2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

DASA: Delay-Adaptive Multi-Agent Stochastic Approximation

Nicolò Dal Fabbro*, Arman Adibi*, H. Vincent Poor, Sanjeev R. Kulkarni, Aritra Mitra and George J. Pappas

2024 IEEE 63rd Conference on Decision and Control (CDC) | 979-8-3503-1633-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/CDC56724.2024.10886380

***Abstract*— We consider a setting in which N agents aim to speedup a common Stochastic Approximation (SA) problem by acting in parallel and communicating with a central server. We assume that the up-link transmissions to the server are subject to asynchronous and potentially unbounded time-varying delays. To mitigate the effect of delays and stragglers while reaping the benefits of distributed computation, we propose DASA, a Delay-Adaptive algorithm for multi-agent Stochastic Approximation. We provide a finite-time analysis of DASA assuming that the agents' stochastic observation processes are independent Markov chains. Significantly advancing existing results, DASA is the first algorithm whose convergence rate depends only on the mixing time τ_{mix} and on the average delay τ_{avg} while jointly achieving an N-fold convergence speedup under Markovian sampling. Our work is relevant for various SA applications, including multi-agent and distributed temporal difference (TD) learning, Q-learning and stochastic optimization with correlated data.**

I. Introduction

Motivated by emerging applications in data-hungry algorithms like distributed large-scale reinforcement learning (RL), we study an instance of Stochastic Approximation (SA) in which multiple agents act in parallel to cooperatively solve a common SA problem. The main goal of splitting the computational effort across N agents is to achieve an N-fold convergence speedup in the sample complexity relative to when a single agent acts alone. Only recently, theoretical studies have shown the beneficial effect of distributed computing in multi-agent and federated SA when agents' observations are temporally correlated (Markov) processes [1]–[3]. However, in real-world large-scale communication networks and client-server architectures, inter-node communications are commonly subject to (potentially large) delays. While several algorithms have been provided to study and mitigate the effect of delays in distributed asynchronous optimization and federated learning [4]–[7] under i.i.d. sampling, very little is known about distributed SA under Markovian sampling. Existing works, indeed, either provide algorithms that are heavily impacted by the maximum delay [8], [9] (which can be potentially very large), or are not able to provide convergence speedups under Markovian sampling [10]. In the recent work in [8], the authors study collaborative gains of parallel learners in RL under Markovian sampling, but they obtain convergence rates affected by the maximum network delay and provide convergence speedups only when the data are collected in an i.i.d. fashion. However, note that the main challenges in proving multi-agent convergence speedups of SA algorithms emerge when agents' samples form temporally correlated trajectories. Indeed, even for single-agent settings, non-asymptotic convergence rates for SA under Markovian sampling have only recently been established [11]–[13]. In our recent work [10], we provide finite-time convergence rates for an SA setting under Markovian sampling in which the SA operator is computed with delayed updates, and we propose a delay-adaptive algorithm that replaces the dependence on the maximum delay τ_{max} with a dependence on the average delay τ_{avg}. However, the work in [10] only considers a single-agent setting. In this paper, we bridge this gap proposing and analyzing a new multi-agent algorithm for distributed asynchronous SA, DASA. The main contributions of this paper are as follows.

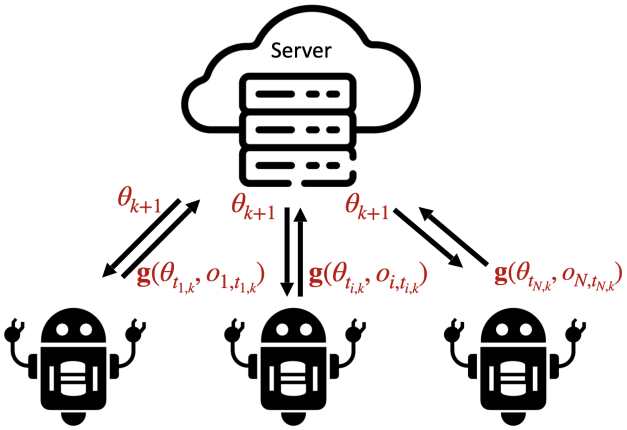


Fig. 1: DASA system model. Agents $1, \dots, N$ cooperatively work to solve the SA problem. At iteration k, the server updates $\boldsymbol{\theta}_k$ using delayed agents' operators $\{\mathbf{g}(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}})\}_{i=1}^N$ and broadcasts $\boldsymbol{\theta}_{k+1}$ to the agents.

Contributions: We propose and provide a finite-time analysis of DASA, that has the following unprecedented features: (i) DASA has a convergence rate that depends on the network delays only via the average delay τ_{avg} and (ii) at the same time, DASA achieves an N-fold linear convergence speedup with the number of agents under Markovian sampling. Providing *jointly* (i) and (ii) presents major technical challenges that we carefully address in our analysis in Section III- IV. Note that we do not need conditions on the maximum delay: it can be as large as the global number of iterations T and our results would still hold. To validate our theoretical findings, we provide simple simulations on a distributed TD learning problem in Section V.

N. Dal Fabbro and G. J. Pappas are with the Electrical and Systems Engineering Department, University of Pennsylvania. A. Adibi, H. V. Poor, and S. R. Kulkarni, are with the Department of Electrical and Computer Engineering, Princeton University. A. Mitra is with the Electrical and Computer Engineering Department, North Carolina State University.

*These authors contributed equally to this work.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

II. Problem Setting and DASA

The goal of SA is to solve a root-finding problem formulated as follows:

$$\text{Find } \boldsymbol{\theta}^* \in \mathbb{R}^m \text{ such that } \bar{\mathbf{g}}(\boldsymbol{\theta}^*) \triangleq \mathbb{E}_{o_k \sim \pi}\left[\mathbf{g}(\boldsymbol{\theta}^*, o_k)\right] = 0, \tag{1}$$

where, for a given approximation parameter $\boldsymbol{\theta} \in \mathbb{R}^m$, the deterministic function $\bar{\mathbf{g}}(\boldsymbol{\theta})$ represents the expectation of a noisy operator $\mathbf{g}(\boldsymbol{\theta}, o_k)$ taken over distribution π, and $\{o_k\}$ denotes a stochastic observation process. It is commonly assumed that $\{o_k\}$ converges in distribution to π [14], [15].

Distributed Asynchronous SA. We consider a setting in which N agents communicate with a central server to collaboratively estimate $\boldsymbol{\theta}^*$. To achieve this goal, the agents compute in parallel their versions of the noisy operator $\mathbf{g}(\boldsymbol{\theta}_k, o_{i,k})$, where $\boldsymbol{\theta}_k$ is the approximation parameter at iteration k and $\{o_{i,k}\}_{k\geq 0}$ is the local observation process of agent i. At each iteration, the agents transmit to a server their local operators via up-link communications subject to asynchronous time-varying delays. We denote by $\tau_{i,k}$ the delay of agent i at iteration k. At each iteration k, the server can use the (delayed) update directions $\{\mathbf{g}(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}})\}_{i=1}^N$ to update the SA parameter $\boldsymbol{\theta}_k$, where $t_{i,k} = k - \tau_{i,k}$ is the iteration at which the SA update direction of agent i was computed. On the other hand, *we assume that the down-link communication from the server to the agents is not subject to delays*. This assumption is practically motivated by the fact that in most client-server architectures the main communication bottleneck comes from up-link transmissions, rather than from down-link broadcasting. *We anticipate that this assumption can be removed and we plan to do so in future work.* The aggregation protocol is illustrated in Figure 1. SA with delayed updates can estimate $\boldsymbol{\theta}^*$, but with significant performance degradation in the rate, as theoretically shown, e.g., in [7], [9], [10] especially in the presence of large delays. To mitigate the impact of stragglers on the performance, we propose DASA, in which, using available delayed operators, the server updates the parameter $\boldsymbol{\theta}_k$ according to a delay-adaptive update rule, which we describe next.

DASA: We now present DASA. The server updates the parameter $\boldsymbol{\theta}_k$ according to the following rule:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \alpha \mathbf{v}_k, \tag{2}$$

where α is a constant step-size/learning rate, and

$$\mathbf{v}_k \triangleq \begin{cases} \frac{1}{|\mathcal{I}_k|} \sum\limits_{i \in \mathcal{I}_k} \mathbf{g}(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}}) & \text{if } \|\boldsymbol{\theta}_{t_{i_{med,k},k}} - \boldsymbol{\theta}_k\| \leq \epsilon \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where $\|\cdot\|$ denotes the Euclidean norm, while $i_{med,k}$, $\mathcal{I}_k$ and ϵ are defined as follows:

$$\begin{aligned} i_{med,k} &\triangleq \arg\text{median}(\{\|\boldsymbol{\theta}_{t_{j,k}} - \boldsymbol{\theta}_k\|\}_{j=1}^N), \\ \mathcal{I}_k &\triangleq \{i : \|\boldsymbol{\theta}_{t_{i,k}} - \boldsymbol{\theta}_k\| \leq \|\boldsymbol{\theta}_{t_{i_{med,k},k}} - \boldsymbol{\theta}_k\|\}, \\ \epsilon &\triangleq \max\{\alpha^{1.5}, \frac{\alpha}{\sqrt{M}}\}, \end{aligned} \tag{4}$$

with $M = |\mathcal{I}_k| = \lceil \frac{N}{2} \rceil, \forall k \geq 0$.

Remark 1. *Note that the size of $\mathcal{I}_k$ is $|\mathcal{I}_k| = \lceil \frac{N}{2} \rceil$. Here, we assume without loss of generality that N is even. Therefore, $|\mathcal{I}_k| = \frac{N}{2}$. The proof is equivalent for odd N.*

Remark 2. *If $i \in \mathcal{I}_k$, then we have $\|\boldsymbol{\theta}_{t_{i,k}} - \boldsymbol{\theta}_k\| \leq \epsilon$.*

The rationale behind DASA is to control the error introduced by delays, using, at each iteration, the information from the set of $M = N/2$ agents in $\mathcal{I}_k$ whose available operators were computed with iterates that are *the least stale at iteration k*, as dictated by the median delay error index $i_{med,k}$. At the same time, averaging $(N/2)$ noisy operators, we wish to benefit from a variance reduction effect, achieving an N-fold convergence speedup thanks to the independence of the different stochastic observation processes. DASA allows the server to control the error of the aggregated operator $\mathbf{v}_k$ that is used to update $\boldsymbol{\theta}_k$, via the threshold parameter ϵ. When carefully chosing ϵ, we obtain the convergence guarantees that we want: *a convergence rate that does not depend on the maximum delay and, jointly, a linear convergence speedup with the number of agents N*. Providing these convergence guarantees presents key technical challenges that we address in the next section.

III. Convergence Analysis

In this section, we provide the convergence analysis of DASA, under standard assumptions that we illustrate next. These assumptions are common in most finite-time analyses of SA algorithms [11]–[13].

Assumption 1. *Problem (1) admits a solution $\boldsymbol{\theta}^*$, and $\exists\, \mu > 0$ such that for all $\boldsymbol{\theta} \in \mathbb{R}^m$, we have*

$$\langle \boldsymbol{\theta} - \boldsymbol{\theta}^*, \bar{\mathbf{g}}(\boldsymbol{\theta}) - \bar{\mathbf{g}}(\boldsymbol{\theta}^*) \rangle \leq -\mu \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|^2. \tag{5}$$

Assumption 2. *There exists $L > 0$ such that for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathbb{R}^m$ and $o \in \{o_t\}$, we have*

$$\|\mathbf{g}(\boldsymbol{\theta}_1, o) - \mathbf{g}(\boldsymbol{\theta}_2, o)\| \leq L \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|. \tag{6}$$

Furthermore, there exists $\sigma \geq 1$ such that for any $\boldsymbol{\theta} \in \mathbb{R}^m$, we have

$$\|\mathbf{g}(\boldsymbol{\theta}, o)\| \leq L(\|\boldsymbol{\theta}\| + \sigma). \tag{7}$$

Assumption 3. *For each agent $i = 1, \ldots, N$, the stochastic observation process $\{o_{i,k}\}$ is an aperiodic and irreducible Markov chain. For each $i \neq j$ and for each iteration k, $o_{i,k}$ and $o_{j,k}$ are statistically independent.*

Assumption 1 ensures the strong monotonicity of the expected operator $\bar{\mathbf{g}}(\cdot)$, which is a standard requirement for proving linear convergence rates. Assumption 2 imposes Lipschitz continuity on the noisy operator $\mathbf{g}(\cdot, \cdot)$, which is also common in SA analyses. Assumption 3 implies that the Markov chain $\{o_k\}$ mixes at a geometric rate.

Definition 1. *Let τ_{mix} be such that, for every $\boldsymbol{\theta} \in \mathbb{R}^m$, for all $i = 1, ..., N$ and time-step $\bar{t}$, we have*

$$\begin{aligned} \|\mathbb{E}\left[\mathbf{g}(\boldsymbol{\theta}, o_{i,t}) | o_{i,\bar{t}}\right] - \bar{\mathbf{g}}(\boldsymbol{\theta})\| \leq \alpha^2 \left(\|\boldsymbol{\theta}\| + \sigma\right), \\ \forall t : t - \bar{t} \geq \tau_{mix}, \forall o_{i,\bar{t}}. \end{aligned} \tag{8}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

Definition 1 introduces the mixing time τ_{mix}, which quantifies the number of steps after which the conditional expectation of $\mathbf{g}(\boldsymbol{\theta}, o_{i,t})$ is close to its stationary expectation $\bar{\mathbf{g}}(\boldsymbol{\theta})$, after conditioning over any past observation $o_{i,\bar{t}}$. Note that, thanks to Assumption 3, for any $i = 1, ..., N$, the Markov chain $\{o_{i,k}\}$ mixes at a geometric rate, and hence there exists some $K \geq 1$ such that $\tau_{mix} \leq K \log(\frac{1}{\alpha})$. Under the assumptions above, we are now ready to state the main result of this paper, which is the non-asymptotic convergence analysis of DASA. We provide the convergence result in the following theorem, whose proof requires a novel and careful analysis, which we defer to the next section. In all of our results, we use the notation $y = O(x)$ to describe the relation between positive terms x and y, meaning that the value y is at most equal to some positive constant multiplying the quantity x. We define $\delta_k \triangleq \|\boldsymbol{\theta}_k - \boldsymbol{\theta}^*\|$, and assume without loss of generality that $\sigma \geq \max\{1, \|\boldsymbol{\theta}^*\|, \|\boldsymbol{\theta}_0\|\}$, $L \geq 1$ and $\mu \leq 1$. We also define an indicator function $\mathcal{X}_k$, that is equal to 1 when the median agent's parameter is not too far away from the server's parameter $\boldsymbol{\theta}_k$, and the average delay τ_{avg}:

$$\begin{aligned} \mathcal{X}_k &\triangleq \begin{cases} 1 & \text{if } \|\boldsymbol{\theta}_{t_{i_{med,k},k}} - \boldsymbol{\theta}_k\| \leq \epsilon \\ 0 & \text{otherwise} \end{cases}, \\ \tau_{avg} &\triangleq \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} \tau_{i,t}. \end{aligned} \tag{9}$$

Theorem 1. *Consider the update rule of* DASA *in* (2) *and let* $T \geq 0$ *be such that* $\mathcal{X}_T = 1$. *There exists a universal constant* $C_1 \geq 1$ *such that for* $\alpha \leq \frac{\mu}{C_1 L^2 \tau_{mix}}$ *the following holds:*

$$\begin{aligned} \mathbb{E}\left[\delta_T^2\right] \leq \exp\left(\frac{-\alpha\mu T}{8\left(\tau_{avg}+1\right)}\right) O\left(\sigma^2\right) \\ + O\left(\frac{\sigma^2 L^2 \alpha \tau_{mix}}{\mu}\right)\left(\frac{1}{N} + \alpha\right) \end{aligned} \tag{10}$$

Main Takeaways: We now discuss the main takeaways from Theorem 1. We first note that, with a constant step size α satisfying the theorem conditions, DASA guarantees mean-square exponential convergence of the approximation parameter $\boldsymbol{\theta}_T$ to a ball around $\boldsymbol{\theta}^*$. We now compare this result against existing related work and comment on the effect of the delays on the convergence rate. We also illustrate explicitly the linear convergence speedup achieved by DASA.

Effect of Delays and Dependence on τ_{avg}. Note that the exponential convergence rate of DASA depends on the average delay τ_{avg}, while neither the rate nor the convergence ball are affected by the maximum delay τ_{max}. The convergence ball, in fact, does not depend on the delay sequence at all. In this regard, note that convergence rates of non-adaptive SA algorithms under delayed updates are impacted by the maximum delay τ_{max} both in the convergence rate and in the noise ball of convergence, *even in a single-agent setting* [10]. At the same time, the step size choice that guarantees the convergence of DASA does not require any knowledge of the delay sequence, while it requires knowledge of the mixing time of the Markov sampling process, which is commonly required for finite-time convergence results in SA [11]–[13]. With respect to the single-agent SA setting with delayed updates in [10], DASA manages to remove the dependence on the maximum delay τ_{max} from both the convergence rate and from the step size choice even in the multi-agent setting *while jointly achieving a linear convergence speedup with the number of agents under Markovian sampling*. We now comment on the convergence speedup.

Linear Convergence Speedup. We now compare the convergence bound of DASA with the convergence bound obtained by the single-agent delay-adaptive algorithm provided in [10]. Note that the noise ball is deflated by $\frac{1}{N} + \alpha$, which, for a proper choice of step size α and for a number of iterations T large enough, provides a linear convergence speedup with the number of agents N. In particular, we can easily show that if we set α and T as follows:

$$\alpha = O\left(\frac{\log(NT)\tau_{avg}}{T\mu}\right), \quad T \geq \frac{CN\log(NT)\tau_{avg}\tau_{mix}L^2}{\mu^2},$$

with $C \geq 2$ a constant, then the mean-square error after T iterations is

$$\mathbb{E}\left[\delta_T^2\right] \leq O\left(\frac{\log(NT)\tau_{avg}\tau_{mix}\sigma^2 L^2}{NT\mu}\right), \tag{11}$$

which implies that DASA is *sample efficient* and that, despite the joint presence of asynchronous delays and Markovian sampling, we can achieve an N-fold linear convergence speedup when N agents act in parallel, without dependence on the maximum delay. DASA *is the first algorithm for multi-agent/distributed SA with these convergence guarantees.*

IV. Proof of Theorem 1

In this section, we prove Theorem 1, providing the complete convergence analysis of DASA. We start by introducing some notation. Let us define the following:

$$\begin{aligned} \mathcal{T}_k &\triangleq \{\tau_{i,k} : i \in \mathcal{I}_k\}, \\ \eta_{i,k}(\boldsymbol{\theta}) &\triangleq \|\mathbb{E}\left[\mathbf{g}_{i,k}(\boldsymbol{\theta}, o_{i,k}) | o_{i,k-\tau_{mix}}\right] - \bar{\mathbf{g}}(\boldsymbol{\theta})\|, \ k \geq \tau_{mix}, \\ \delta_{k,h}^2 &\triangleq \|\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-h}\|^2, k \geq h \geq 0. \end{aligned} \tag{12}$$

Outline of the Analysis: As a first step, we here provide a high-level outline of the convergence analysis of DASA. We start by writing the following recursion, which holds for the iterations $k \geq \tau_{mix}$ in which $\mathcal{X}_k = 1$,

$$\begin{aligned} \mathbb{E}\left[\delta_{k+1}^2\right] &= \mathbb{E}\left[\delta_k^2\right] - 2\alpha\mathbb{E}\left[\langle \mathbf{v}_k, \boldsymbol{\theta}_k - \boldsymbol{\theta}^*\rangle\right] + \alpha^2\mathbb{E}\left[\|\mathbf{v}_k\|^2\right] \\ &= \mathbb{E}\left[\delta_k^2\right] - 2\alpha\mathbb{E}\left[\langle \bar{\mathbf{g}}(\boldsymbol{\theta}_k), \boldsymbol{\theta}_k - \boldsymbol{\theta}^*\rangle\right] \\ &\quad - 2\alpha\mathbb{E}\left[\psi_k\right] + \alpha^2\mathbb{E}\left[\|\mathbf{v}_k\|^2\right], \\ &\overset{(*)}{\leq} (1 - 2\alpha\mu)\mathbb{E}\left[\delta_k^2\right] \\ &\quad - 2\alpha\mathbb{E}\left[\psi_k\right] + \alpha^2\mathbb{E}\left[\|\mathbf{v}_k\|^2\right], \ \text{with} \\ \psi_k &\triangleq \langle \mathbf{v}_k - \bar{\mathbf{g}}(\boldsymbol{\theta}_k), \boldsymbol{\theta}_k - \boldsymbol{\theta}^*\rangle, \end{aligned} \tag{13}$$

where for inequality $(*)$ we used Assumption 1. The main objective of the analysis is to obtain a one-step bound of the

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

following form (for iterations k in which an update occurs, i.e., in which $\mathcal{X}_k = 1$):

$$\begin{aligned}\mathbb{E}\left[\delta_{k+1}^2\right] &\leq \exp\left(-\alpha\mu\right)\mathbb{E}\left[\delta_k^2\right] \\ &+ O\left(\sigma^2 L^2 \alpha^2 \tau_{mix}\right)\left(\frac{1}{N}+\alpha\right),\end{aligned} \tag{14}$$

where we note that there is not any dependence on the maximum delay τ_{max}. In order to obtain this one-step bound, we need to provide bounds on the two terms $\mathbb{E}\left[\|\mathbf{v}_k\|^2\right]$ and $\mathbb{E}\left[\psi_k\right]$ of the following form:

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{v}_k\|^2\right] &\leq O\left(L^2\right)\mathbb{E}\left[\delta_k^2\right] + O\left(L^2\frac{\sigma^2}{N}\right) \\ &+ O\left(\sigma^2\alpha^4\right) + O\left(L^2\epsilon^2\right), \\ \mathbb{E}\left[\psi_k\right] &\leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] \\ &+ O\left(\alpha\tau_{mix}\right)\left(L^2\frac{\sigma^2}{N} + \epsilon^2 L^2 + \sigma^2\alpha^4\right).\end{aligned} \tag{15}$$

Providing these two bounds is the main technical burden of the proof of Theorem 1. The bound on $\mathbb{E}\left[\|\mathbf{v}_k\|^2\right]$ is the first key result that is used then to bound $\mathbb{E}\left[\psi_k\right]$. Once we establish the one-step bound in (14), to conclude the proof, we apply a lower bound on the number of updates performed by DASA, which is a function of τ_{avg} and that we obtain in Lemma 1. In the next paragraph, we provide the auxiliary lemmas needed to obtain the bounds illustrated above.

Auxiliary Lemmas. In this paragraph, we will often use the following two basic results. For any $a, b \in \mathbb{R}, c > 0$, and for $a_i \in \mathbb{R}, i = 1, ..., N$,

$$\begin{aligned}ab &= a\sqrt{c}\frac{b}{\sqrt{c}} \leq \frac{1}{2}\left(ca^2 + \frac{b^2}{c}\right), \\ \left(\sum_{i=1}^N a_i\right)^2 &\leq N\sum_{i=1}^N a_i^2.\end{aligned} \tag{16}$$

For the purpose of the proof and without loss of generality, we define also samples and iterates for negative iterations, defining $\boldsymbol{\theta}_j = \boldsymbol{\theta}_0$ and $o_{i,j} = o_{i,0}$ for $j \leq 0$ and for all $i = 1, ..., N$. We will also use the fact that from the definition of mixing time τ_{mix} (see Definition 1), we have

$$\eta_{i,k}(\boldsymbol{\theta}) \leq \alpha^2(\|\boldsymbol{\theta}\|+\sigma). \tag{17}$$

We start the convergence analysis by proving a lemma, partially inspired by Lemma 4 in [16], that establishes a lower bound on the number of updates that the server performs when executing DASA, which is a function of the average delay τ_{avg}.

Lemma 1. *Let τ_{avg} be the average delay experienced by agents up to iterate T, $\tau_{avg} = \frac{1}{NT}\sum_{t=1}^T\sum_{i=1}^N \tau_{i,t}$. The number of updates that DASA makes is at least $\lceil \frac{T}{8\tau_{avg}+8} \rceil$.*

Proof. Recall the definition of the average delay up to time T, $\tau_{avg} = \frac{1}{NT}\sum_{t=1}^T\sum_{i=1}^N \tau_{i,t}$. Consider $U_{2\tau_{avg}}$, the number of steps t for which the delay $\tau_t = \frac{1}{N}\sum_{i=1}^N \tau_{i,t}$ is at least $2\tau_{avg}$. We must have $U_{2\tau_{avg}} \leq \frac{T}{2}$ (otherwise the total sum of delays exceeds $\tau_{avg}T$, contradicting the definition of τ_{avg}). On the other hand, let m be the number of updates that the algorithm makes. Let $t_1 < t_2 < \ldots < t_m$ be the steps in which an update is made. Denote $t_0 = 0$ and $t_{m+1} = T$. Now, fix i and consider the steps at times $s_n = t_i + n$ for $n \in [1, 2, \ldots, t_{i+1} - t_i - 1]$. In all those steps no update takes place and $\boldsymbol{\theta}_{s_n} = \boldsymbol{\theta}_{t_i}$. We must have $\tau_{i,s_n} > n$ for all n for at least $\frac{N}{2}$ agents (otherwise $\boldsymbol{\theta}_k = \boldsymbol{\theta}_{t_{i,k}}$ for $k = s_n$ for more than $\frac{N}{2}$ agents and an update occurs since the median becomes zero). In particular, we have that $\tau_{i,s_n} \geq 4\tau_{avg}$ for at least $\frac{N}{2}$ of agents in at least $t_{i+1} - t_i - 1 - 4\tau_{avg}$ steps. Which means $\tau_{s_n} = \frac{1}{N}\sum_{i=1}^N \tau_{i,s_n} \geq 2\tau_{avg}$ in at least $t_{i+1} - t_i - 1 - 4\tau_{avg}$ steps. Formally,

$$\begin{aligned}&\#\text{ of steps in } [t_i, t_{i+1}] \text{ with delay bigger or equal to } 2\tau_{avg} \\ &\geq \max\{0, t_{i+1} - t_i - 1 - 4\tau_{avg}\} \\ &\geq t_{i+1} - t_i - 1 - 4\tau_{avg}.\end{aligned} \tag{18}$$

Hence, $U_{2\tau_{avg}} \geq \sum_{i=0}^{m-1}(t_{i+1} - t_i - 1 - 4\tau_{avg}) = T - m(1 + 4\tau_{avg})$. Finally, it follows that $T - m(1+4\tau_{avg}) \leq \frac{T}{2}$ which implies $m \geq \lceil \frac{T}{8(\tau_{avg}+1)} \rceil$. □

In the next lemma, we provide a bound on the expected squared norm of the DASA update direction, $\mathbb{E}\left[\|\mathbf{v}_k\|^2\right]$, obtaining precisely what we outlined in (15).

Lemma 2. *For $k \geq 0$, the following holds,*

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{v}_k\|^2\right] &\leq O\left(L^2\right)\mathbb{E}\left[\delta_k^2\right] + D, \quad \textit{with} \\ D &\triangleq O\left(L^2\frac{\sigma^2}{N}\right) + O\left(\sigma^2\alpha^2\right) + O\left(L^2\epsilon^2\right).\end{aligned} \tag{19}$$

Proof. We write

$$\begin{aligned}\|\mathbf{v}_k\|^2 &\leq \frac{12}{N^2}(V_1 + V_2 + V_3), \quad \text{with} \\ V_1 &= \|\sum_{i\in\mathcal{I}} \mathbf{g}(\boldsymbol{\theta}_k, o_{i,t_{i,k}}) - \mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}})\|^2, \\ V_2 &= \|\sum_{i\in\mathcal{I}} \mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}})\|^2, \\ V_3 &= \|\sum_{i\in\mathcal{I}} \mathbf{g}(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}}) - \mathbf{g}(\boldsymbol{\theta}_k, o_{i,t_{i,k}})\|^2.\end{aligned} \tag{20}$$

Note that the above bound holds regardless of $\mathcal{X}_k = 0$ or $\mathcal{X}_k = 1$. We now bound V_1,

$$\begin{aligned}V_1 &\leq \frac{N}{2}\sum_{i\in\mathcal{I}}\|\mathbf{g}(\boldsymbol{\theta}_k, o_{i,t_{i,k}}) - \mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}})\|^2 \\ &\overset{(6)}{\leq} \frac{N^2L^2}{2}\delta_k^2. \quad \text{Thus,} \\ \mathbb{E}\left[V_1\right] &\leq \frac{N^2L^2}{2}\mathbb{E}\left[\delta_k^2\right].\end{aligned} \tag{21}$$

We now proceed to bound V_2,

$$\begin{aligned}V_2 &= V_{21} + V_{22}, \quad \text{with} \\ V_{21} &= \sum_{i\in\mathcal{I}}\|\mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}})\|^2 \\ V_{22} &= \sum_{\substack{i,j\in\mathcal{I} \\ i\neq j}}\langle \mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}}), \mathbf{g}(\boldsymbol{\theta}^*, o_{j,t_{j,k}})\rangle.\end{aligned} \tag{22}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

We see that, using (7), we get

$$V_{21} \leq 2L^2N(\|\boldsymbol{\theta}^*\|^2+\sigma^2) \leq 4L^2N\sigma^2. \tag{23}$$

Now, using the fact that the observations $o_{i,k}$ and $o_{j,k'}$ are independent for $i \neq j$ and for any $k, k' \geq 0$,

$$\begin{aligned}\mathbb{E}\left[V_{22}\right] = \sum_{\substack{i,j\in\mathcal{I}\\ i\neq j}} \langle \mathbb{E}\left[\mathbb{E}\left[\mathbf{g}(\boldsymbol{\theta}^*, o_{i,t_{i,k}})|o_{i,t_{i,k}-\tau_{mix}}\right]\right], \\ \mathbb{E}\left[\mathbb{E}\left[\mathbf{g}(\boldsymbol{\theta}^*, o_{j,t_{j,k}})|o_{j,t_{j,k}-\tau_{mix}}\right]\right]\rangle.\end{aligned} \tag{24}$$

Using the fact that $\bar{\mathbf{g}}(\boldsymbol{\theta}^*) = 0$, and the Cauchy-Schwarz inequality followed by the Jensen's inequality, we can write

$$\begin{aligned}\mathbb{E}\left[V_{22}\right] &\leq \sum_{\substack{i,j\in\mathcal{I}\\ i\neq j}} \mathbb{E}\left[\eta_{i,t_{i,k}}(\boldsymbol{\theta}^*)\right] \times \mathbb{E}\left[\eta_{j,t_{j,k}}(\boldsymbol{\theta}^*)\right] \\ &\leq \frac{1}{4}N^2\alpha^4(\|\boldsymbol{\theta}^*\|+\sigma)^2 \leq N^2\alpha^4\sigma^2,\end{aligned} \tag{25}$$

where we used the mixing time definition and the consequent inequality (17). For V_3 we have

$$\begin{aligned}V_3 &\leq \frac{N}{2}\sum_{i\in\mathcal{I}}\|\mathbf{g}(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}}) - \mathbf{g}(\boldsymbol{\theta}_k, o_{i,t_{i,k}})\|^2, \\ &\leq \frac{N}{2}\sum_{i\in\mathcal{I}} L^2\|\boldsymbol{\theta}_{t_{i,k}} - \boldsymbol{\theta}_k\|^2, \\ &\leq \frac{N^2L^2\epsilon^2}{4}.\end{aligned} \tag{26}$$

Substituting the above bounds on $\mathbb{E}\left[V_1\right]$, $\mathbb{E}\left[V_2\right]$, and $\mathbb{E}\left[V_3\right]$ into (20), we can conclude the proof of the lemma. □

We need the following result, which provides a bound to the expected drift term $\mathbb{E}\left[\delta^2_{k,\tau_{mix}}\right]$.

Lemma 3. *Let D be as defined in (19). There exists a constant $C' \geq 1$ such that for $\alpha \leq \frac{1}{C'L^2\tau_{mix}}$ we have*

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k,\tau_{mix}}\right] \leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] + B_{\tau_{mix}}, \quad \text{with} \\ B_{\tau_{mix}} = O\left(\alpha^2\tau^2_{mix}D\right).\end{aligned} \tag{27}$$

Proof. We start out by establishing a bound on $\mathbb{E}\left[\delta^2_k\right]$ as a function of $\mathbb{E}\left[\delta^2_{k-\tau_{mix}}\right]$,

$$\begin{aligned}\delta^2_{k+1} &= \mathcal{X}_k\left(\delta^2_k - 2\alpha\langle\mathbf{v}_k, \boldsymbol{\theta}^* - \boldsymbol{\theta}_k\rangle + \alpha^2\mathbb{E}\left[\|\mathbf{v}_k\|^2\right]\right) \\ &\quad + (1-\mathcal{X}_k)\,\delta^2_k, \\ &\leq (1+\alpha)\delta^2_k + 2\alpha\|\mathbf{v}_k\|^2,\end{aligned} \tag{28}$$

where we used (16). Now taking the expectation and using Lemma 2 we get

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k+1}\right] &\leq (1+\alpha)\mathbb{E}\left[\delta^2_k\right] + O\left(\alpha L^2\right)\mathbb{E}\left[\delta^2_k\right] + \alpha D \\ &\leq (1+O\left(\alpha L^2\right))\mathbb{E}\left[\delta^2_k\right] + \alpha D,\end{aligned} \tag{29}$$

where D is defined as in (19). Iterating the above inequality, we get, for all k' such that $k - \tau_{mix} \leq k' \leq k$,

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k'}\right] &\leq \left(1+O\left(\alpha L^2\right)\right)^{\tau_{mix}}\mathbb{E}\left[\delta^2_{k-\tau_{mix}}\right] \\ &\quad + \alpha D\sum_{j=0}^{\tau_{mix}-1}\left(1+O\left(\alpha L^2\right)\right)^j \\ &\leq 2\mathbb{E}\left[\delta^2_{k-\tau_{mix}}\right] + 2\alpha\tau_{mix}D,\end{aligned} \tag{30}$$

where we used the fact that, given a constant $C \geq 1$ large enough, $\left(1+C\alpha L^2\right)^{\tau_{mix}} \leq \exp\left(C\tau_{mix}\alpha L^2\right) \leq \exp\left(0.25\right) \leq 2$ for $\alpha \leq \frac{1}{4C\tau_{mix}L^2}$. Now, note that, for $k \geq 2\tau_{mix}$, we can apply Lemma 2, getting

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k,\tau_{mix}}\right] &\leq \tau_{mix}\sum_{l=k-\tau_{mix}}^{k-1}\mathbb{E}\left[\|\boldsymbol{\theta}_{l+1} - \boldsymbol{\theta}_l\|^2\right] \\ &= \tau_{mix}\alpha^2\sum_{l=k-\tau_{mix}}^{k-1}\mathbb{E}\left[\|\mathbf{v}_l\|^2\right] \\ &\overset{(19)}{\leq} \tau_{mix}\alpha^2\sum_{l=k-\tau_{mix}}^{k-1}O\left(L^2\right)\mathbb{E}\left[\delta^2_l\right] + D \\ &\leq O\left(\alpha^2\tau^2_{mix}D\right) \\ &\quad + O\left(\tau_{mix}\alpha^2L^2\right)\sum_{l=k-\tau_{mix}}^{k-1}\mathbb{E}\left[\delta^2_l\right].\end{aligned} \tag{31}$$

Now, using (30) to upper bound $\mathbb{E}\left[\delta^2_l\right]$,

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k,\tau_{mix}}\right] &\leq O\left(\alpha^2\tau^2_{mix}D\right) + O\left(\tau_{mix}\alpha^2L^2\right) \\ &\quad \times \sum_{l=k-\tau_{mix}}^{k-1}\left(\mathbb{E}\left[\delta^2_{k-\tau_{mix}}\right] + 2\alpha\tau_{mix}D\right) \\ &\leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_{k-\tau_{mix}}\right] \\ &\quad + O\left(\alpha^2\tau^2_{mix}D\right),\end{aligned} \tag{32}$$

where the last inequality holds by setting $\alpha \leq \frac{1}{C'L^2\tau_{mix}}$ for a constant C' large enough. Now, note that $\delta^2_{k-\tau_{mix}} \leq 2\delta^2_k + 2\delta^2_{k,\tau_{mix}}$, from which we can conclude, setting $\alpha \leq \frac{1}{C'L^2\tau_{mix}}$, with C' a constant large enough, noting that $B_{\tau_{mix}} = O\left(\alpha^2\tau^2_{mix}D\right)$. □

We will also need bounds on some further drift terms, that we group together in the following lemma.

Lemma 4. *Let $\tilde{\tau} \in \mathcal{T}_k$ and let $B_{\tau_{mix}}$ be as in (27). The following inequalities hold,*

$$\begin{aligned}&(i)\ \mathbb{E}\left[\delta^2_{k-\tilde{\tau},\tau_{mix}}\right] \leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] + O\left(B_{\tau_{mix}}\right); \\ &(ii)\ \mathbb{E}\left[\delta^2_{k,\tau_{mix}+\tilde{\tau}}\right] \leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] + O\left(B_{\tau_{mix}}\right) + O\left(\epsilon^2\right); \\ &(iii)\ \mathbb{E}\left[\delta^2_{k-\tau_{mix}-\tilde{\tau}}\right] \leq O\left(\mathbb{E}\left[\delta^2_k\right]\right) + O\left(B_{\tau_{mix}}\right) + O\left(\epsilon^2\right).\end{aligned} \tag{33}$$

Proof. (i) Using Lemma 3, we can bound

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k-\tilde{\tau},\tau_{mix}}\right] &\leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_{k-\tilde{\tau}}\right] + B_{\tau_{mix}}, \\ &\leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] \\ &\quad + O\left(\alpha^2\tau^2_{mix}L^2\epsilon^2\right) + B_{\tau_{mix}}, \\ &\leq O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] + O\left(B_{\tau_{mix}}\right).\end{aligned} \tag{34}$$

For (ii), note that, using the above inequality,

$$\begin{aligned}\mathbb{E}\left[\delta^2_{k,\tilde{\tau}+\tau_{mix}}\right] &\leq 2\mathbb{E}\left[\delta^2_{k,\tilde{\tau}}\right] + 2\mathbb{E}\left[\delta^2_{k-\tilde{\tau},\tau_{mix}}\right], \\ &\leq 2\epsilon^2 + O\left(\alpha^2\tau^2_{mix}L^2\right)\mathbb{E}\left[\delta^2_k\right] \\ &\quad + O\left(B_{\tau_{mix}}\right).\end{aligned} \tag{35}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

Finally, we can use the above inequality to prove (iii),

$$\begin{aligned}\mathbb{E}\left[\delta_{k-\tau_{mix}-\tilde{\tau}}^2\right] &\leq 2\mathbb{E}\left[\delta_k^2\right] + 2\mathbb{E}\left[\delta_{k,\tilde{\tau}+\tau_{mix}}^2\right], \\ &\leq O\left(\mathbb{E}\left[\delta_k^2\right]\right) + O\left(\epsilon^2\right) + O\left(B_{\tau_{mix}}\right),\end{aligned} \tag{36}$$

where the last inequality follows for $\alpha \leq \frac{1}{\tau_{mix} L^2 C}$, for some constant $C \geq 1$. □

Equipped with the lemmas above, we can now focus on the more challenging part of the proof of Theorem 1, which consists in bounding $\mathbb{E}[\psi_k]$ with a bound of the form we anticipated in the outline of the proof (see (15)). We proceed in doing so by stating and proving the following lemma.

Lemma 5. *Let D be defined as in (19). There exists a constant $C \geq 1$ such that, for $\alpha \leq \frac{1}{CL^2\tau_{mix}}$,*

$$\mathbb{E}[\psi_k] \leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] + O\left(\alpha\tau_{mix}D\right). \tag{37}$$

Proof. We start by defining $\bar{\mathbf{g}}_{N,k} \triangleq \frac{2}{N}\sum_{i\in\mathcal{I}} \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}}\right)$, and, recalling the definition of ψ_k in (13), we write

$$\begin{aligned}&\psi_k = K_1 + K_2, \ \text{ with} \\ &K_1 = \langle \bar{\mathbf{g}}(\boldsymbol{\theta}_k) - \bar{\mathbf{g}}_{N,k}, \boldsymbol{\theta}_k - \boldsymbol{\theta}^*\rangle \\ &K_2 = \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}^*, \bar{\mathbf{g}}_{N,k} - \mathbf{v}_k\rangle.\end{aligned} \tag{38}$$

We now bound K_1. Using (16) and (6),

$$\begin{aligned}K_1 &\leq \alpha\delta_k^2 + O\left(\frac{1}{\alpha N^2}\right)\left\|\sum_{i\in\mathcal{I}} \bar{\mathbf{g}}(\boldsymbol{\theta}_k) - \bar{\mathbf{g}}(\boldsymbol{\theta}_{t_{i,k}})\right\|^2 \\ &\leq \alpha\delta_k^2 + O\left(\frac{L^2}{\alpha N}\right)\sum_{i\in\mathcal{I}}\left\|\boldsymbol{\theta}_k - \boldsymbol{\theta}_{t_{i,k}}\right\|^2 \\ &\leq \alpha\delta_k^2 + O\left(L^2\frac{\epsilon^2}{\alpha}\right) \\ &\leq \alpha\delta_k^2 + O\left(L^2\max\left\{\alpha^2, \frac{2\alpha}{N}\right\}\right).\end{aligned} \tag{39}$$

We then proceed to bound K_2.

$$\begin{aligned}&K_2 = K_{21} + K_{22} + K_{23}, \ \text{with} \\ &K_{21} = \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}^*, \frac{2}{N}\sum_{i\in\mathcal{I}} \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}}\right) - \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\right)\rangle, \\ &K_{22} = \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}^*, \\ &\qquad \frac{2}{N}\sum_{i\in\mathcal{I}} \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\right) - \mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}}\right)\rangle, \\ &K_{23} = \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}^*, \\ &\qquad \frac{2}{N}\sum_{i\in\mathcal{I}} \mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}}\right) - \mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}}, o_{i,t_{i,k}}\right)\rangle.\end{aligned} \tag{40}$$

Using (6) and (16), K_{21} and K_{23} can be bounded in the same way, writing $K_{21} = \frac{2}{N}\sum_{i\in\mathcal{I}} K_{21,i}$, with

$$K_{21,i} = \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}^*, \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}}\right) - \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\right)\rangle, \tag{41}$$

and note that, recalling that $t_{i,k} = k - \tau_{i,k}$, and that $\tau_{i,k} \in \mathcal{T}_k$, taking the expectation and using (33)-(i), we obtain

$$\begin{aligned}\mathbb{E}[K_{21,i}] &\leq \alpha\tau_{mix}\mathbb{E}\left[\delta_k^2\right] + \frac{L^2\mathbb{E}\left[\delta_{t_{i,k},\tau_{mix}}^2\right]}{\alpha\tau_{mix}} \\ &\leq \alpha\tau_{mix}\mathbb{E}\left[\delta_k^2\right] + O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] \\ &\quad + O\left(\frac{B_{\tau_{mix}}}{\alpha\tau_{mix}}\right) \\ &\leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] + O\left(\alpha\tau_{mix}D\right),\end{aligned} \tag{42}$$

and we can obtain the same bound for $\mathbb{E}[K_{23}]$. We now bound $\mathbb{E}[K_{22}]$. Let $\hat{\tau}_k \triangleq \max\{\tau_{i,k} \in \mathcal{T}_k\}$ and let

$$\Delta G_{i,k} \triangleq \bar{\mathbf{g}}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\right) - \mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}}\right). \tag{43}$$

Note that we can write $K_{22} = \bar{K}_1 + \bar{K}_2$, with

$$\begin{aligned}\bar{K}_1 &= \frac{2}{N}\sum_{i\in\mathcal{I}}\langle \boldsymbol{\theta}_{k-\tau_{mix}-\hat{\tau}_k} - \boldsymbol{\theta}^*, \Delta G_{i,k}\rangle, \\ \bar{K}_2 &= \langle \boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-\tau_{mix}-\hat{\tau}_k}, \frac{2}{N}\sum_{i\in\mathcal{I}}\Delta G_{i,k}\rangle.\end{aligned} \tag{44}$$

Now we write $\bar{K}_1 = \frac{2}{N}\sum_{i\in\mathcal{I}} \bar{K}_{1,i}$, with

$$\bar{K}_{1,i} = \langle \boldsymbol{\theta}_{k-\tau_{mix}-\hat{\tau}_k} - \boldsymbol{\theta}^*, \Delta G_{i,k}\rangle. \tag{45}$$

Let $t_{mix}^{(k)} = k - \hat{\tau}_k - \tau_{mix}$. Taking the expectation,

$$\mathbb{E}\left[\bar{K}_{1,i}\right] = \mathbb{E}\left[\langle \boldsymbol{\theta}_{t_{mix}^{(k)}} - \boldsymbol{\theta}^*, \mathbb{E}\left[\Delta G_{i,k}|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right]\rangle\right]. \tag{46}$$

Now, define the two following sets,

$$\begin{aligned}\Theta_{i,k} &= \{\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}-\tau_{mix}}\}, \\ \bar{\Theta}_{i,k} &= \{\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}-\tau_{mix}}, \boldsymbol{\theta}_{t_{mix}^{(k)}}\}.\end{aligned} \tag{47}$$

Now note that from standard results in probability theory (see [17, Theorem 4.1.13.]),

$$\mathbb{E}\left[\Delta G_{i,k}|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right] = \mathbb{E}\left[\mathbb{E}\left[\Delta G_{i,k}|\bar{\Theta}_k\right]|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right]. \tag{48}$$

Now note further that the following equality holds:

$$\mathbb{E}\left[\Delta G_{i,k}|\bar{\Theta}_{i,k}\right] = \mathbb{E}\left[\Delta G_{i,k}|\Theta_{i,k}\right]. \tag{49}$$

Indeed, since for $i \in \mathcal{I}_k$, $\hat{\tau}_k \geq \tau_{i,k}$, we have $t_{mix}^{(k)} \leq t_{i,k} - \tau_{mix}$. Also note that the parameter $\boldsymbol{\theta}_{t_{mix}^{(k)}}$ is a function of the samples of the Markov chain $o_{i,l}$ only for $l \leq t_{mix}^{(k)} - 1 < t_{i,k} - \tau_{mix}$. Hence, thanks to the Markov property, once we condition on $o_{i,t_{i,k}-\tau_{mix}}$, the random variable $\Delta G_{i,k}$ does not depend on the statistical information contained in $\boldsymbol{\theta}_{t_{mix}^{(k)}}$. In addition, from the definition of mixing time τ_{mix} (see Definition 1) the following inequality holds:

$$\|\mathbb{E}\left[\Delta G_{i,k}|\Theta_{i,k}\right]\| \leq \alpha^2(\|\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\| + \sigma). \tag{50}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

Using these last facts contained in equations (48), (49) and (50), we can now proceed to bound $\mathbb{E}\left[\bar{K}_{1,i}\right]$ as follows:

$$\begin{aligned}\mathbb{E}\left[\bar{K}_{1,i}\right] &= \mathbb{E}\left[\langle \boldsymbol{\theta}_{t_{mix}^{(k)}} - \boldsymbol{\theta}^*, \mathbb{E}\left[\Delta G_{i,k} | \boldsymbol{\theta}_{t_{mix}^{(k)}}\right]\rangle\right] \\ &\overset{(48),(49)}{=} \mathbb{E}\left[\langle \boldsymbol{\theta}_{t_{mix}^{(k)}} - \boldsymbol{\theta}^*, \mathbb{E}\left[\mathbb{E}\left[\Delta G_{i,k}|\Theta_k\right]|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right]\rangle\right] \\ &\leq \mathbb{E}\left[\delta_{t_{mix}^{(k)}} \left\|\mathbb{E}\left[\mathbb{E}\left[\Delta G_{i,k}|\Theta_k\right]|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right]\right\|\right] \\ &\overset{(a)}{\leq} \mathbb{E}\left[\delta_{t_{mix}^{(k)}} \mathbb{E}\left[\|\mathbb{E}\left[\Delta G_{i,k}|\Theta_k\right]\|\right]|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right] \\ &\overset{(50)}{\leq} \mathbb{E}\left[\delta_{t_{mix}^{(k)}} \mathbb{E}\left[\alpha^2\left(\|\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\|+\sigma\right)|\boldsymbol{\theta}_{t_{mix}^{(k)}}\right]\right] \\ &\overset{(b)}{\leq} \mathbb{E}\left[\delta_{t_{mix}^{(k)}} \alpha^2\left(\|\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\|+\sigma\right)\right]\end{aligned} \tag{51}$$

where for (a) we used Jensen's inequality for conditional expectation (see [17, Theorem 4.1.10.]) and (b) follows from the fact that, for random variables X and Y, $\mathbb{E}[XY|X] = X\mathbb{E}[Y|X]$ [17, see Theorem 4.1.14.]. We can now bound $\mathbb{E}\left[\bar{K}_1\right]$ using (16) and (33), recalling that $B_{\tau_{mix}} = O\left(\alpha^2\tau_{mix}^2 D\right)$, as follows:

$$\begin{aligned}\mathbb{E}\left[\bar{K}_{1,i}\right] &\leq \mathbb{E}\left[\delta_{t_{mix}^{(k)}} \alpha^2\left(\|\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}\|+\sigma\right)\right] \\ &\overset{(16)}{\leq} \mathbb{E}\left[\alpha\delta^2_{t_{mix}^{(k)}} + \alpha^3 O\left(\delta^2_{t_{i,k}-\tau_{mix}} + \sigma^2\right)\right] \\ &\overset{(33)}{\leq} O\left(\alpha\mathbb{E}\left[\delta_k^2\right]\right) + O\left(\alpha D\right),\end{aligned} \tag{52}$$

where the last inequality follows because $\alpha \leq \frac{1}{CL^2\tau_{mix}}$. We can then conclude that $\mathbb{E}\left[\bar{K}_1\right] \leq O\left(\alpha\mathbb{E}\left[\delta_k^2\right]\right) + O\left(\alpha D\right)$. Now we proceed to bound $\bar{K}_2$. Using (16),

$$\begin{aligned}\mathbb{E}\left[\bar{K}_2\right] \leq \frac{1}{2\alpha\tau_{mix}}\mathbb{E}\left[\delta^2_{k,\tau_{mix}+\tilde{\tau}}\right] \\ + O\left(\frac{\alpha\tau_{mix}}{N^2}\right)\mathbb{E}\left[\left\|\sum_{i\in\mathcal{I}}\Delta G_{i,k}\right\|^2\right].\end{aligned} \tag{53}$$

Using (33)-(ii), we can get, recalling the choice of ϵ in (4),

$$\begin{aligned}\frac{1}{2\alpha\tau_{mix}}\mathbb{E}\left[\delta^2_{k,\tau_{mix}+\tilde{\tau}}\right] &\leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] \\ &+ O\left(\frac{B_{\tau_{mix}}}{\alpha\tau_{mix}}\right) + O\left(\frac{\epsilon^2}{\alpha\tau_{mix}}\right) \\ &\leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] \\ &+ O\left(\alpha\tau_{mix}D\right).\end{aligned} \tag{54}$$

Finally, we proceed to bound $\mathbb{E}\left[\left\|\sum_{i\in\mathcal{I}}\Delta G_{i,k}\right\|^2\right]$:

$$\begin{aligned}\mathbb{E}\left[\left\|\sum_{i\in\mathcal{I}}\Delta G_{i,k}\right\|^2\right] &\overset{(16)}{\leq} \mathbb{E}\left[G_1\right] + \mathbb{E}\left[G_2\right] \quad \text{with} \\ G_1 &= 2\left\|\sum_{i\in\mathcal{I}}\bar{\mathbf{g}}(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}})\right\|^2 \\ G_2 &= 2\left\|\sum_{i\in\mathcal{I}}\mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}-\tau_{mix}}, o_{i,t_{i,k}}\right)\right\|^2.\end{aligned} \tag{55}$$

Note that, using (5) and (33)-(iii), we have

$$\begin{aligned}\mathbb{E}\left[G_1\right] &\overset{(5)}{\leq} O\left(N^2\mathbb{E}\left[\delta^2_{t_{i,k}-\tau_{mix}}\right]\right) \\ &\overset{(33)}{\leq} O\left(N^2\left(\mathbb{E}\left[\delta_k^2\right] + B_{\tau_{mix}} + \epsilon^2\right)\right).\end{aligned} \tag{56}$$

Note that G_2 can be bounded in the same way in which we bounded V_2 in Lemma 2, obtaining

$$\mathbb{E}\left[G_2\right] \leq O\left(N^2\right)\left(L^2\mathbb{E}\left[\delta_k^2\right] + \alpha^4\sigma^2 + L^2\frac{\sigma^2}{N}\right). \tag{57}$$

Substituting the bounds obtained on G_1 and G_2 we can conclude the bound on $\mathbb{E}\left[\bar{K}_2\right]$, obtaining

$$\mathbb{E}\left[\bar{K}_2\right] \leq O\left(\alpha\tau_{mix}L^2\right)\mathbb{E}\left[\delta_k^2\right] + O\left(\alpha\tau_{mix}D\right). \tag{58}$$

With this last result, we conclude the proof of the lemma. □

Conclusion of the proof. Equipped with the results of the previous paragraph, we can now conclude the proof of Theorem 1. Note that Lemma 2 and Lemma 5 provide the desired bounds on $\mathbb{E}\left[\|\mathbf{v}_k\|^2\right]$ and $\mathbb{E}\left[\psi_k\right]$ that we had showcased in the outline of the proof, in (15). If we substitute these bounds into (13), we obtain the desired one-step improvement of (14), by choosing $\alpha \leq \frac{\mu}{CL^2\tau_{mix}}$, for $C \geq 1$ large enough. Now let $T \geq 0$ be such that $\chi_T = 1$. Note that, thanks to Lemma 1, we know that the number of updates l performed by DASA is $l \geq \frac{T}{8(\tau_{avg}+1)}$. Now noting that for $0 \leq k \leq T-1$ the error $\mathbb{E}\left[\delta_{k+1}^2\right]$ either satisfies the one-step improvement in (14) or it is equal to $\mathbb{E}\left[\delta_k^2\right]$, we see that, unrolling the one-step improvement inequality in (14),

$$\begin{aligned}\mathbb{E}\left[\delta_T^2\right] &\leq \exp\left(-\frac{\alpha\mu T}{8\left(\tau_{avg}+1\right)}\right)\sigma^2 + \gamma, \quad \text{with} \\ \gamma &= O\left(\frac{\alpha\tau_{mix}\sigma^2L^2}{\mu}\right)\left(\frac{1}{N}+\alpha\right).\end{aligned} \tag{59}$$

For a more technical description of the derivation of the above inequality, the interested reader can check the analogous derivation of Eq. (253) in Appendix B.3.5 of [10].

V. Experiments

In this section, we provide simulation results to validate our theoretical analysis and show the superior performance of DASA when compared to a non-adaptive baseline of asynchronous SA. We consider a simple instance of linear SA, i.e., temporal difference (TD) learning with linear function approximation. The setting that we consider is similar to the one adopted for the simulations in [2], [18] of federated TD learning, with the difference that we let the agents' TD update directions be subject to delays, i.e., the server updates the TD approximation parameter with $\mathbf{v}_k$ defined in (3), where $\mathbf{g}\left(\boldsymbol{\theta}_{t_{i,k}}, o_{t_{i,k}}\right)$ is the TD direction computed by agent i at (delayed) time $t_{i,k}$. For the TD learning setting, the SA approximation setting consists of estimating the best approximation $\boldsymbol{\theta}^*$ of a value function associated to a policy μ that induces a Markov chain when applied to a Markov Decision Process (MDP) $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where we consider a finite state space $\mathcal{S}$ of dimension $n = 30$. With

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.

respect to the other notation, $\mathcal{A}$ is a finite set of actions, $\mathcal{P}$ is the set of action-dependent Markov transition kernels, $\mathcal{R}$ is a reward function, and the discount factor is $\gamma = 0.5$. We approximate the value function with a set of $r = 10$ basis vectors. For the considered TD learning setting, we have $\mu = (1-\gamma)\omega$, where $\omega > 0$ is an eigenvalue [11] that we compute explicitly to set the step sizes. We set the maximum delay to $\tau_{max} = 50$. In Figure 2-(a), we consider a distributed SA setting with $N = 10$ agents and compare DASA against (i) a version of distributed SA with no delays, which we refer to as "Non-delayed", and (ii) a non-adaptive version of distributed asynchronous SA, which we simply call "Delayed". Convergence guarantees for "Delayed" have been studied in [7], [19], where it was shown that tight convergence bounds are achievable for a step size α inversely proportional to $\tau_{mix}(\alpha) + \tau_{max}$. The plot clearly shows the advantage of the delay-adaptive solution (DASA) when compared to the non-adaptive SA algorithm: for the same number of updates, the approximation error achieved by DASA is almost two orders of magnitude smaller than that of the non-adaptive algorithm ("Delayed"). We can also notice how the approximation error to which DASA converges is quite close to the approximation error of the "Non-delayed" version. In Figure 2-(b), we illustrate the linear speedup effect that we established theoretically for DASA comparing a single-agent configuration ($N = 1$) to a multi-agent setting ($N = 20$). We include also the performance of the "Non-delayed" version of the algorithm for the same numbers of agents $N = 1, 20$. With the same step size, when $N = 20$ agents are estimating the value function in parallel, the approximation error is improved by more than one order of magnitude, which validates the convergence speedup effect. We can indeed note that DASA and "Non-delayed" obtain a similar speedup effect.

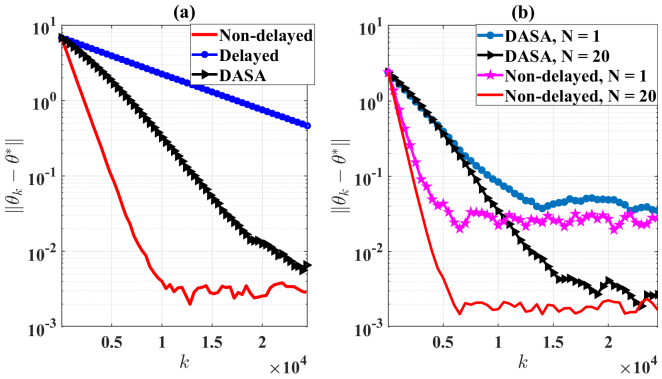


Fig. 2: In (a), we show the superior performance of DASA compared to non-adaptive distributed SA under delayed updates ("Delayed") when the number of agents is $N = 10$. In (b), we show the convergence speedup effect comparing a single-agent ($N = 1$) and a multi-agent ($N = 20$) setting. In all simulations, we set $\tau_{max} = 50$.

VI. Conclusion and Future Work

We introduced DASA, a novel algorithm that effectively mitigates the effect of asynchronous and potentially unbounded network delays and ensures collaborative speedups in multi-agent SA. Future works include the design of more advanced delay-adaptive algorithms, and extensions to fully distributed (peer-to-peer) network architectures.

References

[1] S. Khodadadian, P. Sharma, G. Joshi, and S. T. Maguluri, "Federated reinforcement learning: Linear speedup under Markovian sampling," in *Proceedings of International Conference on Machine Learning*. PMLR, 2022, pp. 10 997–11 057.

[2] N. Dal Fabbro, A. Mitra, and G. J. Pappas, "Federated TD-learning over finite-rate erasure channels: Linear speedup under Markovian sampling," *IEEE Control Systems Letters*, 2023.

[3] C. Zhang, H. Wang, A. Mitra, and J. Anderson, "Finite-time analysis of on-policy heterogeneous federated reinforcement learning," *arXiv preprint arXiv:2401.15273*, 2024.

[4] D. P. Bertsekas and J. N. Tsitsiklis, "Convergence rate and termination of asynchronous iterative algorithms," in *Proceedings of the 3rd International Conference on Supercomputing*, 1989, pp. 461–470.

[5] A. Koloskova, S. U. Stich, and M. Jaggi, "Sharper convergence guarantees for asynchronous SGD for distributed and federated learning," *Advances in Neural Information Processing Systems*, vol. 35, pp. 17 202–17 215, 2022.

[6] S. U. Stich and S. P. Karimireddy, "The error-feedback framework: Better rates for SGD with delayed gradients and compressed updates," *Journal of Machine Learning Research*, vol. 21, no. 1, pp. 9613–9648, 2020.

[7] A. Adibi, A. Mitra, and H. Hassani, "Min-max optimization under delays," *arXiv preprint arXiv:2307.06886*, 2023.

[8] H. Shen, K. Zhang, M. Hong, and T. Chen, "Towards understanding asynchronous advantage actor-critic: Convergence and linear speedup," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2579–2594, 2023.

[9] H. R. Feyzmahdavian, A. Aytekin, and M. Johansson, "A delayed proximal gradient method with linear convergence rate," in *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2014, pp. 1–6.

[10] A. Adibi, N. Dal Fabbro, L. Schenato, S. Kulkarni, H. V. Poor, G. J. Pappas, H. Hassani, and A. Mitra, "Stochastic approximation with delayed updates: Finite-time rates under markovian sampling," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2024, pp. 2746–2754.

[11] J. Bhandari, D. Russo, and R. Singal, "A finite time analysis of temporal difference learning with linear function approximation," in *Proceedings of Conference on learning theory*, 2018.

[12] R. Srikant and L. Ying, "Finite-time error bounds for linear stochastic approximation and TD learning," in *Proceedings of Conference on Learning Theory*. PMLR, 2019, pp. 2803–2830.

[13] Z. Chen, S. Zhang, T. T. Doan, J.-P. Clarke, and S. T. Maguluri, "Finite-sample analysis of nonlinear stochastic approximation with applications in reinforcement learning," *Automatica*, vol. 146, p. 110623, 2022.

[14] V. S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*. Springer, 2009, vol. 48.

[15] S. Meyn, "Stability of Q-Learning through design and optimism," *arXiv preprint arXiv:2307.02632*, 2023.

[16] A. Cohen, A. Daniely, Y. Drori, T. Koren, and M. Schain, "Asynchronous stochastic optimization robust to arbitrary delays," *Advances in Neural Information Processing Systems*, vol. 34, pp. 9024–9035, 2021.

[17] R. Durrett, *Probability: Theory and Examples*. Cambridge university press, 2019, vol. 49.

[18] H. Wang, A. Mitra, H. Hassani, G. J. Pappas, and J. Anderson, "Federated temporal difference learning with linear function approximation under environmental heterogeneity," *Transactions on Machine Learning Research*, 2023.

[19] N. Dal Fabbro, A. Adibi, A. Mitra, and G. J. Pappas, "Finite-time analysis of asynchronous multi-agent TD learning," in *2024 American Control Conference (ACC)*. IEEE, 2024.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 14:27:29 UTC from IEEE Xplore. Restrictions apply.