4(a) and (b) present the agent state trajectories for the **NR-PushSum** algorithm realizations with sequences in Table I rows (i) and (ii) respectively under the additive noise. The inset in Fig. 4(a) and (b) show that the proposed **NR-PushSum** algorithm has strong performance leading to consensus among the agents under additive noise. Moreover, the final consensus value of the agents is close to the initial average of 5.5. Fig. 4(c) and (d) show the consensus error with the parameters in Table I for rows (i) and (ii), respectively. The sequence $\theta(k)$ belongs to the family of sequences given

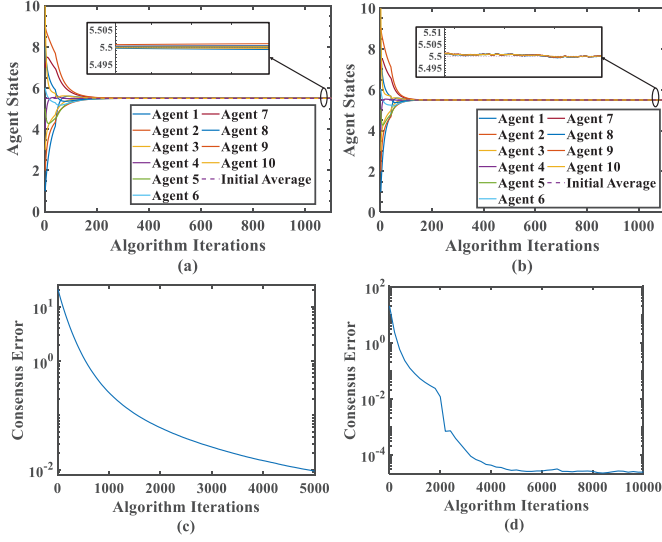


Fig. 4. Performance of the **NR-PushSum** algorithm under noisy communication with different choices of algorithm parameters: (a) Row (i) of Table I, (b) Row (ii) of Table I, (c) Consensus error for parameters in Row (i) of Table I, (d) Consensus error for parameters in Row (ii) of Table I.

by (46). Here, the **NR-PushSum** algorithm achieves a sub-linear decay of consensus error, analogous to Theorem 3.

2) *Different Magnitudes of the Additive Noise*: The effect of the magnitude of the communication noise relative to the initial values of the agents is presented here. The noise is realized as a uniform random variable. For the numerical experiment, the magnitude of the noise realizations is scaled to (i) 5 times the original value, i.e. $\mathcal{U}(-5, 5)$ and (ii) 10 times the original value i.e. $\mathcal{U}(-10, 10)$. We choose the sequence $\beta(k) = 0.35, \forall k < 200, \beta(k) = 1/k^{1.2}, \forall k \geq 200$. The sequence $\theta(k)$ is modified according to the noise upper bound. For the two cases, the following $\theta(k)$ sequences are chosen in accordance to the family (46) for the **NR-PushSum** algorithm: (i) $\theta(k) = 100, \forall k < 200, \theta(k) = 50/k^{1.2}, \forall k \geq 200$ and (ii) $\theta(k) = 100, \forall k < 200, \theta(k) = 100/k^{1.2}, \forall k \geq 200$. Fig. 5(a) and (b) provide the state trajectories of different agents under the **NR-PushSum** algorithm for cases (i) and (ii) respectively. It can be seen that the agent states achieve consensus close to the initial average value of 5.5. Fig. 5(c) and (d) plot the consensus error decay under the noise realizations via $\mathcal{U}(-5, 5)$ and $\mathcal{U}(-10, 10)$, respectively.

3) *Robustness to Noise With Extreme Magnitudes*: The parameters of the proposed **NR-PushSum** algorithm depend on an upper bound on the communication noise (see (46)). To test the applicability of the **NR-PushSum** algorithm in extreme scenarios, here we test the robustness of the **NR-PushSum** algorithm under situations when the communication noise violates the upper bound in Assumption 4. In particular, we design the algorithm parameters assuming that the communication noise is a uniform random variable $\mathcal{U}(-1, 1)$. We test the following two scenarios: (i) communication noise in the worst case becomes 400 times at every 50 iterations, i.e. noise is realized as $\mathcal{U}(-400, 400)$ at every iteration of the form $50i$,

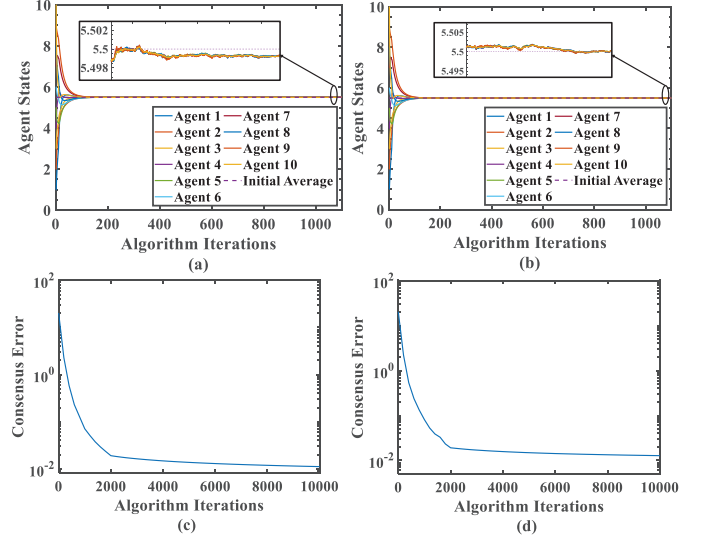


Fig. 5. Performance of the **NR-PushSum** algorithm under the variation of the magnitude of additive communication noise: (a) noise realized via $\mathcal{U}(-5, 5)$, (b) noise realized via $\mathcal{U}(-10, 10)$. In both cases, the state estimates under **NR-PushSum** with appropriate algorithmic parameters converge to the initial average, (c) Consensus error under the noise realization via $\mathcal{U}(-5, 5)$, (d) Consensus error under the noise realization via $\mathcal{U}(-10, 10)$.

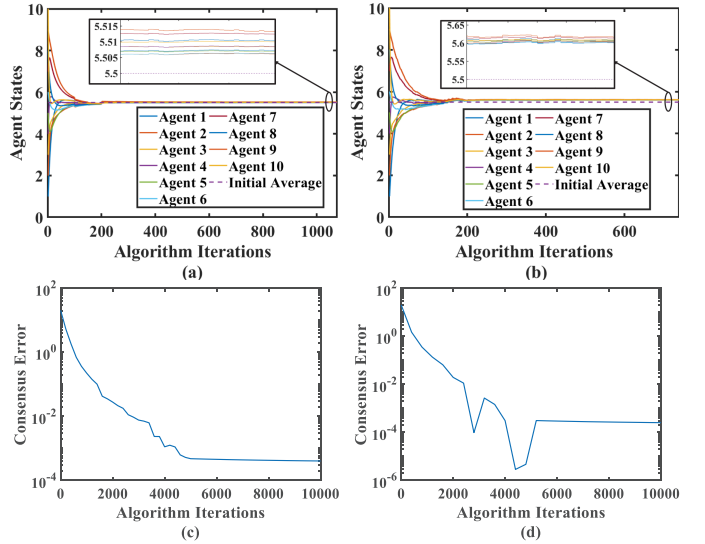


Fig. 6. Performance of the **NR-PushSum** algorithm under scenarios with the noise realizations violating the upper bound utilized to the algorithms parameters. Fig. 6(a) and (b) demonstrate the robustness of the proposed **NR-PushSum** algorithm under extreme violation scenarios, Fig. 6(c) and (d) present the consensus error decay during the iterations of the **NR-PushSum** algorithm.

where, $i = 1, 2, \dots$, (ii) communication noise in the worst case becomes 400 times at every 10 iterations, i.e. noise is realized as $\mathcal{U}(-400, 400)$ at every iteration of the form $10i$, where, $i = 1, 2, \dots$. The two situations are designed to test the effect of the number of violations of the bound in Assumption 4 during run-time on the convergence of the **NR-PushSum** algorithm. Fig. 6(a) and (b) show the agent states for the two test situations. The frequency of the event when the noise value is greater than the original upper bound affects the final converged value

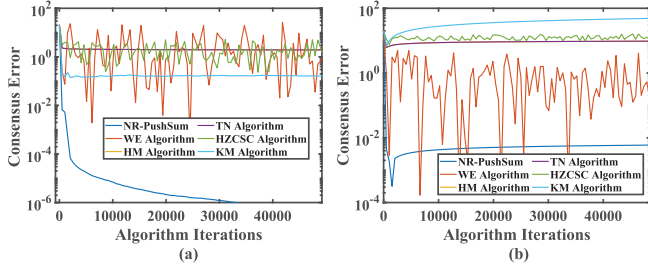


Fig. 7. Performance comparison of the **NR-PushSum** algorithm with existing algorithms in the literature: (a) performance under bounded noise with zero mean and (b) performance under bounded communication noise with non-zero mean.

of the algorithm. In case (i) (Fig. 6(a)) with less number of violations the **NR-PushSum** algorithm converges closer to the initial average value of 5.5, while the estimates in situation (ii) have a larger deviation from the original average value of 5.5 (Fig. 6(b)). However, we emphasize that the proposed **NR-PushSum** algorithm performs well and remains stable under the extreme violations of the design assumptions as shown by the state trajectories in Fig. 6(a) and (b). The consensus error decay for the two test situations is presented in Fig. 6(c) and (d), respectively. Note that the extreme noise magnitudes do affect the consensus error. However, the **NR-PushSum** algorithm is still able to achieve a very small consensus error in both test scenarios.

C. Comparison With Other Algorithms

In this section, we present a performance comparison between the proposed **NR-PushSum** algorithm and existing algorithms in the literature. Here, we refer to the existing algorithms via the initial of the authors' names. We compare with the following algorithms: WE algorithm of [53], HM algorithm in [41], TN algorithm in [50], HZCSC algorithm in [44], and the KM algorithm in [35]. To obtain the consensus error for any other algorithm we replace the $z_i(k)$ estimates with the corresponding estimates in the chosen algorithm. We compare the results under two different noise realizations: (i) we first consider the communication noise realized as a uniform random variable $\mathcal{U}(-1, 1)$. (ii) Next, we compare the performance of all the algorithms when the communication noise has a **non-zero** mean. In particular, we consider the noise realizations to be generated as a non-zero mean uniform random variable $\mathcal{U}(0, 1)$. In both cases we choose the following sequences for the **NR-PushSum** algorithm: $\beta(k) = 0.2, \forall k < 500, \beta(k) = 1/k^{1.5}, \forall k \geq 500$ and $\theta(k) = 100, \forall k < 500, \theta(k) = 10/k^{1.5}, \forall k \geq 500$. To present the performance comparison, we plot the consensus error for all the algorithms with respect to the iteration count under the zero mean and non-zero mean cases in Fig. 7(a) and (b) respectively. It can be seen that the proposed **NR-PushSum** algorithm performs the best in both cases. The **NR-PushSum** algorithm has a consensus error that decays faster compared to the other algorithms when the communication noise has zero mean (see Fig. 7(a)). When the communication noise has a non-zero mean, it can be seen that the consensus error for

the existing algorithms in the literature is greater than 1 and increases with the iteration count of the algorithm. In this case, too the proposed **NR-PushSum** algorithm performs well. The consensus error for the **NR-PushSum** algorithm is less than 10^{-2} for all the iterations.

Summarizing, the simulation results indicate that the proposed **NR-PushSum** algorithm has a good performance under communication noise and shows robustness to uncertain changes in the communication noise. Compared to the existing state-of-the-art algorithms the **NR-PushSum** algorithm has superior performance with respect to the ability to achieve consensus to the initial average under communication noise.

VI. CONCLUSION

In this article, we address the challenge of designing an average consensus algorithm for directed graphs when communication links are impacted by noise. Initially, we demonstrate that the widely used PushSum algorithm, often employed for distributed decision-making schemes over directed graphs, does not perform well in the presence of communication noise. Subsequently, we introduce the **NR-PushSum** algorithm, which involves each agent utilizing its own noise-free weighted initial value for updating estimates during each iteration. Furthermore, each agent governs the amount of noise permissible in its state updates by appropriately weighing the information received from its neighbors through noisy communication links. These chosen weights enable agents to counteract the effects of noisy estimates from neighbors and steer towards consensus. We establish that the proposed **NR-PushSum** algorithm achieves consensus at a geometric rate when communication is noiseless. Under additive noise, we prove that agents' estimates following the **NR-PushSum** algorithm converge to a consensus value almost surely. We validate the effectiveness of the proposed algorithm through numerical case studies encompassing different noise scenarios and varying algorithm parameters. Simulation results highlight the superior performance of the **NR-PushSum** algorithm compared to existing state-of-the-art methods. A potential future research direction for the authors is the utilization of the proposed **NR-PushSum** algorithm to design distributed optimization algorithms for directed graphs in the presence of communication noise.

REFERENCES

- [1] S. Patel, S. Attree, S. Talukdar, M. Prakash, and M. V. Salapaka, "Distributed apportioning in a power network for providing demand response services," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, 2017, pp. 38–44.
- [2] S. Patel, B. Lundstrom, G. Saraswat, and M. V. Salapaka, "Distributed power apportioning with early dispatch for ancillary services in renewable grids," 2020, *arXiv:2007.11715*.
- [3] S. Patel, V. Khatana, G. Saraswat, and M. V. Salapaka, "Distributed detection of malicious attacks on consensus algorithms with applications in power networks," in *Proc. IEEE 7th Int. Conf. Control Decis. Inf. Technol.*, 2020, pp. 397–402.
- [4] M. Ala'raj and M. F. Abbod, "Classifiers consensus system approach for credit scoring," *Knowl.-Based Syst.*, vol. 104, pp. 89–105, 2016.
- [5] A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, Sep. 2004.

- [6] A. Nedic, "Distributed gradient methods for convex machine learning problems in networks: Distributed optimization," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 92–101, May 2020.
- [7] M. H. DeGroot, "Reaching a consensus," *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.
- [8] G. Pasolini, D. Dardari, and M. Kieffer, "Exploiting the agent's memory in asymptotic and finite-time consensus over multi-agent networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 479–490, 2020.
- [9] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003, doi: [10.1109/TAC.2003.812781](https://doi.org/10.1109/TAC.2003.812781).
- [10] G. Soatti, M. Nicoli, S. Savazzi, and U. Spagnolini, "Consensus-based algorithms for distributed network-state estimation and localization," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 2, pp. 430–444, Jun. 2017.
- [11] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *Proc. IEEE 44th Annu. Symp. Found. Comput. Sci.*, 2003, pp. 482–491.
- [12] C. N. Hadjicostis and T. Charalambous, "Average consensus in the presence of delays and dynamically changing directed graph topologies," 2012, *arXiv:1210.4778*.
- [13] G. Saraswat, V. Khatana, S. Patel, and M. V. Salapaka, "Distributed finite-time termination for consensus algorithm in switching topologies," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 1, pp. 489–499, Jan./Feb. 2023.
- [14] C. N. Hadjicostis and T. Charalambous, "Average consensus in the presence of delays in directed graph topologies," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 763–768, Mar. 2014.
- [15] M. Prakash, S. Talukdar, S. Attree, V. Yadav, and M. V. Salapaka, "Distributed stopping criterion for consensus in the presence of delays," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 1, pp. 85–95, Mar. 2020.
- [16] J. Melbourne, G. Saraswat, V. Khatana, S. Patel, and M. V. Salapaka, "On the geometry of consensus algorithms with application to distributed termination in higher dimension," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 2951–2956, 2020.
- [17] J. Melbourne, G. Saraswat, V. Khatana, S. Patel, and M. V. Salapaka, "Convex decreasing algorithms: Distributed synthesis and finite-time termination in higher dimension," 2020, *arXiv:2007.13050*.
- [18] C. Zhao, J. He, and J. Chen, "Resilient consensus with mobile detectors against malicious attacks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 1, pp. 60–69, Mar. 2018.
- [19] A. Nowzari and M. G. Rabbat, "Improved bounds for max consensus in wireless networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 2, pp. 305–319, Jun. 2019.
- [20] G. Muniraju, C. Tepedelenlioglu, and A. Spanias, "Analysis and design of robust max consensus for wireless sensor networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 4, pp. 779–791, Dec. 2019.
- [21] V. Khatana, G. Saraswat, S. Patel, and M. V. Salapaka, "Gradient-consensus method for distributed optimization in directed multi-agent networks," in *Proc. IEEE Amer. Control Conf.*, Denver, CO, USA, 2020, pp. 4689–4694, doi: [10.23919/ACC45564.2020.9147544](https://doi.org/10.23919/ACC45564.2020.9147544).
- [22] V. Khatana, G. Saraswat, S. Patel, and M. V. Salapaka, "GradConsensus: Linearly convergent algorithm for reducing disagreement in multi-agent optimization," *IEEE Trans. Netw. Sci. Eng.*, early access, Oct. 06, 2023, doi: [10.1109/TNSE.2023.3321757](https://doi.org/10.1109/TNSE.2023.3321757).
- [23] V. Khatana and M. V. Salapaka, "D-DistADMM: A $O(1/k)$ distributed ADMM for distributed optimization in directed graph topologies," in *Proc. IEEE 59th Conf. Decis. Control*, Jeju, Korea (South), 2020, pp. 2992–2997, doi: [10.1109/CDC42340.2020.9304086](https://doi.org/10.1109/CDC42340.2020.9304086).
- [24] V. Khatana and M. V. Salapaka, "DC-DistADMM: ADMM algorithm for constrained optimization over directed graphs," *IEEE Trans. Autom. Control*, vol. 68, no. 9, pp. 5365–5380, Sep. 2023.
- [25] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, Mar. 2015.
- [26] A. Nedic, A. Olshevsky, and W. Shi, "Achieving geometric convergence for distributed optimization over time-varying graphs," *SIAM J. Optim.*, vol. 27, no. 4, pp. 2597–2633, 2017.
- [27] S. Liu, L. Xie, and H. Zhang, "Distributed consensus for multi-agent systems with delays and noises in transmission channels," *Automatica*, vol. 47, no. 5, pp. 920–934, 2011.
- [28] T. Li and J.-F. Zhang, "Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929–1936, 2009.
- [29] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comput.*, vol. 67, no. 1, pp. 33–46, 2007.
- [30] R. Carli, G. Como, P. Frasca, and F. Garin, "Average consensus on digital noisy networks," *IFAC Proc. Volumes*, vol. 42, no. 20, pp. 36–41, 2009.
- [31] T. Pan, L. Mo, and X. Cao, "Consensus of discrete-time multi-agent systems with white noise disturbance," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 202–205, 2015.
- [32] S. Dasarathan, C. Tepedelenlioglu, M. K. Banavar, and A. Spanias, "Robust consensus in the presence of impulsive channel noise," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2118–2129, Apr. 2015.
- [33] B. Li, H. Leung, and C. Seneviratne, "Analysis of noise impact on distributed average consensus," in *Signal Processing, Sensor/Information Fusion, and Target Recognition XXVII*, vol. 10646. Bellingham, WA, USA: Int. Soc. Opt. Photonics, 2018, Art. no. 106460L.
- [34] S. Kar and J. M. Moura, "Distributed average consensus in sensor networks with random link failures and communication channel noise," in *Proc. IEEE Conf. Rec. 41st Asilomar Conf. Signals Syst. Comput.*, 2007, pp. 676–680.
- [35] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 355–369, Jan. 2009.
- [36] L. Pescosolido, S. Barbarossa, and G. Scutari, "Average consensus algorithms robust against channel noise," in *Proc. IEEE 9th Workshop Signal Process. Adv. Wireless Commun.*, 2008, pp. 261–265.
- [37] T. Li and J.-F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2043–2057, Sep. 2010.
- [38] R. Rajagopal and M. J. Wainwright, "Network-based consensus averaging with general noisy channels," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 373–385, Jan. 2011.
- [39] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in Ad Hoc WSNs with noisy links—Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.
- [40] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, "Consensus in ad hoc WSNs with noisy links—Part II: Distributed estimation and smoothing of random signals," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1650–1666, Apr. 2008.
- [41] M. Huang and J. H. Manton, "Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 134–161, 2009.
- [42] A. Y. Kibangou, "Finite-time average consensus based protocol for distributed estimation over AWGN channels," in *Proc. IEEE 50th Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 5595–5600.
- [43] M. Zhou, J. He, P. Cheng, and J. Chen, "Discrete average consensus with bounded noise," in *Proc. IEEE 52nd Conf. Decis. Control*, 2013, pp. 5270–5275.
- [44] J. He, M. Zhou, P. Cheng, L. Shi, and J. Chen, "Consensus under bounded noise in discrete network systems: An algorithm with fast convergence and high accuracy," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 2874–2884, Dec. 2016.
- [45] F. F. Rego, Y. Pu, A. Alessandretti, A. P. Aguiar, and C. N. Jones, "A consensus algorithm for networks with process noise and quantization error," in *Proc. 53rd Annu. Allerton Conf. Commun. Control Comput.*, 2015, pp. 488–495.
- [46] A. Jadbabaie and A. Olshevsky, "On performance of consensus protocols subject to noise: Role of hitting times and network structure," in *Proc. IEEE 55th Conf. Decis. Control*, 2016, pp. 179–184.
- [47] G. Chen, L. Y. Wang, C. Chen, and G. Yin, "Critical connectivity and fastest convergence rates of distributed consensus with switching topologies and additive noises," *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6152–6167, Dec. 2017.
- [48] O. Granichin, V. Erofeeva, Y. Ivanskiy, and Y. Jiang, "Simultaneous perturbation stochastic approximation-based consensus for tracking under unknown-but-bounded disturbances," *IEEE Trans. Autom. Control*, vol. 66, no. 8, pp. 3710–3717, Aug. 2021.
- [49] Y. Yazıcıoğlu and A. Speranzon, "High dimensional robust consensus over networks with limited capacity," *IEEE Control Syst. Lett.*, vol. 5, no. 6, pp. 2024–2029, Dec. 2021.
- [50] B. Touri and A. Nedic, "Distributed consensus over network with noisy links," in *Proc. IEEE 12th Int. Conf. Inf. Fusion*, 2009, pp. 146–154.

- [51] J. Wang and N. Elia, "Distributed agreement in the presence of noise," in *Proc. IEEE 47th Annu. Allerton Conf. Commun. Control Comput.*, 2009, pp. 1575–1581.
- [52] J. Wang and N. Elia, "Dynamic average consensus over random networks with additive noise," in *Proc. IEEE 49th Conf. Decis. Control*, 2010, pp. 4789–4794.
- [53] J. Wang and N. Elia, "Distributed averaging algorithms resilient to communication noise and dropouts," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2231–2242, May 2013.
- [54] S. Pu and A. Garcia, "A flocking-based approach for distributed stochastic optimization," *Operations Res.*, vol. 66, no. 1, pp. 267–281, 2018.
- [55] Y. Long, S. Liu, and L. Xie, "Distributed consensus of discrete-time multi-agent systems with multiplicative noises," *Int. J. Robust Nonlinear Control*, vol. 25, no. 16, pp. 3113–3131, 2015.
- [56] S. Sheipak, "Reaching consensus in a variable-topology multiagent system under additive random noise," *Automat. Remote Control*, vol. 81, no. 5, pp. 911–921, 2020.
- [57] Y. Wang, L. Cheng, Z.-G. Hou, M. Tan, C. Zhou, and M. Wang, "Consensus seeking in a network of discrete-time linear agents with communication noises," *Int. J. Syst. Sci.*, vol. 46, no. 10, pp. 1874–1888, 2015.
- [58] Z. Wang, W. Wang, and H. Zhang, "Robust consensus for linear multi-agent systems with noises," *IET Control Theory Appl.*, vol. 10, no. 17, pp. 2348–2356, 2016.
- [59] X. Zong, T. Li, and J.-F. Zhang, "Stochastic consensus of continuous-timemulti-agent systems with additive measurement noises," in *Proc. IEEE 54th Conf. Decis. Control*, 2015, pp. 543–548.
- [60] L. Cheng, Z.-G. Hou, and M. Tan, "A mean square consensus protocol for linear multi-agent systems with communication noises and fixed topologies," *IEEE Trans. Autom. Control*, vol. 59, no. 1, pp. 261–267, Jan. 2014.
- [61] K. Hanada, T. Wada, I. Masubuchi, T. Asai, and Y. Fujisaki, "Stochastic consensus algorithms over general noisy networks," *SICE J. Control Meas. Syst. Integr.*, vol. 13, no. 6, pp. 274–281, 2020.
- [62] D. Mateos-Núñez and J. Cortés, "Noise-to-state exponentially stable distributed convex optimization on weight-balanced digraphs," *SIAM J. Control Optim.*, vol. 54, no. 1, pp. 266–290, 2016.
- [63] M. Huang and J. H. Manton, "Stochastic approximation for consensus seeking: Mean square and almost sure convergence," in *Proc. IEEE 46th Conf. Decis. Control*, 2007, pp. 306–311.
- [64] M. Huang and J. H. Manton, "Stochastic consensus seeking with noisy and directed inter-agent communication: Fixed and randomly varying topologies," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 235–241, Jan. 2010.
- [65] T. C. Aysal and K. E. Barner, "Convergence of consensus models with stochastic disturbances," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 4101–4113, Aug. 2010.
- [66] G. Morral, P. Bianchi, and G. Fort, "Success and failure of adaptation-diffusion algorithms with decaying step size in multiagent networks," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2798–2813, Jun. 2017.
- [67] R. Diestel, *Graph Theory*. Berlin, Germany: Springer-Verlag, 2006.
- [68] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [69] V. Khatana and M. V. Salapaka, "Distributed average consensus over noisy communication links in directed graphs," 2022, *arXiv:2211.10334*.
- [70] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, vol. 39. Berlin, Germany: Springer, 1991.
- [71] J. Wolfowitz, "Products of indecomposable, aperiodic, stochastic matrices," in *Proc. Amer. Math. Soc.*, vol. 14, no. 5, pp. 733–737, 1963. [Online]. Available: <http://www.jstor.org/stable/2034984>
- [72] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.*, vol. 5, no. 1, pp. 17–60, 1960.



Vivek Khatana (Student Member, IEEE) received the bachelor's degree in electrical engineering from the Indian Institute of Technology, Roorkee, India, in 2018. He is currently a Ph.D. Candidate with the Department of Electrical Engineering, the University of Minnesota, Twin Cities, MN, USA. His research interests include distributed optimization, distributed and stochastic control, networks of dynamical systems, data-driven system identification, and control of power systems.



Murti V. Salapaka (Fellow, IEEE) received the B.Tech. degree in mechanical engineering from the Indian Institute of Technology, Madras, Chennai, India, in 1991 and the M.S. and Ph.D. degrees in mechanical engineering from the University of California, Santa Barbara, CA, USA, in 1993 and 1997, respectively. He is currently the Director of Graduate Studies and the Vincentine Hermes Luh Chair Professor with the Electrical and Computer Engineering Department, University of Minnesota, Twin Cities, MN, USA. He was a Faculty Member with the Electrical and Computer Engineering Department, Iowa State University, Ames, IA, USA, from 1997 to 2007. His research interests include control and network science, nanoscience, and single-molecule physics. Dr. Salapaka was the recipient of the NSF CAREER Award and the ISU Young Engineering Faculty Research Award for the years 1998 and 2001.