Robust Dynamic Corrective Consensus over Lossy Networks with Gilbert-Elliott Channels

Sabato Manfredi*°

***Abstract*—In this paper we propose a distributed algorithm to track the average of time-varying reference signals at nodes (i.e. dynamic average consensus) over networks with heterogeneous, asymmetric Gilbert-Elliott (GE) type lossy channels. The algorithm is robust to GE type network disruptions, (re)-initialization errors and agents joining, faulty or leaving the network during time. Necessary and sufficient conditions are formulated to assess almost sure and mean bounded dynamic average consensus according to the characteristics of reference signals. The proposed conditions can be used for design purpose as they relate algorithm step size, network connectivity and channel quality parameters (e.g. recovery and failure rate) to the algorithm convergence. Simulations confirm the effectiveness of the proposed algorithm and design conditions in two representative validation scenarios.**

I. Introduction

In recent years intensive research on consensus problems has led to the formulation of distributed monitoring, control, and optimization algorithms for emerging applicative paradigms such as cyber-physical systems, wireless sensor and surveillance networks (see, e.g., [1] and references therein). The cooperative, distributed nature of these algorithms relies on network connectivity and link reliability to ensure their effectiveness. In this respect the connectivity and packet lossy of communication channels may be governed by a random process, usually assumed as Bernoulli identical independent distribution (i.i.d.). In real scenarios, wireless channels are affected by fading, interference, and burst effects with correlated packet losses, making the Bernoulli model inadequate. A more realistic alternative is the Gilbert-Elliott (GE) model, which represents packet loss as a time-homogeneous two-state Markov chain [2], [3], [4]. When extended to all network links, this model is commonly referred to as Markovian packet loss. This has motivated during last years a fruitful research on consensus algorithms in the presence of Bernoulli and GE/Markovian stochastic lossy channels. Sufficient and/or necessary conditions for almost sure/mean square convergence of static (resp. average) consensus to nodes initial state-depending agreement value (resp. averaged value) has been assessed under i.i.d. graph adjacency matrix process ([5], [6]), Bernoulli algorithm updates [7], identical and nonidentical Markovian dropouts [8], [9], [10], heterogeneous (symmetric/balanced) random i.i.d. link failures [11]. A necessary assumption to guarantee the convergence of the static consensus algorithm to the average initial state values (i.e. static average consensus) is that the graph/links failures have to be symmetric or at least balanced. The packet loss is a stochastic process that may affect the symmetry and balancing property of the graph, yielding an error between the consensus value and the average value of the initial states ([12]). An interesting approach proposed in [13] deals with the graph unbalancing effect due to i.i.d. Bernoulli packet losses in order to guarantee the convergence of the static average consensus algorithm.

In practical applications of cyber-physical monitoring and control systems based on wireless network architecture, the commonly known *dynamic average consensus algorithm* (see, e.g., [14], [1], [15], and references therein) is more appealing, as it allows each node to estimate the average value of measures of interest (e.g., average temperature, centroid of mass of a robot swarm) in a distributed manner. Further, dynamic consensus algorithms are essential building blocks for distributed optimization. In deterministic scenario several robust dynamic consensus algorithms have been proposed to deal with either initialization errors or agents leaving/joining the network [16], [17]. Recently, [18] presented a novel dynamic multi-stage cascade of dynamic consensus filters designed to track the average value of inputs with a tunable error that decreases progressively as the number of stages increases. The filter is robust with respect to initialization errors and a randomized implementation is devised when i.i.d. sequence of edges are selected with uniform probability. In [19] an extended multistage filter is proposed to track slowly varying reference signals in the presence of time delays, heterogeneous independent Bernoulli packet drops, and random noise. It is employed a combination of Kalman filtering, multistage consensus filtering, and predictive control to assess mean-square dynamic consensus robust to initialization errors. A tunable tradeoff between communication/computation cost and stationary-state tracking error is shown. The above multi-stage filters deal with independent Bernoulli type losses. In addition, the implementation of a filter of order m requires a sufficient memory storage and communication capability at node to handle the m local variables received from each of its neighbors. Moreover, the final dynamic consensus error is reduced as far as the filter order m and related memory/communication requirement increase. The underling or physical network graph (with no packet losses) is assumed be connected, symmetric or balanced (e.g. [11], [18], [19]).

With regard to the literature on discrete-time average con-

*Prof. S. Manfredi is with the Department of Electrical Engineering and Information Technology, University of Naples Federico II, Italy Email: sabato.manfredi@unina.it

° This work was supported by project "Resilient Optimization in Distributed Cyber-Physical Control Problems subject to Adversarial Behaviour", funded in part by Italian Ministry of University and Research (MUR) within the Progetti di Rilevante Interesse Nazionale (PRIN) 2022 Program under Project 20225MESZB, and in part by European Union - Next Generation EU, Mission 4 Component 2, CUP E53D23000470006.

sensus algorithms in the presence of stochastic lossy channels, specifically, the paper's contributions are:

(1.) compared to the discrete time static consensus algorithms in the literature ([5]- [13]), a novel consensus-based algorithm is devised to track the average of time-varying reference signals at nodes. The proposed algorithm is resilient to stochastic heterogeneous Gilbert-Elliott type lossy channels and, notable, does not exploit m-th order input differences. In addition, it is robust to initialization errors, agents joining, faulty or leaving the network at different points in time. The approach is inspired by the corrective method originally proposed for static average consensus with i.i.d. Bernoulli losses (see i.e. [13] and references therein);

(2.) in contrast to existing dynamic consensus algorithms that primarily address deterministic (i.e., [1], [14] - [17]) or i.i.d. Bernoulli link disruption scenarios (i.e., [18], [19]), the proposed algorithm extends them to address heterogeneous, asymmetric, and correlated losses. Furthermore, necessary and sufficient conditions are formulated to evaluate stronger and varied types of convergence (i.e. almost sure and mean bounded) towards the dynamic consensus equilibrium. Significantly, these conditions can be utilized for design purposes as they delineate algorithm parameters, topology connectivity, and channel characteristics (i.e., initial loss probability, recovery, and failure rate) necessary to achieve the desired level of final consensus error. Another contribution concerning the algorithm computation aspect is that the proposed strategy requires less computational power and memory compared to the multi-stage filters ([18], [19]), thus making it more viable for implementation on architectures relying on small devices such as wireless sensor networks and low-power IoT networks. Finally, the proposed convergence conditions may handle both deterministic and stochastic reference signals, and their relation with respect to the convergence defined in [11] is highlighted.

The rest of the paper is organized as follows. In Section II the notation and the problem framework are presented. In Section III preliminary lemmas and proposition are introduced and the main convergence results in the case of heterogeneous, symmetric GE lossy channels are formulated. The latter are then extended to the case of heterogeneous channels. A simulation validation is carried out in Section IV, while conclusions and future work are given in Section V.

II. Preliminaries and Problem framework

Notation: We consider all vectors $x = [x_i]$ as columns with i-th entry x_i. $\|x\|$ denotes the standard Euclidean vector norm in L_2, $\|x\| = \sqrt{\sum_i x_i^2}$, while I and $\mathbf{1}$ are the identity matrix and a vector with all entries equal to 1, respectively. $|Y|$ denotes either the absolute value if Y is a complex number or the cardinality if Y is a finite set. For a matrix $W = [w_{ij}]$ with element w_{ij}, W^T represents its matrix transpose. $P(A)$ denotes the probability of an event A and $\mathbb{E}[X]$ is the expected value of a random variable X. 1_a^u is the indicator function with $1_a^u = 1$ if $u \geq a$, $1_a^u = 0$ for $u < a$.

We consider a realistic wireless network characterized by GE lossy channels. The wireless network communications are bidirectional although with asymmetric packet loss probabilities. Therefore for every link, there will be a chance to recover packet drops for both directions with positive (potentially different) probability p. The physical or underling communication network is represented by an undirected (physical) graph $\mathcal{G} = \mathcal{G}(\mathcal{V}, \mathcal{E})$ with the set of nodes $\mathcal{V} = \{1, \ldots, N\}$ and set of links $\mathcal{E} = \{(i,j) : i,j \in \mathcal{V}, p_{ij} > 0\}$. The graph is connected if there is a path between any two nodes. A GE lossy communication link $(i,j) \in \mathcal{E}$ is modelled by a two state Markov chain such that in the *GOOD* state G the link is active, while it is down in the *BAD* state B ([2], [3]). Let $A(t) = [a_{ij}(t)]$ the network adjacency matrix, the elements a_{ij} are affected by the state of the GE channels at each time step t. Specifically, the state $s_{ij}(t)$ of the link (i,j) at step t is related to the element $a_{ij}(t)$ as follows: $P(a_{ij}(t) = 1) = P(s_{ij}(t) = G) = q_{ij}^G(t); P(a_{ij}(t) = 0) = P(s_{ij}(t) = B) = q_{ij}^B(t)$. In addition: $P(s_{ij}(t+1) = G/s_{ij}(t) = G) = 1 - \alpha_{ij}; P(s_{ij}(t+1) = B/s_{ij}(t) = G) = \alpha_{ij}; P(s_{ij}(t+1) = G/s_{ij}(t) = B) = \beta_{ij}; P(s_{ij}(t+1) = B/s_{ij}(t) = B) = 1 - \beta_{ij}$, with α_{ij} is the transition probability from state G to B (failure rate) and β_{ij} from B to G (recovery rate). A generic link (i,j) at time step t is in the GOOD state with a probability [2], [3]:

$$q_{ij}^G(t) = \frac{1}{\alpha_{ij} + \beta_{ij}} \big[q_{ij}^G(0)\big(\beta_{ij} + \alpha_{ij}(1 - \alpha_{ij} - \beta_{ij})^t\big) + \\ q_{ij}^B(0)\big(\beta_{ij} - \beta_{ij}(1 - \alpha_{ij} - \beta_{ij})^t\big)\big], \tag{1}$$

while it is in a BAD state with probability:

$$q_{ij}^B(t) = \frac{1}{\alpha_{ij} + \beta_{ij}} \big[q_{ij}^G(0)\big(\alpha_{ij} - \alpha_{ij}(1 - \alpha_{ij} - \beta_{ij})^t\big) + \\ q_{ij}^B(0)\big(\alpha_{ij} + \beta_{ij}(1 - \alpha_{ij} - \beta_{ij})^t\big)\big], \tag{2}$$

where, without loss of generality, it is assumed $0 < \alpha_{ij} + \beta_{ij} < 1$. Each GE model representing a link can be considered independent of the others and identically distributed, as their transition and recovery probabilities are independent and constant across links. However, within each individual link model, there is an inherent correlation arising from its representation as a two-state Markov chain. Therefore, with the occurrence of packet dropout, the network's topology fluctuates among the set of graphs denoted as $\mathbb{G} = \{\mathbb{G}_1, \ldots, \mathbb{G}_s\}$, which are associated with the s states of the Markov chain corresponding to the overall network, composed from both positive recurrent and transient states. The set of possible states the network may be in is determined based on the underlying physical graph $\mathcal{G}$ and channel characteristics. This process can be accurately represented by a finite and irreducible Markov chain, given the finite number of different topologies in which the network can actually exist. The irreducibility ensures that the chain will return to the positive recurrent states infinitely often with probability 1 (unlike the transient states), regardless of its initial state. We assume that at least one connected undirected graph exists in $\mathbb{G}$ corresponding to a positive recurrent state. Notice that such a graph could be the underlying topology $\mathcal{G} \in \mathbb{G}$ itself, as it is designed to guarantee sufficient connectivity when losses are not present. Hence each link $(i,j) \in \mathcal{E}$ is governed by a different GE model affecting the time-dependent network adjacency matrix $A(t) = [a_{ij}(t)]$. Let $D(t) = diag\{d_i(t)\}$ is the the (input) degree matrix with $d_i(t) = \sum_{j=1}^N a_{ij}(t)$, the Laplacian matrix is defined $L(t) = D(t) - A(t)$. We consider a graph without loops $a_{ii}(t) = 0$ for $i \in \mathcal{V}$ and $t \geq 0$. The Laplacian matrix is semi-definite positive with eigenvalues $0 = \xi_1 \leq \xi_2 \leq \cdots \leq \xi_N$, and ξ_2 is called *algebraic connectivity* [20]. Herein we consider the following discrete-time dynamic consensus protocol at sensor node i-th that is an iterative version of one in [1], [14]:

$$x_i(t+1) = x_i(t) + \epsilon \sum_{j=1}^{N} a_{ij}(t)(x_j(t) - x_i(t)) + \Delta r_i(t), \quad (3)$$

with $x_i \in \mathbb{R}$ is the estimation variable, $\Delta r_i(t) = r_i(t) - r_i(t-1)$ is the variation of the reference signal r_i measured by the sensor node, and ϵ is the step size fulfilling $0 < \epsilon < \frac{1}{max_{i,t} d_i(t)}$. In compact form (3) becomes: $x(t+1) = W(t)x(t) + \Delta r(t)$, with $W(t) = I - \epsilon L(t)$ and $\Delta r(t) = [\Delta r_i(t)]$. If $\epsilon < \frac{1}{max_{i,t} d_i(t)}$ the weight matrix $W(t)$ defines a row-stochastic Perron matrix (i.e., $W(t)\mathbf{1} = \mathbf{1}$) with eigenvalues distribution such that: $0 \leq |\lambda_n| \leq \ldots \leq |\lambda_2| \leq \lambda_1 = 1$, [20]. Being $W(t)$ row stochastic, algorithm (3) may be recast as: $x_i(t+1) = x_i(t) + \sum_{j=1}^{N} w_{ij}(t)(x_j(t) - x_i(t)) + \Delta r_i(t)$. The aim of protocol (3) is to dynamically track in a distributed way the average value of the reference signals sampled by the network nodes, i.e. $\frac{1}{N}\sum_{i=1}^{N} r_i(t)$.

We propose the *Dynamic Corrective Consensus* (shortly DC^2) algorithm in order to guarantee distributed tracking of $\frac{1}{N}\sum_{i=1}^{N} r_i(t)$ by the node variables x_i, despite the presence of losses generated by GE type channels. Similarly to the approach proposed for the static consensus [13], auxiliary variables $\phi_{ij}(t)$ (later defined) are introduced to accumulate disagreement state information. Let $\Delta_{ij}(t) = \phi_{ij}(t) + \phi_{ji}(t)$, the DC^2 algorithm is composed of the following *standard* and *corrective* iterations:

I_1 Standard iteration for $t \neq t_u = u(K+1) - 1$, $u \geq 1$
- $\phi_{ij}(t+1) = \phi_{ij}(t) + w_{ij}(t)(x_j(t) - x_i(t))$
- $x_i(t+1) = \sum_{j=1}^{N} \phi_{ij}(t+1) + r_i(t)$

I_2 Corrective iteration for $t = t_u = u(K+1) - 1$, $u \geq 1$
- $\phi_{ij}(t+1) = \phi_{ij}(t) - \frac{1}{2}\Delta_{ij}(t)$
- $x_i(t+1) = \sum_{j=1}^{N} \phi_{ij}(t+1) + r_i(t)$

In the standard iteration, the variables ϕ_{ij} accumulate link state disagreement, while such information is corrected by the formulation in the corrective algorithm phase. The initialization phase of the DC^2 algorithm involves each node i to distributively set its internal variables $\phi_{ij}(-1) = 0$ and to run a corrective iteration I_2. Each node then repeatedly executes a corrective iteration I_2 after every K standard iterations I_1. For the sake of notation, we denote the time steps t at which the corrective iteration is executed by $t_u = u(K+1) - 1$, where u is a positive integer. The parameter $K \geq 1$ is a fixed design value that defines the frequency of the corrective phase I_2. Consequently, the algorithm cyclically executes K iterations I_1 for $t \in [(u-1)(K+1), u(K+1) - 2] = [t_u - K, t_u - 1]$ followed by one iteration of I_2 at step $t_u = u(K+1) - 1$, for $u \geq 1$. The number of applied corrective iterations u, also represents the number of algorithm loops executed. Fig. 1 shows the timing of execution of algorithm loops composed of

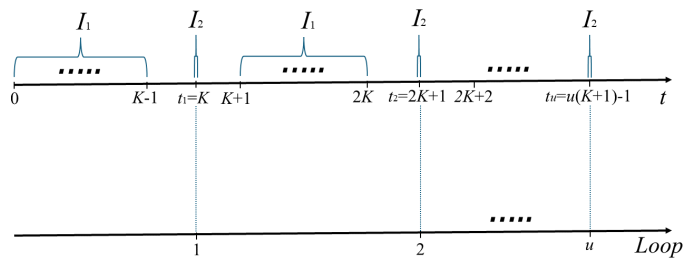

Fig. 1. Timing of the execution of the algorithm loop, which consists of K iterations of I_1 followed by one iteration of I_2

K standard iterations I_1 and one corrective iteration I_2. Such DC^2 algorithm implementation makes it robust to initialization errors and does not exploit input order differences. Moreover, if a neighbor of a node joins/leaves the network, all of its neighbours set the appropriate ϕ_{ij} to zero in order to correctly neutralize the related contribution in the estimation x_i, making the algorithm also robust to re-initialization error. The notion of robustness feature considered herein refers to the algorithm's ability to withstand stochastic link failures and (re)-initialization errors, consistent with the literature (e.g. [1], [16], [17], [13]).

It can be shown that the standard iteration I_1 is equivalent to the standard dynamic consensus protocol (3), [1], while the corrective iteration I_2 may be rewritten as: $x_i(t+1) = x_i(t) - \frac{1}{2}\sum_{j=1}^{N} \Delta_{ij}(t) + \Delta r_i(t)$. For the algorithm convergence analysis purpose, in the following we will refer to the equivalent latter forms.

Let the consensus error $\tilde{x}(t) = x(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T r(t-1)$, the following definitions hold:

***Definition** 1 (Almost-sure dynamic average consensus):* The DC^2 algorithm achieves almost-sure dynamic average consensus if for any initial conditions $x(0)$ it holds: $P\left(\lim_{t\to\infty} \|\tilde{x}(t)\| = 0\right) = 1$.

***Definition** 2 (Mean bounded dynamic average consensus):* The DC^2 algorithm achieves Λ-Mean bounded dynamic average consensus if for any initial conditions $x(0)$ it holds: $\lim_{t\to\infty} \mathbb{E}[\|\tilde{x}(t)\|] \leq \Lambda$, with Λ is a positive constant.

III. Main results

In what follows we present preliminary results, lemmas and proposition that are instrumental in deriving the main results. All analytical proofs of the new presented results will be deferred to the Appendix to enhance the paper's flow and readability.

***Lemma** 1 ([21]):* Let $0 < \beta < 1$, $\{\gamma_t\}$ be a positive scalar sequence such that $\lim_{t\to\infty} \gamma_t = 0$, then $\lim_{t\to\infty} \sum_{l=0}^{t} \beta^{t-l}\gamma_l = 0$.

Our aim is to derive dynamic evaluation and provide a comprehensive bound on the consensus error after multiple loops u of the algorithm (Theorem 1 later defined), according to the timing in Fig. 1. To achieve this, we analyze the consensus error after the first loop of the algorithm for $u = 1$ (Lemma 2 and Proposition 1 later defined), and then extend the result to multiple loops to obtain the final bound. Specifically, the following Lemma 2 is concerned with the evolution of the consensus error over the course of the first K standard iterations, $t = 0, 1, \dots, t_1 - 1 = K - 1$, before applying the first corrective iteration I_2 at step $t_1 = K$ at the end of the first loop $u = 1$ (Fig. 1).

Let $\widetilde{\Delta r}(t) \coloneqq \Delta r(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T \Delta r(t)$, $\bar{\lambda}_2 = \mathbb{E}[|\lambda_2(W(t-1))|]$, to enhance conciseness in the upcoming derivations, we will use the following notation:

$$f(t_u) \coloneqq \sum_{s=t_u-K}^{t_u-1} \bar{\lambda}_2^{t_u-1-s} \mathbb{E}[||\widetilde{\Delta r}(s)||]. \tag{4}$$

***Lemma** 2:* After performing the first K standard iterations I_1 of the DC^2 algorithm, the following inequality holds:

$$\mathbb{E}[\|\tilde{x}(t_1)\|] \leq \bar{\lambda}_2^K \mathbb{E}[\|\tilde{x}(0)\|] + f(t_1) \tag{5}$$

The following Proposition 1 provides the evaluation and bound of consensus error when the first corrective iteration is applied after K standard ones (Fig. 1). The corrective iteration I_2 aims to counteract the accumulated errors from the previous K iterations caused by imbalances in the time varying network topology due to the effects of GE packet loss.

***Proposition** 1:* Let $\mu_i(s) = \max_j q_{ij}^G(s) q_{ij}^B(s)$, $\gamma(s) = \sqrt{\sum_{i=1}^N \mu_i(s)}$, $s \geq 0$, under the application of the first corrective algorithm iteration I_2 at step t_1 it results:

$$\begin{aligned}\mathbb{E}[\|\tilde{x}(t_1+1)\|] \leq \mathbb{E}[\|\tilde{x}(t_1)\|] + \sqrt{2}\epsilon \sum_{s=0}^{t_1-1} \gamma(s) \mathbb{E}[\|\tilde{x}(s)\|] \\ + \mathbb{E}[||\widetilde{\Delta r}(t_1)||]\end{aligned} \tag{6}$$

The following Theorem 1 combines the result from the standard iterations (as described in Lemma 2) and the effect of the corrective iteration (from Proposition 1) during the first loop $u = 1$, to iteratively provide a bound on the consensus error after multiple loops $u \geq 2$. To improve brevity in the upcoming derivations, we will adopt the following notations:

$$v(t_u) \coloneqq \bar{\lambda}_2^K + \sqrt{2}\epsilon \sum_{s=t_u-K}^{t_u-1} \gamma(s) \bar{\lambda}_2^s \tag{7}$$

$$g(t_u) \coloneqq 1 + \sqrt{2}\epsilon \sum_{s=t_u-K}^{t_u-1} \gamma(s) \tag{8}$$

***Theorem** 1:* Let $K \geq 1$ be the number of standard iterations per loop. For any $u \geq 1$, the following inequality holds:

$$\begin{aligned}\mathbb{E}[\|\tilde{x}(t_u+1)\|] \leq \prod_{j=1}^{u} v(t_j) \mathbb{E}[\|\tilde{x}(0)\|] + \mathbb{1}_2^u \cdot \sum_{i=1}^{u-1} g(t_i) f(t_i) \\ \cdot \prod_{j=i+1}^{u} v(t_j) + \mathbb{1}_2^u \cdot \sum_{i=1}^{u-1} \|\widetilde{\Delta r}(t_i)\| \cdot \prod_{j=i+1}^{u} v(t_j) + g(t_u) f(t_u) \\ + \mathbb{E}[||\widetilde{\Delta r}(t_u)||]\end{aligned} \tag{9}$$

where $\mathbb{1}_2^u = 1$ if $u \geq 2$, or $\mathbb{1}_2^u = 0$ if $u < 2$.

To evaluate the asymptotic behavior of the consensus error as $u \to \infty$ in (9), which is crucial for proving the main result of Theorem 2, the following Lemma 3 and Corollary 1 are instrumental in providing bounds and decay rates for the terms on the right-hand side of (9). Specifically, $\mu_{ij}(s)$ is depending on the GOOD and BAD state probabilities of the channel connecting node j-th to node i-th. The next lemma defines lower and upper bound $\check{\mu}_{ij}$ and $\hat{\mu}_{ij}$ for μ_{ij}.

***Lemma** 3:* Let $\mu_{ij}(s) = q_{ij}^G(s) q_{ij}^B(s)$ and defined $\nu_{ij} = (q_{ij}^G(0)\alpha_{ij} - q_{ij}^B(0)\beta_{ij})(\alpha_{ij} - \beta_{ij})$, $\xi_{ij} = (q_{ij}^G(0)\alpha_{ij} - q_{ij}^B(0)\beta_{ij})^2$, $\tau_{ij} = \frac{5}{-\ln(1-\alpha_{ij}-\beta_{ij})}$, then the following conditions hold for all $(i,j) \in \mathcal{E}$:

a) If $\nu_{ij} \geq 2\xi_{ij}$, then $\forall s \geq 0$ it results: $\check{\mu}_{ij}^a \coloneqq \frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2} \leq \mu_{ij}(s) \leq q_{ij}^G(0) q_{ij}^B(0) \coloneqq \hat{\mu}_{ij}^a$;
b) if $2\xi_{ij}\,(1-\alpha_{ij}-\beta_{ij})^{\tau_{ij}} < \nu_{ij} < 2\xi_{ij}$, then $\forall s \geq 0$ it results: $\check{\mu}_{ij}^b \coloneqq \min\left\{\frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2}, q_{ij}^G(0) q_{ij}^B(0)\right\} \leq \mu_{ij}(s) \leq \frac{1}{4} \coloneqq \hat{\mu}_{ij}^b$;
c) if $\nu_{ij} \leq 2\xi_{ij}\,(1-\alpha_{ij}-\beta_{ij})^{\tau_{ij}}$, then $\forall s \geq 0$ it results: $\check{\mu}_{ij}^c \coloneqq q_{ij}^G(0) q_{ij}^B(0) \leq \mu_{ij}(s) \leq \frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2} \coloneqq \hat{\mu}_{ij}^c$.

A direct consequence of the previous lemma and finiteness of μ_{ij} is that for any node i there exist computable constants $\hat{\mu}_i^p$ and $\check{\mu}_i^p$ of type $\hat{\mu}_{ij}^p$ and $\check{\mu}_{ij}^p$, $p \in \{a,b,c\}$, respectively, such that $\max_j \mu_{ij}(s) \leq \hat{\mu}_i^p$ and $\check{\mu}_i^p \leq \min_j \mu_{ij}(s)$ for any $s \geq 0$. In this respect, before of presenting the next result, we need to define the following sets $\hat{N}^p = \{i \in \mathcal{V} : \forall j, (i,j) \in \mathcal{E}, s \geq 0, \max_j \mu_{ij}(s) \leq \hat{\mu}_i^p\}$, $p \in \{a,b,c\}$.

***Corollary** 1:* Let $\hat{N}^p$ detailed as above, $p \in \{a,b,c\}$, defined the network channels-dependent parameter $\hat{\gamma} = \sqrt{\sum_{i\in\hat{N}^a} q_i^G(0) q_i^B(0) + \frac{|\hat{N}^b|}{4} + \sum_{i\in\hat{N}^c} \frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2}}$, then it results: $\gamma(s) \leq \hat{\gamma}$ for all $s \geq 0$. Moreover, If $\mathcal{G}$ is connected and $\rho_s \coloneqq \bar{\lambda}_2^K + \sqrt{2}\epsilon\hat{\gamma}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2} < 1$, it results: $\prod_{j=1}^u v(t_j) \leq \rho_s^u$.

The presented results hold for both stochastic and deterministic reference signals $r_i(t)$. In what follows we present the main result about the DC^2 algorithm almost sure convergence in the case of deterministic signals $r_i(t)$.

***Theorem** 2:* The DC^2 algorithm achieves almost sure dynamic average consensus under symmetric and heterogeneous GE channels if and only if:

A_1 $\mathcal{G}$ is connected
A_2 $\rho_s < 1$
A_3 $\lim_{t\to\infty} ||\widetilde{\Delta r}(t)|| = 0$

***Remark** 1:* Let the network features, i.e. number of nodes N, average connectivity $\bar{\lambda}_2$ (or one its estimation or upper bound), and channels-dependent parameter $\hat{\gamma}$, the parameters K and ϵ of DC^2 algorithm can be designed to fulfill the condition $\rho_s < 1$ of Theorem 2 to assess almost sure dynamic average consensus, for reference signals satisfying assumption A_3. Notice that $\hat{\gamma}$ may be derived using Corollary 1 and the (time-independent) conditions a)-c) of Lemma 3 by the known/measured channel characteristics. Additionally, knowledge of $\mathcal{G}$ or topologies of positive recurrent states enables estimation of $\bar{\lambda}_2$ or its upper bound.

***Remark** 2:* Asymptotically constant or slow varying reference signals satisfy assumption A_3 of Theorem 2: this is also in according with assumptions made in the literature, i.e. [16]. In addition, Assumption A_3 is also fulfilled by reference signals having the same asymptotically variation.

The next Coreollary 2 states conditions for mean-bounded dynamic average consensus as in Definition 2 for bounded reference signals $r_i(t)$.

***Corollary** 2:* If assumptions A_1 and A_2 of Theorem 2 hold and $\limsup_{t\to+\infty} ||\widetilde{\Delta r}(t)|| \leq \theta$, then the DC^2 algorithm achieves Λ-mean bounded dynamic average consensus, i.e. $\lim_{t\to\infty}\mathbb{E}[\|\tilde{x}(t)\|] \leq \Lambda$ with $\Lambda = \theta(1+\frac{\hat{v}}{1-\hat{v}})\big[\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}(1+\sqrt{2}\epsilon\hat{\gamma}(K-1))+1\big]$ and $\hat{v} = \bar{\lambda}_2^K + \sqrt{2}\epsilon\hat{\gamma}\bar{\lambda}_2^{(K+1)}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}$.

The following Corollary 3 extends the previous results to the case of asymmetric and heterogeneous GE channels. In this case, $q_{ij}^G \neq q_{sk}^G$ for $(i,j) \neq (s,k)$ yields $\mu_{ij} \neq \mu_{ji}$. Let $\check{\mu}_i^{re}$ and $\check{\mu}_i^{tr}$ (resp. $\hat{\mu}_i^{re}$ and $\hat{\mu}_i^{tr}$) such that $\check{\mu}_i^{re} \leq \min_j \mu_{ij}(s)$ and $\check{\mu}_i^{tr} \leq \min_j \mu_{ji}(s)$ (resp. $\max_j \mu_{ij}(s) \leq \hat{\mu}_i^{re}$ and $\max_j \mu_{ji}(s) \leq \hat{\mu}_i^{tr}$), $s \geq 0$, depending on the GOOD and BAD probabilities in the reception and transmission of input and output channels at node i, respectively. Defined the following sets: $\check{N}^p = \{i \in \mathcal{V} : \forall s, \check{\mu}_i^p \leq \min[\check{\mu}_i^{tr}, \check{\mu}_i^{re}]\}$ (resp. $\hat{N}^p = \{i \in \mathcal{V} : \forall s, \max[\hat{\mu}_i^{tr}, \hat{\mu}_i^{re}] \leq \hat{\mu}_i^p\}$), $p \in \{a,b,c\}$, then the following result holds:

***Corollary** 3:* The DC^2 algorithm achieves almost sure dynamic average consensus under asymmetric and heterogeneous GE channels if and only if assumptions A_1 and A_3 of Theorem 2 hold and it results: $\rho_a := \bar{\lambda}_2^K + \epsilon\sqrt{\frac{N+1-\check{\gamma}_a+\hat{\gamma}_a}{2}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}} < 1$, with $\check{\gamma}_a = 2\big(\sum_{i\in\check{N}^c} q_i^G(0)q_i^B(0) + \sum_{i\in\check{N}^b}\check{\mu}_i^b + \sum_{i\in\check{N}^a}\frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2}\big)$ and $\hat{\gamma}_a = 2\big(\sum_{i\in\hat{N}^a} q_i^G(0)q_i^B(0) + \sum_{i\in\hat{N}^b}\hat{\mu}_i^b + \sum_{i\in\hat{N}^c}\frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2}\big)$.

***Remark** 3:* The approach extends to non-irreducible finite-state Markov chains by decomposing states into irreducible classes C_i of positively recurrent states. Transitions occur within each class, but not between classes. The method holds as long as each C_i includes at least one state representing a connected undirected graph.

***Remark** 4:* The Λ-mean bounded dynamic average consensus can be also defined in terms of convergence in (c_1, c_2), i.e. $\lim_{t\to\infty} P(\|\tilde{x}(t)\| \leq c_1) \geq c_2$, $c_1 \geq 0$, $c_2 \in [0,1]$ [11]. Indeed, exploiting the Markov inequality, it results: $\lim_{t\to\infty} P(\|\tilde{x}(t)\| \leq c_1) = 1 - \lim_{t\to\infty} P(\|\tilde{x}(t)\| \geq c_1) \geq 1 - \lim_{t\to\infty}\big[\frac{\rho^t\mathbb{E}[\|\tilde{x}(0)\|]+\Lambda}{c_1}\big] \geq 1 - \frac{\Lambda}{c_1}$ with ρ equal to ρ_s or ρ_a for symmetric and asymmetric GE channels, respectively. This assesses convergence in (c_1, c_2), with $c_2 = 1 - \frac{\Lambda}{c_1}$, if $c_1 \geq \Lambda$.

IV. Simulation Validation

We will validate the design conditions in two scenarios with an undirected graph $\mathcal{G}$ of $N = 50$ nodes. Links follow Gilbert-Elliott models with random selected parameters α_{ij}, β_{ij}, $q_{ij}^G(0)$. In the first scenario, each node's reference signal r_i is a randomly generated positive step at $t = 0$ and $t = 500$. Additionally, 30 agents leave the network at $t = 300$ and rejoin at $t = 700$, simulating either planned disconnection or temporary faults. According to the design guidelines outlined in Remark 1, and given the parameters of the asymmetric/heterogeneous GE channels (i.e., α_{ij}, β_{ij}, $q_{ij}^G(0)$), by exploiting Lemma 3 we first derive the values $\check{N}_a = 17$, $\check{N}_b = 25$, $\check{N}_c = 8$, $\hat{N}_a = 26$, $\hat{N}_b = 10$, $\hat{N}_c = 14$. Consequently, by Corollary 3, the bounds $\check{\gamma}_a$ and $\hat{\gamma}_a$ are evaluated. Then, using the topological information, we derive $\bar{\lambda}_2 \leq 0.81$. Finally, based on the values above, the number of standard iterations is designed as $K = 15$ and the sampling time as $\epsilon = 0.024$, ensuring that $\rho_a = 0.79 < 1$, in accordance with Corollary 3. At the top of Fig. 2 the L_2 norm of the estimation error of nodes composing the network during time (i.e. 50 for $t \in [0, 300)$, 20 for $t \in [300, 700)$, 50 for $t \geq 700$) is depicted for both the proposed DC^2 and standard dynamic consensus algorithm (3) [1]. $x_i(0)$ are randomly selected for the DC^2 algorithm to test its robustness, while they are correctly set for the standard dynamic consensus algorithm. The results show that the DC^2 algorithm (solid line) almost surely converges and outperforms the standard approach (dashed line). The DC^2 algorithm proves robustness against (re)-initialization errors, node departures/joins, and disruptions from asymmetric GE channels. Next, we consider bounded periodic signals r_i and symmetric, heterogeneous GE channels. By exploiting Corollary 2, given the final consensus error constraint $\Lambda \leq 3$, along with the evaluated $\hat{\gamma}$ and $\bar{\lambda}_2 \leq 0.77$, we design $K = 24$ and $\epsilon = 0.021$ to achieve $\rho_s = 0.38 < 1$ and $\Lambda = 2.9$. The bottom of Fig. 2 shows the network consensus error (solid line) and the bound Λ (dashed line), demonstrating the design's effectiveness in meeting the constraint.

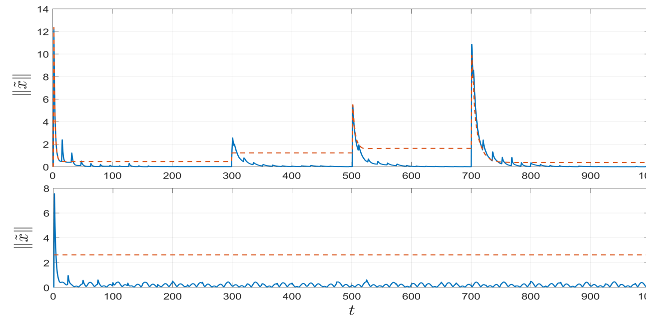

Fig. 2. L_2 norm of the network estimation error. Top: DC^2 algorithm (continuous line) and standard dynamic consensus algorithm (dashed line); Bottom: Mean bounded average consensus convergence of DC^2 algorithm.

Notice that the proposed conditions and design method, as shown in the examples above, leverage link-based bounds (e.g. $\hat{\gamma}$, $\check{\gamma}_a$, $\hat{\gamma}_a$) instead of a global Markov model of the network (e.g. [8]), which simplifies the design, ensures scalability, and reduces computational overhead. They also enable precise tuning of the key DC^2 algorithm and network parameters to ensure either accurate performance (Theorem 2, Corollary 3) or effective consensus error control to fulfill a required application constraint (Corollary 2). These features make the method useful for applications where performance and error constraints are critical.

V. Conclusions and Future Work

The paper introduces a dynamic corrective consensus algorithm to track the average of network reference signals $r_i(t)$. It is robust against re-initialization errors, agents joining/leaving, and disruptions in heterogeneous/asymmetric Gilbert-Elliot channels. Conditions for almost sure and mean-bounded dynamic average consensus are established, depending on algorithm/network parameters and signal characteristics. These conditions facilitate algorithm and network design to attain a desired level of consensus error. While stochastic $r_i(t)$ is addressed, derivations are omitted for brevity. Current research aims to extend the algorithm to a connected directed topology $\mathcal{G}$ and validate it with a network simulator like NS-2.

APPENDIX

A. Proof of Lemma 2

Let $\tilde{x}(t) = x(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T r(t-1)$ and $W(t-1)\frac{1}{N}\mathbf{1}\mathbf{1}^T r(t-2) = \frac{1}{N}\mathbf{1}\mathbf{1}^T r(t-2)$. During the application of the first K standard iterations I_1 (i.e. eq. (3)) at steps $0, \dots, K-1 = t_1 - 1$, the following holds for $t \in [1, t_1]$: $\tilde{x}(t) = x(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T(\Delta r(t-1) + r(t-2)) = W(t-1)(x(t-1) - \frac{1}{N}\mathbf{1}\mathbf{1}^T r(t-2)) + (I - \frac{1}{N}\mathbf{1}\mathbf{1}^T)\Delta r(t-1) = W(t-1)\tilde{x}(t-1) + \widetilde{\Delta r}(t-1)$. To analyze the system, the error is decomposed into the consensus direction ($\mathbf{1}$) and the disagreement directions (orthogonal to $\mathbf{1}$) using a well-known change of coordinate, [1]: $\tilde{e} = [\tilde{e}_1^T \quad \tilde{e}_{2:N}^T]^T = T^T\tilde{x}$, with $T = [\frac{\mathbf{1}}{\sqrt{N}} \quad R]$ and R such that $R^T R = I_{N-1}$, $T^T T = T T^T = I_N$. In the new coordinate the model becomes: $\tilde{e}_1(t) = \tilde{e}_1(t-1)$; $\tilde{e}_{2:N}(t) = R^T W R \tilde{e}_{2:N}(t-1) + R^T\widetilde{\Delta r}(t-1)$. From the algorithm initialization phase, involving a corrective iteration and the setting $\phi_{ij}(-1) = 0$ for a node i and its neighbors j, it follows: $\tilde{e}_1(t) = 0$ for all $t \in [0, K]$, being $\tilde{e}_1(t) = \tilde{e}_1(t-1) = \tilde{e}_1(0) = \frac{\sum_{i=1}^N (x_i(0) - r_i(-1))}{\sqrt{N}} = \frac{\sum_{i=1}^N \sum_{j=1}^N \phi_{ij}(0)}{\sqrt{N}} = \frac{\sum_{i=1}^N \sum_{j=1}^N (\phi_{ij}(-1) - \phi_{ji}(-1))}{2\sqrt{N}} = 0$. Therefore from Cauchy-Schwarz inequality: $||\tilde{e}|| \le ||\tilde{e}_{2:N}|| + |\tilde{e}_1| \le ||R^T W R \tilde{e}_{2:N}(t-1)|| + ||R^T\widetilde{\Delta r}(t-1)||$. Assumed W diagonalizzabile through the orthogonal and unitary matrix R, being $R^{-1} = R^T$ and $||R|| = ||R^T|| = 1$, it results: $||\tilde{e}(t)|| \le ||R^{-1} W R|| ||\tilde{e}_{2:N}(t-1)|| + ||R^T\widetilde{\Delta r}(t-1)|| \le |\lambda_2(W(t-1))| ||\tilde{e}_{2:N}(t-1)|| + ||\widetilde{\Delta r}(t-1)|| \le |\lambda_2(W(t-1))| ||\tilde{e}(t-1)|| + ||\widetilde{\Delta r}(t-1)||$. Due to the norm-preserving property of unitary matrices, it follows that: $||\tilde{e}|| = ||T^T\tilde{x}|| = ||\tilde{x}||$ yielding to: $||\tilde{x}(t)|| \le |\lambda_2(W(t-1))| ||\tilde{x}(t-1)|| + ||\widetilde{\Delta r}(t-1)||$. Taking the expectation yields: $\mathbb{E}[\|\tilde{x}(t)\|] \le \bar{\lambda}_2 \mathbb{E}[\|\tilde{x}(t-1)\|] + \mathbb{E}[\|\widetilde{\Delta r}(t-1)\|]$, with $\bar{\lambda}_2 = \mathbb{E}[|\lambda_2(W(t-1))|]$.

Using the latter inequality, applying the first standard iteration I_1 at $t = 0$ yields: $\mathbb{E}[\|\tilde{x}(1)\|] \le \bar{\lambda}_2 \mathbb{E}[\|\tilde{x}(0)\|] + \mathbb{E}[\|\widetilde{\Delta r}(0)\|]$. Applying the second iteration I_1 at $t = 1$ gives: $\mathbb{E}[\|\tilde{x}(2)\|] \le \bar{\lambda}_2 \mathbb{E}[\|\tilde{x}(1)\|] + \mathbb{E}[\|\widetilde{\Delta r}(1)\|]$. Substituting the previous result for $\mathbb{E}[\|\tilde{x}(1)\|]$, it results: $\mathbb{E}[\|\tilde{x}(2)\|] \le \bar{\lambda}_2^2 \mathbb{E}[\|\tilde{x}(0)\|] + \bar{\lambda}_2 \mathbb{E}[\|\widetilde{\Delta r}(0)\|] + \mathbb{E}[\|\widetilde{\Delta r}(1)\|]$. Hence, applying K iterations I_1 and incorporating the notation (4), we obtain: $\mathbb{E}[\|\tilde{x}(t_1)\|] = \mathbb{E}[\|\tilde{x}(K)\|] \le \bar{\lambda}_2^K \mathbb{E}[\|\tilde{x}(0)\|] + \sum_{s=0}^{K-1} \bar{\lambda}_2^s \mathbb{E}[\|\widetilde{\Delta r}(K-1-s)\|] = \bar{\lambda}_2^K \mathbb{E}[\|\tilde{x}(0)\|] + f(t_1)$, and the result follows.

B. Proof of Proposition 1

The first algorithm corrective iteration I_2 is taken at $t = t_1 = K$ after the initial K standard iterations I_1. By introducing the variables $\delta_{ij}(t) = w_{ij}(t)(\tilde{x}_j(t) - \tilde{x}_i(t)) + w_{ji}(t)(\tilde{x}_i(t) - \tilde{x}_j(t))$ and considering the definition of ϕ_{ij} during the iteration I_1, the first corrective iteration I_2, given by $x_i(t_1+1) = x_i(t_1) - \frac{1}{2}\sum_{j=1}^N \Delta_{ij}(t_1) + \Delta r_i(t_1)$, can be concisely expressed in terms of the tracking error $\tilde{x}$ after some manipulations:

$$\tilde{x}(t_1+1) = \tilde{x}(t_1) - \tfrac{1}{2}\textstyle\sum_{s=0}^{t_1-1} \begin{bmatrix} \sum_{i=1}^N \delta_{i1}(s) \\ \vdots \\ \sum_{i=1}^N \delta_{iN}(s) \end{bmatrix} + \widetilde{\Delta r}(t_1).$$

The application of firstly the L_2-norm operator $\|\cdot\|$ and then of the expectation one $\mathbb{E}[\cdot]$ yields:

$$\mathbb{E}[\|\tilde{x}(t_1+1)\|] \le \mathbb{E}[\|\tilde{x}(t_1)\|] + \frac{1}{2}\sum_{s=0}^{t_1-1} \mathbb{E}\left[\sqrt{\sum_{j=1}^N \left(\sum_{i=1}^N \delta_{ij}(s)\right)^2}\right] + \mathbb{E}[||\widetilde{\Delta r}(t_1)||] \tag{10}$$

Considered the second term of RHS of (10), from the law of total expectation and Jensen's inequality we may infer:

$$\mathbb{E}\left[\sqrt{\sum_{j=1}^N \left(\sum_{i=1}^N \delta_{ij}(s)\right)^2}\right] = \mathbb{E}\left[\mathbb{E}\left[\sqrt{\sum_{j=1}^N \left(\sum_{i=1}^N \delta_{ij}(s)\right)^2} \,\Big|\, \tilde{x}(s)\right]\right] \tag{11}$$

$$\le \mathbb{E}\left[\sqrt{\mathbb{E}\left[\sum_{j=1}^N \left(\sum_{i=1}^N \delta_{ij}(s)\right)^2 \,\Big|\, \tilde{x}(s)\right]}\right]$$

The radicand of the last term in (11) may be recast as:

$$\mathbb{E}\left[\sum_{j=1}^N \left(\sum_{i=1}^N \delta_{ij}(s)\right)^2 \,\Big|\, \tilde{x}(s)\right] = \mathbb{E}\Bigg[\sum_{j=1}^N \sum_{i=1}^N \delta_{ij}^2(s) + \sum_{j=1}^N \sum_{i,l=1, l\neq i}^N \delta_{ij}(s)\delta_{lj}(s) \,\Big|\, \tilde{x}(s)\Bigg] = \sum_{j=1}^N \sum_{i=1}^N \mathbb{E}[\delta_{ij}^2(s)|\tilde{x}(s)]$$

$$= \sum_{j=1}^N \sum_{i=1}^N \mathbb{E}[(w_{ij}(s) - w_{ji}(s))^2](\tilde{x}_j(s) - \tilde{x}_i(s))^2 \tag{12}$$

with: $\mathbb{E}[(w_{ij}(s) - w_{ji}(s))^2] = \epsilon^2 \mathbb{E}[a_{ij}^2(s)] + \epsilon^2 \mathbb{E}[a_{ji}^2(s)] - 2\epsilon^2 \mathbb{E}[a_{ij}(s)]\mathbb{E}[a_{ji}(s)]$. In what follows for the sake of presentation, we assume symmetric heterogeneous GE channels (e.g $q_{ij}^G = q_{ks}^G$, for $(i,j) = (s,k)$, otherwise $q_{ij}^G \neq q_{ks}^G$ for $(i,j) \neq (k,s)$), deferring later the extension to the asymmetric case. Therefore we may infer: $\sum_{j=1}^N \sum_{i=1}^N \mathbb{E}[(w_{ij}(s) - w_{ji}(s))^2](\tilde{x}_j(s) - \tilde{x}_i(s))^2 = \sum_{j=1}^N \sum_{i=1}^N 2\epsilon^2[q_{ij}^G(s) - (q_{ij}^G(s))^2](\tilde{x}_j(s) - \tilde{x}_i(s))^2 = 2\epsilon^2 \sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)(\tilde{x}_j(s) - \tilde{x}_i(s))^2$. From assumption on the symmetry of communication channels, it results: $\sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)(\tilde{x}_j(s) - \tilde{x}_i(s))^2 = \sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)\tilde{x}_j^2(s) + \sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)\tilde{x}_i^2(s) - 2\sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)\tilde{x}_j(s)\tilde{x}_i(s) = \sum_{j=1}^N \sum_{i=1}^N \mu_{ij}(s)\tilde{x}_j^2(s) + \sum_{j=1}^N \sum_{i=1}^N \mu_{ji}(s)\tilde{x}_i^2(s) - 2\sum_{j=1}^N \sum_{i=1}^N \mu_{ji}(s)\tilde{x}_j(s)\tilde{x}_i(s) \le \sum_{i=1}^N \mu_i(s) \sum_{j=1}^N \tilde{x}_j^2(s) + \sum_{j=1}^N \mu_j(s) \sum_{i=1}^N \tilde{x}_i^2(s) - 2\sum_{j=1}^N \mu_j(s) \sum_{i=1}^N \tilde{x}_j(s)\tilde{x}_i(s) = 2\sum_{i=1}^N \mu_i(s)\|\tilde{x}(s)\|^2 - 2\sum_{j=1}^N \mu_j(s) \sum_{i=1}^N \tilde{x}_j(s)\tilde{x}_i(s) \le 4\sum_{i=1}^N \mu_i(s)\|\tilde{x}(s)\|^2$, with

$\mu_i(s) = \max_j \mu_{ij}(s)$. Combining the latter derivations with (11) and (12), it results:

$$\mathbb{E}\left[\sqrt{\sum_{j=1}^{N}\left(\sum_{i=1}^{N}\delta_{ij}(s)\right)^2}\right] \leq \sqrt{8}\epsilon\gamma(s)\mathbb{E}[\|\tilde{x}(s)\|] \quad (13)$$

with $\gamma(s) = \sqrt{\sum_{i=1}^{N}\mu_i(s)}$. Finally, combining (10) and (13) the inequality (6) follows.

C. Proof of Theorem 1

To estimate the consensus error bound along the first algorithm loop ($u = 1$), we can leverage the results of Lemma 2 and Proposition 1. Exploiting (5) of Lemma 2 in (6) of Proposition 1 yields:

$\mathbb{E}[\|\tilde{x}(t_1+1)\|] \leq \bar{\lambda}_2^K\mathbb{E}[\|\tilde{x}(0)\|] + f(t_1) + \sqrt{2}\epsilon\sum_{s=0}^{t_1-1}\gamma(s)\left[\bar{\lambda}_2^s\mathbb{E}[\|\tilde{x}(0)\|] + f(s)\right] + \mathbb{E}[||\widetilde{\Delta r}(t_1)\|] \leq \left[\bar{\lambda}_2^K + \sqrt{2}\epsilon\sum_{s=0}^{t_1-1}\gamma(s)\bar{\lambda}_2^s\right]\mathbb{E}[\|\tilde{x}(0)\|] + \left[1+\sqrt{2}\epsilon\sum_{s=0}^{t_1-1}\gamma(s)\right]f(t_1) + \mathbb{E}\left[||\widetilde{\Delta r}(t_1)||\right]$, where the latter inequality has been derived exploiting the inequality $f(s) \leq f(t_1)$ for $s \in [0, t_1]$ and factoring the terms out. Using the definitions of $f(t_1)$, $v(t_1)$ and $g(t_1)$ from (4), (7) and (8) for $u = 1$, respectively, the above inequality can be compactly written as: $\mathbb{E}[\|\tilde{x}(t_1+1)\|] \leq v(t_1)\mathbb{E}[\|\tilde{x}(0)\|] + g(t_1)f(t_1) + \mathbb{E}[||\widetilde{\Delta r}(t_1)||]$. By similar computation at the second algorithm's loop for $u = 2$ it results:

$\mathbb{E}[\|\tilde{x}(t_2+1)\|] \leq v(t_2)\mathbb{E}[\|\tilde{x}(t_1+1)\|] + g(t_2)f(t_2) + \mathbb{E}[||\widetilde{\Delta r}(t_2)||] \leq v(t_2)v(t_1)\mathbb{E}[\|\tilde{x}(0)\|] + v(t_2)g(t_1)f(t_1) + v(t_2)\mathbb{E}[||\widetilde{\Delta r}(t_1)||] + g(t_2)f(t_2) + \mathbb{E}[||\widetilde{\Delta r}(t_2)||]$.

Finally, applying u algorithm iterations yields inequality (9).

D. Proof of Lemma 3

From expressions (1) and (2) of q_{ij}^G and q_{ij}^B we may derive: $\mu_{ij}(s) = q_{ij}^G(s)q_{ij}^B(s) = \frac{\nu_{ij}}{(\alpha_{ij}+\beta_{ij})^2}(1-\alpha_{ij}-\beta_{ij})^s + \frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2} - \frac{\xi_{ij}}{(\alpha_{ij}+\beta_{ij})^2}(1-\alpha_{ij}-\beta_{ij})^{2s}$ with $\nu_{ij} = (q_{ij}^G(0)\alpha_{ij} - q_{ij}^B(0)\beta_{ij})(\alpha_{ij}-\beta_{ij})$ and $\xi_{ij} = (q_{ij}^G(0)\alpha_{ij} - q_{ij}^B(0)\beta_{ij})^2$. Moreover: $\dot{\mu}_{ij}(s) = (1-\alpha_{ij}-\beta_{ij})^s \cdot \frac{\ln(1-\alpha_{ij}-\beta_{ij})}{(\alpha_{ij}+\beta_{ij})^2}\left[\nu_{ij} - 2\xi_{ij}(1-\alpha_{ij}-\beta_{ij})^s\right]$.

Let $0 < \alpha_{ij}+\beta_{ij} < 1$, and τ_{ij} be sufficiently large, for example, $\tau_{ij} = \frac{5}{-\ln(1-\alpha_{ij}-\beta_{ij})}$, such that the term $(1-\alpha_{ij}-\beta_{ij})^s$ can be regarded as vanishing for $s \geq \tau_{ij}$, and $\mu_{ij}(\tau_{ij})$ reaches a stationary regime, i.e., $\mu_{ij}(\tau_{ij}) \simeq \mu_{ij}(\infty)$ and $\dot{\mu}_{ij}(\tau_{ij}) \simeq 0$. Let the expression of $\dot{\mu}_{ij}(s)$, if $\nu_{ij} \geq 2\xi_{ij}$ then the function $\mu_{ij}(s)$ is decreasing $\forall s \geq 0$ such that $\frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2} \simeq \mu_{ij}(\tau_{ij}) \leq \mu_{ij}(s) \leq \mu_{ij}(0) = q_{ij}^G(0)q_{ij}^B(0)$, and condition a) follows.

If $2\xi_{ij}(1-\alpha_{ij}-\beta_{ij})^{\tau_{ij}} < \nu_{ij} < 2\xi_{ij}$, from the expression of $\dot{\mu}_{ij}(s)$ and the continuity of $\mu_{ij}(s)$ on the compact $[0, \tau_{ij}]$, there exists $s^\star \in (0, \tau_{ij})$ such that $\mu_{ij}(s)$ is increasing for $s < s^\star$ and decreasing for $s > s^\star$, where $s^\star = \frac{\ln\left(\frac{\nu_{ij}}{2\xi_{ij}}\right)}{\ln(1-\alpha_{ij}-\beta_{ij})}$. In addition from the continuity of $\dot{\mu}_{ij}(s)$, $\dot{\mu}_{ij}(s) \to 0$ for $s \to \infty$. Then there is only one change of derivative sign for $s \in [0, \infty)$. Hence $\mu_{ij}(s)$ has the unique maximum at $s^\star$ with value $\mu_{ij}(s^\star) = \frac{1}{4}$. This also yields: $\mu_{ij}(s) \geq \min\{\mu_{ij}(\tau_{ij}), \mu_{ij}(0)\}$, with $\mu_{ij}(\tau_{ij}) \simeq \frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2}$. Hence condition b) follows.

Finally, if $\nu_{ij} \leq 2\xi_{ij}(1-\alpha_{ij}-\beta_{ij})^{\tau_{ij}}$, the function $\mu_{ij}(s)$ is increasing $\forall s \in [0, \tau_{ij}]$ such that $\mu_{ij}(0) \leq \mu_{ij}(s) \leq \mu_{ij}(\tau_{ij}) \simeq \frac{\alpha_{ij}\beta_{ij}}{(\alpha_{ij}+\beta_{ij})^2}$. Furthermore, $\dot{\mu}_{ij}(\tau_{ij}) \simeq 0$ for $s \geq \tau_{ij}$, with $\dot{\mu}_{ij}(s) \to 0$ for $s \to \infty$. Hence condition c) follows.

E. Proof of Corollary 1

Let a $\hat{\mu}_i^p$ for each set N^p, $p \in \{a, b, c\}$, and being $\gamma(s) \leq \sqrt{\sum_{i\in\hat{N}^a}\hat{\mu}_i^a + \sum_{i\in\hat{N}^b}\hat{\mu}_i^b + \sum_{i\in\hat{N}^c}\hat{\mu}_i^c}$, it results for any s: $\gamma(s) \leq \sqrt{\sum_{i\in\hat{N}^a}q_i^G(0)q_i^B(0) + \frac{|\hat{N}^b|}{4} + \sum_{i\in\hat{N}^c}\frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2}} := \hat{\gamma}$ and the first part of corollary follows.

The second part of corollary follows observing that from (7) we may infer:

$\prod_{j=1}^{u}v(t_j) = \prod_{j=1}^{u}\left(\bar{\lambda}_2^K + \sqrt{2}\epsilon\sum_{s=t_j-K}^{t_j-1}\gamma(s)\bar{\lambda}_2^s\right) \leq \prod_{j=1}^{u}\left(\bar{\lambda}_2^K + \sqrt{2}\epsilon\hat{\gamma}\bar{\lambda}_2^{t_j-K}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\right) \leq \left(\bar{\lambda}_2^K + \sqrt{2}\epsilon\hat{\gamma}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\right)^u$. The last inequalities follow accounting for $\gamma(s) \leq \hat{\gamma}$ and observing that: $\sum_{s=t_j-K}^{t_j-1}\bar{\lambda}_2^s = \bar{\lambda}_2^{t_j-K}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2} \leq \frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}$, where $\bar{\lambda}_2 \in (0, 1)$ since the network Markov chain is irreducible and there exists at least one positive recurrent state associated with a connected undirected graph.

F. Proof of Theorem 2

(Sufficiency) Under the assumption of deterministic signals $r_i(t)$, let $\Delta r_M(t_u) := \max_{t\in[t_u-K, t_u-1]}||\widetilde{\Delta r}(t)||$ then from (4) and (8) the following bounds may be derived for any u: $f(t_u) \leq \Delta r_M(t_u)\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}$ and $g(t_u) \leq 1 + \sqrt{2}\epsilon\hat{\gamma}(K-1)$.

Given assumption A_1, applying the derivation for $j = i+1$ as shown in the second part of the proof of Corollary 1 yields the bound: $\prod_{j=i+1}^{u}v(t_j) \leq \hat{v}^{u-i}$, where $\hat{v} = \bar{\lambda}_2^K + \sqrt{2}\epsilon\hat{\gamma}\bar{\lambda}_2^{(K+1)}\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}$. Notice that $\hat{v} \in (0, 1)$, being from assumption A_2: $0 < \hat{v} < \rho_s < 1$. Exploiting the above bounds on f, g, $\prod_{j=i+1}^{u}v(t_j)$ in the second, third and fourth term of the RHS of inequality (9) of Theorem 1 yields:

$$\mathbb{E}[||\tilde{x}(t_u+1)||] \leq \prod_{j=1}^{u}v(t_j)\mathbb{E}[||\tilde{x}(0)||] + \quad (14)$$

$$1_2^u \cdot \left[1+\sqrt{2}\epsilon\hat{\gamma}(K-1)\right]\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\sum_{i=1}^{u-1}\hat{v}^{u-i}\Delta r_M(t_i) \quad (15)$$

$$+ [1+\sqrt{2}\epsilon\hat{\gamma}(K-1)]\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\Delta r_M(t_u) \quad (16)$$

$$+ 1_2^u \cdot \sum_{i=1}^{u-1}\hat{v}^{u-i}||\widetilde{\Delta r}(t_i)|| + ||\widetilde{\Delta r}(t_u)|| \quad (17)$$

We evaluate $\lim_{u\to\infty}\mathbb{E}[||\tilde{x}(t_u+1)||]$. For the term (14) it results $\lim_{u\to\infty}\prod_{j=1}^{u}v(t_j)\mathbb{E}[||\tilde{x}(0)||] = 0$, by assumptions A_1, A_2 and Corollary 1. From A_3, $\lim_{u\to\infty}||\widetilde{\Delta r}(t_u)|| = 0$ and $\lim_{u\to\infty}\Delta r_M(t_u) = 0$. Hence, the term (16) and the last term in (17) are vanishing as $u \to \infty$. As $\hat{v} \in (0, 1)$ and A_3 holds, the term (15) and the first term of (17) are asymptotically null by Lemma 1. Hence, it results: $\lim_{u\to\infty}\mathbb{E}[\|\tilde{x}(t_u+1)\|]=0$. Due to the generality of the chosen subsequence starting point, the above convergence result also holds for all K subsequences $\{\mathbb{E}[\|\tilde{x}(t_u+h)\|]\}_u$, obtained for $h=2,\dots,K+1$. Hence $\lim_{t\to\infty}\mathbb{E}[\|\tilde{x}(t)\|]=0$. Finally, by Fatou's Lemma, $0\le\mathbb{E}[\lim_{t\to\infty}\|\tilde{x}(t)\|]\le\lim_{t\to\infty}\mathbb{E}[\|\tilde{x}(t)\|]=0$. Hence, $P(\lim_{t\to\infty}\|\tilde{x}(t)\|=0)=1$, and the result follows.

(Necessity) Assume the physical graph $\mathcal{G}$ is disconnected. Then, at least one agent does not exchange information with others. Consequently, there exists an initial state vector $x(0)$ for which dynamic average consensus fails with non-zero probability. Hence, the state does not converge in probability, and thus almost surely, to the average of the reference signals.

G. Proof of Corollary 2

By the assumption $\limsup_{t\to+\infty}||\widetilde{\Delta r}(t)||\le\theta$ and arguments from Lemma 1, we infer (similar to the proof in [21], omitted for brevity): $\lim_{u\to\infty}\sum_{i=1}^{u-1}\hat{v}^{u-i}\Delta r_M(t_i)\le\theta\frac{\hat{v}}{1-\hat{v}}$ and $\lim_{u\to\infty}\sum_{i=1}^{u-1}\hat{v}^{u-i}||\widetilde{\Delta r}(t_i)||\le\theta\frac{\hat{v}}{1-\hat{v}}$. Exploiting the above bounds in (14)-(17) of Theorem 2 it results: $\lim_{u\to\infty}\mathbb{E}[\|\tilde{x}(t_u+1)\|]\le\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\theta(1+\sqrt{2}\epsilon\hat{\gamma}(K-1))\frac{\hat{v}}{1-\hat{v}}+\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\theta(1+\sqrt{2}\epsilon\hat{\gamma}(K-1))+\theta\frac{\hat{v}}{1-\hat{v}}+\theta=\frac{1-\bar{\lambda}_2^K}{1-\bar{\lambda}_2}\big(1+\sqrt{2}\epsilon\hat{\gamma}(K-1)\big)\theta(1+\frac{\hat{v}}{1-\hat{v}})+\theta(1+\frac{\hat{v}}{1-\hat{v}}):=\Lambda$. Since this also holds for the subsequences $\{\mathbb{E}[\|\tilde{x}(t_u+h)\|]\}_u$ for $h=2,\dots,K+1$, the result follows.

H. Proof of Corollary 3

Due to space constraints, we provide only a brief sketch of the proof for the result analogous to Proposition 1, tailored to the asymmetric/heterogeneous GE channels case. Being $q_{ij}^G\neq q_{ji}^G$, for the term in (12) it results: $\mathbb{E}[(w_{ij}(s)-w_{ji}(s))^2]=\epsilon^2[q_{ij}^G+q_{ji}^G-2q_{ij}^Gq_{ji}^G]=\epsilon^2[q_{ij}^G+q_{ji}^G-2q_{ij}^G(1-q_{ji}^B)]\le\epsilon^2[-q_{ij}^G+q_{ji}^G+(q_{ij}^G)^2+(q_{ji}^B)^2]=\epsilon^2[-q_{ij}^G(1-q_{ij}^G)+(1-q_{ji}^B)+(q_{ji}^B)^2]=\epsilon^2[-q_{ij}^Gq_{ij}^B+1-q_{ji}^B(1-q_{ji}^B)]=\epsilon^2[1-\mu_{ij}(s)-\mu_{ji}(s)]$. Therefore, (12) can be upper bounded as: $\sum_{j=1}^N\sum_{i=1}^N\mathbb{E}[(w_{ij}(s)-w_{ji}(s))^2](\tilde{x}_j(s)-\tilde{x}_i(s))^2\le\sum_{j=1}^N\sum_{i=1}^N\epsilon^2[1-\mu_{ij}(s)\;-\mu_{ji}(s)](\tilde{x}_j(s)-\tilde{x}_i(s))^2=\epsilon^2\sum_{j=1}^N\sum_{i=1}^N(\tilde{x}_j(s)-\tilde{x}_i(s))^2-\epsilon^2\sum_{j=1}^N\sum_{i=1}^N\mu_{ij}(s)(\tilde{x}_j(s)-\tilde{x}_i(s))^2-\epsilon^2\sum_{j=1}^N\sum_{i=1}^N\mu_{ji}(s)(\tilde{x}_j(s)-\tilde{x}_i(s))^2\le2\epsilon^2(N+1)\|\tilde{x}_i(s)\|^2-2\epsilon^2\sum_{i=1}^N(\check{\mu}_i^{tr}+\check{\mu}_i^{re})\|\tilde{x}_i(s)\|^2+2\epsilon^2\sum_{i=1}^N(\hat{\mu}_i^{tr}+\hat{\mu}_i^{re})\|\tilde{x}_i(s)\|^2=2\epsilon^2(N+1-\check{\gamma}_1+\hat{\gamma}_1)\|\tilde{x}_i(s)\|^2$, with: $\check{\gamma}_1=\sum_{i=1}^N(\check{\mu}_i^{tr}+\check{\mu}_i^{re})$ (resp. $\hat{\gamma}_1=\sum_{i=1}^N(\hat{\mu}_i^{tr}+\hat{\mu}_i^{re})$), $\check{\mu}_i^{re}\le\min_j\mu_{ij}(s)$ (resp. $\max_j\mu_{ij}(s)\le\hat{\mu}_i^{re}$) and $\check{\mu}_i^{tr}\le\min_j\mu_{ji}(s)$ (resp. $\max_j\mu_{ji}(s)\le\hat{\mu}_i^{tr}$), $s\ge0$. By Lemma 3 and definition of $\check{N}^p$ (resp. $\hat{N}^p$), we get a lower bound $\check{\gamma}_a=2(\sum_{i\in\check{N}^c}q_i^G(0)q_i^B(0)+\sum_{i\in\check{N}^b}\check{\mu}_i^b+\sum_{i\in\check{N}^a}\frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2})$ (resp. upper bound $\hat{\gamma}_a$) such that $\check{\gamma}_a\le\check{\gamma}_1$ (resp. $\hat{\gamma}_1\le\hat{\gamma}_a$). Applying (11) and (10) from Proposition 1, we obtain an inequality analogous to (6) for this case: $\mathbb{E}[\|\tilde{x}(t_1+1)\|]\le\mathbb{E}[\|\tilde{x}(t_1)\|]+\epsilon\sum_{s=0}^{t_1-1}\sqrt{\frac{N+1-\check{\gamma}_a+\hat{\gamma}_a}{2}}\mathbb{E}[\|\tilde{x}(s)\|]+||\widetilde{\Delta r}(t_1)||$. Using the latter inequality, the proof follows similar arguments as in Theorem 1 and Theorem 2.

References

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

[2] E. N. Gilbert, "Capacity of a burst-noise channel," *The Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[3] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell Sys. Tech. Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[4] T. Farjam, H. Wymeersch, and T. Charalambous, "Distributed channel access for control over known and unknown gilbert–elliott channels," *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 7405–7419, 2023.

[5] B. Touri and A. Nedic, "On ergodicity, infinite flow, and consensus in random models," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1593–1605, 2011.

[6] A. Tahbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 3, pp. 791–795, 2008.

[7] G. Shi, B. D. O. Anderson, and K. H. Johansson, "Consensus over random graph processes: Network borel–cantelli lemmas for almost sure convergence," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5690–5707, 2015.

[8] R. C. Ballam, A. McFadyen, and D. E. Quevedo, "Local averaging for consensus over communication links with random dropouts," *IEEE Control Systems Letters*, vol. 7, pp. 1387–1392, 2023.

[9] Y. Zhang and Y.-P. Tian, "Consentability and protocol design of multi-agent systems with stochastic switching topology," *Automatica*, vol. 45, no. 5, pp. 1195–1201, 2009.

[10] I. Matei, N. Martins, and J. S. Baras, "Almost sure convergence to consensus in markovian random graphs," in *2008 47th IEEE Conference on Decision and Control*, 2008, pp. 3535–3540.

[11] Z. Chen, L. Cai, and C. Zhao, "Convergence time of average consensus with heterogeneous random link failures," *Automatica*, vol. 127, p. 109496, 2021.

[12] F. Fagnani and S. Zampieri, "Average consensus with packet drop communication," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 102–133, 2009.

[13] Y. Chen, R. Tron, A. Terzis, and R. Vidal, "Corrective consensus with asymmetric wireless links," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 6660–6665.

[14] P. S. Demetri, R. Olfati-Saber, and M. M. Richard, "Dynamic consensus on mobile networks," in *IFAC world congress*, 2005.

[15] S. Manfredi and D. Angeli, "Asymptotic consensus on the average of a field for time-varying nonlinear networks under almost periodic connectivity," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2389–2404, 2018.

[16] E. Montijano, J. I. Montijano, C. Sagüés, and S. Martínez, "Robust discrete time dynamic average consensus," *Automatica*, vol. 50, no. 12, pp. 3131–3138, 2014.

[17] J. George and R. A. Freeman, "Robust dynamic average consensus algorithms," *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4615–4622, 2019.

[18] M. Franceschelli and A. Gasparri, "Multi-stage discrete time and randomized dynamic average consensus," *Automatica*, vol. 99, pp. 69–81, 2019.

[19] F. Chen, C. Chen, G. Guo, C. Hua, and G. Chen, "Delay and packet-drop tolerant multistage distributed average tracking in mean square," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9535–9545, 2022.

[20] R. Merris, "Laplacian matrices of graphs: a survey," *Linear Algebra and its Applications*, vol. 197-198, pp. 143–176, 1994.

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