IEEE Open Journal of Signal Processing

Received 9 August 2023; revised 2 December 2023; accepted 6 December 2023. Date of publication 18 December 2023; date of current version 19 January 2024. The review of this article was arranged by Associate Editor U. A Khan.

Digital Object Identifier 10.1109/OJSP.2023.3344052

Distributed Adaptive Learning Under Communication Constraints

MARCO CARPENTIERO [1,2] (Graduate Student Member, IEEE), VINCENZO MATTA [1,2] (Senior Member, IEEE), AND ALI H. SAYED [3] (Fellow, IEEE)

[1]Department of Information and Electrical Engineering and Applied Mathematics (DIEM), University of Salerno, I-84084 Fisciano, Italy
[2]National Inter-University Consortium for Telecommunications (CNIT), Italy
[3]School of Engineering, École Polytechnique Fédérale de Lausanne EPFL, CH-1015 Lausanne, Switzerland

CORRESPONDING AUTHOR: MARCO CARPENTIERO (e-mail: mcarpentiero@unisa.it)

An earlier version of this paper was presented at ICASSP 2022 [DOI: 10.1109/ICASSP43922.2022.9747154].

ABSTRACT We consider a network of agents that must solve an online optimization problem from continual observation of *streaming* data. To this end, the agents implement a distributed *cooperative* strategy where each agent is allowed to perform *local* exchange of information with its neighbors. In order to cope with communication constraints, the exchanged information must be compressed to reduce the communication load. We propose a distributed diffusion strategy nicknamed as ACTC (Adapt-Compress-Then-Combine), which implements the following three operations: adaptation, where each agent performs an individual stochastic-gradient update; compression, which leverages a recently introduced class of *stochastic compression operators*; and combination, where each agent combines the *compressed* updates received from its neighbors. The main elements of novelty of this work are as follows: *i*) *adaptive* strategies, where constant (as opposed to diminishing) step-sizes are critical to infuse the agents with the ability of responding in real time to nonstationary variations in the observed model; *ii*) *directed*, i.e., non-symmetric combination policies, which allow us to enhance the role played by the network topology in the learning performance; *iii*) *global strong convexity*, a condition under which the individual agents might feature even non-convex cost functions. Under this demanding setting, we establish that the iterates of the ACTC strategy fluctuate around the exact global optimizer with a mean-square-deviation on the order of the step-size, achieving remarkable savings of communication resources. Comparison against up-to-date learning strategies with compressed data highlights the benefits of the proposed solution.

INDEX TERMS Distributed optimization, adaptation and learning, diffusion strategies, stochastic quantizers.

I. INTRODUCTION

In the last decades, the steady progress of statistical learning and network science led to great interest in *distributed optimization* strategies implemented by multi-agent networks [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. Primary advantages of distributed strategies include scalability, possibility of working with reduced-size and spatially dispersed datasets, and robustness against failures. Distributed optimization allows each network agent to overcome its individual limitations through cooperation with its neighbors, and to deliver superior performance with respect to single-agent strategies [21], [22].

Several distributed implementations have been proposed and examined in great detail in previous works. These implementations can be categorized in terms of different attributes. For example, we have distributed gradient descent and stochastic gradient descent algorithms; constant as opposed to diminishing step-size implementations; consensus and diffusion strategies [21]. In all these cases, distributed optimization and learning algorithms require exchange of information among spatially dispersed agents. Accordingly, the exchanged data must be compressed to reduce the communication load. For this reason, *data compression* lies at the core of any distributed implementation. The main purpose of this work is to carry out a detailed analysis of the learning

behavior of distributed strategies operating under communication constraints.

The article is organized as follows. In Section II, we provide a summary of the related literature, while Section III contains the background on the distributed optimization problem and the network graph descriptors. These two sections can be skipped by the expert reader, who can directly go to Section IV, where we illustrate the strategy proposed in this work. Section V introduces the equations and tools necessary to study the dynamics of the proposed strategy, while in Section VI we provide a qualitative illustration for its behavior. Sections VII and VIII collect the results about mean-square stability and transient analysis, respectively. In Section IX, we show the outcomes from a series of numerical experiments, while conclusions follow in Section X.

II. RELATED WORK

A. DISTRIBUTED QUANTIZATION

The earlier works on quantization for distributed inference considered: *i*) network architectures with a single fusion center; *ii*) a static setting where inference is performed in a single step (e.g., no gradient descent or other iterative algorithms are used); and *iii*) known parametric models for the underlying distributions of the data. Under this setting, there are several quantization strategies available in the literature, with reference to different inferential and learning tasks, such as decentalized estimation [23], [24], [25], [26], decentralized detection and classification [27], [28], [29], [30], [31], [32], [33], [34], also with many useful variations including cross-layer [35], [36] or censoring approaches [37].

One critical difficulty in data compression for inferential problems is the lack of knowledge about the underlying data distributions, which depend on the same *unknown* parameters that the agents are trying to learn. This lack of knowledge complicates severely the tuning of the quantizers' thresholds, since these thresholds would depend on the unknown parameters to be estimated. A breakthrough in the solution to this problem was given by the introduction of *stochastic quantizers* in the distributed learning strategies. In the context of data compression (with single-agent and not for inference applications) the usage of stochastic quantizers can be dated back to the seminal works on probabilistic analysis of quantizers [38] and dithering [39]. In comparison, one of the earliest appearances of stochastic quantizers in distributed estimation was provided in [40], which inspired subsequent applications to distributed learning [41], [42]. These works rely on a universal, fully data-driven approach, where the thresholds of the stochastic quantizers do not require knowledge of the underlying distribution. In fact, these thresholds are randomly adapted to the measurements observed locally by the agents, in such a way that the resulting quantizer is *unbiased on average*, despite the lack of knowledge about the underlying distribution. This property is critical to make the inferential strategies capable of learning well.

B. COMPRESSION FOR DISTRIBUTED OPTIMIZATION

The works mentioned in the previous section refer to static strategies (where inference is performed in a single step), in optimization and learning problems there is often the necessity of employing iterative strategies to converge to the desired solution. In these settings, the data compression issue becomes even more critical, since the quantization error can accumulate over successive iterations, destroying the convergence of the pertinent optimization algorithms.

This issue motivated the recent introduction of randomized quantizers for distributed optimization algorithms such as gradient descent or stochastic gradient descent algorithms. One recent solution that attracted attention in this context is the general class of randomized compression strategies introduced in [43]. This class relies on two properties that are critical to guarantee proper learning, namely, *unbiasedness* and *bounded variance*.

More specifically, in [43] it was assumed that each agent broadcasts compressed versions of its locally observed gradients to all other agents in the network. However, calling upon the theory of predictive quantization (e.g., the sigma-delta modulation adopted in PCM [44]), we see that the impact of quantization errors on convergence can be reduced by properly leveraging the inherent memory arising in recursive implementations such as gradient descent implementations. Two canonical paradigms to achieve this goal are *error-feedback management* [45], [46], [47], [48] and *differential quantization* [49], [50], which, perhaps surprisingly, have been applied to distributed optimization quite recently.

Under error-feedback management, at each time step: first, the current gradient plus an error stored from the previous iteration is compressed; and then, the compression error is compensated in the subsequent iteration. In comparison, differential quantization leverages the memory present in the iterative algorithm to reduce the error variance, by compressing only the difference between subsequent iterates. For a fixed bit budget, it is indeed more convenient to compress the difference (i.e., the innovation) between consecutive samples, rather than the samples themselves. This is because: *i*) the innovation typically exhibits a reduced range as compared to the entire sample; and *ii*) owing to the correlation between consecutive samples, quantizing the entire sample will waste resources by transmitting redundant information. The information-theoretic fundamental limits of (non-stochastic) gradient descent under differential quantization have been recently established in [50].

However, the aforementioned works on error compensation and differential quantization refer either to a *federated learning* scheme where all agents communicate with a fusion center (as in [49]) or to *fully connected* networks where each agent is connected to all other agents. In this work, we focus instead on the more challenging setting where optimization must be *fully decentralized*.

C. FULLY DECENTRALIZED COMPRESSED LEARNING

Under the fully decentralized scenario, each agent is responsible for its own inference, which is obtained by successive steps of local interaction with its neighbors. Typical strategies for fully distributed optimization *without* compression are consensus or diffusion strategies [21], [51].

One early characterization of adaptive diffusion with compressed data was provided in [52], where the compression errors were modeled as noise over the communication channel. There are other works that address the quantization issue by focusing on the explicit encoder structure. Some useful results are available for the case of *exact, i.e., non-stochastic gradient-type algorithms*. In [53], uniform quantization of the iterates is considered for a distributed implementation of the subgradient descent algorithm. In this scheme, the agents update locally their state variables by averaging the state variables received from their neighbors, and then follow the subgradient descent direction. More recently, additional convergence results were presented in [54], where random (dithered) quantization is applied, along with a weighting scheme to give more or less importance to the analog local state and the quantized averaged state of the neighbors. In [55], the randomized quantizers proposed in [43] are considered for a distributed gradient descent implementation using consensus with compressed iterates and an update rule similar to the one adopted in [54].

The aforementioned works on fully decentralized schemes rely on the availability of the *exact* gradient. In the present work, we focus instead on the *adaptive* setting where the agents collect noisy streaming data to evaluate a *stochastic instantaneous approximation* of the actual gradient, and must be endowed with online algorithms capable to respond in real time to drifts in the underlying conditions Useful communication-constrained implementations that can be applied to stochastic gradients were recently proposed in [56], [57]. However, these works rely on symmetric combination matrices and assume that each individual cost function is strongly convex. In the next section we explain why it is useful to overcome these limitations.

D. ADVANCES IN RELATION TO PRIOR WORK

— *Diffusion:* The Adapt-Compress-Then-Combine (ACTC) strategy proposed in this work belongs to the family of *diffusion* strategies [21], while the available works on distributed optimization under communication constraints focus on consensus strategies [56], [57]. Along with many commonalities, one fundamental difference between diffusion and consensus resides in the asynchrony of the latter strategy in the combination step (where the updated state of an agent is combined with the previous states of its neighbors) [21]. This asynchrony has an effect both in terms of stability and learning performance. In fact, it has been shown that consensus algorithms can feature smaller range of stability as compared to diffusion strategies and a slightly worse learning performance [21]. For these reasons, in this work we opt for a diffusion scheme. Starting from the traditional (uncompressed) Adapt-Then-Combine (ATC) diffusion strategy detailed in [21], we allow for local exchange of *compressed* variables by means of *stochastic quantizers*. We will see that, thanks to the diffusion mechanism, the ACTC scheme will be able to adapt and learn well, and in particular it will outperform previous distributed quantized strategies based on consensus.

— *Adaptation:* We focus on a dynamic setting where the agents are called to learn by continually collecting *streaming data* from the environment. Under an adaptive setting, once the distributed learning algorithm starts, we want the agents to learn virtually forever, by automatically adapting their behavior in face of nonstationary drifts in the streaming data. To this end, *stochastic* gradient algorithms with *constant* step-size are necessary. These algorithms have been shown to tradeoff well learning and adaptation. On the learning side, each agent resorts to some instantaneous approximation of the cost function (which is not perfectly known in practice) and tries to learn with increasing precision by leveraging the increasing information coming from the streaming data. On the adaptation side, the constant step-size leaves a persistent amount of "noise" in the algorithm (the "gradient noise") which automatically infuses the algorithm with the ability of promptly reacting to drifts. In contrast, over diminishing step-size implementations, the gradient noise is progressively annihilated over time. As a consequence, diminishing step-size algorithms learn infinitely better as time progresses under stationary conditions. At the same time, if the minimizer changes (e.g., because of drifts in the underlying distribution) diminishing step-size algorithms get stuck on the previously computed minimizer, exhibiting a sort of "elephant's memory", i.e., requiring a time to get out from a local minimizer that is at least proportional to the time the algorithm needed to approach that minimizer. One example of the distinct behavior characterizing adaptive and non-adaptive implementations is offered in Fig. 1. This figure provides a preview of the good learning/adaptation trade-off offered by the ACTC strategy proposed in this work: the ACTC strategy converges to a small error *before* the drift occurring at time $i = 4000$, concurrently ensuring prompt reaction, i.e., fast convergence to a small error *after* the drift. In comparison, we see that the non-adaptive version of the ACTC strategy employing diminishing step-sizes reacts very slowly, with an error about 20 dB higher than the error of the adaptive strategy even after 4000 iterations from the change.

— *Left-Stochastic combination policies:* In many applications, the exchange and pooling of information between agents need not be symmetric. For example, two agents can be allowed to communicate only in one direction or, even if bi-directional communication is permitted, two agents might decide to give different importance to the data they exchange with each other. As we will explain in the next section, this general setting is described by a weighted *directed* graph associated with a *left-stochastic* combination matrix. For the important class of directed graphs, useful optimization algorithms, also for the adaptive case with constant step-size, have

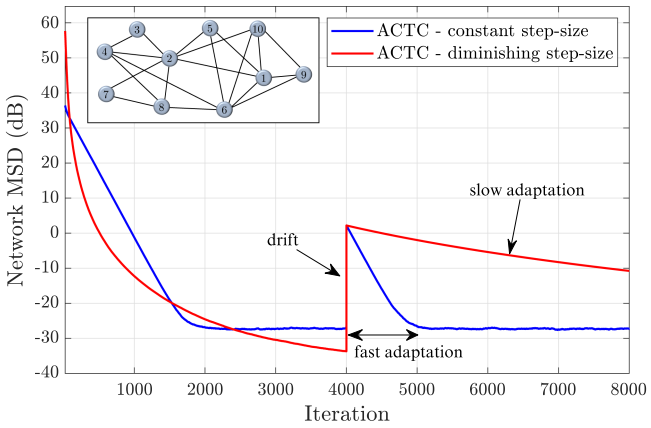

FIGURE 1. **This plot shows the network mean-square-deviation (the error quantifying the distance of the minimizer of agent k at iteration i from the correct theoretical minimizer, averaged over all agents in the network shown in the inset plot) as a function of the iteration i, achieved by two types of distributed learning strategies operating under communication constraints. The blue curve refers to the ACTC strategy proposed in this work, which employs a constant step-size in the stochastic-gradient update. The red curve refers to a variation of the ACTC strategy that employs instead a diminishing step-size. We consider a nonstationary setting where a drift occurs at iteration $i = 4000$, where the correct minimizer changes. We see that the strategy with constant step-size (set at $\mu = 0.01$) is able to adapt to the change and track the drift. In sharp contrast, the strategy with diminishing step-size (set at $\mu(i) = 5/(i+1)$ following standard prescriptions in stochastic optimization [21]), while achieving a lower error *before* the change, see around $i \approx 2500$, requires an unacceptably long time to converge *after* the change, since after 4000 iterations from the change it still has an error significantly higher than the error of the adaptive strategy. The details of the simulation are: regression problem of Section IX, with dimensionality $M = 20$ and data generated according to the example in Fig. 2; uniform averaging combination rule [21]; randomized quantizer from [43] with 2 bits; ACTC algorithm with stability parameter $\zeta = 0.1$.**

been proposed for the traditional setup without communication constraints [58], [59], [60]. In contrast, the existing works on distributed optimization under communication constraints focus on *symmetric and doubly-stochastic* combination policies.

In this work we consider the more general setting of left-stochastic combination policies, which allow us to represent a richer variety of distributed interactions, where the network topology plays a fundamental role. For example, the limiting Perron eigenvector of left-stochastic matrices is not uniform, a property that can be exploited to compensate for non-uniform agents' behavior [21]. Note that, differently from doubly-stochastic matrices, left-stochastic matrices can be constructed in practice without requiring any coordination across the agents.

— *Global strong convexity:* Convergence of the stochastic gradient *iterates* is typically examined under the assumption that the gradients are Lipschitz and the cost functions are strongly convex. In the distributed setting, the latter property is usually translated into assuming that all the local cost functions pertaining to the individual agents are strongly convex [56], [57]. Sometimes the additional assumption of uniform gradient boundedness is adopted (e.g., in [56]), which can however hold only approximately in the Lipschitz and strongly convex setting. For Lipschitz gradient and strongly convex local functions, without the uniform boundedness approximation, convergence results were recently obtained for distributed primal-dual algorithms with compressed communication [57].

Notably, in the present work we relax the aforementioned assumptions, since we do not rely on any uniform boundedness approximation, and require only a *global* cost function (i.e., a linear combination of the local cost functions) to be strongly convex. This extension allows us to cover relevant cases where the individual agents have non-convex costs. Moreover, global, as opposed to local strong convexity, can be exploited to implement distributed regularization procedures where the limitations of the individual agents can be overcome by cooperation. For example, $N-1$ agents might be unable to learn the true minimizer (local unidentifiability), but they can nevertheless compensate their limited view by cooperating with a single farsighted agent. One example of this type will be considered in Section IX.

E. MAIN CONTRIBUTIONS

The main achievement of this work is the characterization of the learning behavior of the ACTC strategy, namely, of an adaptive diffusion strategy for learning under communication constraints. An essential building block to prove the result is given by the unifying description and mathematical tools developed in [14], [21], which allow us to decouple the learning dynamics over two main components by means of a suitable network coordinate transformation. These two components are: i) a *centralized* stochastic-gradient component that converges to a unique solution common to all agents; ii) a deviation component that dies out as time elapses, and which takes into account the initial discrepancy between the trajectories of the individual agents.

Finally, we compare our algorithm against two up-to-date solutions, namely, the algorithms CHOCO-SGD [56] and DUAL-SGD [57], showing that both algorithms are outperformed by our ACTC strategy. As we will carefully explain in Section IX: i) the improvement on CHOCO-SGD arises primarily from using a diffusion-type strategy as opposed to the consensus strategy employed by CHOCO-SGD; and ii) the improvement on DUAL-SGD arises mainly from the advantage of primal-domain strategies (like the proposed ACTC) over primal-dual distributed strategies (like DUAL-SGD).

Notation: We use boldface letters to denote random variables, and normal font letters for their realizations. Capital letters refer to matrices, small letters to both vectors and scalars. Sometimes we violate the latter convention, for instance, we denote the total number of network agents by N. All vectors are column vectors. In particular, the symbol $\mathbb{1}_L$ denotes an $L \times 1$ vector whose entries are identically equal to 1. Likewise, the identity matrix of size L is denoted by I_L. For two matrices X and Y, the notation $X \geq Y$ signifies that $X - Y$ is positive semi-definite. In comparison, the notation $X \geq Y$ signifies that the individual entries of $X - Y$ are nonnegative.

For a vector x, the symbol $\|x\|$ denotes the ℓ_2 norm of x. For a matrix X, the ℓ_2 induced matrix norm is accordingly $\|X\|$. Other norms will be characterized by adding the pertinent subscript. For example $\|x\|_1$ will denote the ℓ_1 norm of x, and $\|X\|_1$ the ℓ_1 induced matrix norm (maximum absolute column sum of X). The symbol $\otimes$ denotes the Kronecker product. The symbol $*$ denotes complex conjugation. $X^\top$ is the transpose of matrix X, whereas X^{H} is the Hermitian (i.e., conjugate) transpose of a complex matrix X. The symbol $\mathbb{E}$ denotes the expectation operator. For $\mu > 0$ and a nonnegative function $f(\mu)$, the notation $f(\mu) = O(\mu)$ signifies that there exist a constant $C > 0$ and a value $\mu_0 > 0$ such that $f(\mu) \leq C\mu$ for all $\mu \leq \mu_0$.

III. BACKGROUND

We consider a network of N agents solving a distributed optimization problem. Each individual agent $k = 1, 2, \ldots, N$ is assigned a *local* cost or risk function:

$$J_k(w) : \mathbb{R}^M \to \mathbb{R}. \tag{1}$$

The local cost functions are assumed to satisfy the following regularity condition.

Assumption 1 (Individual cost function smoothness): For all $w \in \mathbb{R}^M$, each cost function $J_k(w)$ is twice-differentiable and its Hessian matrix $\nabla^2 J_k(w)$ satisfies the following Lipschitz condition, for some positive constants $\{\eta_k\}$:

$$\nabla^2 J_k(w) \leq \eta_k I_M. \tag{2}$$

□

In practice, it is seldom the case that the cost functions are perfectly known to the agents. In contrast, each agent usually has access to a *stochastic* approximation of the true cost function. For example, in the adaptation and learning theory the cost functions are often modeled as the expected value of a loss function $L_k(w; \boldsymbol{x}_k)$, namely,

$$J_k(w) = \mathbb{E}[L_k(w; \boldsymbol{x}_k)], \tag{3}$$

where the expectation is taken with respect to a random variable $\boldsymbol{x}_k$ that can represent, e.g., some training data observed by agent k. In many scenarios of interest, the statistical characterization of $\boldsymbol{x}_k$ is not available to the agents and, hence, $J_k(w)$ is not known and is rather approximated by the stochastic quantity $L_k(w; \boldsymbol{x}_k)$. Moreover, if different data $\boldsymbol{x}_{k,i}$ are collected over time by agent k, the stochastic approximation takes the form of an *instantaneous* approximation $L_k(w; \boldsymbol{x}_{k,i})$ depending on time index i.

More generally, whether or not the cost function is defined through (3), in the following treatment we assume that agent k at time i is able to approximate the true gradient $\nabla J_k(w)$ through a *stochastic instantaneous approximation* $\boldsymbol{g}_{k,i}(w)$ which, without loss of generality, can be written as the true gradient plus a *gradient noise* term $\boldsymbol{n}_{k,i}(w)$, namely,

$$\boldsymbol{g}_{k,i}(w) = \nabla J_k(w) + \boldsymbol{n}_{k,i}(w). \tag{4}$$

A. CLASSICAL ATC DIFFUSION STRATEGY

The Adapt-Then-Combine (ATC) diffusion strategy is a popular distributed mechanism that consists of iterated application of the following two steps, for $i = 1, 2, \ldots$

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} - \mu_k \boldsymbol{g}_{k,i}(\boldsymbol{w}_{k,i-1}) & \text{[Adapt]} \\ \boldsymbol{w}_{k,i} = \sum_{\ell=1}^N a_{\ell k} \boldsymbol{\psi}_{\ell,i} & \text{[Combine]} \end{cases} \tag{5}$$

In (5), agents $k = 1, 2, \ldots, N$ evolve over time i by producing a sequence of iterates $\boldsymbol{w}_{k,i} \in \mathbb{R}^M$. The adaptation step is a *self-learning* step, where each agent k at time i computes its own instantaneous stochastic approximation $\boldsymbol{g}_{k,i}(\cdot)$ of the local gradient $\nabla J_k(\cdot)$, evaluated at the previous iterate $\boldsymbol{w}_{k,i-1}$. Such an approximation is scaled by a small step-size $\mu_k > 0$ and used to update the previous iterate $\boldsymbol{w}_{k,i-1}$ following the (stochastic) gradient descent. The maximum step-size across the agents will be denoted by:

$$\mu \triangleq \max_{k=1,2,\ldots,N} \mu_k, \tag{6}$$

giving rise to the *scaled* step-sizes:

$$\alpha_k \triangleq \frac{\mu_k}{\mu}. \tag{7}$$

The combination step is a *social learning* step, where agent k aims at realigning its descent direction with the rest of the network by combining its local update $\boldsymbol{\psi}_{k,i}$ with the other agents' updates scaled by some nonnegative scalars $\{a_{\ell k}\}$, which are referred to as *combination weights*. The support graph of the combination matrix $A = [a_{\ell k}]$ describes the connections between agents, i.e., the topology of a network whose vertices correspond to the agents, and whose edges represent directional links between agents. According to this model, when no communication link exists between agents ℓ and k, the combination weights $a_{\ell k}$ and $a_{k\ell}$ must be equal to zero. Likewise, when information can flow only from ℓ to k, we will have $a_{\ell k} > 0$ and $a_{k\ell} = 0$. In summary, the combination process is a local process where only neighboring agents interact.

It is useful to introduce the neighborhood of agent k:

$$\mathcal{N}_k \triangleq \{\ell = 1, 2, \ldots, N : a_{\ell k} > 0\}, \tag{8}$$

which is a *directed* neighborhood that accounts for the incoming flow of information from ℓ to k (possibly including the self-loop $\ell = k$).

We will work under the following standard regularity conditions on the network.

Assumption 2 (Strongly Connected Network): The network is strongly connected, which means that, given any pair of nodes (ℓ, k), a path with nonzero weights exists in both directions, i.e., from ℓ to k and vice versa (the two paths need

not be the same) and that at least one agent k in the entire network has a self-loop ($a_{kk} > 0$). □

Assumption 3 (Stochastic combination matrix): For each agent $k = 1, ..., N$ the following conditions hold:

$$a_{\ell k} \geq 0, \quad \sum_{\ell \in \mathcal{N}_k} a_{\ell k} = 1, \quad a_{\ell k} = 0 \text{ for } \ell \notin \mathcal{N}_k, \tag{9}$$

which imply that the combination matrix $A = [a_{\ell k}]$ is a left-stochastic matrix. □

Under Assumptions 2 and 3, the combination matrix A is a primitive matrix, and thus satisfies the Perron-Frobenius theorem, which in particular implies the existence of the Perron vector $\pi = [\pi_1, \pi_2, \ldots, \pi_N]^\top$, a vector with all strictly positive entries satisfying the following relationships:

$$A\pi = \pi, \qquad \mathbb{1}_N^\top \pi = 1. \tag{10}$$

For later use, it is convenient to introduce the vector p that mixes the topological information encoded in the Perron eigenvector with the scaled step-sizes $\{\alpha_k\}$, namely,

$$p = [\alpha_1\pi_1, \alpha_2\pi_2, \ldots, \alpha_N\pi_N]^\top. \tag{11}$$

We are now ready to introduce the global strong convexity assumption that will be required for our results to hold.

Assumption 4 (Global Strong Convexity): Let $p_k = \alpha_k\pi_k$ be the k-th entry of the scaled Perron eigenvector in (11). The (twice differentiable) aggregate cost function:

$$J(w) = \sum_{k=1}^{N} p_k J_k(w) \tag{12}$$

is ν-strongly convex, namely, a positive constant ν exists such that, for all $w \in \mathbb{R}^M$ we have:

$$\sum_{k=1}^{N} p_k \nabla^2 J_k(w) \geq \nu I_M. \tag{13}$$

□

We remark that, for Assumption 4 to hold, the local cost functions are allowed to be non-convex, provided that (13) is satisfied.

For later use, it is useful to notice that Assumptions 1 and 4 entail a relationship between the Lipschitz constants $\{\eta_k\}$ and the strong convexity constant ν. Specifically, using (2) and (13) and applying the triangle inequality we can write:

$$\nu \leq \sum_{k=1}^{N} p_k \|\nabla^2 J_k(w)\| \leq \sum_{k=1}^{N} p_k \eta_k \triangleq \eta. \tag{14}$$

The adaptation and learning performance of the ATC strategy has been examined in great detail in previous works [14], [15], [21]. The major conclusion stemming from these works is that, under Assumptions 1–3 (plus some classical assumptions on the gradient noise — see Assumption 7 further ahead), the ATC strategy is able to drive each agent toward a close neighborhood of the minimizer $w^\star$ of the *global* cost function (12). We remark that in the considered framework each cost function can be different, having a specific minimizer w_k^0 (or even multiple minimizers), which may not coincide with the *unique* global network minimizer $w^\star$ of the aggregate cost function $J(w)$.

We notice also that the structure of (12) allows us to solve different optimization problems where the objective function can be expressed as the linear combination of local cost functions, including the special case where the weights p_k are all uniform, which can be obtained when the step-sizes are all equal (i.e., $\alpha_k = 1$ for all k) and the combination matrix is doubly-stochastic (since the Perron eigenvector of a doubly-stochastic matrix has entries $p_\ell = 1/N$ for all $\ell = 1, 2, \ldots, N$). In addition, and remarkably, it was shown in [21], [62] that the minimizer of (12) corresponds to a Pareto solution to the multi-objective optimization problem:

$$\min_w \{J_1(w), J_2(w), \ldots, J_N(w)\}, \tag{15}$$

where the choice of the weights $\{p_k\}$ drives convergence to a particular Pareto solution.

IV. ATC WITH COMPRESSED COMMUNICATION

The ATC diffusion scheme (5) is designed under the assumption that agents exchange over the communication channel their intermediate updates $\boldsymbol{\psi}_{k,i}$. However, in a realistic environment the information shared by the agents must be necessarily *compressed*. This necessity gives rise to at least two fundamental questions. First, is it possible to design diffusion strategies that preserve the adaptation and learning capabilities of the ATC strategy despite the presence of data compression? Assume the answer to the first question is in the affirmative. Then it is natural to ask whether there is a limit on the amount of compression, since it is obviously desirable for the agents to save as much bandwidth and energy as possible. Our analysis will give precise elements to address these important questions. To start with, we introduce a compressed version of the ATC strategy.

There exist obviously several possibilities to perform data compression. In order to select one particular strategy, we need to consider the fundamental limitations of our setting, in particular: lack of knowledge of the underlying statistical model and correlation across subsequent iterates. As explained in Section II-A, an excellent tool to overcome these limitations is *stochastic differential quantization*, which will be embedded into the ATC recursion giving rise to the Adapt-Compress-Then-Combine (ACTC) diffusion strategy, which can be described as follows.

The time-varying variables characterizing the ACTC recursion are: an intermediate update $\boldsymbol{\psi}_{k,i}$, a quantized state $\boldsymbol{q}_{k,i}$, and the current minimizer $\boldsymbol{w}_{k,i}$. At time $i = 0$ each agent k is initialized with an arbitrary state value $\boldsymbol{q}_{k,0}$ (with finite second moment). Then, agent k receives the initial states $\boldsymbol{q}_{\ell,0}$ from its neighbors $\ell \in \mathcal{N}_k$ and computes an initial minimizer $\boldsymbol{w}_{k,0} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{q}_{\ell,0}$.

Then, for every $i > 0$, the agents perform the following four operations. First, each agent k performs *locally* the same

adaptation step as in the ATC strategy:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} - \mu_k \boldsymbol{g}_{k,i}(\boldsymbol{w}_{k,i-1}) \quad \text{[Adapt]} \tag{16}$$

Second, each agent k compresses the *difference* between the update $\boldsymbol{\psi}_{k,i}$ and the previous quantized state $\boldsymbol{q}_{k,i-1}$, through a compression function $\boldsymbol{Q}_k : \mathbb{R}^M \to \mathbb{R}^M$:

$$\boldsymbol{Q}_k(\boldsymbol{\psi}_{k,i} - \boldsymbol{q}_{k,i-1}) \quad \text{[Compress]} \tag{17}$$

The bold notation for the compression function highlights that *randomized* functions are permitted, as we will explain more carefully in Section IV-A. Then, agent k receives from its neighbors $\ell \in \mathcal{N}_k$ the *compressed* values $\boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1})$. Since the quantization operation acts on differences, the quantized states $\boldsymbol{q}_{\ell,i}$ must be updated by adding the quantized difference to the previous value $\boldsymbol{q}_{\ell,i-1}$. Specifically, agent k updates *the quantized values corresponding to all* $\ell \in \mathcal{N}_k$:

$$\boldsymbol{q}_{\ell,i} = \boldsymbol{q}_{\ell,i-1} + \zeta\, \boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1}), \tag{18}$$

where $\zeta \in (0, 1)$ is a design parameter that will be useful to tune the stability of the algorithm, as we will carefully explain in due time. Finally, agent k combines the updated states corresponding to its neighbors as usual:

$$\boldsymbol{w}_{k,i} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{q}_{\ell,i} \quad \text{[Combine]} \tag{19}$$

It is important to remark that, in order to perform the update step in (18), agent k must possess the variables $\boldsymbol{q}_{\ell,i-1}$ from its neighbors $\ell \in \mathcal{N}_k$. This might appear problematic at first glance, since we have just seen that only the differences $\boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1})$ are actually received by k from a neighboring agent ℓ. On the other hand, since at $i = 0$ agent k knows the initial quantized states $\{\boldsymbol{q}_{\ell,0}\}_{\ell \in \mathcal{N}_k}$, and since the update rule in (18) depends on $\{\boldsymbol{q}_{\ell,i-1}\}_{\ell \in \mathcal{N}_k}$ and the quantized innovation $\{\boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1})\}_{\ell \in \mathcal{N}_k}$, we conclude that sharing of the quantized differences along with the initial states $\{\boldsymbol{q}_{\ell,0}\}_{\ell \in \mathcal{N}_k}$ is enough for agent k to implement (18) at every instant i, by keeping memory only of the last neighboring variables $\{\boldsymbol{q}_{\ell,i}\}_{\ell \in \mathcal{N}_k}$.

In summary, the ACTC scheme can be compactly described as in Algorithm 1.

Before concluding this section, we remark that the combination step of the ACTC strategy considers only quantized variables, even if, in principle, agent k might also combine its *analog* state $\boldsymbol{\psi}_{k,i}$ in place of the quantized counterpart $\boldsymbol{q}_{k,i}$. As a general principle, the more we compress, the less we spend (even in terms of local processing). In this respect, showing that the ACTC strategy learns properly by combining only quantized variables has its own interest. Moreover, considering only quantized variables gives to the algorithm a symmetric structure that avoids adding further complexity to the mathematical formalization that is necessary to support the analysis.

A. COMPRESSION OPERATORS

In order to implement the ACTC diffusion strategy in Algorithm 1, it is necessary to specify the compression operators

Algorithm 1: ACTC Diffusion Strategy.

for $i = 1, 2, \ldots$ **do**
  for $k = 1, 2, \ldots, N$ **do**
    Adaptation step
    Evaluate the instantaneous stochastic approximation $\boldsymbol{g}_{k,i}(\boldsymbol{w}_{k,i-1})$ and compute

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} - \mu_k \boldsymbol{g}_{k,i}(\boldsymbol{w}_{k,i-1}). \tag{20}$$

    Compression step
    – Transmit the compressed difference

$$\boldsymbol{Q}_k(\boldsymbol{\psi}_{k,i} - \boldsymbol{q}_{k,i-1}).$$

    – Receive $\boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1})$ from $\ell \in \mathcal{N}_k \setminus \{k\}$
    – Decode by computing

$$\boldsymbol{q}_{\ell,i} = \boldsymbol{q}_{\ell,i-1} + \zeta\, \boldsymbol{Q}_\ell(\boldsymbol{\psi}_{\ell,i} - \boldsymbol{q}_{\ell,i-1}), \quad \forall \ell \in \mathcal{N}_k. \tag{21}$$

    Combination step

$$\boldsymbol{w}_{k,i} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{q}_{\ell,i}. \tag{22}$$

  end for
end for

$\boldsymbol{Q}_k(\cdot)$. We focus on the class of compression operators considered in the majority of prior works on distributed optimization and learning [43], [56], [57], [61]. This class is described by Assumption 5. We remark that our analysis can be readily extended to other classes, such as the recent variant proposed in [65].

Assumption 5 (Compression operators): A randomized compression operator associates to an input x a stochastic output $\boldsymbol{Q}(x)$ whose distribution is governed by a conditional probability measure $\mathbb{P}(\cdot|x)$ fulfilling the following two conditions:

$$\mathbb{E}\left[\boldsymbol{Q}(x)\right] = x \quad \text{[unbiasedness]} \tag{23}$$

$$\mathbb{E}\,\|\boldsymbol{Q}(x) - x\|^2 \le \omega\,\|x\|^2 \quad \text{[bounded variance]} \tag{24}$$

where expectations are relative to $\mathbb{P}(\cdot|x)$. □

By "randomized operator" we mean that, given a *deterministic* input x, the output $\boldsymbol{Q}(x)$ is *randomly* chosen (two meaningful ways to perform such random choice will be illustrated in Appendix K).

Two main observations are important to capture the meaning of Assumption 5. The first one concerns the role of parameter ω, which quantifies the amount of compression. Small values of ω correspond to small amount of compression, i.e., finely quantized data. Large values of ω are instead representative of severe compression. The second observation concerns properties (23) and (24). It will emerge from the technical analysis that property (23) enables the possibility that the quantization errors arising during the ACTC evolution

average to zero as time elapses. Property (24) will be critical to guarantee that the variance of the quantization errors does not blow-up over time.

Remarkably, the class of randomized quantizers introduced in Assumption 5 is fairly general and flexible, including a broad variety of compression paradigms. Some of these paradigms are particularly tailored to our setting, where we have to compress vectors with possibly large dimensionality M. For example, one useful paradigm is the *sparse compression paradigm*, where a small subset of the M components of the input vector x is sent with arbitrarily large precision, while the remaining entries are set to zero. In this context, the cost of communication is usually evaluated in terms of the reduction of dimensionality, i.e., in terms of the number of nonzero components as done, e.g., for the lossless case in the compressed sensing or analog compression framework [63], [64]. Another compression scheme that meets Assumption 5 is the randomized quantization scheme recently proposed in [43]. Both these popular schemes can be found in earlier works, see, e.g. [43], [57]. To make the paper self-contained and for the benefit of the reader we report their description in Appendix K.

In order to avoid misunderstanding, we stress that in the forthcoming treatment we will use interchangeably the terminology compression or quantization to indicate a general compression operator fulfilling Assumption 5. Reference to a specific class of compression operators, such as the randomized quantizers and the randomized sparsifiers illustrated in Appendix K, will be made when needed (e.g., in Section IX).

Since we focus on a distributed setting, we allow different agents to employ different compression operators. Accordingly, for a given input x, each agent k generates the stochastic output $\boldsymbol{Q}_k(x)$ according to a probability measure $\mathbb{P}_k(\cdot|x)$ satisfying Assumption 5 with compression parameter ω_k. The largest (i.e., worst-case) compression parameter is denoted by:

$$\Omega \triangleq \max_{k=1,2,\ldots,N} \omega_k. \tag{25}$$

As is standard in distributed optimization with data compression, we assume that the compression operators are memoryless and independent across agents.

Assumption 6 (Conditions on compression operators): Let

$$\boldsymbol{e}_{k,i} \triangleq \boldsymbol{Q}_k(\boldsymbol{\delta}_{k,i}) - \boldsymbol{\delta}_{k,i} \tag{26}$$

be the compression error. The compression process is independent over time and space, i.e., given the past history formed by all the iterates generated by the ACTC algorithm before applying the i-th compression step, namely,

$$\boldsymbol{h}_i \triangleq \left\{ \{\boldsymbol{q}_{k,j}\}_{j=0}^{i-1}, \{\boldsymbol{\psi}_{k,j}\}_{j=1}^{i} \right\}_{k=1}^{N}, \tag{27}$$

the vector $\boldsymbol{Q}_k(\boldsymbol{\delta}_{k,i})$ depends only on the input $\boldsymbol{\delta}_{k,i}$, and is generated, independently across agents, as:

$$\boldsymbol{\delta}_{k,i} \overset{\mathbb{P}_k(\cdot|\boldsymbol{\delta}_{k,i})}{\longrightarrow} \boldsymbol{Q}_k(\boldsymbol{\delta}_{k,i}). \tag{28}$$

In view of Assumption 5, from these properties we have:

$$\mathbb{E}\left[\boldsymbol{e}_{k,i} \,|\, \boldsymbol{h}_i\right] = \mathbb{E}\left[\boldsymbol{e}_{k,i} \,|\, \boldsymbol{\delta}_{k,i}\right] = 0, \tag{29}$$

$$\mathbb{E}\left[\|\boldsymbol{e}_{k,i}\|^2 \,|\, \boldsymbol{h}_i\right] = \mathbb{E}\left[\|\boldsymbol{e}_{k,i}\|^2 \,|\, \boldsymbol{\delta}_{k,i}\right] \leq \omega_k \|\boldsymbol{\delta}_{k,i}\|^2, \tag{30}$$

$$\mathbb{E}\left[\boldsymbol{e}_{\ell,i}\boldsymbol{e}_{k,i}^{\top} \,|\, \boldsymbol{h}_i\right] = \mathbb{E}\left[\boldsymbol{e}_{\ell,i}\boldsymbol{e}_{k,i}^{\top} \,|\, \boldsymbol{\delta}_{k,i}\right] = 0, \ \forall \ell \neq k. \tag{31}$$

□

We will show in the remainder of the article that the mean-square-error approaches $O(\mu)$ for small step-sizes, i.e., the algorithm is mean-square-error stable even in the presence of quantization errors and gradient noise. The derivations are demanding and challenging due to the nonlinear and coupled nature of the network dynamics, as is clear from the arguments in the appendices. Nevertheless, when all is said and done, we arrive at the reassuring conclusion that the diffusion strategy is able to learn well in quantized/compressed environments.

V. NETWORK ERROR DYNAMICS

In this section we illustrate the main formalism that will be exploited to conduct our analysis. Since we are interested in computing the deviation of the ACTC iterates from the global minimizer $w^\star$, it is expedient to introduce the following centered variables:

$$\begin{cases} \widetilde{\boldsymbol{w}}_{k,i} \triangleq \boldsymbol{w}_{k,i} - w^\star \\ \widetilde{\boldsymbol{\psi}}_{k,i} \triangleq \boldsymbol{\psi}_{k,i} - w^\star \\ \widetilde{\boldsymbol{q}}_{k,i} \triangleq \boldsymbol{q}_{k,i} - w^\star \end{cases} \tag{32}$$

It is also convenient to rewrite the adaptation step in order to make explicit the role of the *true* cost functions $J_k(w)$. To this end, we must exploit the *gradient noise* introduced in (4), which quantifies the discrepancy between the approximate and true gradients. Exploiting (4), the adaptation step in (20) can be rewritten as:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} - \mu_k \nabla J_k(\boldsymbol{w}_{k,i-1}) - \mu_k \boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1}). \tag{33}$$

Examining (33), we see that the gradient noise contains an additional source of randomness given by its argument $\boldsymbol{w}_{k,i-1} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{q}_{\ell,i-1}$, whose randomness comes accordingly from the previous-step quantized iterates $\{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^N$. We assume the following standard regularity properties for the gradient noise process.

Assumption 7 (Gradient Noise): For all $k = 1, 2, \ldots, N$ and all $i > 0$, the gradient noise meets the following conditions:

$$\mathbb{E}\left[\boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1}) \,\middle|\, \{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^N\right] = 0, \tag{34}$$

$$\mathbb{E}\left[\|\boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1})\|^2 \,\middle|\, \{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^N\right] \leq \beta_k^2 \|\widetilde{\boldsymbol{w}}_{k,i-1}\|^2 + \sigma_k^2, \tag{35}$$

for some constants β_k and σ_k. □

The bounded gradient-noise variance is one standard assumption in optimization theory. It is possible to provide some intuition on (35) as done in [21]. Consider first a bound of the form:

$$\mathbb{E}\left[\|\boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1})\|^2 \,\middle|\, \{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^N\right] \leq \bar{\beta}_k^2 \|\boldsymbol{w}_{k,i-1}\|^2 + \bar{\sigma}_k^2, \tag{36}$$

where the non-centered iterate $\boldsymbol{w}_{k,i-1}$ appears on both sides. Note that $\boldsymbol{w}_{k,i-1}$ is a deterministic function of $\{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^{N}$, and the expectation is conditioned on $\{\boldsymbol{q}_{\ell,i-1}\}_{\ell=1}^{N}$. Therefore, the value $\boldsymbol{w}_{k,i-1}$ can be considered fixed, and condition (36) is a classical boundedness condition where the conditional variance of the gradient noise is bounded by a constant term plus a term proportional to the squared norm of the input value $\boldsymbol{w}_{k,i-1}$. This condition is met for popular models, including the regression model and the logistic, or regularized logistic models — see, e.g., [21]. Then, we obtain (35) from (36) by applying Jensen's inequality to get:

$$\|\widetilde{\boldsymbol{w}}_{k,i-1}\|^2 = \|\boldsymbol{w}_{k,i-1} - w^\star\|^2 \le 2\|\boldsymbol{w}_{k,i-1}\|^2 + 2\|w^\star\|^2 \tag{37}$$

such that (35) holds with the choices: $\beta_k^2 = 2\bar{\beta}_k^2$ and $\sigma_k^2 = \bar{\sigma}_k^2 + 2\bar{\beta}_k^2\|w^\star\|^2$.

With reference to the actual gradient in (4), from the mean-value theorem one has (we recall that $\widetilde{\boldsymbol{w}}_{k,i-1} = \boldsymbol{w}_{k,i-1} - w^\star$) [21]:

$$\nabla J_k(\boldsymbol{w}_{k,i-1}) = \nabla J_k(w^\star) + \left[\int_0^1 \nabla^2 J_k(w^\star + t\widetilde{\boldsymbol{w}}_{k,i-1})dt\right] \widetilde{\boldsymbol{w}}_{k,i-1}, \tag{38}$$

where the integral of a matrix is intended to operate entrywise. Introducing the bias of agent k,

$$b_k \triangleq \alpha_k \nabla J_k(w^\star), \tag{39}$$

and the following random and time-varying matrix defined in terms of the Hessian of the cost function of agent k:

$$\boldsymbol{H}_{k,i-1} \triangleq \alpha_k \int_0^1 \nabla^2 J_k(w^\star + t\widetilde{\boldsymbol{w}}_{k,i-1})dt, \tag{40}$$

Eq. (38) yields:

$$\mu_k \nabla J_k(\boldsymbol{w}_{k,i-1}) = \mu\left(b_k + \boldsymbol{H}_{k,i-1}\widetilde{\boldsymbol{w}}_{k,i-1}\right). \tag{41}$$

For notational convenience, the scaled step-size α_k has been embodied in the definitions of the bias and the matrix $\boldsymbol{H}_{k,i-1}$. Likewise, it is useful to introduce the *scaled* gradient noise vector:

$$\boldsymbol{s}_{k,i} \triangleq \alpha_k \boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1}). \tag{42}$$

Using now (32), (33) and (41) in the three steps (20)–(22) from Algorithm 1, the ACTC recursion can be recast in the form:

$$\begin{cases} \widetilde{\boldsymbol{\psi}}_{k,i} = (I_M - \mu\boldsymbol{H}_{k,i-1})\widetilde{\boldsymbol{w}}_{k,i-1} - \mu\,\boldsymbol{s}_{k,i} - \mu\,b_k \\ \widetilde{\boldsymbol{q}}_{k,i} = \widetilde{\boldsymbol{q}}_{k,i-1} + \zeta\,\boldsymbol{Q}_k(\widetilde{\boldsymbol{\psi}}_{k,i} - \widetilde{\boldsymbol{q}}_{k,i-1}) \\ \widetilde{\boldsymbol{w}}_{k,i} = \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\widetilde{\boldsymbol{q}}_{\ell,i} \end{cases} \tag{43}$$

We remark that in the second equation of (43), the argument of the compression function is expressed in terms of centered variables by adding and subtracting $w^\star$.

In order to assess the goodness of an individual agent's estimate $\boldsymbol{w}_{k,i}$ we will focus on the mean-square-deviation:

$$\mathbb{E}\|\boldsymbol{w}_{k,i} - w^\star\|^2 = \mathbb{E}\|\widetilde{\boldsymbol{w}}_{k,i}\|^2. \tag{44}$$

Since the ACTC iterates $\widetilde{\boldsymbol{w}}_{k,i}$ are convex combinations of the quantized iterates $\{\widetilde{\boldsymbol{q}}_{k,i}\}$, the characterization of the mean-square behavior of $\widetilde{\boldsymbol{q}}_{k,i}$ will enable immediate characterization of the mean-square behavior of $\widetilde{\boldsymbol{w}}_{k,i}$. It is therefore convenient to incorporate the third step of the ACTC strategy into the first step, and focus on the behavior of $\widetilde{\boldsymbol{q}}_{k,i}$, obtaining:

$$\begin{cases} \widetilde{\boldsymbol{\psi}}_{k,i} = (I_M - \mu\boldsymbol{H}_{k,i-1})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\widetilde{\boldsymbol{q}}_{\ell,i} - \mu\,\boldsymbol{s}_{k,i} - \mu\,b_k \\ \widetilde{\boldsymbol{q}}_{k,i} = \widetilde{\boldsymbol{q}}_{k,i-1} + \zeta\,\boldsymbol{Q}_k(\widetilde{\boldsymbol{\psi}}_{k,i} - \widetilde{\boldsymbol{q}}_{k,i-1}) \end{cases} \tag{45}$$

Moreover, it is convenient to introduce the difference variable:

$$\boldsymbol{\delta}_{k,i} \triangleq \widetilde{\boldsymbol{\psi}}_{k,i} - \widetilde{\boldsymbol{q}}_{k,i-1} = \boldsymbol{\psi}_{k,i} - \boldsymbol{q}_{k,i-1}. \tag{46}$$

Subtracting $\widetilde{\boldsymbol{q}}_{k,i-1}$ from the first equation in (45), we finally obtain:

$$\begin{cases} \boldsymbol{\delta}_{k,i} = (I_M - \mu\boldsymbol{H}_{k,i-1})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\widetilde{\boldsymbol{q}}_{\ell,i} \\ \qquad\qquad - \widetilde{\boldsymbol{q}}_{k,i-1} - \mu\boldsymbol{s}_{k,i} - \mu\,b_k \\ \widetilde{\boldsymbol{q}}_{k,i} = \widetilde{\boldsymbol{q}}_{k,i-1} + \zeta\,\boldsymbol{Q}_k(\boldsymbol{\delta}_{k,i}) \end{cases} \tag{47}$$

A. RECURSIONS IN EXTENDED FORM

Since we are interested in a network-oriented analysis, it is useful to introduce a notation where the N agents' vectors of size $M \times 1$ are stacked into the following $MN \times 1$ vectors:

$$\boldsymbol{\delta}_i \triangleq \begin{bmatrix} \boldsymbol{\delta}_{1,i} \\ \boldsymbol{\delta}_{2,i} \\ \vdots \\ \boldsymbol{\delta}_{N,i} \end{bmatrix},\ \widetilde{\boldsymbol{q}}_i \triangleq \begin{bmatrix} \widetilde{\boldsymbol{q}}_{1,i} \\ \widetilde{\boldsymbol{q}}_{2,i} \\ \vdots \\ \widetilde{\boldsymbol{q}}_{N,i} \end{bmatrix},\ \boldsymbol{s}_i \triangleq \begin{bmatrix} \boldsymbol{s}_{1,i} \\ \boldsymbol{s}_{2,i} \\ \vdots \\ \boldsymbol{s}_{N,i} \end{bmatrix},\ b \triangleq \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}. \tag{48}$$

Likewise, it is useful to consider the extended compression operator $\boldsymbol{\mathcal{Q}}(\cdot)$ that applies the compression operation to each $M \times 1$ block of its input, and stacks the results as follows:

$$\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i) \triangleq \begin{bmatrix} \boldsymbol{Q}_1(\boldsymbol{\delta}_{1,i}) \\ \boldsymbol{Q}_2(\boldsymbol{\delta}_{2,i}) \\ \vdots \\ \boldsymbol{Q}_N(\boldsymbol{\delta}_{N,i}) \end{bmatrix}. \tag{49}$$

Finally, in order to express the recursion (43) in terms of the joint evolution of the extended vectors we need to introduce the extended matrices:

$$\mathcal{A} \triangleq A \otimes I_M,\ \ \boldsymbol{\mathcal{H}}_{i-1} \triangleq \mathsf{diag}\{\boldsymbol{H}_{1,i-1}, \boldsymbol{H}_{2,i-1}, ..., \boldsymbol{H}_{N,i-1}\}, \tag{50}$$

where $\otimes$ denotes the Kronecker product. It is now possible to describe compactly the ACTC strategy of the individual agents in (43) as:

$$\begin{cases} \boldsymbol{\delta}_i = [(I_{MN} - \mu\,\boldsymbol{\mathcal{H}}_{i-1})\mathcal{A}^\top - I_{MN}]\widetilde{\boldsymbol{q}}_{i-1} - \mu\boldsymbol{s}_i - \mu\,b \\ \widetilde{\boldsymbol{q}}_i = \widetilde{\boldsymbol{q}}_{i-1} + \zeta\,\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i) \end{cases} \tag{51}$$

B. NETWORK COORDINATE TRANSFORMATION

By means of a proper linear transformation of the network evolution it is possible to separate the two fundamental mechanisms that characterize the learning behavior of the ACTC

strategy. The first mechanism characterizes the *coordinated evolution* enabled by the social learning phenomenon. This is a desired behavior that will be critical to let each individual agent agree and converge to a small neighborhood of the global minimizer $w^\star$. In comparison, the second mechanism represents the departure of the agents' evolution from the coordinated evolution, which arises from the distributed nature of the system (agents need some time to reach agreement through successive local exchange of information). We will see later how the ACTC diffusion strategy blends these two mechanisms so as to achieve successful learning.

The enabling tool to develop the aforementioned network transformation is the Jordan canonical decomposition of the (transposed) combination matrix A [66]:[1]

$$A^\top \triangleq V^{-1} J_{\rm tot} V. \tag{52}$$

The matrix $J_{\rm tot}$ is the Jordan matrix of $A^\top$ (equivalently, of A, since A and $A^\top$ are similar matrices [66]) arranged in canonical form, i.e., made up of the Jordan blocks corresponding to the eigenvalues of A as detailed in Appendix A [66]. In particular, the single[2] eigenvalue equal to 1 corresponds to a 1×1 block, and the other Jordan blocks can be collected into the $(N-1)\times(N-1)$ reduced Jordan matrix J — see (108):

$$J_{\rm tot} = \begin{bmatrix} 1 & 0 \\ 0 & J \end{bmatrix}. \tag{53}$$

The columns of V^{-1} collect the generalized right-eigenvectors of $A^\top$ (hence, generalized left-eigenvectors of A) [67], [68]. Likewise, the rows of V collect the generalized left-eigenvectors of $A^\top$. Recalling that π is a right-eigenvector of A, and $\mathbb{1}_N$ a left-eigenvector of A, the matrices V and V^{-1} can be conveniently block-partitioned as follows:

$$V = \begin{bmatrix} \pi^\top \\ V_R \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} \mathbb{1}_N & V_L \end{bmatrix}, \tag{54}$$

where the subscript R is associated with generalized right-eigenvectors of A. The same applies to subscript L as regards generalized left-eigenvectors.

We are now ready to detail the network coordinate transformation relevant to our analysis. To this end, we introduce the extended matrix:

$$\mathcal{V} \triangleq V \otimes I_M\,, \tag{55}$$

and the *transformed* extended vector $\widehat{\boldsymbol{q}}_i$:

$$\widehat{\boldsymbol{q}}_i \triangleq \mathcal{V}\widetilde{\boldsymbol{q}}_i = \begin{bmatrix} (\pi^\top \otimes I_M)\widetilde{\boldsymbol{q}}_i \\ (V_R \otimes I_M)\widetilde{\boldsymbol{q}}_i \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{q}}_i \\ \check{\boldsymbol{q}}_i \end{bmatrix}. \tag{56}$$

As we see, the transformed vector $\widehat{\boldsymbol{q}}_i$ has been partitioned in two blocks: an $M\times 1$ vector $\bar{\boldsymbol{q}}_i$ and an $M(N-1)\times 1$ vector $\check{\boldsymbol{q}}_i$. These two vectors admit a useful physical interpretation. Vector $\bar{\boldsymbol{q}}_i$ is a linear combination, through the Perron weights, of the N vectors composing $\widetilde{\boldsymbol{q}}_i$, i.e., of the vectors corresponding to all agents. As we will see, such combination reflects a "coordinated" evolution that will apply to all agents. In contrast, vector $\check{\boldsymbol{q}}_i$ is representative of the departure of the individual agent's behavior from the coordinated behavior. In the following, we will sometimes refer to $\bar{\boldsymbol{q}}_i$ as the *coordinated-evolution component*, and to $\check{\boldsymbol{q}}_i$ as the *network-error component*. It is worth noticing that the coordinated-evolution component $\bar{\boldsymbol{q}}_i$ is real-valued, whereas the network-error component $\check{\boldsymbol{q}}_i$ is in general *complex-valued*, since the matrix V_R can contain complex-valued eigenvectors.

In order to highlight the role of the aforementioned two components on the individual agents, we can apply the inverse transformation $\mathcal{V}^{-1}$ to $\widehat{\boldsymbol{q}}_i$, obtaining:

$$\widetilde{\boldsymbol{q}}_{k,i} = \bar{\boldsymbol{q}}_i + \underbrace{([V_L]_k \otimes I_M)}_{\triangleq \mathcal{T}_k}\check{\boldsymbol{q}}_i, \tag{57}$$

where $[V_L]_k$ denotes the k-th row of matrix V_L. Equation (57) reveals that agent k progresses over time by combining the coordinated evolution $\bar{\boldsymbol{q}}_i$ (which is equal for *all* agents) and a perturbation vector $\mathcal{T}_k\check{\boldsymbol{q}}_i$, which quantifies the specific discrepancy of agent k (since matrix $\mathcal{T}_k$ depends on k) from the coordinated behavior. From the distributed optimization perspective, the goal is to reach agreement among all agents and, hence, it is necessary that the perturbation term is washed out as time elapses, letting all agents converge to the same coordinated behavior. Establishing that this is the case will be the main focus of our analysis.

For later use, it is convenient to introduce also the transformed versions of $\boldsymbol{\delta}_i$, $\boldsymbol{s}_i$ and b:

$$\widehat{\boldsymbol{\delta}}_i \triangleq \mathcal{V}\boldsymbol{\delta}_i = \begin{bmatrix} (\pi^\top \otimes I_M)\boldsymbol{\delta}_i \\ (V_R \otimes I_M)\boldsymbol{\delta}_i \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\delta}}_i \\ \check{\boldsymbol{\delta}}_i \end{bmatrix}, \tag{58}$$

$$\widehat{\boldsymbol{s}}_i \triangleq \mathcal{V}\boldsymbol{s}_i = \begin{bmatrix} (\pi^\top \otimes I_M)\boldsymbol{s}_i \\ (V_R \otimes I_M)\boldsymbol{s}_i \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{s}}_i \\ \check{\boldsymbol{s}}_i \end{bmatrix}, \tag{59}$$

$$\widehat{b} \triangleq \mathcal{V}b = \begin{bmatrix} (\pi^\top \otimes I_M)b \\ (V_R \otimes I_M)b \end{bmatrix} = \begin{bmatrix} 0 \\ \check{b} \end{bmatrix}. \tag{60}$$

The zero entry in the transformed bias vector arises from the fact that, in view of (39), we have:

$$(\pi^\top \otimes I_M)b = \sum_{k=1}^{N} \pi_k b_k = \sum_{k=1}^{N} p_k\,\nabla J_k(w^\star) = 0, \tag{61}$$

since the Perron-weighted sum of the gradients computed at the limit point $w^\star$ corresponds to the *exact* minimizer of the global cost function in (12).

[1] More often, the similarity transform in (52) is written as $VJ_{\rm tot}V^{-1}$. We opt for the alternative form in (52) since, as we will see soon, with this choice the matrix V will correspond to the *direct* network coordinate transformation that will be relevant to our treatment (and its inverse V^{-1} will be the corresponding *inverse* transformation).

[2] The existence of a single (and maximum magnitude) eigenvalue equal to 1 is guaranteed by Assumptions 2 and 3.

VI. QUALITATIVE ILLUSTRATION

The learning behavior of the ACTC strategy will be characterized by studying the time evolution of the transformed vector $\widehat{\boldsymbol{q}}_i$ introduced in (56). In particular, since we are interested in mean-square-errors, it is particularly convenient to work in terms of the average energy operator defined in Appendix A2 — see (114). Applying the energy operator to $\widehat{\boldsymbol{q}}_i$, we obtain the average energy vector:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] = \begin{bmatrix} \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 \\ \mathscr{P}[\check{\boldsymbol{q}}_i] \end{bmatrix}, \tag{62}$$

where $\mathscr{P}[\check{\boldsymbol{q}}_i]$ is an $(N-1)\times 1$ vector. Theorem 3 reported in Appendix B establishes that the energy vectors evolve according to the following inequality recursion:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq T\ \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + x. \tag{63}$$

In (63):

- T is a matrix that admits the following representation:

$$T \preceq \underbrace{\begin{bmatrix} 1-\mu\,\zeta\,\nu & O(1)\mathbb{1}_{N-1}^{\top} \\ 0 & E \end{bmatrix}}_{T_0} + O(\mu^2)\,\mathbb{1}_N\mathbb{1}_N^{\top}, \tag{64}$$

  where E is a matrix independent of μ, introduced in (153).
- x is a vector with entries proportional to μ^2, which are determined by the three sources of error present in the system, namely, the gradient noise, the bias, and the compression error — see (155) for a detailed description.

By inspection of (64), we see that the transfer matrix T can be written as the sum of an upper-diagonal matrix T_0 and a rank-one perturbation of order μ^2. It is tempting to conclude from (64) that the rank-one perturbation can be neglected as $\mu \to 0$. Were the matrix T_0 independent of μ, this conclusion would be obvious. However, since T_0 does depend upon μ, proving that for small μ the recursion in (63) can be examined by replacing T with T_0 is not necessarily true. We will be able to show that this is actually the case, but to this end we need to carry out the demanding technical analysis reported in the appendices.

Nevertheless, to gain insight on how recursion (63) is relevant to the mean-square behavior of the ACTC strategy, let us simply assume for now that we can replace T with T_0 and examine instead the recursion:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq T_0\ \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + x. \tag{65}$$

The upper-diagonal structure of T_0 implies that the evolution relating the network error components $\mathscr{P}[\check{\boldsymbol{q}}_{i-1}]$ to $\mathscr{P}[\check{\boldsymbol{q}}_i]$ takes place only through matrix E. Accordingly, such matrix will be referred to as the network error matrix. In particular, matrix E is critical to characterize the *network transient*, which is the phase of the learning process necessary for the agents to reach a coordinated evolution that drives them to converge to a small neighborhood of the true minimizer $w^\star$.

By developing the inequality recursion (65), we would arrive at the following inequality:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq T_0^i\ \mathscr{P}[\widehat{\boldsymbol{q}}_0] + \sum_{j=0}^{i-1} T_0^j x. \tag{66}$$

Assume that T_0 is stable. Then, from (66) we would have:

$$\begin{aligned} \limsup_{i\to\infty} \mathscr{P}[\widehat{\boldsymbol{q}}_i] &= \limsup_{i\to\infty} \begin{bmatrix} \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 \\ \mathscr{P}[\check{\boldsymbol{q}}_i] \end{bmatrix} \\ &\preceq (I-T_0)^{-1}x \preceq \begin{bmatrix} O(1/\mu) & O(1/\mu)\,\mathbb{1}_{N-1}^{\top} \\ 0 & O(1)\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^{\top} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \check{x} \end{bmatrix}, \end{aligned} \tag{67}$$

where we exploited the upper triangular shape of T_0 from (64) to evaluate the inverse $(I-T_0)^{-1}$ [66]. Since all the entries of x scale as μ^2, from (67) we obtain the following important conclusion regarding the mean-square stability of the ACTC strategy. The limit superior of $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2$ scales as $O(\mu)$, whereas the limit superior of $\mathscr{P}[\check{\boldsymbol{q}}_i]$ is a higher-order correction scaling as $O(\mu^2)$. In other words, for sufficiently small step-sizes the learning behavior of the ACTC strategy is dominated by the coordinated-evolution component, while the network-error component becomes negligible.

Further insights can be obtained by contrasting (67) with the behavior of the classical ATC diffusion strategy, which operates *without compression*. In this case, the (1,2)-block of $(I-T_0)^{-1}$ is $O(1)$ in place of $O(1/\mu)$ [14], [21]. This implies that for the ATC strategy the network-error component $\check{x}$ induces on the coordinated-evolution term only a higher-order $O(\mu^2)$ correction. Therefore, the term $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2$ is influenced only by the component $\bar{x}$. Moreover, for the ATC strategy the component $\bar{x}$ is the same gradient-noise term that would characterize a centralized evolution. In contrast, the component $\check{x}$ is a network-error term arising from propagation of the gradient noise across the network.

Let us see what happens instead for the ACTC strategy. Examining (67), we see that the term $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2$ is influenced by both components $\bar{x}$ and $\check{x}$. As can be verified by (155) and Table 2, in addition to the gradient noise terms, these components contain the bias and the compression parameters. We conclude that, for the ACTC strategy, backpropagation of the compression error lets additional components of the gradient noise and the bias seep into the ACTC evolution, determining an increase of the mean-square-deviation. The qualitative arguments illustrated in this section will be made rigorous in the theorems presented in the next two sections and in the pertinent appendices.

VII. MEAN-SQUARE STABILITY

In order to state the theorem establishing mean-square stability, it is necessary to introduce a useful function that will be critical to evaluate the mean-square stability of the ACTC strategy. Following the canonical Jordan decomposition illustrated in Appendix A, we denote by λ_n the eigenvalue of A associated with the n-th Jordan block of A (with the eigenvalues being ordered in descending order of magnitude — see

TABLE 1. Useful Symbols Used Throughout the Article.

Step-sizes μ_k	$\mu = \max_{k=1,2,\ldots,N} \mu_k$, $\quad \alpha_k = \mu_k/\mu$, $\quad C_\alpha = \mathsf{diag}(\alpha_1^2, \alpha_2^2, \ldots, \alpha_N^2)$
Combination matrix A	π is the Perron eigenvector, $A\pi = \pi$, $\quad p = [\alpha_1\pi_1, \alpha_2\pi_2, \ldots, \alpha_N\pi_N]^\top$ Jordan form: $A^\top = V^{-1} J_{\rm tot} V$, $\quad V = \begin{bmatrix} \pi^\top \\ V_R \end{bmatrix}$, $\quad \lambda_2$ is the 2nd largest magnitude eigenvalue of A V_{2R} is the $(N-1)\times N$ matrix with (ℓ,k)-entry equal to the squared magnitude of the corresponding entry in V_R
Convexity/Lipschitz	ν is the global-strong-convexity constant in (13), $\quad \sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}$ are the constants in Lemma 1
Gradient noise	Bounding constants at agent k: $\{\beta_k\}$ and $\{\sigma_k^2\}$ — see (35) $C_\beta = \mathsf{diag}(\beta_1^2, \beta_2^2, \ldots, \beta_N^2)$, $\quad C_\sigma = \mathsf{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$
Bias $b_k = \alpha_k \nabla J_k(w^\star)$	$\breve{b} = (V_R \otimes I_M) b$, $\quad \mathscr{P}[\breve{b}] = \left[\|\breve{b}_1\|^2, \|\breve{b}_2\|^2, \ldots, \|\breve{b}_{N-1}\|^2\right]^\top$ (the equality holds since the bias is not random)
Network error matrix	$J = \Lambda + U$ is the reduced Jordan matrix in (111), $\quad E_0 = \left((1-\zeta) I_{N-1} + \zeta \frac{\Lambda\Lambda^*}{\lvert\lambda_2\rvert}\right) + \frac{2\zeta}{1-\lvert\lambda_2\rvert} U$
Compression parameters	ω_k is the compression factor of agent k, $\quad \bar{\Delta} = \|V^{-1}\|^2 \max_{k=1,2,\ldots,N} \pi_k^2 \omega_k$, $\quad \breve{\Delta} = \|V^{-1}\|^2 \max_{\substack{\ell=2,3,\ldots,N \\ k=1,2,\ldots,N}} \lvert v_{\ell k}\rvert^2 \omega_k$

TABLE 2. Useful Matrices and Vectors Appearing in Lemmas 2, 3, and 4. The Symbols Used in the Present Table are Defined in Table 1.

$[T_s]_{11} = \mu^2 \|V^{-1}\|^2 \sum_{\ell=1}^N \pi_\ell \alpha_\ell^2 \beta_\ell^2$	$[T_\delta]_{11} = 2\mu^2 \sigma_{11}^2$	$[T_q]_{11} = 1 - \mu\zeta\nu$
$[T_s]_{12} = \mu^2 \|V^{-1}\|^2 \sum_{\ell=1}^N \pi_\ell \alpha_\ell^2 \beta_\ell^2 \mathbb{1}_{N-1}^\top$	$[T_\delta]_{12} = 2\mu^2 \sigma_{12}^2 \mathbb{1}_{N-1}^\top$	$[T_q]_{12} = \mu\zeta \frac{\sigma_{12}^2}{\nu} \mathbb{1}_{N-1}^\top$
$[T_s]_{21} = \mu^2 N \|V^{-1}\|^2 V_{2R} C_\alpha C_\beta \mathbb{1}_N$	$[T_\delta]_{21} = 8\mu^2 \sigma_{21}^2 \mathbb{1}_{N-1}$	$[T_q]_{21} = \mu^2\zeta \frac{6\sigma_{21}^2}{1-\lvert\lambda_2\rvert} \mathbb{1}_{N-1}$
$[T_s]_{22} = \mu^2 N \|V^{-1}\|^2 V_{2R} C_\alpha C_\beta \mathbb{1}_N \mathbb{1}_{N-1}^\top$	$[T_\delta]_{22} = 8(I_{N-1} + U) + 8\mu^2 \sigma_{22}^2 \mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top$	$[T_q]_{22} = E_0 + \mu^2\zeta \frac{6\sigma_{22}^2 \mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top}{1-\lvert\lambda_2\rvert}$
$\bar{x}_s = \mu^2 \sum_{\ell=1}^N \pi_\ell \alpha_\ell^2 \sigma_\ell^2$	$\bar{x}_\delta = 0$	$\bar{x}_q = 0$
$\breve{x}_s = \mu^2 N V_{2R} C_\alpha C_\sigma \mathbb{1}_N$	$\breve{x}_\delta = 8\mu^2 \mathscr{P}[\breve{b}]$	$\breve{x}_q = \mu^2\zeta \frac{6}{1-\lvert\lambda_2\rvert} \mathscr{P}[\breve{b}]$

(105) in Appendix A), by L_n the dimension of this block, and by B the number of blocks. Let

$$a_n \triangleq \frac{2|\lambda_2|}{1-|\lambda_2|} \frac{1}{|\lambda_2| - |\lambda_n|^2}, \qquad 2 \le n \le B. \tag{68}$$

We note that $a_n \ge 2$ and introduce the function:

$$\gamma(A) \triangleq \sum_{n=2}^{B} \frac{|\lambda_2|}{|\lambda_2| - |\lambda_n|^2} \left(\frac{a_n^{L_n+1} - 1}{(a_n - 1)^2} - \frac{L_n + 1}{a_n - 1} \right). \tag{69}$$

Theorem 1 (Mean-Square Stability): Let

$$\zeta < \frac{1}{16\breve{\Delta}\gamma(A)}, \tag{70}$$

and let

$$\mu < \frac{2}{\zeta(\eta + \nu)}, \quad \mu < \mu^\star, \tag{71}$$

with $\mu^\star$ being the positive[3] root of the equation:

$$\mu^2 \left(1 + \frac{\sigma_{12}^2}{\nu^2}\right) \varphi + \mu\zeta \frac{1 + 16\bar{\Delta}\varphi}{\nu} - \frac{1}{\phi} = 0, \tag{72}$$

where

$$\varphi \triangleq \frac{\gamma(A)}{1 - 16\zeta\,\breve{\Delta}\,\gamma(A)}, \tag{73}$$

$\bar{\Delta}$ and $\breve{\Delta}$ are defined in Table 1, and ϕ is a positive scalar that embodies the constants appearing in the μ^2-terms of matrices T_s, T_δ and T_q in Table 2. The evaluation of ϕ is rather cumbersome and is detailed in Appendix F. Then the ACTC strategy is mean-square stable, namely,

$$\limsup_{i\to\infty} \mathbb{E}\|\boldsymbol{w}_{k,i} - w^\star\|^2 < \infty. \tag{74}$$

[3]Equation (72) is in the form $a\mu^2 + b\mu - c = 0$, where a, b and c are positive. Thus, the equation admits two real-valued roots of opposite sign.

Moreover, in the small step-size regime the mean-square-deviation is of order μ, namely,

$$\limsup_{i\to\infty} \mathbb{E}\|\boldsymbol{w}_{k,i} - w^{\star}\|^2 = O(\mu) \text{ as } \mu \to 0. \tag{75}$$

■

Proof: See Appendix I.

A. INSIGHTS FROM THEOREM 1

Conditions (70) and (71) reveal how the parameter ζ and the step-size μ are used to control the mean-square stability of the ACTC algorithm. Regarding ζ, this parameter is not present (i.e., $\zeta = 1$) in the classical ATC algorithm, which operates without compression. However, Eq. (70) shows that in the presence of compression the value $\zeta = 1$ does not grant stability in general. In particular, Eq. (70) pertains to the stability of the matrix E introduced in (153) (see Lemma 6 in Appendix G), whereas Eq. (71) controls the stability of the matrix T introduced in (151) (see Lemma 7 in Appendix H).

Examining the structure of $\gamma(A)$ in (69), we see that the RHS of (70) depends on two main elements, namely, *i*) a constant $\breve{\Delta}$ that contains the compression parameters ω_k; *ii*) the eigenstructure of A, and particularly the second largest magnitude eigenvalue λ_2. Let us examine in detail the role played by each of these elements.

Regarding the quantization constant $\breve{\Delta}$, we see that poorer resolutions (i.e., higher values of $\breve{\Delta}$) go against stability, an effect that can be compensated by choosing smaller values for ζ. This means that ζ is useful to compensate for the quantization error that seeps into the recursion of the individual agent's errors.

Regarding the eigenstructure of A, its fundamental role for mean-square stability is summarized by the function $\gamma(A)$ in (69). This function provides an accurate stability threshold by capturing the full eigenstructure of A through the eigenvalues λ_n, the size and the number of Jordan blocks. In this way, we are given the flexibility of providing accurate stability thresholds for different types of combination matrices. Let us illustrate these useful features in relation to existing results.

A detailed stability analysis of the classical (i.e., uncompressed) ATC diffusion strategy under the general assumptions considered in this work was originally carried out in [14], [21]. Such analysis relies on a more general Jordan decomposition (with arbitrary ϵ replacing the ones on the superdiagonal) which provides a sufficient condition for mean-square stability. One feature of this sufficient condition is that the stability range for μ scales exponentially as ϵ^{-N}, with $\epsilon < 1$, a condition that becomes stringent for large-scale networks.

The stability analysis conducted in this work is different. Our analysis is conducted in the transformed (complex) z-domain exploiting the formalism of resolvent matrices. This turns out to be a powerful approach that allows us to get a necessary and sufficient condition for mean-square stability that accounts for the entire eigenspectrum of A through the function $\gamma(A)$ in (69). We will now explain how such eigenstructure plays a role for different types of combination matrices, and how the actual conditions for stability are in fact milder than the aforementioned exponential scaling.

Diagonalizable Matrices: For simplicity, let us consider the case that, excluding $\lambda_1 = 1$, all the remaining eigenvalues are equal, namely, $\lambda_n = \lambda_2$ for $n > 1$. Note that this yields $a_n = const.$ in (68). If A is diagonalizable, we have $B = N$ and $L_n = 1$ for all $n > 1$, which, using (69), yields:

$$\gamma(A) \propto N, \tag{76}$$

i.e., $\gamma(A)$ scales linearly with N.[4]

"Very" Non-Diagonalizable Matrices: Consider the opposite case where $B = 2$ and $L_n = N - 1$, namely, apart from the first Jordan block (i.e., the one associated with the single eigenvalue $\lambda_1 = 1$), we have only another block associated with $\lambda_n = \lambda_2$. Under this setting, we see from (69) that $\gamma(A)$ grows exponentially with N as:

$$\gamma(A) \sim \left(\frac{2}{(1-|\lambda_2|)^2}\right)^N. \tag{77}$$

Typical Non-Diagonalizable Matrices: The exponential scaling observed in (77) is clearly not desirable for stability, since, in light of (70), it would significantly reduce the stability range.

However, typical non-diagonalizable matrices adopted in distributed optimization applications seldom feature the extreme eigenstructure described above. As a matter of fact, if we perform the Jordan decomposition of typical combination matrices, we see that the size of the Jordan blocks is usually modest, and in any case it does not increase linearly with N. This means that the exponents in (69) would be determined by the maximum Jordan block size, and not by the network size.

We note also that larger values of $|\lambda_n|$, and particularly of $|\lambda_2|$, go against the stability of the network error matrix E, which means that ζ is useful to regulate the stability when the network component evolves more slowly, i.e., when $|\lambda_2|$ is closer to 1. In the latter case, the threshold in (77) is particularly penalized. However, taking into account the fact that for typical combination matrices many eigenvalue magnitudes are considerably smaller than the second largest magnitude, the stability thresholds obtained through (69) are expected to be significantly less restrictive than (77).

In summary, the network error convergence depends upon the eigenstructure of A and the quantizer's resolution. While the role of the eigenstructure of A (and, hence, of the network connectivity) is common to the standard ATC strategy, one distinguishing feature of the ACTC strategy is represented by the fact that the spectral radius of E increases due to the presence of the additional factor $\breve{\Delta}$. This means that the agreement among agents slows down due to backpropagation of the

[4]We remark that, in the case of diagonalizable A, the linear scaling is a tight estimate, since a known result about rank-one perturbations of diagonal matrices allows us to evaluate the spectral radius in an exact manner as the sum of the spectral radius of the unperturbed matrix plus N times the size of the perturbation [69].

quantization error. However, and remarkably, this slowdown does not preclude the possibility of accurate performance given sufficient time for learning. Paralleling classical coding results from Shannon's theory, we could say that *the price of compression is not the impossibility of learning, rather a slowdown in the convergence*.

VIII. TRANSIENT ANALYSIS

The next result refines the mean-square stability result in Theorem 1 by characterizing the learning dynamics of the ACTC strategy, i.e., the transient phases of the algorithm before it converges to the steady state. Preliminarily, we recall the main differences, in terms of convergence behavior, between diminishing and constant step-sizes for traditional (i.e., without compression) stochastic gradient algorithms.

Under stationary conditions, algorithms with diminishing step-size converge to the correct minimizer as the number of iterations i grows. In particular, the mean-square-error converges to zero at the rate $O(1/i)$. In comparison, under stationary conditions, algorithms with constant step-size converge to a small *steady-state* error on the order of the step-size. The mean-square-error reaches the steady-state error at a geometric rate ρ^i for some positive $\rho < 1$. In Section II-D we have already explained why, under *nonstationary* conditions, tolerating a small error in place of a vanishing error is rewarding in terms of adaptation.

In the next theorem we provide a convergence analysis for the ACTC strategy by obtaining an upper bound on the mean-square-error of each individual agent, which holds for any finite number of iterations i. The analysis reveals that, despite the *compression* constraint, the steady-state error is reached at a *geometric* rate. Specifically, we establish that the convergence is ruled by the superposition of two terms that vanish geometrically with i at different rates, highlighting the presence of two transient phases.

Theorem 2 (ACTC Learning Behavior): Let

$$\rho_{\text{cen}} \triangleq (1-\mu\,\zeta\,\nu)^2, \tag{78}$$

where ν is the global strong convexity constant in (13), and assume that

$$\zeta < \frac{1}{16\check{\Delta}\gamma(A)}. \tag{79}$$

Let $\rho(E)$ be the spectral radius of the matrix E in (153). Under assumption (79) we have $\rho(E) < 1$, and we set $\epsilon > 0$ such that:

$$\rho_{\text{net}} = \rho(E) + \epsilon < 1. \tag{80}$$

Using the definitions of the compression factor Ω in (25), of the entries $\{\pi_\ell\}$ of the Perron eigenvector in (10), of the scaled step-sizes $\{\alpha_\ell\}$ in (7), and of the gradient-noise variances $\{\sigma_\ell^2\}$ in (35), in the small-μ regime the evolution of the mean-square-deviation of the individual agent k can be cast in the form, for all $i > 0$:

$$\mathbb{E}\|\boldsymbol{w}_{k,i} - w^\star\|^2 \le O(1)\,\rho_{\text{net}}^i + O(1)\,\rho_{\text{cen}}^i + O(\mu)\,\rho_{\text{cen}}^{i/4} + \mu\,\zeta\left(\frac{\sum_{\ell=1}^N \pi_\ell\,\alpha_\ell^2\sigma_\ell^2}{2\nu} + c_q\,\Omega\,(1+\Omega)\right) + O(\mu^{3/2}), \tag{81}$$

where the Big-O terms depend in general on the particular agent k, except for the $O(1)$ term multiplying ρ_{cen}^i, and c_q is a positive constant independent of μ and i. ■

Proof: See Appendix J.

Thanks to (81), we arrive at a sharp and revealing description of the learning behavior of compressed diffusion strategies over adaptive networks. In fact, the individual terms arising in (81) admit the following physical interpretation:

$$\underbrace{O(1)\,\rho_{\text{net}}^i}_{\substack{\text{network convergence } \textbf{to} \\ \text{the centralized solution}}} + \underbrace{O(1)\,\rho_{\text{cen}}^i}_{\substack{\text{convergence } \textbf{of} \text{ the} \\ \text{centralized solution}}} + \underbrace{O(\mu)\,\rho_{\text{cen}}^{i/4}}_{\substack{\text{higher-order correction} \\ \text{relative to the transient phase}}}$$

$$+ \overbrace{\mu\,\zeta\left(\underbrace{\frac{\sum_{\ell=1}^N \pi_\ell\,\alpha_\ell^2\sigma_\ell^2}{2\nu}}_{\text{uncompressed ACTC}} + \underbrace{c_q\,\Omega\,(1+\Omega)}_{\text{compression loss}}\right)}^{\text{steady-state error}} + O(\mu^{3/2}), \tag{82}$$

which allows us to examine closely the distinct learning phases as detailed in the following remarks.

— *Transient Phases:* First, we notice that the network rate ρ_{net} depends only on the stability parameter ζ, and on the network connectivity properties through the eigenspectrum of the combination matrix A. As a result, we see that for sufficiently small μ we have that $\rho_{\text{cen}} > \rho_{\text{net}}$. Accordingly, for small step-sizes μ, the transient associated with the convergence of the network solution toward the centralized solution dies out earlier (Phase I). After this initial transient, a second transient dominates (Phase II), which is relative to the slower (since $\rho_{\text{cen}} > \rho_{\text{net}}$) process that characterizes the convergence of the centralized solution to the steady state. Remarkably, these two distinct phases of adaptive diffusion learning have already been identified in the context of adaptive learning over networks *without communication constraints* [14], [15].

— *Compression Loss:* We see from (81) that, after transient Phase II, the ACTC strategy reaches a *steady-state* mean-square-deviation that is upper bounded by the quantity:

$$\mu\,\zeta\left(\frac{\sum_{\ell=1}^N \pi_\ell\,\alpha_\ell^2\sigma_\ell^2}{2\nu} + c_q\,\Omega\,(1+\Omega)\right) + O(\mu^{3/2}). \tag{83}$$

Once we set a prescribed convergence rate ρ_{cen}, the product $\mu\,\zeta$ is determined from (78) as $\mu\,\zeta = (1-\rho_{\text{cen}}^{1/2})/\nu$, such that, using (83) and neglecting higher-order corrections, we can introduce the error:

$$\overline{\mathsf{MSD}}_{\text{ACTC}} \triangleq \frac{1-\rho_{\text{cen}}^{1/2}}{\nu}\left(\frac{\sum_{\ell=1}^N \pi_\ell\,\alpha_\ell^2\sigma_\ell^2}{2\nu} + c_q\,\Omega\,(1+\Omega)\right), \tag{84}$$

which provides an upper bound (in the regime of small step-sizes) on the steady-state mean-square-deviation for a given convergence rate ρ_{cen}. We see that $\overline{\mathsf{MSD}}_{\text{ACTC}}$ is composed of two main terms. The first term does not depend on the

amount of compression, and is proportional to an average over the Perron weights $\{\pi_\ell\}$ of the scaled gradient noise powers $\{\alpha_\ell^2\sigma_\ell^2\}$, further divided by the global strong convexity constant ν. The second term on the RHS of (84) is the *compression loss*, which is in fact an increasing function of the compression factor Ω. The limiting case $\Omega = 0$ corresponds to the setting *without compression*, i.e., to the ACTC diffusion strategy in Algorithm 1 where the compression operator is the identity operator, yielding, for small step-sizes, the *excess error* arising from compression:

$$\mathscr{E} \triangleq \overline{\mathsf{MSD}}_{\mathrm{ACTC}} - \overline{\mathsf{MSD}}_{\mathrm{unc.\,ACTC}} = \frac{1-\rho_{\mathrm{cen}}^{1/2}}{\nu}\, c_q\, \Omega\,(1+\Omega), \tag{85}$$

where we referred to the strategy with $\Omega = 0$ as to the *uncompressed* ACTC.

Moreover, from Algorithm 1 it is readily seen that the *uncompressed* ACTC strategy with $\zeta = 1$ coincides with the classical ATC strategy. This implies that Eq. (85) holds also for the ATC strategy, namely, we can write:

$$\overline{\mathsf{MSD}}_{\mathrm{ACTC}} - \overline{\mathsf{MSD}}_{\mathrm{ATC}} = \mathscr{E}. \tag{86}$$

— *Equation (86) for Relevant Compression Operators:* Choosing as compression operators the randomized quantizers examined in Appendix K1, we can obtain an explicit connection between the mean-square-deviation and the bit-rate. In fact, from (315) we can write:

$$\Omega\,(1+\Omega) \leq \frac{2M}{(2^{r_{\min}}-1)^2}, \tag{87}$$

where we denoted by $r_{\min}$ the minimum bit-rate across agents. From (85) and (87) we conclude that the increase in mean-square-deviation from the uncompressed to the standard ACTC is upper bounded as:

$$\mathscr{E} \leq \frac{1-\rho_{\mathrm{cen}}^{1/2}}{\nu}\,\frac{2\,c_q M}{(2^{r_{\min}}-1)^2}, \tag{88}$$

which reveals the following remarkable analogy with the fundamental laws of high-resolution quantization: for small step-sizes, the *excess of mean-square-deviation* due to quantization scales exponentially with the bit-rate as $2^{-2r_{\min}}$ [44].

Likewise, from (320) in Appendix K2, we know that for the randomized sparsifier the maximum compression parameter is given by:

$$\Omega = \frac{M}{S_{\min}} - 1, \tag{89}$$

where $S_{\min}$ is the minimum number of transmitted components across agents. Substituting (89) into (85) yields:

$$\mathscr{E} \leq \frac{1-\rho_{\mathrm{cen}}^{1/2}}{\nu}\, c_q \left(\frac{M}{S_{\min}}\right)^2. \tag{90}$$

We see from (90) that the excess error due to compression is given by the sparsification error $(M/S_{\min})^2$, which quantifies the discrepancy between the vector dimensionality M and the reduced dimensionality $S_{\min}$ corresponding to the (minimum) number of transmitted components.

IX. ILLUSTRATIVE EXAMPLES

The ACTC diffusion strategy can be applied to learning problems that can be solved using stochastic gradient approximations, such as regression problems involving quadratic risks or classification problems involving logistic risks. More broadly, the cost $J_k(w)$ can represent the expected loss computed by using estimation or decision functions arising from complex learning architectures, such as a multi-layer neural network, with the parameter w representing the neural network weights. For our technical results to hold, we assume that the cost functions fulfill classical smoothness/convexity conditions. As is typical in optimization theory, when these assumptions are not met, the learning algorithms offer different types of guarantees, for example, convergence of the gradients in place of convergence of the iterates — see, e.g., [51], [61].

To illustrate the behavior of the ACTC algorithm and to highlight a number of interesting phenomena in a structured manner, we find it convenient to focus on the following distributed regression problem. We consider N agents interconnected through a network satisfying Assumptions 2 and 3. At time i, agent k observes data $\boldsymbol{d}_{k,i} \in \mathbb{R}$ and regressors $\boldsymbol{u}_{k,i} \in \mathbb{R}^M$, which obey the linear regression model:

$$\boldsymbol{d}_{k,i} = \boldsymbol{u}_{k,i}^{\top} w^{\star} + \boldsymbol{v}_{k,i} \quad k = 1,\ldots,N, \tag{91}$$

where $w^{\star} \in \mathbb{R}^M$ is an unknown (deterministic) parameter vector and $\boldsymbol{v}_{k,i} \in \mathbb{R}$ acts as noise. We assume that processes $\{\boldsymbol{u}_{k,i}\}$ and $\{\boldsymbol{v}_{k,i}\}$ have mean equal to zero, are independent both over time and space (i.e., across the agents), with second-order statistics given by, respectively:

$$R_{u,k} = \mathbb{E}[\boldsymbol{u}_{k,i}\boldsymbol{u}_{k,i}^{\top}], \qquad \mathbb{E}\boldsymbol{v}_{k,i}^2 = \sigma_{v,k}^2. \tag{92}$$

The goal is to to estimate the unknown $w^{\star}$, which, by applying straightforward manipulations to (91), can be seen to obey the relationship:

$$r_{du,k} = R_{u,k} w^{\star}, \tag{93}$$

where $r_{du,k} = \mathbb{E}[\boldsymbol{d}_{k,i}\boldsymbol{u}_{k,i}]$. In principle, each agent could perform estimation of $w^{\star}$ by solving the optimization problem:

$$\min_{w\in\mathbb{R}^M} \mathbb{E}\left(\boldsymbol{d}_{k,i} - \boldsymbol{u}_{k,i}^{\top} w\right)^2, \tag{94}$$

which in turn corresponds to adopting the following quadratic loss and cost functions:

$$L_k(w;\{\boldsymbol{d}_{k,i},\boldsymbol{u}_{k,i}\}) = (\boldsymbol{d}_{k,i} - \boldsymbol{u}_{k,i}^{\top} w)^2, \tag{95}$$

$$J_k(w) = \mathbb{E}\left[L_k(w;\{\boldsymbol{d}_{k,i},\boldsymbol{u}_{k,i}\})\right]. \tag{96}$$

There are several reasons why the agents can be interested in solving the regression problem in a cooperative fashion. First of all, it was shown in [21] that, under suitable design, cooperation is beneficial in terms of inference performance. Even more remarkably, in many cases the local regression problem (94) can be ill-posed if the agents' regressors do

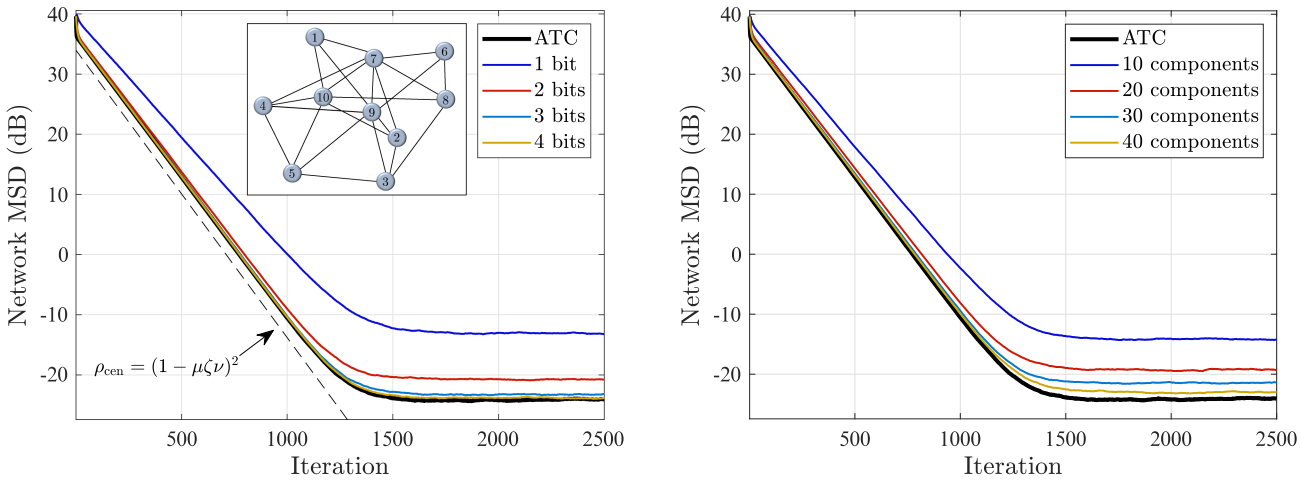

FIGURE 2. ***Left plot.*** **ACTC network mean-square-deviation in (99) as a function of the iteration i, for different values of the bit-rate r. We consider the distributed regression problem of Section IX, with dimensionality $M = 50$, Gaussian regressors $u_{k,i}$ with diagonal matrices and variances drawn as independent realizations from a uniform distribution in (1, 4), and Gaussian disturbances $v_{k,i}$ with variances drawn as independent realizations from a uniform distribution in (0.25, 1). The ACTC algorithm is run with stability parameter $\zeta = 0.25$, and with equal step-sizes $\mu_k = \mu = 4 \times 10^{-3}$. All agents use the randomized quantizer in Appendix K1 with bit-rate r. The inset plot shows the network topology, on top of which we build a Metropolis combination matrix to run the ACTC algorithm. All nodes have a self-loop, not shown in the figure. The mean-square-deviations are estimated by means of 10^2 Monte Carlo runs.** ***Right plot.*** **Same setting of the left plot, with the randomized sparsifier in Appendix K2 in place of the randomized quantizers.**

not contain sufficient information. This is an issue classically known as *collinearity*, which basically implies that the regression covariance matrix $R_{u,k}$ is singular, and many w exist that solve (94). This behavior can be easily grasped by noticing that

$$\nabla J_k(w) = 2(R_{u,k}w - r_{du,k}), \tag{97}$$

and, examining (93), we see that if $R_{u,k}$ is not invertible, $w^\star$ is a solution to the optimization problem, but not the only one. Accordingly, reliable inference about w is impaired by collinearity. Technically, when the regression covariance matrix of agent k is singular, the cost function $J_k(w)$ is *not* strongly convex, and the true $w^\star$ is *one among* the minimizers of (94).

However, if we now replace (94) with its global counterpart:

$$\min_{w\in\mathbb{R}^M} \mathbb{E}\left[\sum_{k=1}^{N} p_k \left(\boldsymbol{d}_{k,i} - \boldsymbol{u}_{k,i}^\top w\right)^2\right], \tag{98}$$

it is readily seen that a single agent with a non-singular $R_{u,k}$ (i.e., with a strongly convex cost function) is able to enable successful inference. Notably, such a minimal requirement is sufficient to our ACTC strategy to solve the regression problem in a *distributed* way and under *communication constraints*.

A. ROLE OF COMPRESSION DEGREE

In Fig. 2, we examine the learning performance of the ACTC strategy as a function of the iteration i. The left plot refers to the randomized quantizer described in Appendix K1, whereas the right plot refers to the randomized sparsifier in Appendix K2. In both plots, different curves refer to different communication budgets, namely, number of bits for the randomized quantizer and number of transmitted components for the randomized sparsifier.

The simulations were run under the following setting. The regression problem has dimensionality $M = 50$. The covariance matrices $R_{u,k}$ are all diagonal, and the associated regressors' variances are drawn as independent realizations from a uniform distribution with range (1, 4). The noise variances $\sigma^2_{v,k}$ are drawn as independent realizations from a uniform distribution with range (0.25, 1). The network is made of $N = 10$ agents, connected through the topology displayed in Fig. 2, and with Metropolis combination policy. Under this setting, we run the ACTC algorithm with stability parameter $\zeta = 0.25$, and with equal step-sizes $\mu_k = \mu = 4 \times 10^{-3}$. All errors are estimated by means of 10^2 Monte Carlo runs. For the example pertaining to randomized quantizers, all agents use the same number of bits r. For the example pertaining to randomized sparsifiers, all agents use the same number of transmitted components S.

As a first performance index, we examine the *network* ACTC learning performance, i.e., the mean-square-deviation averaged over all agents:

$$\frac{1}{N}\mathbb{E}\|\boldsymbol{w}_i - w^\star\|^2. \tag{99}$$

The behavior observed in Fig. 2 summarizes sharply the essential characteristics of the ACTC algorithm, as captured by Theorem 2: i) for all quantizer's resolutions (left plot) and for all levels of sparsity (right plot), the mean-square-deviation has a transient that is essentially governed by the predicted rate $\rho_{\text{cen}} = (1 - \mu\,\zeta\,\nu)^2$ (dashed line); ii) some higher-order discrepancies are absorbed in an initial (much

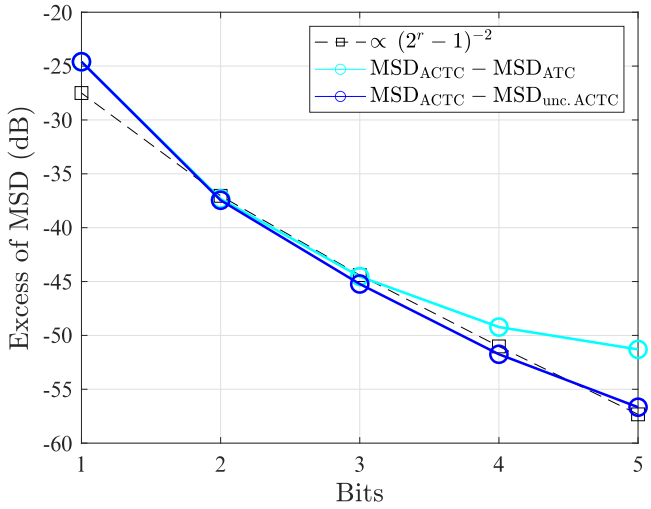

FIGURE 3. Difference between the network mean-square-deviation of the ACTC strategy and the mean-square-deviation of the uncompressed strategies, estimated over 5×10^3 Monte Carlo runs. The label "uncompressed ACTC" stems for the ACTC strategy in Algorithm 1 with $\mathcal{Q}(x) = x$, whereas "ATC" stems for the classical ATC in [21]. The dashed curve is obtained by depicting a curve proportional to $(2^r - 1)^{-2}$, with proportionality constant set so as to match the second point of the uncompressed ACTC curve. The relevant parameters of the ACTC strategy and of the distributed regression problem are set as in the left plot in Fig. 2, but for the dimensionality $M = 10$.

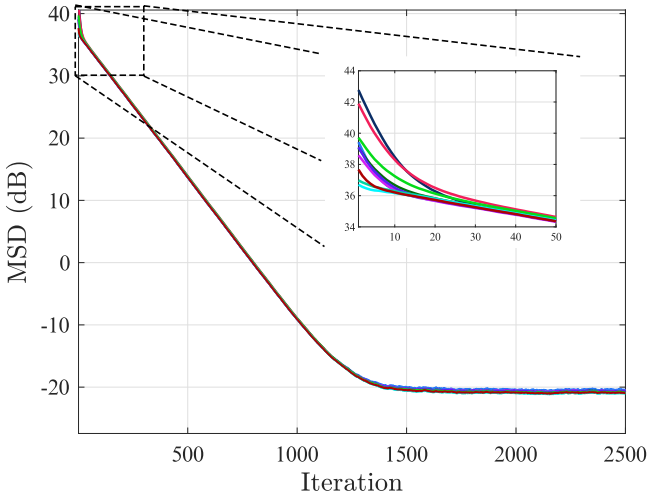

FIGURE 4. ACTC mean-square-deviation of the *individual* agents achieved using $r = 2$ bits, under the same setting of the left plot in Fig. 2. The inset plot zooms in on the faster initial transient needed by the agents to reach a coordinated evolution.

faster) transient; *iii*) the ACTC errors corresponding to different communication budgets (number of bits for the quantizers and number of components for the sparsifiers) converge to different steady-state error values, approaching the performance of the ATC (i.e., uncompressed) strategy as the number of bits or components increases.

With reference to the same setting of the left plot in Fig. 2, but for a smaller dimensionality $M = 10$, in Fig. 3 we display the *excess of mean-square-deviation* of the ACTC strategy with respect to the uncompressed ACTC. We remark that the curves in Fig. 3 *do not represent mean-square-deviations*, but the difference between the mean-square-deviation attained by the ACTC diffusion strategy, and the mean-square-deviation that would be attainable in the absence of data compression by the uncompressed ACTC and the classical ATC. Accordingly, such excess error summarizes only the effect of data compression, and is therefore expected to reduce as the bit-rate increases (while the overall mean-square-deviation cannot vanish since we are in a stochastic-gradient environment). We see that the curves scale with the number of bits as $(2^r - 1)^{-2}$. This result is in perfect accordance with (88). Moreover, we see that the uncompressed ACTC strategy features a slight increase in mean-square-deviation with respect to the ATC strategy. This difference is due to the higher-order terms $O(\mu^{3/2})$, which are in general different for the two strategies. In particular, lower values of the stabilization parameter ζ that is required by the ACTC strategy (and is instead equal to 1 for the ATC strategy) typically entails a slight increase in the mean-square-deviation that is incorporated in the $O(\mu^{3/2})$ correction.[5] Such small increase becomes visible only when comparable to the quantization error, i.e., when the quantization error becomes negligible. In the particular example of Fig. 3, this happens when the quantization error is ≈ -55 dB.

In Fig. 4, we continue by examining the performance of the *individual* agents. In accordance with our results, *all agents behave similarly*, both in terms of transient and steady-state behavior. As shown by (81), initial discrepancies between the agents (see the inset plot) are absorbed into a faster network transient, after which all agents act in a coordinated manner, and converge to the steady-state value.

Finally, in Fig. 5 we examine the joint role of the step-size μ and of the stabilizing parameter ζ. Again, the theoretical predictions are confirmed, since we see that by keeping the product $\mu\,\zeta$ constant, all curves behave equally in terms of rate $(1 - \mu\,\zeta\,\nu)^2$ and steady-state error, with ζ playing some role only in the faster initial network transient (see inset plot).

It is useful to evaluate the saving, in terms of bits, achieved with the ACTC strategy. To this end, we must recall that the randomized quantizers in Appendix K1 compress finely (say, with machine precision 32 bits) the norm of the vectors to be quantized, send one additional bit for the sign of each entry, and then apply random quantization with r bits to each entry — see (311). Accordingly, given a dimensionality M, a number of iterations $i_{\max}$, and a number of quantization bits r, the overall bit expense of each agent is:

$$r_{\mathrm{tot}} = (32 + M \times (r+1))\, i_{\max}. \tag{100}$$

Applying this formula to the setting of the left plot in Fig. 2, we see that, for the time necessary to enter reliably the steady

[5]These types of higher-order differences among similar forms of distributed implementations are commonly encountered in the pertinent literature, for example, when one compares the ATC strategy against the Combine-Then-Adapt strategy or against the consensus strategy [21].

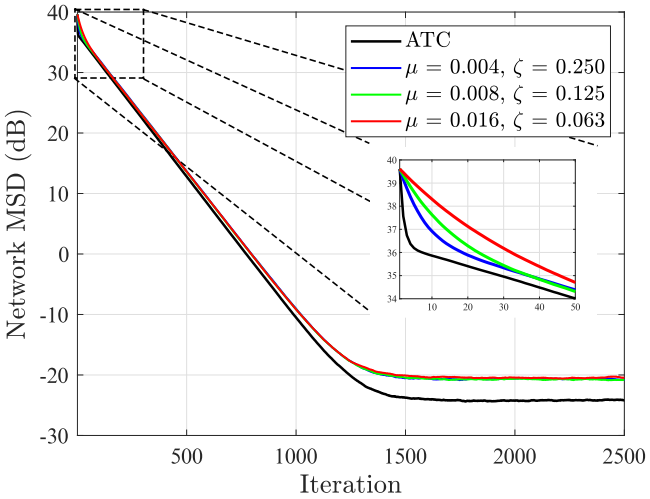

FIGURE 5. Network mean-square-deviation of the ACTC strategy, for different (μ, ζ) pairs guaranteeing the same value of the product $\mu\,\zeta$ (and, hence, the same value of convergence rate ρ_{cen}). The simulation setting is the same as in the left plot in Fig. 2. The inset plot highlights the impact of ζ on the faster initial transient needed by the agents to reach a coordinated evolution.

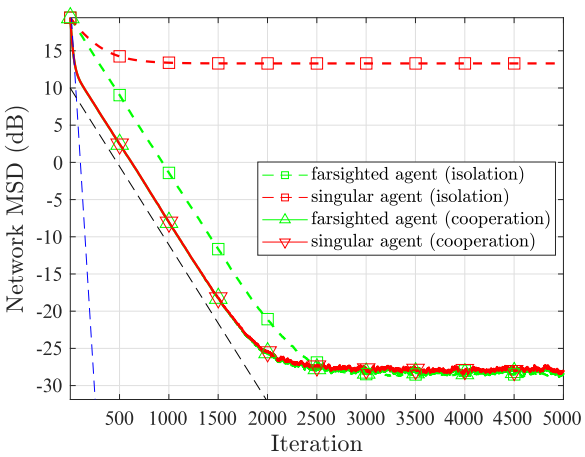

FIGURE 6. ACTC network mean-square-deviation in (99) as a function of the iteration i, for the case of a *doubly-stochastic* combination matrix obtained with the Metropolis rule. We consider the distributed regression problem of Section IX-B with $N-1$ singular agents having locally unidentifiable regression problems, and one farsighted agent having a locally well-posed regression problem. The dimensionality is $M = 10$, and the parameters of the Gaussian regressors $u_{k,i}$ and of the Gaussian disturbances $v_{k,i}$ are drawn as in Fig. 2, but for the fact that the regressors' matrices of the singular agents have two equal columns (which yields local unidentifiability at these agents). The ACTC algorithm is run with stability parameter $\zeta = 0.8$, and with equal step-sizes $\mu_k = \mu = 5 \times 10^{-3}$. All agents use the randomized quantizer in Appendix K1 with bit-rate $r = 3$. All errors are estimated by means of 10^2 Monte Carlo runs.

state ($i_{\max} \approx 2500$), and referring to the coarser scheme that uses only 2 bits, we get:

$$r_{\text{ACTC}}^{(\text{tot})} = (32 + 50 \times 2) \times 2500 = 455 \text{ kbit}. \tag{101}$$

This value should be compared against the expense required by the plain ATC strategy, where each entry of the vector to be quantized is represented by 32 bits, yielding:

$$r_{\text{ATC}}^{(\text{tot})} = 32 \times 50 \times 2500 = 4 \text{ Mbit}, \tag{102}$$

implying a remarkable gain of about one order of magnitude in terms of bit rate. This gain should be evaluated in relation to the loss, in terms of mean-square-deviation, arising from data compression. Inspecting the left plot in Fig. 2, we see that we loose ≈ 4 dB, which is definitely tolerable, especially in the light of the remarkable bit-rate savings.

Similar comments apply to the ACTC strategy with randomized sparsifiers considered in the right plot in Fig. 2. Here we see that 4 dB are lost with respect to the plain ATC strategy by approximately halving the number of transmitted components.

B. UNIDENTIFIABLE PROBLEM

We now move on to consider the challenging case where only one agent (say, agent 1) has a locally identifiable regression problem.

Technically, the regressors' matrix $R_{u,1}$ is invertible, whereas the remaining matrices $R_{u,k}$, for $k > 1$, are singular. In particular, agents $k = 2, 3, \ldots, N$ solve a regression problem with two linearly dependent features. In the following, agent 1 will be referred to as "farsighted agent", whereas the other agents as "singular agents". Moreover, we assume that all agents have the same regressor and noise variances σ_u^2 and σ_v^2, and we consider equal step-sizes at all agents, such that the vector p coincides with the Perron eigenvector.

Under this setting the local cost functions of the singular agents are *not* strongly convex, whereas the aggregate cost function in (98) is strongly convex, with constant ν given by:

$$\nu = 2p_1\sigma_u^2, \tag{103}$$

where p_1 is the entry of the Perron eigenvector corresponding to the farsighted agent.

In Fig. 6, we consider a doubly-stochastic combination policy, namely, the Metropolis rule. Five main conclusions arise. First, agent 1 in isolation is able to learn fairly well, with a steady-state error ≈ -26 dB, while the other agents, when in isolation, are unable to learn properly, with steady-state errors ≈ 18 dB.[6] Second, when organized into a network, *all agents are able to learn properly and even with only 3 bits per iteration.* Third, after the initial network transient, an intermediate transient arises (with convergence rate shown in dashed blue line), which comes before the final transient leading to the coordinated evolution (with rate shown in dashed black line). The intermediate transient can be shown to represent the time needed by the singular agents to align with the farsighted one.[7] Fourth, despite the fact that $N-1$ agents have a singular

[6] Actually, these errors depend on the initial conditions since the singular agents have a *continuum* of minimizers.

[7] This further transient is not visible in our formulas, since it is absorbed in the upper-bound corresponding to the slower transient dominated by $(1 - \mu\,\zeta\,\nu)^2$. By examining the considered example, it is possible to evaluate analytically this transient, but the analysis is beyond the scope of this work.

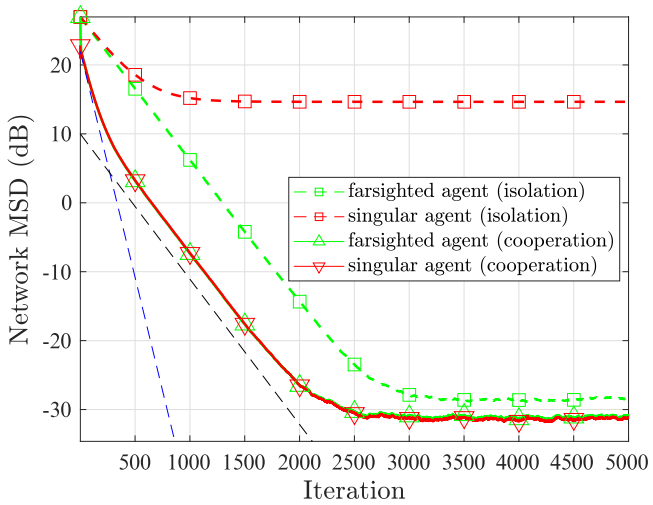

FIGURE 7. ACTC network mean-square-deviation in (99) as a function of the iteration i, for a *left-stochastic* combination matrix obtained through the Metropolis-Hastings procedure to match a target Perron vector that optimizes the (unquantized) ATC performance. In order to match the convergence rate ρ_{cen} corresponding to Fig. 6, we here used $\zeta = 0.8$, and equal step-sizes $\mu_k = \mu = 16 \times 10^{-3}$, while leaving all other system parameters unchanged.

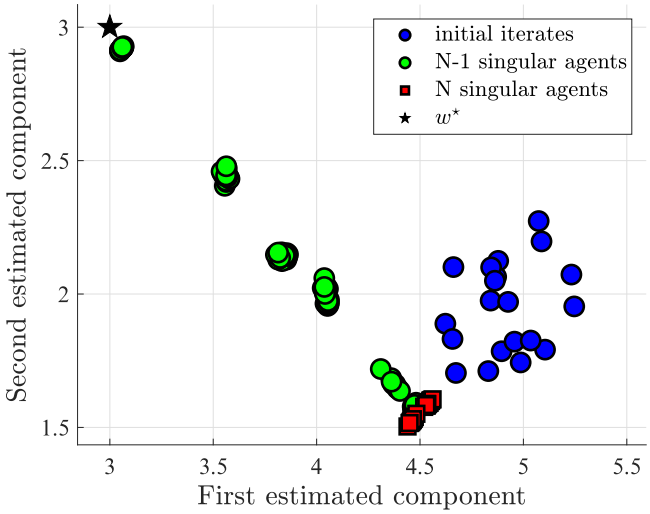

FIGURE 8. Illustrative example of time evolution of the individual agents' iterates for the ACTC strategy.

regressors' matrix, they contribute to accelerate convergence to the steady state. However, we see that the cooperative steady-state performance is equivalent to the individual (i.e., non-cooperative) performance of the farsighted agent. We will now show that this conclusion is not a general conclusion, and depends on the particular combination policy.

To this end, in Fig. 7 we consider the same setting of Fig. 6, but for the choice of the combination policy, which is now left-stochastic. Specifically, we use the Metropolis-Hastings rule to construct a combination matrix with a target Perron eigenvector — see the procedure explained in [21, p. 630]. The target Perron eigenvector is chosen so as to minimize the steady-state error of the unquantized ATC strategy. Clearly, this does not guarantee that we are choosing the best Perron eigenvector for the ACTC strategy, but we will now see that this turns out to be a meaningful choice. In fact, we see from Fig. 7 that, with the optimized combination policy, network cooperation achieves a twofold goal. As in the case of a doubly-stochastic policy, cooperation is beneficial to the singular agents. Moreover, it is also beneficial to the farsighted agent, which is now able to improve on the steady-state performance achieved without cooperation.

In summary, the conducted experiments lead to a revealing conclusion as regards the role of topology on the learning performance. By suitable design of the combination matrix, the regularization action played by agent 1 makes the singular agents capable of contributing more fully to the optimization problem, allowing all agents to achieve a mean-square-deviation that outperforms the non-cooperative performance achievable by the farsighted agent in isolation.

Finally, the convergence behavior of the ACTC strategy is visually illustrated in Fig. 8, with reference to a simple example with dimensionality $M = 2$, and $N = 20$ agents. Let us consider first the case where the N agents are all singular (red squares). In this case, we see that, moving from the initial iterates (blue circles), the singular agents follow wrong paths converging around the wrong point $(4.5, 1.5) \neq w^\star$. In contrast, the ACTC strategy (green circles) allows all agents to converge well to a small neighborhood of the true minimizer, after an initial transient where they need to coordinate with each other.

C. COMPARISON WITH EXISTING STRATEGIES

As we have illustrated in Section II-A, the present work generalizes the existing works on compressed distributed implementations under several aspects, including left-stochastic combination policies, lack of local strong convexity, diffusion strategies. For this reason, the existing theoretical results cannot cover the challenging setting considered in the present work, which required instead a significant additional effort. Nevertheless, even if formulated and studied under alternative settings, some of the existing algorithms can be practically applied to our setting. In particular, we select from the existing algorithms two particular up-to-date implementations that, as far as we know, constitute the actual benchmark performance, namely, CHOCO-SGD [56] and its dual version, DUAL-SGD [57]. Notably, the latter two algorithms have more tuning parameters than our algorithm. Even if the necessity itself of tuning more parameters might be considered a disadvantage of these strategies, in order to ensure a fair comparison we performed a fine tuning of all the parameters to guarantee best performance of CHOCO-SGD and DUAL-SGD. The shaded areas shown in Fig. 9 correspond to the range of mean-square-deviations spanned by a subset of the parameters explored during the tuning phase.

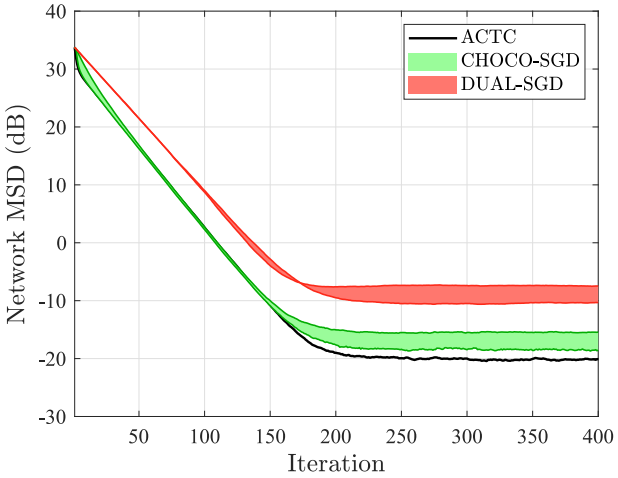

FIGURE 9. Comparison of the proposed ACTC strategy against known strategies, namely, CHOCO-SGD [56] and DUAL-SGD [57]. The shaded areas correspond to the best subsets of tuning parameters of CHOCO-SGD and DUAL-SGD explored during our experiments. The ACTC algorithm is run with stability parameter $\zeta = 0.8$, and with equal step-sizes $\mu_k = \mu = 1.25 \times 10^{-2}$. All agents use the randomized quantizer in Appendix K1 with bit-rate $r = 2$. The network topology is the same shown in the left plot of Fig. 2. The regression problem has dimensionality $M = 10$, with all other parameters being set as in Fig. 2.

Fig. 9 displays the comparison involving the proposed ACTC strategy and the aforementioned two strategies. Remarkably, for the same value of the transient time, the ACTC strategy outperforms both CHOCO-SGD and DUAL-SGD at steady state. In particular, we see that DUAL-SGD performs appreciably worser than the implementations in the primal domain. This conclusion is in perfect agreement with was shown in [70] for the uncompressed case. In fact, the core of DUAL-SGD is a primal-dual distributed strategy of Arrow-Hurwicz type, which was shown in [70, Corollary 3] to converge, despite being a distributed cooperative strategy, at most to the non-cooperative performance. In contrast, both the ACTC and CHOCO-SGD strategies are able to exploit fully the distributed cooperation, which explains the performance improvement exhibited in Fig. 9.[8]

We move on to examine the improvement of the proposed ACTC strategy on the existing CHOCO-SGD strategy. Also in this case, this improvement can be neatly explained in the light of known behavior observed in the uncompressed case, since the improvement matches well similar gains achievable when using diffusion (as ACTC does) as opposed to consensus (as CHOCO-SGD does). In fact, as observed in distributed optimization without compression [21], diffusion strategies can outperform consensus strategies, and, remarkably, from our experiments we observe the same behavior when these types of strategies are called to operate under communication constraints.

[8]In [57], with reference to a regularized logistic regression example, it is shown that DUAL-SGD and CHOCO-SGD perform similarly, but this comparison is in terms of function values and not in terms of iterates.

X. CONCLUSION

We considered a network of agents tasked to solve a certain distributed optimization problem from continual aggregation of *streaming observations*. Fundamental features of our setting are *adaptation*, *local cooperation* and *data compression*. By adaptation we mean that the agents must be able to react promptly to drifts in the operational conditions, so as to adapt their inferential solution quickly. In this regard, stochastic-gradient algorithms with *constant step-size* become critical. By local-cooperation we mean that each individual agent is allowed to implement a *distributed* algorithm by exchanging information with its neighbors. Finally, data compression comes from the need of communicating information at a finite rate, owing to energy/bandwidth constraints.

We introduced a novel strategy nicknamed as Adaptive-Compress-Then-Combine (ACTC), whose core is an *adaptive diffusion* strategy properly twinned with a *differential stochastic compression* strategy. Our analysis is conducted under the challenging setting where: *i*) communication is allowed to be unidirectional (i.e., over *directed* graphs); and *ii*) the cost functions at the individual agents are allowed to be *non-convex*, provided that a global cost function obtained as linear combination of the local cost functions is *strongly convex*. We obtained the following main results. First, we established that the proposed ACTC scheme is mean-square stable, and in particular that *each individual agent* is able to infer well the value of the parameter to be estimated, with a mean-square-deviation vanishing proportionally to the step-size. Second, we characterized the learning behavior of each individual agent, obtaining analytical solutions that highlight the existence of two main transient phases, one (faster) relative to convergence of all agents to a *coordinated evolution*, the other (slower) relative to convergence of the coordinated estimate to the steady-state solution. Notably, these distinct learning phases were shown to emerge in diffusion strategies *without data compression* [14]. Therefore, our result implies that these distinct phases are preserved *despite the presence of compressed data, for any degree of compression*. Moreover, there are also distinguishing features arising from data compression, and the obtained analytical solutions are able to reflect well the role of the compression degree (e.g., in terms of quantization bits or sparsity level) in the final learning behavior. A remarkable conclusion stemming from our analysis is that, for sufficiently small step-sizes, small errors are achievable for *any compression degree*. This behavior brings an interesting analogy with classical information-theoretic results: the information-rate limitation does not preclude the possibility of learning, but involves a reduction in the speed of convergence.

Another useful parallel can be drawn with the recently introduced paradigm of *exact diffusion* [71], [72], where no compression is present, and the *true* (i.e., non stochastic) gradient is available. Under this paradigm, diffusion strategies with constant step-size μ are enriched with an error compensation step, which allows them to attain a zero, rather than $O(\mu)$, mean-square-deviation [71], [72]. Inspired by the

structure of the ACTC diffusion strategy in Algorithm 1, it could be worth including in the exact diffusion algorithm the parameter ζ and a general nonlinearity $\boldsymbol{Q}_k(\cdot)$, and tuning these two quantities to speed up convergence.

There is still a lot of work to be done in the context of distributed adaptive learning under communication constraints. One advance regards a steady-state performance analysis aimed at obtaining exact formulas for the mean-square-deviation. Another advance concerns the generalization of the analysis to quantizers that do not fulfill the unbiasedness assumption. A further contribution is to extend results available under non-convex environments [73], [74], [75] to the case of compressed data. Under such setting, the traditional difficulties arising from the lack of convexity (e.g., the evolution of the stochastic-gradient iterates, their mean-square stability and steady-state performance) will be complicated by the complexity arising from the introduction of the nonlinear compression operator. Finally, we are currently investigating open problems where the techniques and results of the present work can be exploited, regarding the trade-off between network attributes and communication resources (i.e., how to perform jointly the design of the topology and the allocation of the communication budget to maximize the performance) and the analysis of other classes of compression operators that do not assume high-precision representation of some variables.

APPENDIX A
COLLECTION OF USEFUL BACKGROUND RESULTS

1. JORDAN REPRESENTATION

Let J_{tot} be the matrix associated with the canonical Jordan decomposition of the combination matrix A, which can be represented as [66]:

$$J_{\text{tot}} = \mathsf{diag}\{J_1, J_2, \ldots, J_B\}, \tag{104}$$

where B is the number of Jordan blocks. As usual, the individual blocks can have different size, and the unspecified off-diagonal terms arising after block-diagonal concatenation are automatically set to zero. For $n = 1, 2, \ldots, B$, we denote by λ_n the eigenvalue associated with block J_n, and, without loss of generality, we assume that the eigenvalues are sorted in descending order of magnitude, namely,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_B|. \tag{105}$$

Each Jordan block takes on the form:

$$J_n \triangleq \begin{bmatrix} \lambda_n & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_n \end{bmatrix}, \tag{106}$$

and can be accordingly written as:

$$J_n = \lambda_n I_{L_n} + U_{L_n}, \tag{107}$$

where L_n is the dimension of the n-th block, and U_{L_n} is a square matrix of size L_n that has all zero entries, but for the first diagonal above the main diagonal, which has entries equal to 1.

In view Assumptions 2 and 3, the combination matrix A has a unique largest magnitude eigenvalue that is $\lambda_1 = 1$, i.e., the first Jordan block is $J_1 = 1$. The remaining $B - 1$ Jordan blocks can be conveniently arranged in the reduced matrix:

$$J = \mathsf{diag}\{J_2, \ldots, J_B\}. \tag{108}$$

Moreover, letting

$$\Lambda \triangleq \mathsf{diag}\{\lambda_2 I_{L_2}, \lambda_3 I_{L_3}, \ldots, \lambda_B I_{L_B}\}, \tag{109}$$

and

$$U \triangleq \mathsf{diag}\{U_{L_2}, U_{L_3}, \ldots, U_{L_B}\}, \tag{110}$$

we end up with the following useful representation:

$$J = \Lambda + U. \tag{111}$$

2. ENERGY OPERATORS

Definition 1 (Energy Vector Operator): Let $x_1, x_2, \ldots, x_N$ be N vectors of size $M \times 1$, and let

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \tag{112}$$

be the block vector of size $MN \times 1$ obtained by concatenating these vectors. The energy vector operator, $\mathsf{P} : \mathbb{C}^{MN} \to \mathbb{R}^N$, is defined as:

$$\mathsf{P}[x] = \begin{bmatrix} \|x_1\|^2 \\ \|x_2\|^2 \\ \vdots \\ \|x_N\|^2 \end{bmatrix}. \tag{113}$$

□

For a *random* vector $\boldsymbol{x}$, we introduce the following average energy operator:

$$\mathscr{P}[\boldsymbol{x}] = \mathbb{E}\mathsf{P}[\boldsymbol{x}] = \begin{bmatrix} \mathbb{E}\|\boldsymbol{x}_1\|^2 \\ \vdots \\ \mathbb{E}\|\boldsymbol{x}_N\|^2 \end{bmatrix}. \tag{114}$$

The block-matrix counterpart of operator $\mathsf{P}[\cdot]$ is defined as follows.

Definition 2 (Norm Matrix Operator): Let $\{X_{kn}\}$, with $k = 1, 2, \ldots, K$ and $n = 1, 2, \ldots, N$, a doubly-indexed sequence of $M \times M$ matrices, and consider the $MK \times MN$ matrix X whose (k, n)-block is $\{X_{kn}\}$. The norm matrix operator, $\bar{\mathsf{P}}$:

$\mathbb{C}^{MK\times MN} \to \mathbb{R}^{K\times N}$, is defined as:

$$\bar{\mathsf{P}}[X] \triangleq \begin{bmatrix} \|X_{11}\| & \dots & \|X_{1N}\| \\ \vdots & & \vdots \\ \|X_{K1}\| & \dots & \|X_{KN}\| \end{bmatrix}. \tag{115}$$

□

The operators (113) and (115) are equipped with several useful properties. We now list those properties that will be exploited in the forthcoming proofs, and refer the reader to [14] for the proof of these properties.

Property 1 (Energy Vector Operator and Norm Matrix Operator Properties): Let x and y be two block column vectors of size $MN \times 1$ constructed as in Definition 1, and let X and Y be two block matrices of size $MK \times MN$ constructed as in Definition 2. The following properties hold.

P1) Nonnegativity: $\mathsf{P}[x] \succeq 0$, $\bar{\mathsf{P}}[X] \succeq 0$.

P2) Scaling: For any scalar $a \in \mathbb{C}$, $\mathsf{P}[ax] = |a|^2\,\mathsf{P}[x]$ and $\bar{\mathsf{P}}[aX] = |a|\,\bar{\mathsf{P}}[X]$.

P3) Convexity: given a set of L vectors $\{x^{(1)}, \dots, x^{(L)}\}$ having the same structure of x, and a set of convex coefficients $\{a_1, \dots, a_L\}$, then

$$\mathsf{P}[a_1x^{(1)} + \dots + a_Lx^{(L)}] \preceq a_1\mathsf{P}[x^{(1)}] + \dots + a_L\mathsf{P}[x^{(L)}]. \tag{116}$$

P4) Additivity under orthogonality: Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two block column random vectors of size $MN \times 1$ constructed as in Definition 1. If blocks $\boldsymbol{x}_k$ and $\boldsymbol{y}_k$ are orthogonal for all $k = 1, 2, \dots, N$, namely if

$$\mathbb{E}\left[\boldsymbol{x}_k^{\mathsf{H}}\boldsymbol{y}_k\right] = 0, \tag{117}$$

then we have that:

$$\mathscr{P}[\boldsymbol{x} + \boldsymbol{y}] = \mathscr{P}[\boldsymbol{x}] + \mathscr{P}[\boldsymbol{y}]. \tag{118}$$

P5) Relation to Euclidean norm: $\mathbb{1}_N^{\top}\mathsf{P}[x] = \|x\|^2$.

P6) Linear transformations: Applying the energy operator to the linear transformation Yx we obtain:

$$\mathsf{P}[Yx] \preceq \|\bar{\mathsf{P}}[Y]\|_{\infty}\,\bar{\mathsf{P}}[Y]\,\mathsf{P}[x] \tag{119}$$

$$\preceq \|\bar{\mathsf{P}}[Y]\|_{\infty}^2\,\mathbb{1}_K\mathbb{1}_N^{\top}\,\mathsf{P}[x]. \tag{120}$$

P7) Stable Kronecker Jordan operator: Consider the extended version $\mathcal{J} = J \otimes I_M$ of the reduced Jordan matrix $J = \Lambda + U$ in (111). Then, for any two block column vectors x' and y' of size $M(N-1) \times 1$ constructed as in Definition 1, we have that:

$$\mathsf{P}\left[\mathcal{J}x' + y'\right] \preceq \left(\frac{\Lambda\Lambda^*}{|\lambda_2|} + \frac{2\,U}{1-|\lambda_2|}\right)\mathsf{P}[x'] + \frac{2}{1-|\lambda_2|}\mathsf{P}[y']. \tag{121}$$

For the benefit of the reader, we note that this property was not stated in this form in [14]. Equation (121) is readily obtained from [14, Eq. (198)] by exploiting the diagonal structure of Λ.

3. REPRESENTATION IN TRANSFORMED NETWORK COORDINATES

Before proving all the pertinent lemmas and theorems, it is useful to write down the ACTC strategy in terms of the transformed variables introduced in Section V-B. Regarding the transformed quantized vector $\widehat{\boldsymbol{q}}_i$, applying definition (56) in the second step of (51), we readily get:

$$\widehat{\boldsymbol{q}}_i = \mathcal{V}\widetilde{\boldsymbol{q}}_i = \widehat{\boldsymbol{q}}_{i-1} + \zeta\,\mathcal{V}\,\boldsymbol{\mathcal{Q}}\left(\mathcal{V}^{-1}\widehat{\boldsymbol{\delta}}_i\right). \tag{122}$$

We note in passing that the term $\mathcal{V}\,\boldsymbol{\mathcal{Q}}(\mathcal{V}^{-1}\widehat{\boldsymbol{\delta}}_i)$ reflects well the inherent nonlinear behavior of the compression operators. In fact, the linear transformation $\mathcal{V}$ and the nonlinear operator $\boldsymbol{\mathcal{Q}}(\cdot)$ do not commute and, hence, the direct and inverse network transformation, $\mathcal{V}$ and $\mathcal{V}^{-1}$, do *not* compensate perfectly with each other.

Let us switch to the transformed quantization-error vector $\widehat{\boldsymbol{\delta}}_i$, and focus accordingly on the first step in (51). We introduce the extended Jordan matrix:

$$\mathcal{J}_{\text{tot}} \triangleq J_{\text{tot}} \otimes I_M, \tag{123}$$

which, in view of (52), allows us to write the extended combination matrix $\mathcal{A}$ in (50) as:

$$\mathcal{A}^{\top} = \mathcal{V}^{-1}\mathcal{J}_{\text{tot}}\mathcal{V}. \tag{124}$$

Therefore, in view of (51) we can write:

$$(I_{MN} - \mu\,\boldsymbol{\mathcal{H}}_{i-1})\mathcal{A}^{\top} - I_{MN} = \mathcal{V}^{-1}\,(\mathcal{J}_{\text{tot}} - \mu\,\boldsymbol{\mathcal{G}}_{i-1} - I_{MN})\,\mathcal{V}, \tag{125}$$

where we introduced the matrix:

$$\boldsymbol{\mathcal{G}}_i \triangleq \mathcal{V}\,\boldsymbol{\mathcal{H}}_i\,\mathcal{A}^{\top}\mathcal{V}^{-1}. \tag{126}$$

Substituting now (125) into the first step of (51) and applying the network transformation, we get:

$$\widehat{\boldsymbol{\delta}}_i = \mathcal{V}\boldsymbol{\delta}_i = (\mathcal{J}_{\text{tot}} - I_{MN} - \mu\,\boldsymbol{\mathcal{G}}_{i-1})\widehat{\boldsymbol{q}}_{i-1} - \mu\,\widehat{\boldsymbol{s}}_i - \mu\,\widehat{b}. \tag{127}$$

Furthermore, by using (50) and (55) in (126), the matrix $\boldsymbol{\mathcal{G}}_{i-1}$ can be written as:

$$\begin{aligned}\boldsymbol{\mathcal{G}}_{i-1} &= (V \otimes I_M)\boldsymbol{\mathcal{H}}_{i-1}(A^{\top} \otimes I_M)(V^{-1} \otimes I_M) \\ &= (V \otimes I_M)\boldsymbol{\mathcal{H}}_{i-1}(A^{\top}V^{-1} \otimes I_M),\end{aligned} \tag{128}$$

where we used the property of the Kronecker product $(X \otimes Z)(Y \otimes Z) = XY \otimes Z$, holding for any three matrices X, Y, Z with compatible dimensions. Exploiting now the partitioned structure of V and V^{-1} in (54), we can write:

$$V \otimes I_M = \begin{bmatrix} \pi^{\top} \otimes I_M \\ V_R \otimes I_M \end{bmatrix}, \tag{129}$$

$$A^{\top}V^{-1} \otimes I_M = \begin{bmatrix} \mathbb{1}_N \otimes I_M & A^{\top}V_L \otimes I_M \end{bmatrix}, \tag{130}$$

where in the last matrix we used the equality $A^{\top}\mathbb{1}_N = \mathbb{1}_N$, holding since A is a left-stochastic matrix. Using (129) and (130) in (128) we obtain the following block-decomposition

for $\mathcal{G}_i$:

$$\mathcal{G}_i = \begin{bmatrix} G_{11,i} & G_{12,i} \\ G_{21,i} & G_{22,i} \end{bmatrix}, \tag{131}$$

where

$$G_{11,i} = \sum_{k=1}^{N} \pi_k H_{k,i}, \tag{132}$$

$$G_{12,i} = (\pi^\top \otimes I_M)\, \mathcal{H}_i\, (A^\top V_L \otimes I_M), \tag{133}$$

$$G_{21,i} = (V_R \otimes I_M)\, \mathcal{H}_i\, (\mathbb{1}_N \otimes I_M), \tag{134}$$

$$G_{22,i} = (V_R \otimes I_M)\, \mathcal{H}_i\, (A^\top V_L \otimes I_M). \tag{135}$$

Combining now (127) with (131), we obtain:

$$\begin{bmatrix} \bar{\delta}_i \\ \check{\delta}_i \end{bmatrix} = \begin{bmatrix} -\mu G_{11,i-1} & -\mu G_{12,i-1} \\ -\mu G_{21,i-1} & \mathcal{J} - I_{M(N-1)} - \mu G_{22,i-1} \end{bmatrix} \times \begin{bmatrix} \bar{q}_{i-1} \\ \check{q}_{i-1} \end{bmatrix} - \mu \begin{bmatrix} \bar{s}_i \\ \check{s}_i \end{bmatrix} - \mu \begin{bmatrix} 0 \\ \check{b} \end{bmatrix}. \tag{136}$$

We conclude this section with a lemma that will be repeatedly used in the forthcoming proofs.

Lemma 1 (Characterization of $\mathcal{G}_i$): The blocks of matrix $\mathcal{G}_i$ in (131) satisfy the following bounds. First, the $M \times M$ symmetric matrix $G_{11,i}$ in (132) fulfills the bounds:

$$\nu I_M \le G_{11,i} \le \eta I_M, \tag{137}$$

where ν is the global-strong-convexity constant introduced in (13) and η is the average Lipschitz constant in (14).

Second, a positive constant σ_{12} exists such that:

$$\|G_{12,i}\| \le \sigma_{12}. \tag{138}$$

Finally, the matrices $\bar{\mathsf{P}}[G_{21,i}]$ and $\bar{\mathsf{P}}[G_{22,i}]$ obtained by applying the norm matrix operator to the matrices in (134) and (135), have bounded norm, in particular we have:

$$\|\bar{\mathsf{P}}[G_{21,i}]\|_\infty \le \sigma_{21}, \;\; \|\bar{\mathsf{P}}[G_{22,i}]\|_\infty \le \sigma_{22}, \tag{139}$$

for some positive constants σ_{21}, and σ_{22}.

Proof: The proof relies basically on the properties of the matrices $H_{k,i}$, which arise from Assumptions 1 and 4. Let us focus on (137). Using (13) and (40) in (132) we readily obtain:

$$G_{11,i} = \sum_{k=1}^{N} p_k \int_0^1 \nabla^2 J_k(w^\star + t\widetilde{w}_{k,i})dt \ge \nu I_M, \tag{140}$$

which proves the lower bound in (137). The upper bound is obtained by observing that:

$$\begin{aligned} \|G_{11,i}\| &\le \sum_{k=1}^{N} p_k \left\| \int_0^1 \nabla^2 J_k(w^\star + t\widetilde{w}_{k,i})dt \right\| \\ &\le \sum_{k=1}^{N} p_k \int_0^1 \left\| \nabla^2 J_k(w^\star + t\widetilde{w}_{k,i}) \right\| dt \le \eta, \end{aligned} \tag{141}$$

where the first inequality is the triangle inequality, the intermediate inequality is the mean-value inequality, and the last inequality follows by (14).

We continue by proving (138) and (139). First, we note that $G_{12,i}$, $G_{21,i}$, and $G_{22,i}$ have the following common structure:

$$(X \otimes I_M)\mathsf{diag}\{H_{1,i}, H_{2,i}, \ldots, H_{N,i}\}(Y \otimes I_M), \tag{142}$$

for a suitable choice of the matrices X and Y, having made explicit the definition of $\mathcal{H}_i$ in (50). The bound in (138) follows readily from the Lipschitz property in (2). Regarding (139), we observe that we can write:

$$\begin{aligned} &(X \otimes I_M)\mathsf{diag}\{H_{1,i}, H_{2,i}, \ldots, H_{N,i}\} \\ &= \begin{bmatrix} x_{11}H_{1,i} & x_{12}H_{2,i} & \cdots & x_{1N}H_{N,i} \\ x_{21}H_{1,i} & x_{22}H_{2,i} & \cdots & x_{2N}H_{N,i} \\ \vdots & & \vdots & \\ x_{N1}H_{1,i} & x_{N2}H_{2,i} & \cdots & x_{NN}H_{N,i} \end{bmatrix}. \end{aligned} \tag{143}$$

Therefore, for $\ell, \ell' = 1, 2, \ldots, N$, the (ℓ, ℓ')-block of the matrix $(X \otimes I_M)\mathsf{diag}\{H_{1,i}, H_{2,i}, \ldots, H_{N,i}\}(Y \otimes I_M)$ can be written as:

$$\sum_{k=1}^{N} x_{\ell k}\, y_{k\ell'}\, H_{k,i}, \tag{144}$$

showing that each $M \times M$ block of $G_{21,i}$ or $G_{22,i}$ is a linear combination of the matrices $H_{k,i}$. Now, by applying the Lipschitz property in (2) to the individual matrices $H_{k,i}$ in (40) we see that each of these matrices has bounded norm. Applying the norm matrix operator to $G_{21,i}$ or $G_{22,i}$, we conclude that all entries of $\bar{\mathsf{P}}[G_{21,i}]$ and $\bar{\mathsf{P}}[G_{22,i}]$ are bounded, and, hence, so are $\|\bar{\mathsf{P}}[G_{21,i}]\|_\infty$ and $\|\bar{\mathsf{P}}[G_{22,i}]\|_\infty$, which concludes the proof of the lemma. ■

APPENDIX B
AUXILIARY RESULTS

In this appendix we collect four auxiliary results (Lemmas 2, 3, 4, and Theorem 3) that will be used to prove Theorems 1 and 2. In order to facilitate the reading, each of these results is proved in a separate appendix.

As we explained before, it is instrumental to work in terms of the quantized iterates $q_{k,i}$ and, more precisely, in terms of the transformed vectors $\widehat{q}_i = \mathcal{V}\widetilde{q}_i$. In order to characterize the mean-square evolution of the transformed vectors, it is particularly convenient to adopt the formalism of the energy operators introduced in Appendix A2.

In particular, when we apply the average energy operator in (114) to one of our transformed vectors, e.g., to $\widehat{q}_i$ in (56), we obtain the following block decomposition:

$$\mathscr{P}[\widehat{q}_i] = \begin{bmatrix} \mathbb{E}\|\bar{q}_i\|^2 \\ \\ \mathscr{P}[\check{q}_i] \end{bmatrix}, \tag{145}$$

where $\mathscr{P}[\check{q}_i]$ is an $(N-1) \times 1$ vector. Examining the time evolution of the energy vectors in (145) is critical because the

individual agents' errors $\mathbb{E}\|\boldsymbol{q}_{k,i}\|^2$ can be related to these energy vectors through the inverse network transformation $\mathcal{V}^{-1}$ in (57). In particular, we will be able to show that the quantity $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2$ plays a dominant role in determining the (common) individual agents' steady-state mean-square behavior, whereas the quantity $\mathscr{P}[\check{\boldsymbol{q}}_i]$ plays the role of a network error quantity that dies out during the transient phase.

We start with three lemmas that characterize the interplay over time (in terms of energy) of three main quantities in the network transformed domain: the gradient noise $\widehat{\boldsymbol{s}}_i$ in (59), the quantization error $\widehat{\boldsymbol{\delta}}_i$ in (58), and the quantized iterates $\widehat{\boldsymbol{q}}_i$ in (56).

The first lemma relates the gradient noise to the quantized iterates.

Lemma 2 (Gradient Noise Energy Transfer): The average energy of the transformed gradient noise extended vector $\widehat{\boldsymbol{s}}_i$ evolves over time according to the following inequality:

$$\mathscr{P}[\mu\,\widehat{\boldsymbol{s}}_i] \preceq T_s\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + x_s, \tag{146}$$

where the transfer matrix T_s and the driving vector $x_s = [\bar{x}_s, \check{x}_s]^\top$ are defined in Table 2.

Proof: See Appendix C.

The second lemma relates the quantization error to the quantized iterates and to the gradient noise. ■

Lemma 3 (Quantization Error Energy Transfer): The average energy of the transformed quantization-error extended vector $\widehat{\boldsymbol{\delta}}_i$ evolves over time according to the following inequality:

$$\mathscr{P}[\widehat{\boldsymbol{\delta}}_i] \preceq T_\delta\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + \mathscr{P}[\mu\,\widehat{\boldsymbol{s}}_i] + x_\delta, \tag{147}$$

where the transfer matrix T_δ and the driving vector $x_\delta = [\bar{x}_\delta, \check{x}_\delta]^\top$ are defined in Table 2.

Proof: See Appendix D. ■

The third lemma relates the quantized iterates to the quantization error and the gradient noise.

Lemma 4 (Quantized Iterates Energy Transfer): Let

$$\mu\,\zeta < \frac{2}{\eta + \nu}, \tag{148}$$

where ν is the global-strong-convexity constant introduced in (13) and η is the average Lipschitz constant in (14). Let

$$\Delta = \begin{bmatrix} \bar{\Delta} & \\ & \check{\Delta}\mathbb{1}_{N-1} \end{bmatrix}, \tag{149}$$

where the constants $\bar{\Delta}$ and $\check{\Delta}$ are defined in Table 1. Then, the average energy of the transformed extended vector $\widehat{\boldsymbol{q}}_i$ evolves over time according to the following inequality:

$$\begin{aligned}\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq\; & T_q\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] \\ & + \zeta^2\Delta\,\mathbb{1}_N^\top\,\mathscr{P}[\widehat{\boldsymbol{\delta}}_i] + \zeta^2\,\mathscr{P}[\mu\,\widehat{\boldsymbol{s}}_i] + x_q,\end{aligned} \tag{150}$$

where the transfer matrix T_q and the driving vector $x_q = [\bar{x}_q, \check{x}_q]^\top$ are defined in Table 2.

Proof: See Appendix E. ■

Combining Lemmas 2, 3, and 4 we arrive at a recursion on the quantized iterates $\widehat{\boldsymbol{q}}_i$, as stated in the next theorem. Unless otherwise specified, all matrices, vectors and constants in the statement of the theorem can be found in Tables 1 and 2.

Theorem 3 (Recursion on Quantized Iterates): Let

$$T = \underbrace{\begin{bmatrix} \tau & \tau_{12}\mathbb{1}_{N-1}^\top \\ 0 & E \end{bmatrix}}_{T_0} + \upsilon_{\mu,\zeta}\,\mathbb{1}_N^\top, \tag{151}$$

where:

$$\tau \triangleq 1 - \mu\,\zeta\,\nu, \quad \tau_{12} \triangleq 16\,\zeta^2\,\bar{\Delta} + \mu\,\zeta\,\frac{\sigma_{12}^2}{\nu}, \tag{152}$$

$$E = E_0 + 16\,\zeta^2\,\check{\Delta}\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top, \tag{153}$$

and

$$\upsilon_{\mu,\zeta} = \phi \begin{bmatrix} \mu^2\,\zeta^2 \\ \mu^2\,\zeta\,\mathbb{1}_{N-1} \end{bmatrix}, \tag{154}$$

with ϕ being a positive scalar that embodies the constants appearing in the μ^2-terms of the transfer matrices T_s, T_δ and T_q in Table 2. The evaluation of ϕ is rather cumbersome and is detailed in Appendix F. Let also

$$x = \begin{bmatrix} \bar{x} \\ \check{x} \end{bmatrix} = x_q + \zeta^2\Delta\,\mathbb{1}_N^\top\,(x_\delta + x_s) + \zeta^2 x_s, \tag{155}$$

where the driving vectors x_s, x_δ and x_q are defined in Table 2. Then, the average energy of the extended vector $\widehat{\boldsymbol{q}}_i$ obeys the following inequality:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq T\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + x. \tag{156}$$

Proof: See Appendix F. ■

APPENDIX C
PROOF OF LEMMA 2

Proof: Each component of $\boldsymbol{s}_i$ in (42) fulfills the following chain of inequalities:

$$\begin{aligned}\mathbb{E}\,\|\boldsymbol{s}_{k,i}\|^2 &= \alpha_k^2\,\mathbb{E}\,\|\boldsymbol{n}_{k,i}(\boldsymbol{w}_{k,i-1})\|^2 \\ &\overset{(a)}{\leq} \alpha_k^2\,\beta_k^2\,\mathbb{E}\,\|\widetilde{\boldsymbol{w}}_{k,i-1}\|^2 + \alpha_k^2\,\sigma_k^2 \\ &= \alpha_k^2\,\beta_k^2\,\mathbb{E}\,\Big\|\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\widetilde{\boldsymbol{q}}_{\ell,i-1}\Big\|^2 + \alpha_k^2\,\sigma_k^2 \\ &\overset{(b)}{\leq} \alpha_k^2\,\beta_k^2\sum_{\ell=1}^{N} a_{\ell k}\mathbb{E}\,\|\widetilde{\boldsymbol{q}}_{\ell,i-1}\|^2 + \alpha_k^2\,\sigma_k^2 \\ &\leq \alpha_k^2\,\beta_k^2\,\mathbb{E}\,\|\widetilde{\boldsymbol{q}}_{i-1}\|^2 + \alpha_k^2\,\sigma_k^2 \\ &= \alpha_k^2\,\beta_k^2\,\mathbb{E}\,\|\mathcal{V}^{-1}\,\mathcal{V}\widetilde{\boldsymbol{q}}_{i-1}\|^2 + \alpha_k^2\,\sigma_k^2 \\ &\overset{(c)}{\leq} \alpha_k^2\,\beta_k^2\,\|V^{-1}\|^2\,\mathbb{E}\,\|\widehat{\boldsymbol{q}}_{i-1}\|^2 + \alpha_k^2\,\sigma_k^2\end{aligned}$$

$$\overset{(d)}{=} \alpha_k^2\,\beta_k^2\,\|V^{-1}\|^2\,\mathbb{1}_N^\top\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + \alpha_k^2\,\sigma_k^2, \qquad (157)$$

where (a) follows by (35); (b) follows by Jensen's inequality since the weights $\{a_{\ell k}\}$ are convex; (c) exploits the fact that $\widehat{\boldsymbol{q}}_{i-1} = \mathcal{V}\widetilde{\boldsymbol{q}}_{i-1}$ and the equality $\|\mathcal{V}^{-1}\|=\|(V\otimes I_M)^{-1}\|=\|V^{-1}\|$; and (d) follows by property P5) of the energy operators.

Recalling that the transformed gradient noise vector $\widehat{\boldsymbol{s}}_i$ is equal to $(V\otimes I_M)\boldsymbol{s}_i$, the ℓ-th block $\widehat{\boldsymbol{s}}_{\ell,i}$, for $\ell = 1, 2, \ldots, N$, is:

$$\widehat{\boldsymbol{s}}_{\ell,i} = \sum_{k=1}^{N} v_{\ell k}\,\boldsymbol{s}_{k,i}, \qquad (158)$$

where $v_{\ell k}$ is the (ℓ, k)-entry of matrix V. It is useful to examine separately the coordinated-evolution component $\bar{\boldsymbol{s}}_i = \widehat{\boldsymbol{s}}_{1,i}$ and the remaining components $\widehat{\boldsymbol{s}}_{\ell,i}$, for $\ell = 2, 3, \ldots, N$. To this end, we exploit the block decomposition of matrix V in (54). Regarding $\bar{\boldsymbol{s}}_i$, since $v_{1k} = \pi_k$, from (158) we have that:

$$\mathbb{E}\|\bar{\boldsymbol{s}}_i\|^2 = \mathbb{E}\|\widehat{\boldsymbol{s}}_{1,i}\|^2 = \mathbb{E}\left\|\sum_{k=1}^{N}\pi_k\,\boldsymbol{s}_{k,i}\right\|^2 \le \sum_{k=1}^{N}\pi_k\,\mathbb{E}\|\boldsymbol{s}_{k,i}\|^2$$
$$\le \|V^{-1}\|^2\left(\sum_{k=1}^{N}\pi_k\,\alpha_k^2\,\beta_k^2\right)\mathbb{1}_N^\top\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + \sum_{k=1}^{N}\pi_k\,\alpha_k^2\,\sigma_k^2, \qquad (159)$$

where the first inequality is Jensen's inequality with convex weights $\{\pi_k\}$, whereas the second inequality comes from (157). Likewise, for $\ell = 2, 3, \ldots, N$, from (158) we can write:

$$\mathbb{E}\|\widehat{\boldsymbol{s}}_{\ell,i}\|^2 = N^2\mathbb{E}\left\|\frac{1}{N}\sum_{k=1}^{N}v_{\ell k}\,\boldsymbol{s}_{k,i}\right\|^2 \le N\sum_{k=1}^{N}|v_{\ell k}|^2\,\mathbb{E}\|\boldsymbol{s}_{k,i}\|^2$$
$$\le N\,\|V^{-1}\|^2\left(\sum_{k=1}^{N}|v_{\ell k}|^2\,\alpha_k^2\,\beta_k^2\right)\mathbb{1}_N^\top\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}]$$
$$+ N\sum_{k=1}^{N}|v_{\ell k}|^2\,\alpha_k^2\,\sigma_k^2, \qquad (160)$$

where the first inequality is Jensen's inequality with uniform weights $1/N$, whereas the second inequality comes from (157). Finally, by introducing the "squared" counterpart of the complex matrix V_R, whose entries, for $\ell = 1, 2, \ldots, N-1$ and $k = 1, 2, \ldots, N$ are:

$$[V_{2R}]_{\ell k} = \big|[V_R]_{\ell k}\big|^2, \qquad (161)$$

and recalling the definition of the diagonal matrices C_α, C_β, and C_σ in Table 1, from (159) and (160) it is readily seen that the claim in (146) has been in fact proved, with the characterization of matrix T_s and of the quantities $\bar{x}_s$ and $\check{x}_s$ as given in Table 2. ■

APPENDIX D
PROOF OF LEMMA 3

Proof: Applying the average energy operator $\mathscr{P}[\cdot]$ to (136), we obtain:

$$\mathbb{E}\|\bar{\boldsymbol{\delta}}_i\|^2 = \mu^2\,\mathbb{E}\,\|\bar{\boldsymbol{s}}_i\|^2 + \mu^2\,\mathbb{E}\left\|\boldsymbol{G}_{11,i-1}\bar{\boldsymbol{q}}_{i-1} + \boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\right\|^2 \qquad (162)$$

$$\mathscr{P}[\check{\boldsymbol{\delta}}_i] = \mu^2\,\mathscr{P}[\check{\boldsymbol{s}}_i] + \mathscr{P}\big[(\mathcal{J} - I_{M(N-1)} - \mu\,\boldsymbol{G}_{22,i-1})\check{\boldsymbol{q}}_{i-1}$$
$$- \mu\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1} - \mu\,\check{\boldsymbol{b}}\big], \qquad (163)$$

where the energy terms corresponding to the gradient noise are additive in view of property (34) and property P4) of the energy operator.

Let us consider the second term on the RHS of (162), for which we can write:

$$\|\boldsymbol{G}_{11,i-1}\bar{\boldsymbol{q}}_{i-1} + \boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2$$
$$\le 2\|\boldsymbol{G}_{11,i-1}\bar{\boldsymbol{q}}_{i-1}\|^2 + 2\|\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2$$
$$\le 2\sigma_{11}^2\|\bar{\boldsymbol{q}}_{i-1}\|^2 + 2\sigma_{12}^2\|\check{\boldsymbol{q}}_{i-1}\|^2$$
$$= 2\sigma_{11}^2\|\bar{\boldsymbol{q}}_{i-1}\|^2 + 2\sigma_{12}^2\mathbb{1}_{N-1}^\top\mathsf{P}[\check{\boldsymbol{q}}_{i-1}], \qquad (164)$$

where the first inequality is an application of Jensen's inequality, the second inequality comes from (137) and (138), and the final equality comes from property P5) of the energy operator $\mathsf{P}[\cdot]$. Taking expectations in (164) and using the result in (162) we obtain:

$$\mathbb{E}\|\bar{\boldsymbol{\delta}}_i\|^2 \le 2\,\mu^2\,\sigma_{11}^2\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2$$
$$+ 2\,\mu^2\,\sigma_{12}^2\mathbb{1}_{N-1}^\top\mathscr{P}[\check{\boldsymbol{q}}_{i-1}] + \mu^2\,\mathbb{E}\,\|\bar{\boldsymbol{s}}_i\|^2. \qquad (165)$$

Let us move on to examine (163). First of all, we appeal to the Jordan matrix representation in (111) to write:

$$\mathcal{J} - I_{M(N-1)} = (\Lambda - I_{N-1})\otimes I_M + U\otimes I_M \triangleq \mathcal{D} + \mathcal{U}. \qquad (166)$$

Then, the following chain of inequalities holds:

$$\mathsf{P}\big[(\mathcal{J} - I_{M(N-1)} - \mu\boldsymbol{G}_{22,i-1})\check{\boldsymbol{q}}_{i-1} - \mu\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1} - \mu\,\check{\boldsymbol{b}}\big]$$
$$= \mathsf{P}\big[(\mathcal{D} + \mathcal{U} - \mu\boldsymbol{G}_{22,i-1})\check{\boldsymbol{q}}_{i-1} - \mu\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1} - \mu\,\check{\boldsymbol{b}}\big]$$
$$\overset{(a)}{\preceq} 2\,\mathsf{P}[\mathcal{D}\,\check{\boldsymbol{q}}_{i-1}]$$
$$+ 2\,\mathsf{P}\big[(\mathcal{U} - \mu\boldsymbol{G}_{22,i-1})\check{\boldsymbol{q}}_{i-1} - \mu\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1} - \mu\,\check{\boldsymbol{b}}\big]$$
$$\overset{(b)}{\preceq} 2\,\mathsf{P}[\mathcal{D}\,\check{\boldsymbol{q}}_{i-1}] + 8\,\mathsf{P}[\mathcal{U}\,\check{\boldsymbol{q}}_{i-1}]$$
$$+ 8\mu^2\,\mathsf{P}[\boldsymbol{G}_{22,i-1}\check{\boldsymbol{q}}_{i-1}] + 8\mu^2\,\mathsf{P}[\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1}] + 8\mu^2\,\mathsf{P}[\,\check{\boldsymbol{b}}\,]$$
$$\overset{(c)}{\preceq} 2\,\|\bar{\mathsf{P}}[\mathcal{D}]\|_\infty\,\bar{\mathsf{P}}[\mathcal{D}]\,\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 8\,\|\bar{\mathsf{P}}[\mathcal{U}]\|_\infty\,\bar{\mathsf{P}}[\mathcal{U}]\,\mathsf{P}[\check{\boldsymbol{q}}_{i-1}]$$
$$+ 8\mu^2\,\|\bar{\mathsf{P}}[\boldsymbol{G}_{21,i-1}]\|_\infty^2\,\mathbb{1}_{N-1}\mathsf{P}[\bar{\boldsymbol{q}}_{i-1}]$$
$$+ 8\mu^2\,\|\bar{\mathsf{P}}[\boldsymbol{G}_{22,i-1}]\|_\infty^2\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 8\mu^2\,\mathsf{P}[\,\check{\boldsymbol{b}}\,]$$
$$\overset{(d)}{\preceq} 8(I_{N-1} + U)\mathsf{P}[\check{\boldsymbol{q}}_{i-1}]$$

$$
\begin{aligned}
&+ 8\mu^2 \,\|\bar{\mathsf{P}}[\boldsymbol{G}_{21,i-1}]\|_\infty^2 \,\mathbb{1}_{N-1}\mathsf{P}[\bar{\boldsymbol{q}}_{i-1}] \\
&+ 8\mu^2 \,\|\bar{\mathsf{P}}[\boldsymbol{G}_{22,i-1}]\|_\infty^2 \,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top \mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 8\mu^2\,\mathsf{P}[\,\check{b}\,] \\
&\overset{(e)}{\preceq} 8(I_{N-1}+U)\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 8\,\mu^2\sigma_{21}^2\mathbb{1}_{N-1}\mathsf{P}[\bar{\boldsymbol{q}}_{i-1}] \\
&+ 8\,\mu^2\sigma_{22}^2\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 8\,\mu^2\mathsf{P}[\,\check{b}\,],
\end{aligned}
\tag{167}
$$

where (a) follows by the convexity property P3) of the energy operator applied with weights $1/2$; (b) follows by the same property applied with weights $1/4$; (c) follows by property P6) of the energy operator, respectively in form (119) as regards the first two terms, and in form (120) as regards the remaining terms; (d) follows by observing that, due to the peculiar shape of $\mathcal{D}$ and $\mathcal{U}$, one has the identities:

$$
\|\bar{\mathsf{P}}[\mathcal{D}]\|_\infty = \|I_{N-1}-\Lambda\| = \max_{n=2,3,\ldots,N}|1-\lambda_n(A)|, \tag{168}
$$

$$
\bar{\mathsf{P}}[\mathcal{D}] \preceq \|I_{N-1}-\Lambda\|\, I_{N-1}, \tag{169}
$$

$$
\|\bar{\mathsf{P}}[\mathcal{U}]\|_\infty = 1, \quad \bar{\mathsf{P}}[\mathcal{U}] = U, \tag{170}
$$

and by the inequality:

$$
\max_{n=2,3,\ldots,N}|1-\lambda_n(A)| \le 1+|\lambda_2| < 2. \tag{171}
$$

Finally, the inequality in (e) follows by the bounds in (139).

Taking expectations in (167) and then using (163) we get:

$$
\begin{aligned}
&\mathscr{P}[\check{\boldsymbol{\delta}}_i] \preceq 8\,\mu^2\sigma_{21}^2\mathbb{1}_{N-1}\mathscr{P}[\bar{\boldsymbol{q}}_{i-1}] + 8(I_{N-1}+U)\mathscr{P}[\check{\boldsymbol{q}}_{i-1}] \\
&+ 8\,\mu^2\sigma_{22}^2\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\mathscr{P}[\check{\boldsymbol{q}}_{i-1}] + \mu^2\,\mathscr{P}[\check{\boldsymbol{s}}_i] + 8\,\mu^2\mathsf{P}[\,\check{b}\,].
\end{aligned}
\tag{172}
$$

Examining jointly (165) and (172), we see that we have in fact proved (147), with the matrix T_δ and the quantities $\bar{x}_\delta$ and $\check{x}_\delta$ as given in Table 2. ■

APPENDIX E
PROOF OF LEMMA 4

We start with an auxiliary lemma that will be then used to prove Lemma 4.

Lemma 5 (Quantized State Decomposition): Let

$$
\bar{\Delta} \triangleq \|V^{-1}\|^2 \max_{k=1,2,\ldots,N}\pi_k^2\,\omega_k, \tag{173}
$$

$$
\check{\Delta} \triangleq \|V^{-1}\|^2 \max_{\substack{\ell=2,3,\ldots,N\\ k=1,2,\ldots,N}}|v_{\ell k}|^2\,\omega_k, \tag{174}
$$

where $v_{\ell k}$ is the (ℓ,k)-entry of the transformation matrix V in (54). Then, for any $\zeta\in(0,1)$ we have that:

$$
\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}+\zeta\widehat{\boldsymbol{\delta}}_i] + \zeta^2\Delta\mathbb{1}_N^\top\mathscr{P}[\widehat{\boldsymbol{\delta}}_i], \tag{175}
$$

where

$$
\Delta \triangleq \begin{bmatrix} \bar{\Delta} & \\ & \check{\Delta}\mathbb{1}_{N-1} \end{bmatrix}. \tag{176}
$$

Proof: Adding and subtracting $\zeta\widehat{\boldsymbol{\delta}}_i$ in (122), and recalling that $\boldsymbol{\delta}_i = \mathcal{V}^{-1}\widehat{\boldsymbol{\delta}}_i$ we can write:

$$
\widehat{\boldsymbol{q}}_i = \underbrace{\widehat{\boldsymbol{q}}_{i-1}+\zeta\,\widehat{\boldsymbol{\delta}}_i}_{x} + \underbrace{\zeta\,\mathcal{V}\,(\boldsymbol{\mathcal{Q}}\,(\boldsymbol{\delta}_i)-\boldsymbol{\delta}_i)}_{y}. \tag{177}
$$

Now, we observe that:

$$
\begin{aligned}
\mathbb{E}\left[\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i)\middle|\widehat{\boldsymbol{q}}_{i-1},\widehat{\boldsymbol{\delta}}_i\right] &\overset{(a)}{=} \mathbb{E}\left[\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i)\middle|\widehat{\boldsymbol{q}}_{i-1},\widehat{\boldsymbol{\delta}}_i,\boldsymbol{\delta}_i\right] \\
&\overset{(b)}{=} \mathbb{E}\left[\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i)\middle|\boldsymbol{\delta}_i\right] \overset{(c)}{=} \boldsymbol{\delta}_i,
\end{aligned}
\tag{178}
$$

where (a) holds because $\boldsymbol{\delta}_i$ is a deterministic function of $\widehat{\boldsymbol{\delta}}_i$, while (b) and (c) follow from (29), once noting that $\widehat{\boldsymbol{q}}_{i-1}$ and $\widehat{\boldsymbol{\delta}}_i$ are deterministic functions of the past history $\boldsymbol{h}_i$ defined in (27). Using (178) in (177) we conclude that:

$$
\mathbb{E}\left[\boldsymbol{x}^{\mathsf{H}}\boldsymbol{y}\right] = 0, \tag{179}
$$

which allows us to apply property P4) of the energy operator in (177), yielding:

$$
\mathscr{P}\left[\widehat{\boldsymbol{q}}_i\right] = \mathscr{P}\left[\widehat{\boldsymbol{q}}_{i-1}+\zeta\widehat{\boldsymbol{\delta}}_i\right] + \zeta^2\mathscr{P}[\mathcal{V}\,(\boldsymbol{\mathcal{Q}}\,(\boldsymbol{\delta}_i)-\boldsymbol{\delta}_i)]\,. \tag{180}
$$

On the other hand, we can write:

$$
\begin{aligned}
&\mathcal{V}\,(\boldsymbol{\mathcal{Q}}(\boldsymbol{\delta}_i)-\boldsymbol{\delta}_i) \\
&= \begin{bmatrix} v_{11}I_M & \ldots & v_{1N}I_M \\ & \vdots & \\ v_{N1}I_M & \ldots & v_{NN}I_M \end{bmatrix}\begin{bmatrix} \boldsymbol{\mathcal{Q}}_1(\boldsymbol{\delta}_{1,i})-\boldsymbol{\delta}_{1,i} \\ \vdots \\ \boldsymbol{\mathcal{Q}}_N(\boldsymbol{\delta}_{N,i})-\boldsymbol{\delta}_{N,i} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^N v_{1k}\left(\boldsymbol{\mathcal{Q}}_k(\boldsymbol{\delta}_{k,i})-\boldsymbol{\delta}_{k,i}\right) \\ \vdots \\ \sum_{k=1}^N v_{Nk}\left(\boldsymbol{\mathcal{Q}}_k(\boldsymbol{\delta}_{k,i})-\boldsymbol{\delta}_{k,i}\right) \end{bmatrix}.
\end{aligned}
\tag{181}
$$

The expected energy of the ℓ-th block in (181) is:

$$
\begin{aligned}
&\mathbb{E}\left\|\sum_{k=1}^N v_{\ell k}\left(\boldsymbol{\mathcal{Q}}_k(\boldsymbol{\delta}_{k,i})-\boldsymbol{\delta}_{k,i}\right)\right\|^2 \\
&\overset{(a)}{=} \sum_{k=1}^N |v_{\ell k}|^2\,\mathbb{E}\left\|\boldsymbol{\mathcal{Q}}_k(\boldsymbol{\delta}_{k,i})-\boldsymbol{\delta}_{k,i}\right\|^2 \\
&\overset{(b)}{\le} \sum_{k=1}^N |v_{\ell k}|^2\,\omega_k\,\mathbb{E}\,\|\boldsymbol{\delta}_{k,i}\|^2,
\end{aligned}
\tag{182}
$$

where (a) follows from (31) and (b) from (30). Recalling that the first row of matrix V is the (transposed) Perron eigenvector, the first entry ($\ell=1$) in (182) can be upper bounded by:

$$
\max_{k=1,2,\ldots,N}\pi_k^2\,\omega_k\,\mathbb{E}\|\boldsymbol{\delta}_i\|^2 = \mathbb{E}\|\mathcal{V}^{-1}\widehat{\boldsymbol{\delta}}_i\|^2 \le \bar{\Delta}\,\mathbb{E}\|\widehat{\boldsymbol{\delta}}_i\|^2, \tag{183}
$$

where the last inequality follows by the definition of $\bar{\Delta}$ in (173). Likewise, the other entries ($\ell\neq 1$) in (182) can be

upper bounded by:

$$\max_{\substack{\ell=2,3,\ldots,N\\ k=1,2,\ldots,N}} |v_{\ell k}|^2\,\omega_k\;\mathbb{E}\|\boldsymbol{\delta}_i\|^2 \le \check{\Delta}\,\mathbb{E}\|\widehat{\boldsymbol{\delta}}_i\|^2, \tag{184}$$

having used the definition of $\check{\Delta}$ in (174). The claim of the lemma follows by joining (180), (182), (183) and (184), and using property P5) of the energy operator. ■

Proof of Lemma 4: The first term on the RHS in (175) can be represented in block form as follows:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}+\zeta\widehat{\boldsymbol{\delta}}_i] = \begin{bmatrix} \mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}+\zeta\bar{\boldsymbol{\delta}}_i\|^2 \\ \\ \mathscr{P}[\check{\boldsymbol{q}}_{i-1}+\zeta\check{\boldsymbol{\delta}}_i] \end{bmatrix}. \tag{185}$$

Let us start by examining the first block in (185). Exploiting the block decomposition in (136) we can write:

$$\begin{aligned}&\bar{\boldsymbol{q}}_{i-1}+\zeta\bar{\boldsymbol{\delta}}_i = \\ &= (I_M-\mu\,\zeta\,\boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}-\mu\,\zeta\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}-\mu\,\zeta\,\bar{\boldsymbol{s}}_i.\end{aligned} \tag{186}$$

First of all, using (34) we have the equality:

$$\begin{aligned}&\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}+\zeta\bar{\boldsymbol{\delta}}_i\|^2 = \mu^2\zeta^2\mathbb{E}\|\bar{\boldsymbol{s}}_i\|^2 \\ &+\mathbb{E}\|(I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}-\mu\zeta\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2.\end{aligned} \tag{187}$$

Let us now examine the spectral radius of $I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}$. Using (137) we can write:

$$\|I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}\|^2 \le \max\left\{(1-\mu\zeta\eta)^2,(1-\mu\zeta\nu)^2\right\}. \tag{188}$$

We have the following chain of equivalent relationships:[9]

$$\begin{aligned}(1-\mu\zeta\nu)^2 &> (1-\mu\zeta\eta)^2\\ &\Updownarrow\\ (\mu\zeta\nu)^2-2\mu\zeta\nu &> (\mu\zeta\eta)^2-2\mu\zeta\eta\\ &\Updownarrow\\ \mu\zeta\nu^2-2\nu &> \mu\zeta\eta^2-2\eta\\ &\Updownarrow\\ \mu\zeta\left(\eta^2-\nu^2\right) &< 2\,(\eta-\nu)\\ &\Updownarrow\\ &\text{Eq. (148)},\end{aligned} \tag{189}$$

where the last implication is true because $\nu\le\eta$ in view of (14). Since all implications in (189) hold in both directions, we have in fact proved that:

$$\|I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}\|^2 \le (1-\mu\zeta\nu)^2. \tag{190}$$

Moreover, since $\nu\le\eta$, we also have:

$$\frac{2}{\eta+\nu} < \frac{1}{\nu}, \tag{191}$$

which, using (148), implies that $1-\mu\zeta\nu>0$, finally yielding, in view of (190):

$$\|I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}\| \le 1-\mu\zeta\nu. \tag{192}$$

This upper bound will be useful in characterizing the last term in (187), which can be manipulated as follows, for $0<t<1$:

$$\begin{aligned}&\|(I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}-\mu\zeta\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2\\ &= \left\|(1-t)\frac{I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}}{1-t}\bar{\boldsymbol{q}}_{i-1}-t\frac{\mu\zeta\boldsymbol{G}_{12,i-1}}{t}\check{\boldsymbol{q}}_{i-1}\right\|^2\\ &\le \frac{\|(I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}\|^2}{1-t}+\mu^2\zeta^2\frac{\|\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2}{t}\\ &\le (1-\mu\zeta\nu)^2\frac{\|\bar{\boldsymbol{q}}_{i-1}\|^2}{1-t}+\mu^2\zeta^2\sigma_{12}^2\frac{\|\check{\boldsymbol{q}}_{i-1}\|^2}{t}\\ &= (1-\mu\zeta\nu)\|\bar{\boldsymbol{q}}_{i-1}\|^2+\mu\zeta\frac{\sigma_{12}^2}{\nu}\|\check{\boldsymbol{q}}_{i-1}\|^2,\end{aligned} \tag{193}$$

where the first inequality is an application of Jensen's inequality, whereas the second inequality follows by setting $t=\mu\zeta\nu$, and by using (138). Taking expectations in (193) and using the result in (187) we get:

$$\begin{aligned}\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}+\zeta\bar{\boldsymbol{\delta}}_i\|^2 &\le (1-\mu\zeta\nu)\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2\\ &\quad+\mu\zeta\frac{\sigma_{12}^2}{\nu}\mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2+\mu^2\zeta^2\mathbb{E}\|\bar{\boldsymbol{s}}_i\|^2.\end{aligned} \tag{194}$$

We continue by examining the second block in (185). Using the block decomposition in (136) we can write:

$$\begin{aligned}&\check{\boldsymbol{q}}_{i-1}+\zeta\check{\boldsymbol{\delta}}_i = (1-\zeta)\check{\boldsymbol{q}}_{i-1}\\ &+\zeta\left(\underbrace{-\mu\,\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1}+(\mathcal{J}-\mu\,\boldsymbol{G}_{22,i-1})\check{\boldsymbol{q}}_{i-1}-\mu\,\check{b}}_{\boldsymbol{y}}-\mu\,\check{\boldsymbol{s}}_i\right),\end{aligned} \tag{195}$$

and, using (34) along with property P4) of the energy operator, we get:

$$\mathscr{P}[\check{\boldsymbol{q}}_{i-1}+\zeta\check{\boldsymbol{\delta}}_i] = \mu^2\zeta^2\mathscr{P}[\check{\boldsymbol{s}}_i]+\mathscr{P}[(1-\zeta)\check{\boldsymbol{q}}_{i-1}+\zeta\boldsymbol{y}]. \tag{196}$$

On the other hand, by the convexity property P3) of the energy operator we have:

$$\mathsf{P}[(1-\zeta)\check{\boldsymbol{q}}_{i-1}+\zeta\boldsymbol{y}] \le (1-\zeta)\mathsf{P}[\check{\boldsymbol{q}}_{i-1}]+\zeta\,\mathsf{P}[\boldsymbol{y}]. \tag{197}$$

[9] Condition (148) is not the tightest condition one can use to guarantee stability of $I_M-\mu\zeta\boldsymbol{G}_{11,i-1}$. Some examples of how to get a better constant can be found in [14], [21], and, with straightforward algebra, we can get the refined upper bound $\|I_M-\mu\zeta\,\boldsymbol{G}_{11,i-1}\|\le 1-\mu\zeta\nu+1/2(\mu\zeta\eta)^2$. However, in our analysis the additional $O(\mu^2)$ term is expected to bring little information. In fact, as we will see in Lemma 7 further ahead, a number of $O(\mu^2)$ terms will be collected into a large correction constant ϕ that characterizes the stability analysis, with a stability threshold $\mu^\star$ being usually smaller than the factor $2/(\eta+\nu)$ that will be obtained from our characterization of $I_M-\mu\zeta\boldsymbol{G}_{11,i-1}$.

Making explicit the definition of $\boldsymbol{y}$ in (195), we can write the following chain of inequalities:

$$\begin{aligned}
\mathsf{P}[\boldsymbol{y}] &= \mathsf{P}\left[\mathcal{J}\check{\boldsymbol{q}}_{i-1} - \mu\,\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_{i-1} - \mu\,\boldsymbol{G}_{22,i-1}\check{\boldsymbol{q}}_{i-1} - \mu\,\check{b}\right] \\
&\overset{(a)}{\preceq} \underbrace{\left(\frac{\Lambda\Lambda^*}{|\lambda_2|} + \frac{2\,U}{1-|\lambda_2|}\right)}_{\Gamma}\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] \\
&\quad + \frac{2\mu^2}{1-|\lambda_2|}\mathsf{P}\left[\boldsymbol{G}_{21,i-1}\bar{\boldsymbol{q}}_i + \boldsymbol{G}_{22,i-1}\check{\boldsymbol{q}}_i + \check{b}\right] \\
&\overset{(b)}{\preceq} \Gamma\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + \frac{2\mu^2}{1-|\lambda_2|} \\
&\quad \times \left(3\|\bar{\mathsf{P}}[\boldsymbol{G}_{21,i-1}]\|_\infty^2\,\mathbb{1}_{N-1}\|\bar{\boldsymbol{q}}_{i-1}\|^2\right. \\
&\quad \left. + 3\|\bar{\mathsf{P}}[\boldsymbol{G}_{22,i-1}]\|_\infty^2\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + 3\mathsf{P}[\check{b}]\right) \\
&\overset{(c)}{\preceq} \Gamma\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + \frac{6\mu^2}{1-|\lambda_2|} \\
&\quad \times \left(\sigma_{21}^2\mathbb{1}_{N-1}\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \sigma_{22}^2\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\mathsf{P}[\check{\boldsymbol{q}}_{i-1}] + \mathsf{P}[\check{b}]\right),
\end{aligned} \tag{198}$$

where step (a) applies property P7) of the energy operator, step (b) applies property P3) with weights $1/3$ and property P6), and step (c) uses (139).

Letting

$$E_0 = \left((1-\zeta)I_{N-1} + \zeta\,\frac{\Lambda\Lambda^*}{|\lambda_2|}\right) + \frac{2\zeta}{1-|\lambda_2|}\,U, \tag{199}$$

taking expectations in (198), and using the result in (197) and then in (195), we obtain:

$$\begin{aligned}
\mathscr{P}[\check{\boldsymbol{q}}_{i-1} + \zeta\check{\boldsymbol{\delta}}_i] &\preceq \mu^2\zeta\,\frac{6\,\sigma_{21}^2}{1-|\lambda_2|}\,\mathbb{1}_{N-1}\,\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 \\
&\quad + \left(E_0 + \mu^2\zeta\,\frac{6\,\sigma_{22}^2}{1-|\lambda_2|}\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top\right)\mathscr{P}[\check{\boldsymbol{q}}_{i-1}] \\
&\quad + \mu^2\zeta\,\frac{6}{1-|\lambda_2|}\mathsf{P}[\check{b}] + \mu^2\zeta^2\mathscr{P}[\check{\boldsymbol{s}}_i].
\end{aligned} \tag{200}$$

Calling upon Lemma 5 along with (194) and (200), we see that (150) holds true, with the matrix T_q and the quantities $\bar{x}_q$ and $\check{x}_q$ as given in Table 2. ■

APPENDIX F
PROOF OF THEOREM 3

Proof: In (150), we can replace $\mathscr{P}[\widehat{\boldsymbol{\delta}}_i]$ with the RHS of (147) to get an inequality that relates $\mathscr{P}[\widehat{\boldsymbol{q}}_i]$ to $\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}]$ and $\mathscr{P}[\mu\,\widehat{\boldsymbol{s}}_i]$. Replacing now $\mathscr{P}[\mu\,\widehat{\boldsymbol{s}}_i]$ with the RHS of (146) we get the inequality recursion in (156), with the driving vector x defined in (155) and with the matrix T replaced by the matrix:

$$T_{\text{tmp}} = T_q + \zeta^2\Delta\,\mathbb{1}_N^\top\,T_\delta + \zeta^2\left(\Delta\,\mathbb{1}_N^\top + I_N\right)T_s. \tag{201}$$

The claim of the theorem will be proved if we show that the matrix T_{tmp} is upper bounded by the matrix T appearing in (151).

The term T_q can be upper bounded as:

$$T_q \preceq \begin{bmatrix} 1-\mu\,\zeta\,\nu & \mu\,\zeta\,\dfrac{\sigma_{12}^2}{\nu}\mathbb{1}_{N-1}^\top \\ \phi^{(q)}\,\mu^2\,\zeta\,\mathbb{1}_{N-1} & E_0 + \phi^{(q)}\,\mu^2\,\zeta\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top \end{bmatrix}, \tag{202}$$

where for brevity we introduced the bounding constant $\phi^{(q)}$.

Likewise, concerning the term involving T_s we can write:

$$\left(\Delta\,\mathbb{1}_N^\top + I_N\right)T_s \preceq \phi^{(s)}\mu^2\,\mathbb{1}_N\mathbb{1}_N^\top, \tag{203}$$

where $\phi^{(s)}$ is a suitable constant.

Regarding the term involving T_δ, by exploiting Table 2 we have:

$$\begin{aligned}
\Delta\mathbb{1}_N^\top\,T_\delta &= \begin{bmatrix} \bar{\Delta} & \bar{\Delta}\,\mathbb{1}_{N-1}^\top \\ \check{\Delta}\,\mathbb{1}_{N-1} & \check{\Delta}\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top \end{bmatrix}\begin{bmatrix} [T_\delta]_{11} & [T_\delta]_{12} \\ [T_\delta]_{21} & [T_\delta]_{22} \end{bmatrix} \\
&\preceq \begin{bmatrix} 0 & 8\,\bar{\Delta}\,\mathbb{1}_{N-1}^\top(I_{N-1}+U) \\ 0 & 8\,\check{\Delta}\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top(I_{N-1}+U) \end{bmatrix} + \phi^{(\delta)}\mu^2\,\mathbb{1}_N\mathbb{1}_N^\top \\
&\preceq \begin{bmatrix} 0 & 16\,\bar{\Delta}\,\mathbb{1}_{N-1}^\top \\ 0 & 16\,\check{\Delta}\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top \end{bmatrix} + \phi^{(\delta)}\mu^2\,\mathbb{1}_N\mathbb{1}_N^\top,
\end{aligned} \tag{204}$$

where we used the bound:

$$\mathbb{1}_{N-1}^\top(I_{N-1}+U) \preceq 2\,\mathbb{1}_{N-1}^\top, \tag{205}$$

and where $\phi^{(\delta)}$ is a suitable constant that upper bounds the terms of order μ^2.

If we now introduce the maximal constant

$$\phi = \max\left\{\phi^{(s)}, \phi^{(\delta)}, \phi^{(q)}\right\}, \tag{206}$$

and the rank-one perturbed version of E_0:

$$E = E_0 + 16\,\zeta^2\,\check{\Delta}\,\mathbb{1}_{N-1}\mathbb{1}_{N-1}^\top, \tag{207}$$

by using (202), (203), and (204) in (201), we end up with the following bound:

$$T_{\text{tmp}} \preceq \begin{bmatrix} \tau & \tau_{12}\mathbb{1}_{N-1}^\top \\ 0 & E \end{bmatrix} + v_{\mu,\zeta}\,\mathbb{1}_N^\top, \tag{208}$$

where the quantities τ, τ_{12}, and $v_{\mu,\zeta}$ are defined in (152) and (154), respectively. The claim of the theorem now follows from the definition of T in (151). ■

APPENDIX G
STABILITY OF E

Lemma 6 (Stability of E): Let

$$a_n \triangleq \frac{2|\lambda_2|}{(1-|\lambda_2|)}\,\frac{1}{|\lambda_2|-|\lambda_n|^2}. \tag{209}$$

We note that $a_n \geq 2$. Let further

$$\gamma(A) \triangleq \sum_{n=2}^{B} \frac{|\lambda_2|}{|\lambda_2|-|\lambda_n|^2}\left(\frac{a_n^{L_n+1}-1}{(a_n-1)^2}-\frac{L_n+1}{a_n-1}\right). \quad (210)$$

Then, the matrix E defined in (153) has spectral radius less than 1 if, and only if:

$$\zeta < \frac{1}{16\check{\Delta}\gamma(A)}. \quad (211)$$

Proof: We introduce the resolvent of matrix E:

$$\mathcal{R}_E(z) = (zI_{N-1}-E)^{-1}, \quad (212)$$

which is well-posed for z distinct from the eigenvalues of E. The stability of E will be proved if we show that $zI_{N-1}-E$ is invertible for all $|z| \geq 1$. To this end, we examine first the resolvent of the *unperturbed* matrix E_0 in Table 1 — see also (199), namely,

$$\mathcal{R}_{E_0}(z) = (zI_{N-1}-E_0)^{-1}. \quad (213)$$

Since E_0 is upper triangular and all its diagonal elements are positive values strictly less than one, we conclude that E_0 is stable, which further implies that the resolvent $\mathcal{R}_{E_0}(z)$ is bounded for $|z| \geq 1$.

We continue by relating the resolvent of E to the resolvent of the *unperturbed* matrix E_0. Exploiting the structure of E in (207), we see that E is given by E_0 plus an additive rank-one perturbation. Since we have shown that in the range of interest $|z| \geq 1$ the matrix $zI_{N-1}-E_0$ is invertible, we can apply the Sherman-Morrison identity to $zI_{N-1}-E$, obtaining [66]:

$$\mathcal{R}_E(z) = \mathcal{R}_{E_0}(z) + 16\,\zeta^2\check{\Delta}\,\frac{\mathcal{R}_{E_0}(z)\mathbb{1}_{N-1}\mathbb{1}_{N-1}^{\top}\mathcal{R}_{E_0}(z)}{1-16\,\zeta^2\,\check{\Delta}\,\mathbb{1}_{N-1}^{\top}\mathcal{R}_{E_0}(z)\mathbb{1}_{N-1}}, \quad (214)$$

where the identity holds if, and only if, the denominator on the RHS of (214) is not zero. In particular, if the denominator is zero $zI_{N-1}-E$ is not invertible. Therefore, to prove that $zI_{N-1}-E$ is invertible, we must examine the behavior of the complex scalar function:

$$g_E(z) \triangleq \mathbb{1}_{N-1}^{\top}\mathcal{R}_{E_0}(z)\mathbb{1}_{N-1}, \quad (215)$$

over the entire range $|z| \geq 1$. To this end, it is critical to characterize the resolvent of E_0. Exploiting (109), (110), and (111), we can represent E_0 as:

$$E_0 = \mathsf{diag}\left(E_2, E_3, \ldots, E_B\right), \quad (216)$$

where, for $n = 2, 3, \ldots, B$, we introduced the $L_n \times L_n$ matrices:

$$E_n = \underbrace{\left((1-\zeta)+\zeta\,\frac{|\lambda_n|^2}{|\lambda_2|}\right)}_{r_n} I_{L_n} + \underbrace{\frac{2\zeta}{1-|\lambda_2|}}_{u_n} U_{L_n}. \quad (217)$$

By computing the inverse of the block-diagonal matrix $zI_{N-1}-E_0$ as the block-diagonal matrix of the individual inverse matrix-blocks, we have that:

$$\mathcal{R}_{E_0}(z) = \mathsf{diag}\left(\mathcal{R}_{E_2}(z), \mathcal{R}_{E_3}(z), \ldots, \mathcal{R}_{E_B}(z)\right). \quad (218)$$

On the other hand, for Jordan-type matrices like $zI_{N-1}-E_n$, the inverse is known to be in the form:

$$\mathcal{R}_{E_n}(z) = \frac{1}{z-r_n}\begin{bmatrix} 1 & \frac{u_n}{z-r_n} & \ldots & \left(\frac{u_n}{z-r_n}\right)^{L_n-1} \\ 0 & 1 & \ldots & \left(\frac{u_n}{z-r_n}\right)^{L_n-2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}. \quad (219)$$

Using (218) and (219) in (215), we obtain:

$$g_E(z) = \sum_{n=2}^{B}\frac{1}{z-r_n}\sum_{j=0}^{L_n-1}(L_n-j)\left(\frac{u_n}{z-r_n}\right)^j. \quad (220)$$

Applying the triangle inequality in (220) we get:

$$|g_E(z)| \leq \sum_{n=2}^{B}\frac{1}{|z-r_n|}\sum_{j=0}^{L_n-1}(L_n-j)\left(\frac{u_n}{|z-r_n|}\right)^j. \quad (221)$$

In view of the reverse triangle inequality we have:

$$|z-r_n| \geq \big|\,|z|-r_n\,\big| = |z|-r_n \geq 1-r_n, \quad (222)$$

where the last inequality holds because $|z| \geq 1$ in the considered range, whereas the equality follows by considering also that $r_n \leq 1$. Using (222) in (221) we can write the following inequality (in the range $|z| \geq 1$):

$$|g_E(z)| \leq \sum_{n=2}^{B}\frac{1}{1-r_n}\sum_{j=0}^{L_n-1}(L_n-j)\left(\frac{u_n}{1-r_n}\right)^j = g_E(1). \quad (223)$$

Equation (223) implies that a sufficient condition for the stability of E is:

$$16\,\zeta^2\,\check{\Delta}g_E(1) < 1. \quad (224)$$

We now show this is also a necessary condition by reductio ad absurdum. Assume that (224) is violated, namely that (we rule out the equality since it obviously correspond to instability):

$$16\zeta^2\check{\Delta}g_E(1) > 1. \quad (225)$$

Were (225) true, there would certainly exist one value $z^\star \in \mathbb{R}$, with $z^\star > 1$, such that the denominator on the RHS of (214) is equal to zero. This is because the function $g_E(z)$ is analytic over the domain $|z| \geq 1$ and, in particular, it is continuous on the real axis and vanishes as $z \to \infty$. This implies that (224) is a necessary and sufficient condition for the stability of E.

Let us now recast (224) in a more explicit form. To this end, we introduce the quantity:

$$a_n = \frac{u_n}{1-r_n} = \frac{2|\lambda_2|}{(1-|\lambda_2|)}\,\frac{1}{|\lambda_2|-|\lambda_n|^2}, \quad (226)$$

which is readily verified to be greater than or equal to 2. The inner summation in (223) can be written as:

$$\sum_{j=0}^{L_n-1}(L_n-j)a_n^j=\frac{a_n^{L_n+1}-1}{(a_n-1)^2}-\frac{L_n+1}{a_n-1}, \tag{227}$$

where the equality follows from standard results on geometric series. Using (227) in (223), and making explicit the definition of r_n, we have that:

$$g_E(1)=\frac{\gamma(A)}{\zeta}, \tag{228}$$

where $\gamma(A)$ is defined in (210). In view of (214), (215), and (224), Eq. (228) implies that the matrix E is stable if, and only if, Eq. (211) is verified. ■

APPENDIX H
STABILITY OF T

Lemma 7 (Stability of T): Let

$$\zeta<\frac{1}{16\check{\Delta}\gamma(A)},\quad \mu<\frac{2}{\zeta\,(\eta+\nu)}, \tag{229}$$

and let $\mu^\star$ be the positive root of the equation:

$$\mu^2\left(1+\frac{\sigma_{12}^2}{\nu^2}\right)\varphi+\mu\,\zeta\,\frac{1+16\bar{\Delta}\varphi}{\nu}-\frac{1}{\phi}=0, \tag{230}$$

where

$$\varphi\triangleq\frac{\gamma(A)}{1-16\,\zeta\,\check{\Delta}\,\gamma(A)}. \tag{231}$$

Then, the matrix T in (151) has spectral radius less than 1 if, and only if, $\mu<\mu^\star$.

Proof: The matrix T_0 in (151) is stable since $\tau<1$ in view of the second inequality in (229), and E is stable in view of the first inequality in (229) and Lemma 6. Then, the eigenvalues of T_0 lie all strictly within the unit disc, and, hence, the resolvent $\mathcal{R}_{T_0}(z)$ exists. Accordingly, considering the resolvent of matrix T:

$$\mathcal{R}_T(z)=(zI_N-T)^{-1}, \tag{232}$$

and exploiting the structure of T in (151) (i.e., T_0 plus an additive rank-one perturbation), from the Sherman-Morrison identity we have [66]:

$$\mathcal{R}_T(z)=\mathcal{R}_{T_0}(z)+\frac{\mathcal{R}_{T_0}(z)\,v_{\mu,\zeta}\,\mathbb{1}_N^\top\mathcal{R}_{T_0}(z)}{1-\mathbb{1}_N^\top\mathcal{R}_{T_0}(z)\,v_{\mu,\zeta}}, \tag{233}$$

where the formula is valid if, and only if, the denominator on the RHS of (233) is not zero. Moreover, if the denominator is zero, then (zI_N-T) is not invertible. Therefore, the stability of T will be proved if we show that the denominator on the RHS of (233) is not zero over the range $|z|\geq 1$. To this end, we will now examine the resolvent of the unperturbed matrix T_0.

From (151) we see that T_0 is block upper-triangular, which implies that we can compute the inverse $\mathcal{R}_{T_0}(z)=(zI_N-T_0)^{-1}$ as:

$$\mathcal{R}_{T_0}(z)=\begin{bmatrix}\dfrac{1}{z-\tau} & \dfrac{\tau_{12}\mathbb{1}_{N-1}^\top\mathcal{R}_E(z)}{z-\tau}\\[2ex] 0 & \mathcal{R}_E(z)\end{bmatrix}, \tag{234}$$

and, exploiting the definition of $v_{\mu,\zeta}$ in (154) we get:

$$g_T(z)\triangleq\mathbb{1}_N^\top\mathcal{R}_{T_0}(z)\,v_{\mu,\zeta}=\phi\mu^2\times\left(\frac{\zeta^2}{z-\tau}+\frac{\zeta\,\tau_{12}\mathbb{1}_{N-1}^\top\mathcal{R}_E(z)\mathbb{1}_{N-1}}{z-\tau}+\zeta\,\mathbb{1}_{N-1}^\top\mathcal{R}_E(z)\mathbb{1}_{N-1}\right). \tag{235}$$

Using (214) and (215), we have the following identity:

$$\begin{aligned}\mathbb{1}_{N-1}^\top\mathcal{R}_E(z)\mathbb{1}_{N-1}&=g_E(z)+16\,\zeta^2\,\check{\Delta}\,\frac{g_E^2(z)}{1-16\,\zeta^2\,\check{\Delta}\,g_E(z)}\\&=\frac{g_E(z)}{1-16\,\zeta^2\,\check{\Delta}\,g_E(z)},\end{aligned} \tag{236}$$

which, applied in (235), yields:

$$g_T(z)=\phi\mu^2\left\{\frac{\zeta^2}{z-\tau}+\zeta\left(\frac{\tau_{12}}{z-\tau}+1\right)\frac{g_E(z)}{1-16\,\zeta^2\,\check{\Delta}\,g_E(z)}\right\}. \tag{237}$$

Accordingly, by triangle inequality we have that:

$$\begin{aligned}|g_T(z)|&\leq\phi\mu^2\\&\quad\times\left\{\frac{\zeta^2}{|z-\tau|}+\zeta\left(\frac{\tau_{12}}{|z-\tau|}+1\right)\frac{|g_E(z)|}{|1-16\,\zeta^2\,\check{\Delta}\,g_E(z)|}\right\}\\&\leq\phi\mu^2\left\{\frac{\zeta^2}{|z|-\tau}+\zeta\left(\frac{\tau_{12}}{|z|-\tau}+1\right)\frac{|g_E(z)|}{1-16\,\zeta^2\,\check{\Delta}\,|g_E(z)|}\right\}.\end{aligned} \tag{238}$$

where we used the known inequality $|z-\tau|\geq||z|-\tau|$ (which turns into $|z-\tau|\geq|z|-\tau$ since $|z|\geq 1$ over the considered range and $\tau<1$ in view of (229)) and the fact that, since (211) is verified by hypothesis, the denominator of the last fraction in (238) is positive in view of (223) and (228). On the other hand, we know from (223) that $|g_E(z)|\leq g_E(1)$ in the range $|z|\geq 1$, which, applied in (238), allows us to write:

$$\begin{aligned}|g_T(z)|&\leq g_T(1)\\&=\phi\mu^2\left\{\frac{\zeta^2}{1-\tau}+\zeta\left(\frac{\tau_{12}}{1-\tau}+1\right)\frac{g_E(1)}{1-16\,\zeta^2\,\check{\Delta}\,g_E(1)}\right\}.\end{aligned} \tag{239}$$

Reasoning as done in the proof of Lemma 6, a necessary and sufficient condition for the stability of T is:

$$g_T(1)<1. \tag{240}$$

To this aim, let us apply (152), (228), and (231) in (239), to obtain:

$$g_T(1) = \phi \left\{ \mu^2 \left(1 + \frac{\sigma_{12}^2}{\nu^2} \right) \varphi + \mu\, \zeta\, \frac{1 + 16 \bar{\Delta} \varphi}{\nu} \right\}. \quad (241)$$

In view of (241), inequality (240) is true if, and only if, the (positive) step-size μ is smaller than the positive root of the equation in (230), and the proof of the lemma is complete. ■

1. BOUNDS ON THE POWERS OF T

The stability established in Lemma 7 allows to conclude that the matrix powers T^i can be uniformly bounded with respect to i. However, the bound would depend on the matrix T, and, in particular, would depend on the step-size μ. Since we are interested in characterizing the small-μ behavior of the ACTC mean-square-deviation, it is essential to establish how such bound behaves as $\mu \to 0$. To this end, Lemma 7 alone does not provide enough information, and we need to resort to the powerful framework of Kreiss stability [76].

Preliminarily, it is necessary to introduce the concept of *Kreiss constant*. Given the resolvent $\mathcal{R}_X(z)$ associated with a matrix X, the Kreiss constant relative to X is defined as follows [76]:

$$K_X \triangleq \sup_{z \in \mathbb{C}: |z|>1} (|z| - 1)\, \|\mathcal{R}_X(z)\|, \quad (242)$$

and it is useful to bound (from above and from below) the norm of matrix powers as follows:

$$K_X \leq \sup_{i \geq 0} \|X^i\| \leq N e K_X, \quad (243)$$

where e is Euler's number. In the next lemma we exploit the Kreiss constant to characterize the small-μ behavior of the powers of T.

Lemma 8 (Bound on the Powers of T:) Let

$$\zeta < \frac{1}{16 \check{\Delta} \gamma(A)}, \quad \mu < \min \left\{ \frac{2}{\zeta\, (\eta + \nu)},\, \mu^\star \right\}, \quad (244)$$

where $\mu^\star$ is the positive root of the equation in (230). Then we have that:

$$\sup_{i \geq 0} \|T^i\|_1 \leq K(\mu), \quad (245)$$

where $K(\mu)$ is a function of μ, independent of i, with:

$$K(\mu) = O(1) \text{ as } \mu \to 0. \quad (246)$$

Proof: Let us evaluate a bound on the Kreiss constant associated with matrix T. Accordingly, we will examine the behavior of the function

$$f(z) = (|z| - 1) \|\mathcal{R}_T(z)\|_1, \quad (247)$$

over the range $|z| \geq 1$. We have seen in the proof of Lemma 7 — see the argument following (233) — that under (244) it is legitimate to use the representation in (233). Applying now the triangle inequality to (233) we have:

$$\|\mathcal{R}_T(z)\|_1 \leq \|\mathcal{R}_{T_0}(z)\|_1 + \frac{\|\mathcal{R}_{T_0}(z)\|_1^2 \times \|v_{\mu,\zeta} \mathbb{1}_N^\top\|_1}{|1 - \mathbb{1}_N^\top \mathcal{R}_{T_0}(z)\, v_{\mu,\zeta}|}$$

$$\leq \|\mathcal{R}_{T_0}(z)\|_1 + O(\mu^2)\, \frac{\|\mathcal{R}_{T_0}(z)\|_1^2}{1 - g_T(1)}, \quad (248)$$

where the $O(\mu^2)$ term comes from the behavior of the perturbation vector $v_{\mu,\zeta}$ in (154), whereas the bound involving the term $g_T(1)$ comes from (235) and (239). Therefore, from (247) and (248) we conclude that:

$$f(z) \leq (|z| - 1) \|\mathcal{R}_{T_0}(z)\|_1 + O(\mu^2)\, \frac{(|z| - 1) \|\mathcal{R}_{T_0}(z)\|_1^2}{1 - g_T(1)}. \quad (249)$$

Let us examine the behavior of $\|\mathcal{R}_{T_0}(z)\|_1$. Exploiting the structure in (234) and applying the triangle inequality, we can write:

$$\|\mathcal{R}_{T_0}(z)\|_1 \leq \max \left\{ \frac{1}{|z - \tau|},\, \frac{N \tau_{12} \|\mathcal{R}_E(z)\|_1}{|z - \tau|} + \|\mathcal{R}_E(z)\|_1 \right\}. \quad (250)$$

First, we examine the resolvent of matrix E. Since by assumption Eq. (211) is verified, the spectral radius of E is strictly less than 1 in view of Lemma 6. This implies that all eigenvalues of E lie strictly inside in the unit disc, which in turn guarantees the existence of a constant C_E such that [66]:

$$\sup_{z \in \mathbb{C}: |z|>1} \|\mathcal{R}_E(z)\|_1 = C_E < \infty. \quad (251)$$

Moreover, the stability of E implies that all powers of E are bounded, which, in view of the lower bound in (243), implies the existence of a finite Kreiss constant:[10]

$$\sup_{z \in \mathbb{C}: |z|>1} (|z| - 1) \|\mathcal{R}_E(z)\|_1 = K_E' < \infty. \quad (252)$$

We remark that both constants C_E and K_E' are independent of μ, since so is matrix E.

Let us focus on the second term in (249). Using (250), we can write:

$$(|z| - 1) \|\mathcal{R}_{T_0}(z)\|_1^2 \leq (|z| - 1) \times \max \left\{ \frac{1}{|z - \tau|^2},\, \left(\frac{N \tau_{12} \|\mathcal{R}_E(z)\|_1}{|z - \tau|} + \|\mathcal{R}_E(z)\|_1 \right)^2 \right\}. \quad (253)$$

Let us examine the first argument of the maximum in (253). Applying, as done before, the inequality $|z - \tau| \geq |z| - \tau$, we can write:

$$\frac{|z| - 1}{|z - \tau|^2} \leq \underbrace{\frac{|z| - 1}{|z| - \tau}}_{\leq 1 \text{ since } \tau < 1} \times \underbrace{\frac{1}{|z| - \tau}}_{\substack{\leq 1/(1-\tau) \\ \text{since} |z| \geq 1}} \leq \frac{1}{1 - \tau} = \frac{1}{\mu\, \zeta\, \nu}, \quad (254)$$

where the equality comes from the definition of τ in (152).

[10] We use the symbol K_E' in place of K_E since we are working with the ℓ_1-norm, while the definition of Kreiss constant uses the ℓ_2-norm.

Let us switch to the analysis of the second term in (253), which, by expanding the square, yields:

$$N^2\tau_{12}^2 \underbrace{\frac{|z|-1}{|z-\tau|^2}}_{\substack{\leq 1/(\mu\zeta\nu)\\ \text{see (254)}}} \underbrace{\|\mathcal{R}_E(z)\|_1^2}_{\leq C_E^2,\text{see (251)}} + \underbrace{(|z|-1)\,\|\mathcal{R}_E(z)\|_1^2}_{\leq C_E K_E',\text{see (251) and (252)}}$$

$$+ 2N\tau_{12} \underbrace{\frac{|z|-1}{|z-\tau|}}_{\leq 1} \underbrace{\|\mathcal{R}_E(z)\|_1^2}_{\leq C_E^2}$$

$$\leq \frac{N^2\tau_{12}^2 C_E^2}{\mu\,\zeta\,\nu} + C_E K_E' + 2N\tau_{12}C_E^2 = O(1/\mu). \tag{255}$$

Using (254) and (255) in (253), we conclude that:

$$\sup_{z\in\mathbb{C}:|z|>1} (|z|-1)\,\|\mathcal{R}_{T_0}(z)\|_1^2 = O(1/\mu). \tag{256}$$

Reasoning along the same lines we can show that:

$$\sup_{z\in\mathbb{C}:|z|>1} (|z|-1)\,\|\mathcal{R}_{T_0}(z)\|_1 = O(1). \tag{257}$$

Applying now (256) and (257) in (249), we get:

$$f(z) = O(1) + \frac{O(\mu)}{1-g_T(1)}, \tag{258}$$

which implies, in view of (242), the existence of a function $K(\mu)$ such that, under assumption (244), we are allowed to write:

$$\sup_{i\geq 0} \|T^i\|_1 \leq K(\mu) = O(1) + \frac{O(\mu)}{1-g_T(1)}. \tag{259}$$

From the properties of $g_T(1)$ examined in Lemma 7, we know that $g_T(1) < 1$ in the range of μ permitted by (244), and $g_T(1) \to 0$ as $\mu \to 0$, which implies the claim of the lemma (246). ■

APPENDIX I
PROOF OF THEOREM 1

Proof: Developing the recursion in (156) we have:

$$\mathscr{P}[\widehat{\boldsymbol{q}}_i] \preceq T^i\,\mathscr{P}[\widehat{\boldsymbol{q}}_0] + (I_N - T)^{-1}x, \tag{260}$$

and by application of the triangle inequality:

$$\mathbb{E}\|\widehat{\boldsymbol{q}}_i\|^2 = \|\mathscr{P}[\widehat{\boldsymbol{q}}_i]\|_1 \leq \|T^i\|_1\,\|\mathscr{P}[\widehat{\boldsymbol{q}}_0]\|_1 + \|(I_N-T)^{-1}x\|_1, \tag{261}$$

where the first equality comes from the definition of average energy operator in (114). In view of Lemma 7, under the assumptions of the theorem the matrix T has spectral radius strictly less than 1, which, in view of (261), implies:

$$\limsup_{i\to\infty} \mathbb{E}\|\widehat{\boldsymbol{q}}_i\|^2 \leq \|(I_N-T)^{-1}x\|_1 < \infty. \tag{262}$$

On the other hand, from the network coordinate transformation in (56), we have $\widetilde{\boldsymbol{q}}_i = \mathcal{V}^{-1}\widehat{\boldsymbol{q}}_i$, which, in view of (262), implies:

$$\limsup_{i\to\infty} \mathbb{E}\|\widetilde{\boldsymbol{q}}_{k,i}\|^2 < \infty, \tag{263}$$

where we used the fact that the squared norm of the *extended* vector $\widetilde{\boldsymbol{q}}_i$ is the sum of the norms of the N individual vectors $\|\widetilde{\boldsymbol{q}}_{k,i}\|^2$. The claim in (74) follows from the fact that $\widetilde{\boldsymbol{w}}_{k,i}$ is a convex combination of $\{\widetilde{\boldsymbol{q}}_{\ell,i}\}_{\ell\in\mathcal{N}_k}$ — see (43).

We move on to prove (75), for which we need to examine the small-μ behavior of $(I_N - T)^{-1}$ in (260). To this end, let us specialize (233) to the case $z = 1$:

$$(I_N - T)^{-1} = \mathcal{R}_T(1) = (I_N - T_0)^{-1} + \frac{(I_N-T_0)^{-1}\,v_{\mu,\zeta}\,\mathbb{1}_N^\top\,(I_N-T_0)^{-1}}{1-\mathbb{1}_N^\top\,(I_N-T_0)^{-1}\,v_{\mu,\zeta}}. \tag{264}$$

On the other hand, specializing (234) to the case $z = 1$, it is immediate to see that:

$$(I_N - T_0)^{-1} \preceq \begin{bmatrix} O(1/\mu) & O(1/\mu)\,\mathbb{1}_{N-1}^\top \\ 0 & O(1)\,\mathbb{1}_{N-1}\,\mathbb{1}_{N-1}^\top \end{bmatrix}. \tag{265}$$

Therefore we can write:

$$(I_N - T_0)^{-1}\,v_{\mu,\zeta} \preceq \begin{bmatrix} O(\mu) \\ O(\mu^2)\,\mathbb{1}_{N-1} \end{bmatrix}, \tag{266}$$

$$\mathbb{1}^\top\,(I_N - T_0)^{-1} \preceq \begin{bmatrix} O(1/\mu) & O(1/\mu)\mathbb{1}_{N-1}^\top \end{bmatrix}, \tag{267}$$

which further implies:

$$(I_N - T_0)^{-1}\,v_{\mu,\zeta}\,\mathbb{1}_N^\top\,(I_N-T_0)^{-1} \preceq \begin{bmatrix} O(1) & O(1)\,\mathbb{1}_{N-1}^\top \\ O(\mu)\,\mathbb{1}_{N-1} & O(\mu)\,\mathbb{1}_{N-1}\,\mathbb{1}_{N-1}^\top \end{bmatrix}. \tag{268}$$

Combining now (265) and (268) we obtain:

$$(I_N - T)^{-1} = \begin{bmatrix} O(1/\mu) & O(1/\mu)\,\mathbb{1}_{N-1}^\top \\ O(\mu)\,\mathbb{1}_{N-1} & O(1)\,\mathbb{1}_{N-1}\,\mathbb{1}_{N-1}^\top \end{bmatrix}. \tag{269}$$

On the other hand, examining the entries of vector x in (155) with the help of Table 2 we readily see that:

$$x \preceq O(\mu^2)\,\mathbb{1}_N. \tag{270}$$

Combining (260) with (270) we get:

$$\|(I_N - T)^{-1}x\|_1 = O(\mu), \tag{271}$$

which completes the proof. ■

APPENDIX J
PROOF OF THEOREM 2

In the following, we make repeated use of the following known equality, holding for any two nonzero scalars with

$a \neq b$:

$$\sum_{j=0}^{i-1} a^j b^{i-1-j} = b^{i-1} \sum_{j=0}^{i-1} (a/b)^j = b^{i-1} \frac{1-(a/b)^i}{1-(a/b)} = \frac{b^i - a^i}{b-a}. \tag{272}$$

Proof: In view of assumption (79), we can use Lemma 8 in (261) to conclude that, for all $i \geq 1$, and for sufficiently small μ:

$$\|\mathscr{P}[\widehat{\boldsymbol{q}}_i]\|_1 \leq K(\mu)\|\mathscr{P}[\widehat{\boldsymbol{q}}_0]\|_1 + \|(I_N - T)^{-1}x\|_1 = O(1), \tag{273}$$

where $K(\mu)$ is $O(1)$ in view of (246) while the order of magnitude of $\|(I_N - T)^{-1}\,x\|_1$ is $O(\mu)$ in view of (271).

Let us develop the recursion in (156) by separating the role of the unperturbed matrix T_0 and the rank-one perturbation $v_{\mu,\zeta}\,\mathbb{1}_N^\top$ in (151), yielding:

$$\begin{aligned}\mathscr{P}[\widehat{\boldsymbol{q}}_i] &\preceq T_0\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + v_{\mu,\zeta}\,\mathbb{1}_N^\top \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + x \\ &\preceq T_0\,\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + O(\mu^2)\,\mathbb{1}_N,\end{aligned} \tag{274}$$

where the second term on the RHS is $O(\mu^2)$ because so are $v_{\mu,\zeta}$ and x, while

$$\mathbb{1}_N^\top \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] = \|\mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}]\|_1, \tag{275}$$

which is $O(1)$ in view of (273). Developing the recursion in (274) we have:

$$\begin{aligned}\mathscr{P}[\widehat{\boldsymbol{q}}_i] &\preceq T_0^i\,\mathscr{P}[\widehat{\boldsymbol{q}}_0] + O(\mu^2)\,(I_N - T_0)^{-1}\,\mathbb{1}_N \\ &\preceq T_0^i\,\mathscr{P}[\widehat{\boldsymbol{q}}_0] + \begin{bmatrix} O(\mu) \\ \\ O(\mu^2)\,\mathbb{1}_{N-1}\end{bmatrix},\end{aligned} \tag{276}$$

where the estimate of the last term comes from (265).

Let us now evaluate the i-th power of T_0. Since T_0 is block upper-triangular, its i-th power admits the representation [14]:

$$T_0^i = \begin{bmatrix} \tau^i & \tau_{12}\mathbb{1}_{N-1}^\top (\tau I_{N-1} - E)^{-1}(\tau^i I_{N-1} - E^i) \\ \\ 0 & E^i \end{bmatrix}, \tag{277}$$

where we note that in the small-μ regime the matrix $(\tau I_{N-1} - E)^{-1}$ is certainly well-defined and has nonnegative entries, since as $\mu \to 0$ we have $\tau = 1 - \mu\,\zeta\,\nu > \rho(E)$ (and since E has nonnegative entries). Therefore, in the small-μ regime we can write:

$$\begin{aligned}&\tau_{12}\mathbb{1}_{N-1}^\top(\tau I_{N-1} - E)^{-1}(\tau^i I_{N-1} - E^i) \\ &\preceq \tau_{12}\mathbb{1}_{N-1}^\top(\tau I_{N-1} - E)^{-1}\,\tau^i.\end{aligned} \tag{278}$$

Considering the evolution of the first component $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2$ of $\mathscr{P}[\check{\boldsymbol{q}}_i]$ as dictated by (276), from (277) and (278) we get:[11]

$$\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 \leq O(1)\,\tau^i + O(\mu). \tag{279}$$

Likewise, exploiting (276), (277), and (278) to get the evolution relative to the network-error component we can write:

$$\mathscr{P}[\check{\boldsymbol{q}}_i] \preceq E^i\,\mathscr{P}[\check{\boldsymbol{q}}_0] + O(\mu^2)\mathbb{1}_{N-1}, \tag{280}$$

which allows us to write [66]:

$$\mathbb{E}\|\check{\boldsymbol{q}}_i\|^2 \leq O(1)\,\rho_{\text{net}}^i + O(\mu^2), \tag{281}$$

where by definition $\rho_{\text{net}} = \rho(E) + \epsilon < 1$.

We now exploit the exponential bounds in (279) and (281) to get a new recursion that would allow us to obtain refined estimates of both the transient and the steady-state errors. To this end, we revisit the norm on the LHS of (193), and manage to obtain a better bound by computing explicitly the norm and then applying the Cauchy-Schwarz inequality, yielding:

$$\begin{aligned}&\|(I_M - \mu\zeta \boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1} - \mu\zeta \boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2 \\ &= \|(I_M - \mu\zeta \boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\|\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad - 2\mu\zeta\left[(I_M - \mu\zeta \boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}\right]^\top \boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1} \\ &\leq (1-\mu\zeta\nu)^2\,\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + 2\mu\zeta\,\|(I_M - \mu\zeta \boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1}\| \times \|\boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\| \\ &\leq (1-\mu\zeta\nu)^2\,\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + 2\mu\zeta(1-\mu\zeta\nu)\sigma_{12}\|\bar{\boldsymbol{q}}_{i-1}\| \times \|\check{\boldsymbol{q}}_{i-1}\|.\end{aligned} \tag{282}$$

Taking expectations and applying the Cauchy-Schwarz inequality for random variables, we obtain:

$$\begin{aligned}&\mathbb{E}\|(I_M - \mu\zeta \boldsymbol{G}_{11,i-1})\bar{\boldsymbol{q}}_{i-1} - \mu\zeta \boldsymbol{G}_{12,i-1}\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\leq (1-\mu\zeta\nu)^2\,\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + 2\mu\zeta(1-\mu\zeta\nu)\sigma_{12}\sqrt{\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 \times \mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2} \\ &\leq (1-\mu\zeta\nu)^2\,\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + O(\mu)\sqrt{\left(O(1)\,\tau^i + O(\mu)\right)\left(O(1)\,\rho_{\text{net}}^i + O(\mu^2)\right)} \\ &\leq (1-\mu\zeta\nu)^2\,\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + O(\mu)\,\left(O(1)\,\tau^{i/2} + O(\sqrt{\mu})\right)\left(O(1)\,\rho_{\text{net}}^{i/2} + O(\mu)\right) \\ &= (1-\mu\zeta\nu)^2\,\mathbb{E}\|\bar{\boldsymbol{q}}_{i-1}\|^2 + \mu^2\zeta^2\,\sigma_{12}^2\,\mathbb{E}\|\check{\boldsymbol{q}}_{i-1}\|^2 \\ &\quad + \underbrace{O(\mu^2)\,\tau^{i/2} + O(\mu)\,\rho_{\text{net}}^{i/2} + O(\mu^{5/2})}_{\chi_i},\end{aligned} \tag{283}$$

where we applied (279) and (281), along with the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$.

[11] The quantity $\mathbb{1}_{N-1}(\tau I_{N-1} - E)^{-1}$ is $O(1)$ as $\mu \to 0$. This can be seen, e.g., by noticing that $(\tau I_{N-1} - E)^{-1} \to \mathcal{R}_E(1)$ as $\mu \to 0$, and we have already shown that $\mathcal{R}_E(1)$ is bounded.

Using (283), we can rearrange the first row of the transfer matrix T_q and of the vector x_q appearing in Lemma 4, obtaining a new matrix T_q' and a new vector x_q' defined by:

$$[T_q']_{11} = \tau^2, \;\; [T_q']_{12} = \mu^2 \zeta^2 \sigma_{12}^2 \, \mathbb{1}_{N-1}^\top, \;\; \bar{x}_q' = \begin{bmatrix} \bar{x}_q + \chi_i \\ \\ \check{x}_q \end{bmatrix}, \tag{284}$$

which allows us to replace the matrix T appearing in Theorem 3 with a matrix T' of the following form:

$$T' = \underbrace{\begin{bmatrix} \tau^2 & \tau_{12}' \, \mathbb{1}_{N-1}^\top \\ \\ 0 & E \end{bmatrix}}_{T_0'} + O(\mu^2) \, \mathbb{1}_N \, \mathbb{1}_N^\top, \tag{285}$$

where we defined:

$$\tau_{12}' = 16 \, \zeta^2 \, \bar{\Delta}, \tag{286}$$

with the μ^2-term appearing in $[T_q']_{12}$ being conveniently embodied in the overall $O(\mu^2)$ rank-one perturbation. Likewise, we construct a new driving vector x' by replacing x_q in (155) with x_q' in (284). Replacing now T and x with T' and x' in the recursion (156), we get:

$$\begin{aligned} \mathscr{P}[\widehat{\boldsymbol{q}}_i] &\preceq T_0' \, \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + O(\mu^2) \, \mathbb{1}_N \underbrace{\mathbb{1}_N^\top \, \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}]}_{\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 + \mathbb{E}\|\check{\boldsymbol{q}}_i\|^2} + x' \\ &\preceq T_0' \, \mathscr{P}[\widehat{\boldsymbol{q}}_{i-1}] + \underbrace{O(\mu^2) \left(O(1) \, \tau^i + O(1) \, \rho_{\text{net}}^i + O(\mu) \right)}_{\chi_i'} \mathbb{1}_N \\ &\quad + \begin{bmatrix} \bar{x} + \chi_i \\ \\ \check{x} \end{bmatrix}, \end{aligned} \tag{287}$$

where in the last step we used (279) and (281). Developing the recursion in (287) we get:

$$\begin{aligned} \mathscr{P}[\widehat{\boldsymbol{q}}_i] &\preceq (T_0')^i \, \mathscr{P}[\widehat{\boldsymbol{q}}_0] + \sum_{j=0}^{i-1} \chi_{i-j}' \, (T_0')^j \mathbb{1}_N \\ &\quad + \sum_{j=0}^{i-1} (T_0')^j \begin{bmatrix} \chi_{i-j} \\ \\ 0 \end{bmatrix} + (I_N - T_0')^{-1} \begin{bmatrix} \bar{x} \\ \\ \check{x} \end{bmatrix}. \end{aligned} \tag{288}$$

Applying now (277) (with τ^2 in place of τ, and τ_{12}' in place of τ_{12}), and reasoning as done to obtain the bound in (278), we conclude that the first row of matrix $(T_0')^i$ can be upper bounded as:

$$(T_0')^i \preceq \begin{bmatrix} O(1) \, \tau^{2i} \, \mathbb{1}_N^\top \\ \\ \text{not needed} \end{bmatrix}. \tag{289}$$

Moreover, we can evaluate $(I_N - T_0')^{-1}$ through (234) applied with τ^2 in place of τ, and τ_{12}' in place of τ_{12}, obtaining:

$$(I_N - T_0')^{-1} = \begin{bmatrix} \dfrac{1}{1-\tau^2} & \dfrac{\tau_{12}' \, \mathbb{1}_{N-1}^\top (I_{N-1} - E)^{-1}}{1-\tau^2} \\ \\ 0 & (I_{N-1} - E)^{-1} \end{bmatrix}. \tag{290}$$

Using (289) and (290) in (288), we can obtain an inequality recursion on the first entry of $\mathscr{P}[\widehat{\boldsymbol{q}}_i]$:

$$\begin{aligned} \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 &\le O(1) \, \tau^{2i} \, \mathbb{E}\|\widehat{\boldsymbol{q}}_0\|^2 + O(1) \sum_{j=0}^{i-1} \tau^{2j} \, \chi_{i-j} \\ &\quad + \frac{\bar{x}}{1-\tau^2} + \frac{\tau_{12}' \, \mathbb{1}_{N-1}^\top (I_{N-1} - E)^{-1}}{1-\tau^2} \, \check{x}, \end{aligned} \tag{291}$$

where we have ignored the term χ_i', since comparing this term against the term χ_i in (283), we see that χ_i' is dominated by χ_i as $\mu \to 0$ and, hence, can be formally embodied into χ_i through the Big-O notation. Applying the definition of χ_i in (283) we can write:

$$\begin{aligned} \sum_{j=0}^{i-1} \tau^{2j} \, \chi_{i-j} &= O(\mu^2) \sum_{j=0}^{i-1} \tau^{2j} \tau^{(i-j)/2} \\ &\quad + O(\mu) \sum_{j=0}^{i-1} \tau^{2j} \rho_{\text{net}}^{(i-j)/2} + O(\mu^{5/2}) \sum_{j=0}^{i-1} \tau^{2j} \\ &\le O(\mu^2) \frac{\tau^{i/2}}{\tau^{1/2} - \tau^2} + O(\mu) \, \frac{\tau^{2i}}{\tau^2 - \rho_{\text{net}}^{1/2}} + \frac{O(\mu^{5/2})}{1-\tau^2}, \end{aligned} \tag{292}$$

where in the last inequality we exploited the geometric summation in (272).

Now we examine the small-μ behavior of the three terms appearing on the RHS of (292). The first term is $O(\mu) \, \tau^{i/2}$, since we have that:

$$\lim_{\mu \to 0} \frac{\mu}{\tau^{1/2} - \tau^2} = \lim_{\mu \to 0} \frac{\mu}{(1 - \mu\zeta\nu)^{1/2} - (1 - \mu\zeta\nu)^2} = \frac{2}{3\zeta\nu}. \tag{293}$$

The second term on the RHS of (292) is $O(\mu) \, \tau^{2i}$, since:

$$\lim_{\mu \to 0} \tau^2 - \rho_{\text{net}}^{1/2} = 1 - \rho_{\text{net}}^{1/2} > 0. \tag{294}$$

Finally, the third term is $O(\mu^{3/2})$ since:

$$\frac{1}{1-\tau^2} = \frac{1}{\mu \, \zeta \, \nu} \, \frac{1}{2 - \mu \, \zeta \, \nu} = \frac{1}{\mu \, \zeta \, \nu} \left(\frac{1}{2} + O(\mu) \right). \tag{295}$$

As a result, we can use (292) in (291) and substitute the estimated orders of the aforementioned three terms to obtain:

$$\begin{aligned} \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 &\le O(1) \, \tau^{2i} + O(\mu) \, \tau^{i/2} + O(\mu^{3/2}) \\ &\quad + \frac{\bar{x}}{1-\tau^2} + \frac{\tau_{12}' \, \mathbb{1}_{N-1}^\top (I_{N-1} - E)^{-1}}{1-\tau^2} \, \check{x}. \end{aligned} \tag{296}$$

It remains to examine the small-μ behavior of the last two terms in (296). Exploiting the definition of x in (155), the small-μ behavior of the components of x in Table 2, and the definitions of $\bar{\Delta}$ and $\check{\Delta}$ in Table 1, we can write:

$$\begin{aligned}\bar{x} &= \zeta^2 \bar{x}_s + \zeta^2 \bar{\Delta}\left(\bar{x}_s + \mathbb{1}_{N-1}^{\top}(\check{x}_s + \check{x}_\delta)\right)\\ &\leq \mu^2 \zeta^2 \left(\sum_{k=1}^{N} \pi_k \alpha_k^2 \sigma_k^2 + \kappa_1 \Omega\right),\end{aligned} \tag{297}$$

where κ_1 is a suitable constant (i.e., independent of μ) and Ω is the maximum compression factor in (25). Using (295), the first fraction on the RHS of (296) is bounded as:

$$\frac{\bar{x}}{1-\tau^2} \leq \mu\,\zeta \left(\frac{\sum_{k=1}^{N} \pi_k \alpha_k^2 \sigma_k^2}{2\nu} + \frac{\kappa_1}{2\nu}\,\Omega\right) + O(\mu^2). \tag{298}$$

Using similar arguments we can bound the last term in (296). From (286) and (155) we get:

$$\tau_{12}' \leq \kappa_2\,\zeta^2\,\Omega, \quad \check{x} \preceq (\kappa_3 + \kappa_4\,\Omega)\,\mu^2 \mathbb{1}_{N-1}, \tag{299}$$

for some constants κ_2, κ_3, and κ_4. Using now (295) and (299), the last term in (296) can be upper bounded by $\mu\,\zeta\,\kappa_5\Omega(1+\Omega) + O(\mu^2)$, for a suitable constant κ_5. Joining this representation with (297), we can finally write (296) as:

$$\begin{aligned}\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 &\leq O(1)\,\tau^{2i} + O(\mu)\,\tau^{i/2} + O(\mu^{3/2})\\ &\quad + \mu\,\zeta\left(\frac{\sum_{k=1}^{N} \pi_k \alpha_k^2 \sigma_k^2}{2\nu} + c_q\,\Omega\,(1+\Omega)\right),\end{aligned} \tag{300}$$

where c_q is a constant independent of μ.

It remains to characterize the behavior of the mean-square-deviation at an individual agent k. To this end, we evaluate the individual entry of the extended vector $\widetilde{\boldsymbol{q}}_i$ though (57), which allows us to write:

$$\begin{aligned}\mathbb{E}\|\boldsymbol{q}_{k,i}\|^2 &= \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 + \mathbb{E}\|\mathcal{T}_k\check{\boldsymbol{q}}_i\|^2 + 2\mathbb{E}[\bar{\boldsymbol{q}}_i^{\top}\mathcal{T}_k\check{\boldsymbol{q}}_i]\\ &\leq \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 + \mathbb{E}\|\mathcal{T}_k\check{\boldsymbol{q}}_i\|^2 + 2\big|\mathbb{E}[\bar{\boldsymbol{q}}_i^{\top}\mathcal{T}_k\check{\boldsymbol{q}}_i]\big|\\ &\leq \mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 + \|\mathcal{T}_k\|^2\,\mathbb{E}\|\check{\boldsymbol{q}}_i\|^2\\ &\quad + 2\|\mathcal{T}_k\|\sqrt{\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2\,\mathbb{E}\|\check{\boldsymbol{q}}_i\|^2},\end{aligned} \tag{301}$$

where we resorted again to the Cauchy-Schwarz inequality for random variables. Now, from (300) we can write $\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2 \leq O(1)\,\tau^{2i} + O(\mu)$, which, used along with (281), yields:

$$\begin{aligned}&\sqrt{\mathbb{E}\|\bar{\boldsymbol{q}}_i\|^2\,\mathbb{E}\|\check{\boldsymbol{q}}_i\|^2}\\ &\leq \sqrt{\big(O(1)\tau^{2i} + O(\mu)\big)\,\big(O(1)\rho_{\text{net}}^{i} + O(\mu^2)\big)}\\ &\leq \underbrace{O(1)\tau^{i}\rho_{\text{net}}^{i/2} + O(\mu^{1/2})\rho_{\text{net}}^{i/2}}_{O(1)\rho_{\text{net}}^{i/2}} + \underbrace{O(\mu)\tau^{i}}_{\leq O(\mu)\tau^{i/2}} + O(\mu^{3/2}).\end{aligned} \tag{302}$$

Further noticing that from (281) we can write $\mathbb{E}\|\check{\boldsymbol{q}}_i\|^2 \leq O(1)\,\rho_{\text{net}}^{i/2} + O(\mu^2)$, the claim of the theorem follows by using (300) and (302) in (301). ■

APPENDIX K
COMPRESSION OPERATORS

1. RANDOMIZED QUANTIZERS FROM [43]

- The Euclidean norm $\|x\|$ of the input vector x is represented with high resolution h, e.g., with machine precision. Then, each entry x_m of x is separately quantized.
- One bit is used to encode the sign of x_m.
- Then, we encode the absolute value of the m-th entry x_m. Since $\|x\|$ is transmitted with high precision, we can focus on the scaled value:

$$\xi_m \triangleq \frac{|x_m|}{\|x\|} \in [0,\,1]. \tag{303}$$

The interval $[0,\,1]$ is partitioned into L equal-size intervals — see the illustrative example in Fig. 10. The size of each interval is:

$$\vartheta \triangleq \frac{1}{L}, \tag{304}$$

such that the intervals' endpoints can be accordingly represented as:

$$y_0 = 0, \quad y_1 = \vartheta, \quad y_2 = 2\vartheta, \quad \ldots, \quad y_L = 1. \tag{305}$$

In order to avoid confusion, we stress that the quantization scheme will require to transmit one of the $L+1$ indices corresponding to the endpoints. This differs from classical quantization schemes where the index of the interval (instead of the endpoint) is transmitted. Accordingly, the bit-rate r is equal to:

$$r = \log_2(L+1) \Leftrightarrow L = 2^r - 1. \tag{306}$$

- In view of (305), the index of the (lower) endpoint of the interval the scaled entry ξ_m belongs to, is computed as:

$$j(\xi_m) = \left\lfloor \frac{\xi_m}{\vartheta} \right\rfloor, \tag{307}$$

and the corresponding endpoint is:

$$y(\xi_m) = j(\xi_m)\,\vartheta. \tag{308}$$

Then, we *randomize* the quantization operation since we choose randomly to transmit the lower endpoint index $j(\xi_m)$ or the upper endpoint index $j(\xi_m)+1$. Specifically, the probability of transmitting one endpoint index is proportional to the distance of ξ_m from that endpoint. In other words, the closer we are to one endpoint, the higher the probability of transmitting that endpoint will be. Formally, the random transmitted index $\boldsymbol{j}_{\text{tx}}(\xi_m)$ is:

$$\boldsymbol{j}_{\text{tx}}(\xi_m) = \begin{cases} j(\xi_m)+1, & \text{with prob.} \quad \dfrac{\xi_m - y(\xi_m)}{\vartheta}\\ j(\xi_m), & \text{otherwise.}\end{cases} \tag{309}$$

- Once the index $\boldsymbol{j}_{\text{tx}}(\xi_m)$ is received, the unquantized value ξ_m is rounded to the lower or upper endpoint depending on the realization of the transmitted index $\boldsymbol{j}_{\text{tx}}(\xi_m)$, and then the information about the norm $\|x\|$ and the sign of

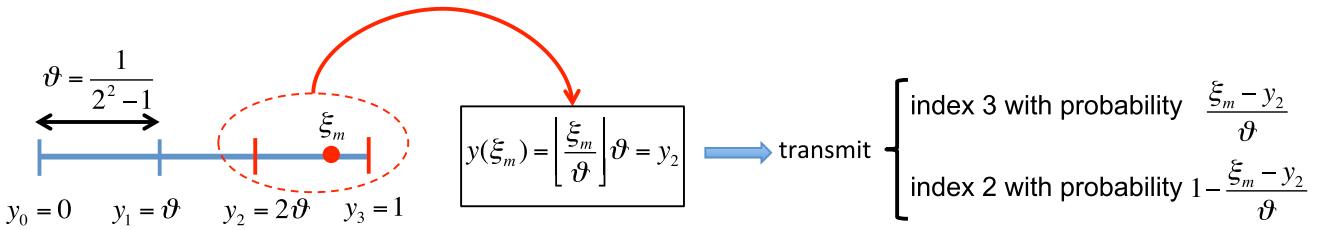

FIGURE 10. Sketch of the randomized quantizer described in Appendix K1, for the case where the bit-rate is $r = 2$.

x_m is recovered, finally yielding the m-th component of the quantized vector $\boldsymbol{Q}(x)$:

$$[\boldsymbol{Q}(x)]_m = \|x\| \, \mathsf{sign}(x_m) \, \boldsymbol{j}_{\mathrm{tx}}(\xi_m) \, \vartheta, \quad m = 1, 2, \ldots, M. \tag{310}$$

- Accounting for the h bits spent for representing the norm $\|x\|$ and the single bit for representing the sign of each of the M entries of x, the total bit-rate is:

$$h + M\,(r+1). \tag{311}$$

It was shown in [43] that the value of the compression factor ω can be computed as:

$$\omega = \min\left\{ \frac{M}{L^2}, \frac{\sqrt{M}}{L} \right\}. \tag{312}$$

Equation (312) provides useful insight on the practical meaning of ω for the considered type of quantizers. We see that, for fixed dimensionality M, the parameter ω decays exponentially fast with the number of bits ($\approx 2^{-2r}$), whereas for fixed number of bits it grows as $\sqrt{M}$. Finally, a useful bound exploited in our analysis is obtained as follows. Note that we have either:

$$\frac{M}{L^2} \leq 1 \Rightarrow \omega = \frac{M}{L^2} \Rightarrow \omega + \omega^2 \leq 2\,\omega = \frac{2M}{L^2}, \tag{313}$$

or

$$\frac{M}{L^2} > 1 \Rightarrow \omega = \frac{\sqrt{M}}{L} \Rightarrow \omega + \omega^2 \leq 2\,\omega^2 = \frac{2M}{L^2}. \tag{314}$$

In summary, in both cases we can write:

$$\omega\,(1+\omega) \leq \frac{2M}{L^2} \leq \frac{2M}{(2^r-1)^2}. \tag{315}$$

2. RANDOMIZED SPARSIFIERS FROM [57]

Another popular compression operator fulfilling Assumption 5 is the randomized sparsifier, which selects S entries of the vector $x \in \mathbb{R}^M$, and transmits only these components. The decoder sets the remaining entries to 0. Moreover, the transmitted entries are scaled by M/S since, as we promptly show, such scaling makes the quantizer unbiased.

Formally, letting e_m be the standard basis vector in $\mathbb{R}^M$, having m-th entry equal to 1 and the remaining entries equal to 0, we have:

$$\boldsymbol{Q}(x) = \frac{M}{S} \sum_{m=1}^{M} x_m e_m \boldsymbol{t}_m \tag{316}$$

where $\boldsymbol{t}_m$ is a random variable equal to 1 if the m-th entry is transmitted, and 0 otherwise. Since the transmitted entries are chosen uniformly at random, we have that:

$$p \triangleq \mathbb{P}[\boldsymbol{t}_m = 1] = \mathbb{E}[\boldsymbol{t}_m = 1] = \frac{S}{M}. \tag{317}$$

Accordingly, the expectation of $\boldsymbol{Q}(x)$ is:

$$\mathbb{E}[\boldsymbol{Q}(x)] = \frac{M}{S} \sum_{m=1}^{M} x_m e_m \underbrace{\mathbb{E}[\boldsymbol{t}_m]}_{p=S/M} = \sum_{m=1}^{M} x_m e_m = x. \tag{318}$$

Likewise, the mean-square-error is evaluated as:

$$\begin{aligned}
\mathbb{E}\|\boldsymbol{Q}(x) - x\|^2 &= \mathbb{E}\left\| \frac{M}{S} \sum_{m=1}^{M} x_m e_m \boldsymbol{t}_m - \sum_{m=1}^{M} x_m e_m \right\|^2 \\
&= \mathbb{E}\left\| \sum_{m=1}^{M} x_m e_m \left(\frac{M}{S} \boldsymbol{t}_m - 1 \right) \right\|^2 \\
&= \mathbb{E} \sum_{m=1}^{M} x_m^2 \left(\frac{M}{S} \boldsymbol{t}_m - 1 \right)^2 \\
&= \sum_{m=1}^{M} x_m^2 \mathbb{E}\left[\left(\frac{M}{S} \boldsymbol{t}_m - 1 \right)^2 \right] \\
&= \|x\|^2 \left[p \left(\frac{M}{S} - 1 \right)^2 + (1-p) \right],
\end{aligned} \tag{319}$$

which, using (317), with straightforward algebra gives:

$$\mathbb{E}\|\boldsymbol{Q}(x) - x\|^2 = \left(\frac{M}{S} - 1 \right) \|x\|^2. \tag{320}$$

REFERENCES

[1] M. Carpentiero, V. Matta, and A. H. Sayed, "Adaptive diffusion with compressed communication," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Singapore, May 2022, pp. 5672–5676.

[2] U. A. Khan, W. U. Bajwa, A. Nedić, M. G. Rabbat, and A. H. Sayed, "Optimization for data-driven learning and control," *Proc. IEEE*, vol. 108, no. 11, pp. 1863–1868, Nov. 2020.

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

[4] A. Nedić and D. P. Bertsekas, "Incremental subgradient methods for nondifferentiable optimization," *SIAM J. Optim.*, vol. 12, no. 1, pp. 109–138, 2001.

[5] A. Nedić and A. Ozdaglar, "Cooperative distributed multi-agent optimization," in *Convex Optimization in Signal Processing and Communications*, Y. Eldar and D. Palomar Eds., Cambridge, U.K.: Cambridge Univ. Press, 2010, pp. 340–386.

[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[7] S. Lee and A. Nedić, "Distributed random projection algorithm for convex optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 221–229, Apr. 2013.

[8] C. Xi and U. A. Khan, "Distributed subgradient projection algorithm over directed graphs," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3986–3992, Aug. 2016.

[9] C. Xi, V. S. Mai, R. Xin, E. Abed, and U. A. Khan, "Linear convergence in optimization over directed graphs with row-stochastic matrices," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3558–3565, Oct. 2018.

[10] M. G. Rabbat and A. Ribeiro, "Multiagent distributed optimization," in *Cooperative and Graph Signal Processing*, P. Djuric and C. Richard, Eds. New York, NY, USA: Elsevier, 2018, pp. 147–167.

[11] M. Nokleby and W. U. Bajwa, "Stochastic optimization from distributed streaming data in rate-limited networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 1, pp. 152–167, Mar. 2019.

[12] A. H. Sayed, S. Y. Tu, J. Chen, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 155–171, May 2013.

[13] A. H. Sayed, "Adaptive networks," *Proc. IEEE*, vol. 102, no. 4, pp. 460–497, Apr. 2014.

[14] J. Chen and A. H. Sayed, "On the learning behavior of adaptive networks — Part I: Transient analysis," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3487–3517, Jun. 2015.

[15] J. Chen and A. H. Sayed, "On the learning behavior of adaptive networks — Part II: Performance analysis," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3518–3548, Jun. 2015.

[16] M. H. DeGroot, "Reaching a consensus," *J. Amer. Statist. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[17] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, Sep. 2004.

[18] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

[19] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Nov. 2010.

[20] A. Nedić, A. Olshevsky, and M. G. Rabbat, "Network topology and communication-computation tradeoffs in decentralized optimization," *Proc. IEEE*, vol. 106, no. 5, pp. 953–976, May 2018.

[21] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Found. Trends Mach. Learn.*, vol. 7, no. 4/5, pp. 311–801, 2014.

[22] V. Matta and A. H. Sayed, "Estimation and detection over adaptive networks," in *Cooperative and Graph Signal Processing*, P. Djuric and C. Richard, Eds., New York, NY, USA: Elsevier, 2018, pp. 69–106.

[23] Wai-Man Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Trans. Commun.*, vol. 41, no. 11, pp. 1602–1605, Nov. 1993.

[24] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.

[25] S. Marano, V. Matta, and P. Willett, "Asymptotic design of quantizers for decentralized MMSE estimation," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5485–5496, Nov. 2007.

[26] P. Venkitasubramaniam, L. Tong, and A. Swami, "Quantization for maximin ARE in distributed estimation," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3596–3605, Jul. 2007.

[27] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 241–255, Mar. 1990.

[28] J. N. Tsitsiklis,, "Decentralized detection," *Adv. Stat. Signal Process.*, vol. 2, pp. 297–344, 1993.

[29] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors Part I. Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[30] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors Part II. Advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.

[31] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[32] V. Saligrama, M. Alanyali, and O. Savas, "Distributed detection in sensor networks with packet losses and finite capacity links," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4118–4132, Nov. 2006.

[33] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.

[34] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[35] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1081–1092, Mar. 2007.

[36] S. Marano, V. Matta, L. Tong, and P. Willett, "A likelihood-based multiple access for estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5155–5166, Nov. 2007.

[37] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[38] B. Widrow, "A study of rough amplitude quantization by means of Nyquist sampling theory," *IRE Trans. Circuit Theory*, vol. 3, no. 4, pp. 266–276, Dec. 1956.

[39] R. M. Gray and T. G. Stockham, "Dithered quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 805–812, May 1993.

[40] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, Jun. 2005.

[41] J. B. Predd, S. B. Kulkarni, and H. V. Poor, "Distributed learning in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 56–69, Jul. 2006.

[42] S. Marano, V. Matta, and P. Willett, "Nearest-Neighbor distributed learning by ordered transmissions," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5217–5230, Nov. 2013.

[43] D. Alistarh, D. Grubic, J. Z. Li, R. Tomioka, and M. Vojnovic, "QSGD: Communication-efficient SGD via gradient quantization and encoding," in *Proc. Neural Inf. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, pp. 1707–1718.

[44] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. New York, NY, USA: Springer Science, 2001.

[45] F. Seide, H. Fu, J. Droppo, G. Li, and D. Yu, "1-Bit stochastic gradient descent and its application to data-parallel distributed training of speech DNNs," in *Proc. Conf. Int. Speech Commun. Assoc.*, Singapore, 2014, pp. 1058–1062.

[46] J. Wu, W. Huang, J. Huang, and T. Zhang, "Error compensated quantized sgd and its applications to large-scale distributed optimization," in *Proc. Int. Conf. Mach. Learn.*, Stockholm, Sweden, 2018, pp. 5235–5333.

[47] S. U. Stich, J.-B. Cordonnier, and M. Jaggi, "Sparsified SGD with Memory," in *Proc. Neural Inf. Process. Syst.*, Montréal, Canada, 2018, pp. 4447–4458.

[48] S. P. Karimireddy, Q. Rebjock, S. U. Stich, and M. Jaggi, "Error feedback Fixes signSGD and other gradient compression schemes," in *Proc. Int. Conf. Mach. Learn.*, Long Beach, CA, USA, 2019, pp. 3252–3261.

[49] A. Reisizadeh, A. Mokhtari, H. Hassani, A. Jadbabaie, and R. Pedarsani, "FedPAQ: A communication-efficient federated learning method with periodic averaging and quantization," in *Proc. Int. Conf. Artif. Intell. Statist.*, Palermo, Italy, 2020, pp. 2021–2031.

[50] C.-Y. Lin, V. Kostina, and B. Hassibi, "Differentially quantized gradient descent," in *Proc. IEEE Int. Symp. Inf. Theory*, Melbourne, VIC, Australia, 2021, pp. 1200–1205.

[51] A. H. Sayed, *Inference and Learning from Data: Volume 3*. Cambridge, U.K.: Cambridge Univ. Press, 2022.

[52] X. Zhao, S.-Y. Tu, and A. H. Sayed, "Diffusion adaptation over networks under imperfect information exchange and non-stationary data," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3460–3475, Jul. 2012.

[53] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. Decis. Control*, Cancun, Mexico, 2008, pp. 4177–4184.

[54] T. Doan, S. T. Maguluri, and J. Romberg, "Convergence rates of distributed gradient methods under random quantization: A stochastic approximation approach," *IEEE Trans. Autom. Control*, vol. 66, no. 10, pp. 4469–4484, Oct. 2021.
[55] A. Reisidazeh, A. Mokhtari, H. Hassani, and R. Pedarsani, "An exact quantized decentralized gradient descent algorithm," *IEEE Trans. Signal Process.*, vol. 67, no. 19, pp. 4934–4947, Oct. 2019.
[56] A. Koloskova, S. U. Stich, and M. Jaggi, "Decentralized stochastic optimization and gossip algorithms with compressed communication," in *Proc. Int. Conf. Mach. Learn.*, Long Beach, CA, USA, 2019, pp. 3478–3487.
[57] D. Kovalev, A. Koloskova, M. Jaggi, P. Richtárik, and S. U. Stich, "A linearly convergent algorithm for decentralized optimization: Sending less bits for free!," in *Proc. Int. Conf. Artif. Intell. Statist.*, San Diego, CA, USA, Apr. 2021, pp. 4087–4095.
[58] M. I. Qureshi, R. Xin, S. Kar, and U. A. Khan, "S-ADDOPT: Decentralized stochastic first-order optimization over directed graphs," *IEEE Control Syst. Lett.*, vol. 5, no. 3, pp. 953–958, Jul. 2021.
[59] M. I. Qureshi, R. Xin, S. Kar, and U. A. Khan, "Push-SAGA: A decentralized stochastic algorithm with variance reduction over directed graphs," *IEEE Control Syst. Lett.*, vol. 6, pp. 1202–1207, 2022.
[60] S. Pu and A. Nedić, "Distributed stochastic gradient tracking methods," *Math. Program.*, vol. 187, pp. 409–457, 2021.
[61] A. Koloskova, T. Lin, S. U. Stich, and M. Jaggi, "Decentralized deep learning with arbitrary communication compression," in *Proc. Int. Conf. Learn. Representations*, Addis Ababa, Ethiopia, 2020, pp. 1–22.
[62] J. Chen and A. H. Sayed, "Distributed Pareto optimization via diffusion strategies," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 205–220, Apr. 2013.
[63] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3721–3748, Aug. 2010.
[64] G. Alberti, H. Bölcskei, C. D. Lellis, G. Koliander, and E. Riegler, "Lossless analog compression," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7480–7513, Nov. 2019.
[65] N. Michelusi, G. Scutari, and C. -S. Lee, "Finite-bit quantization for distributed algorithms with linear convergence," *IEEE Trans. Inf. Theory*, vol. 68, no. 11, pp. 7254–7280, Nov. 2022.
[66] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY, USA: Cambridge Univ. Press, 1985.
[67] G. H. Golub and J. H. Wilkinson, "Ill-conditioned eigensystems and the computation of the Jordan canonical form," *SIAM Rev.*, vol. 18, no. 4, pp. 578–619, 1976.
[68] R. Bronson, *Matrix Methods: An Introduction*. Houston: Gulf Professional Publishing, 1991.
[69] J. Moro, J. V. Burke, and M. L. Overton, "On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure," *SIAM J. Matrix Anal. Appl.*, vol. 18, no. 4, pp. 793–817, Oct. 1997.
[70] Z. J. Towfic and A. H. Sayed, "Stability and performance limits of adaptive primal-dual networks," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2888–2903, Jun. 2015.
[71] K. Yuan, B. Ying, X. Zhao, and A. H. Sayed, "Exact diffusion for distributed optimization and learning–Part I: Algorithm development," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 708–723, Feb. 2019.
[72] K. Yuan, B. Ying, X. Zhao, and A. H. Sayed, "Exact diffusion for distributed optimization and learning–Part II: Convergence analysis," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 724–739, Feb. 2019.
[73] S. Vlaski and A. H. Sayed, "Distributed learning in non-convex environments — Part I: Agreement at a linear rate," *IEEE Trans. Signal Process.*, vol. 69, pp. 1242–1256, 2021.
[74] S. Vlaski and A. H. Sayed, "Distributed learning in non-convex environments — Part II: Polynomial escape from saddle-points," *IEEE Trans. Signal Process.*, vol. 69, pp. 1257–1270, 2021.
[75] S. Vlaski and A. H. Sayed, "Second-order guarantees of stochastic gradient descent in non-convex optimization," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6489–6504, Dec. 2022.
[76] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra. The Behavior of Nonnormal Matrices and Operators*. New Jersey, NJ, USA: Princeton Univ. Press, 2005.

MARCO CARPENTIERO (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees (*cum laude*) in information engineering in 2018 and 2020, respectively, from the University of Salerno, Fisciano, Italy, where he is currently working toward the Ph.D. degree in information engineering with the Department of Information and Electrical Engineering and Applied Mathematics. His research interests include in the field of signal processing, focusing on adaptation and learning over networks, distributed inference, and remote sensing.

VINCENZO MATTA (Senior Member, IEEE) is currently a Full Professor in telecommunications with the Department of Information and Electrical Engineering and Applied Mathematics (DIEM), University of Salerno, Fisciano, Italy. He is the author of more than 140 articles published in international journals and proceedings of international conferences. His research interests include adaptation and learning over networks, social learning, statistical inference on graphs, and security in communication networks. Dr. Matta is an Associate Editor for the IEEE OPEN JOURNAL OF SIGNAL PROCESSING. He was the Senior Area Editor of the IEEE SIGNAL PROCESSING LETTERS, and an Associate Editor for IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, the IEEE SIGNAL PROCESSING LETTERS, and IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS. He was a member of the Sensor Array and Multichannel Technical Committee of the Signal Processing Society (SPS), and was an IEEE SPS Steering Committee Representative on the IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS.

ALI H. SAYED (Fellow, IEEE) is currently the Dean of Engineering at EPFL, Switzerland, where he also leads the Adaptive Systems Laboratory. He was a Distinguished Professor and the Chair of electrical engineering, University of California, Los Angeles, Los Angeles, CA, USA. He is a member of the U.S. National Academy of Engineering (NAE) and The World Academy of Sciences (TWAS). He was the President of the IEEE Signal Processing Society in 2018 and 2019. His work has been recognized with several awards including the 2022 IEEE Fourier Technical Field Award, the 2020 IEEE Norbert Wiener Society Award, and several Best Paper Awards. He is a Fellow of EURASIP, and AAAS.

Open Access funding provided by 'Università degli Studi di Salerno' within the CRUI CARE Agreement