

**A PARAMETER-FREE ZERO-ORDER ALGORITHM
FOR DECENTRALIZED STOCHASTIC CONVEX
OPTIMIZATION****J. CHEN** , **A.V. ROGOZIN** 

*Communicated by the Programme Committee
of the XXV International Conference
"Mathematical Optimization Theory and Operations Research" (MOTOR 2026)*

Abstract: We study decentralized stochastic convex optimization on connected network, in which gradients of agents are unavailable and each agent can query only noisy function values of its own local objective. The goal is to minimize the average objective over a compact convex domain using only local two point zeroth-order oracles and peer-to-peer communication. We propose a decentralized POEM method (D-POEM) that combines symmetric two point smoothing with adaptive radius and stepsize rules, thereby avoiding prior knowledge of the Lipschitz constant and diameter. For convex Lipschitz continuous objectives, we prove an upper bound that separates a centralized optimization term from a network disagreement term. We further conduct the numerical experiments to show POEM outperforms existing distributed zeroth-order method.

Keywords: Gradient free optimization, Parameter free method, Decentralized convex optimization.

CHEN, J., ROGOZIN, A.V., A PARAMETER-FREE ZERO-ORDER ALGORITHM FOR DECENTRALIZED STOCHASTIC CONVEX OPTIMIZATION.

© 2026 CHEN, J., ROGOZIN, A.V..

The research was supported by Russian Science Foundation [project No. 23-11-00229-II](#).

Received July, 10, 2026, Published September, 14, 2026.

1 Introduction

Decentralized zeroth-order optimization studies multi-agent problems in which each node can query only noisy function values of its local objective and can communicate only with graph neighbors. This setting appears in simulation-based control[1], bandit feedback[2] and privacy-constrained learning[3], and it is challenging because the agents must optimize collaboratively without direct gradient information.

Classical stochastic zeroth-order methods based on randomized smoothing and finite differences show that convex optimization from function values is possible with near-optimal dimension dependence, especially with two-point feedback [4, 5, 6]. Their performance, however, typically depends on tuned stepsizes and smoothing radii. In the centralized setting, POEM replaces manual tuning parameter with an adaptive schedule for stepsize and smoothing radius [7]. It is closely related to the first order DoG principle [8], where the stepsize is adapted from traveled distance and accumulated gradient norms.

Recent works in decentralized optimization has shown that parameter-free adaptation is possible in several first-order settings, including convex [9], strongly convex [10], and nonconvex problems [11]. However, these methods do not address the zeroth-order setting, where the algorithm must handle smoothing, estimator variance, and communication over the network at the same time. On the other hand, existing decentralized zeroth-order methods typically rely on prescribed stepsize and smoothing schedules [12, 13], and therefore do not provide a fully parameter-free design. In this sense, our work lies at the intersection of two previously separate directions: parameter-free decentralized optimization and decentralized zeroth-order optimization.

The main contributions of this work are the following.

- We propose D-POEM, a decentralized parameter-free stochastic zeroth-order method that integrates a two-point zeroth-order oracle, and a fully decentralized adaptive scheduler.
- We derive a function gap bound that matches the centralized counterpart up to network dependent terms, with a separation between the optimization error and the consensus error.
- We show that D-POEM outperforms existing distributed zeroth-order baselines on experiments.

2 Problem Setup

We study the decentralized stochastic convex problem

$$\min_{x \in \mathcal{X}} f(x) := \frac{1}{n} \sum_{i=1}^n f_i(x), \quad f_i(x) := \mathbb{E}_{\xi_i} [F_i(x; \xi_i)]. \quad (1)$$

Each agent $i \in \{1, \dots, n\}$ knows only its own stochastic oracle F_i , can query only function values, and can communicate only with its graph neighbors.

Let $x^* \in \arg \min_{x \in \mathcal{X}} f(x)$ denote a minimizer of (1). All oracle randomness, including ξ_i and the random directions used by the algorithm, is defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $W = (w_{ij}) \in \mathbb{R}^{n \times n}$ be a nonnegative, symmetric, doubly stochastic mixing matrix adapted to the communication graph. Write $\sigma := \|W - J\|_2 < 1$, where $J := \frac{1}{n} \mathbf{1}\mathbf{1}^\top$ is the averaging matrix. The matrix J averages over the network, while σ measures how fast contract of disagreement via one communication step. Smaller σ means a better connected graph.

Assumption 1. *The set $\mathcal{X} \subset \mathbb{R}^d$ is compact and convex with diameter $D_{\mathcal{X}} := \max_{x, y \in \mathcal{X}} \|x - y\|$.*

Assumption 2. *There exists a convex set $\mathcal{X}^+ \supseteq \mathcal{X}$ that contains all oracle query points used by the algorithm. For every realization of ξ_i , the map $x \mapsto F_i(x; \xi_i)$ is convex and L -Lipschitz on $\mathcal{X}^+$.*

Assumption 3. *For every $x \in \mathcal{X}$, $\mu > 0$, and $u \sim \text{Unif}(\mathbb{S}^{d-1})$, agent i can query the pair*

$$F_i(x + \mu u; \xi_i), \quad F_i(x - \mu u; \xi_i),$$

using the same sample ξ_i .

Assumption 1 prevents the iterates from escaping to infinity. Assumption 2 controls how fast each oracle can change. Assumption 3 is the standard zeroth-order requirement that lets us build a symmetric two-point gradient estimator with reduced noise.

3 Method

We describe D-POEM, a decentralized parameter-free stochastic zeroth-order method shown in Algorithm 1. D-POEM extends the main ideas of POEM to a decentralized network setting by combining adaptive zeroth-order updates with two average-consensus steps.

In the centralized setting, POEM follows the projected SGD-type recursion

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta_t g_t),$$

where g_t is a gradient estimator obtained from random gradient estimation, and the stepsize η_t is chosen in a DoG-style adaptive manner. More precisely, the stepsize is defined as the ratio between a radius term, measuring the distance traveled from the initial point, and the square root of the accumulated squared norms of the stochastic gradient g_t .

In the decentralized setting, each agent i maintains its own local iterate $x_{i,k}$, gradient estimator $g_{i,k}$, and adaptive quantities. The local radius is first updated as $\hat{r}_{i,k} := \max\{\bar{r}_{i,k-1}, \|x_{i,k} - x_{i,0}\|\}$, then a global radius proxy is built with average consensus $\bar{r}_{i,k} := \sum_{j=1}^n w_{ij} \hat{r}_{j,k}$. The accumulated gradient norm is updated by $G_{i,k} = G_{i,k-1} + \|g_{i,k}\|^2$, and the corresponding adaptive

Algorithm 1 Decentralized POEM (D-POEM)

Require: Number of agents n , mixing matrix $W = [w_{ij}]$, initial point $x_{i,0} = x_0 \in \mathcal{X}$ for all $i \in [n]$, radius floor $r_\epsilon \in (0, D_{\mathcal{X}}]$, horizon $T \geq 1$

- 1: $\bar{r}_{i,-1} \leftarrow r_\epsilon, \quad G_{i,-1} \leftarrow r_\epsilon, \quad \forall i \in [n]$
- 2: **for** $k = 0, 1, \dots, T-1$ **do**
- 3: **for all** agents $i \in [n]$ **in parallel do**
- 4: $\hat{r}_{i,k} \leftarrow \max\{\bar{r}_{i,k-1}, \|x_{i,k} - x_{i,0}\|\}$
- 5: Receive $\{\hat{r}_{j,k}\}_{j \in \mathcal{N}_i}$ from neighbors
- 6: $\bar{r}_{i,k} \leftarrow \sum_{j=1}^n w_{ij} \hat{r}_{j,k}$
- 7: $\mu_{i,k} \leftarrow \bar{r}_{i,k} \sqrt{\frac{d}{k+1}}$
- 8: Sample $v_{i,k} \sim \text{Unif}(\mathbb{S}^{d-1})$ and a fresh oracle sample $\xi_{i,k}$
- 9: $g_{i,k} \leftarrow \frac{d}{2\mu_{i,k}} (F_i(x_{i,k} + \mu_{i,k}v_{i,k}; \xi_{i,k}) - F_i(x_{i,k} - \mu_{i,k}v_{i,k}; \xi_{i,k}))v_{i,k}$
- 10: $G_{i,k} \leftarrow G_{i,k-1} + \|g_{i,k}\|^2$
- 11: $\eta_{i,k} \leftarrow \frac{\bar{r}_{i,k}}{\sqrt{G_{i,k}}}$
- 12: Receive $\{x_{j,k}\}_{j \in \mathcal{N}_i}$ from neighbors
- 13: $z_{i,k} \leftarrow \sum_{j=1}^n w_{ij} x_{j,k}$
- 14: $x_{i,k+1} \leftarrow \Pi_{\mathcal{X}}(z_{i,k} - \eta_{i,k} g_{i,k})$
- 15: **end for**
- 16: **end for**
- 17: **Output:** iterate sequence $\{x_{i,k}\}_{i \in [n], 0 \leq k \leq T}$

stepsize is given by

$$\eta_{i,k} = \frac{\bar{r}_{i,k}}{\sqrt{G_{i,k}}}.$$

Before applying the projected update, each agent performs a consensus step on the primal iterates, $z_{i,k} = \sum_{j=1}^n w_{ij} x_{j,k}$, so that the decentralized update takes the form $x_{i,k+1} = \Pi_{\mathcal{X}}(z_{i,k} - \eta_{i,k} g_{i,k})$. Thus, D-POEM can be viewed as a decentralized analogue of POEM in which both the adaptive radius signal and the primal iterates are coordinated through local communication.

4 Main Result

We analyze the weighted output $\tilde{x}_t$ defined bellow

$$\bar{x}_t := \frac{1}{n} \sum_{i=1}^n x_{i,t}, \quad \bar{r}_t := \frac{1}{n} \sum_{i=1}^n \bar{r}_{i,t}, \quad \tilde{x}_t := \frac{\sum_{k=0}^{t-1} \bar{r}_k \bar{x}_k}{R_t}, \quad R_t := \sum_{k=0}^{t-1} \bar{r}_k.$$

The first step is a decomposition of the weighted function gap. By convexity of f and the definition of $\tilde{x}_t$, for every $t \geq 1$,

$$f(\tilde{x}_t) - f(x^*) \leq \frac{1}{\sum_{k=0}^{t-1} \bar{r}_k} \sum_{k=0}^{t-1} \bar{r}_k (f(\bar{x}_k) - f(x^*)).$$

For each k , Lipschitz continuity gives

$$\begin{aligned} \bar{r}_k (f(\bar{x}_k) - f(x^*)) &\leq \frac{\bar{r}_k}{n} \sum_{i=1}^n (f_{i,\mu_{i,k}}(x_{i,k}) - f_{i,\mu_{i,k}}(x^*)) \\ &\quad + \frac{\bar{r}_k}{n} \sum_{i=1}^n (2L\mu_{i,k} + L\|x_{i,k} - \bar{x}_k\|). \end{aligned}$$

Here we used $f = (1/n) \sum_i f_i$ and $|f_{i,\mu}(x) - f_i(x)| \leq L\mu$, where $f_{i,\mu}(x)$ is the random smooth of f_i with smooth radius μ . By convexity of $f_{i,\mu_{i,k}}$ and smoothed gradient $\nabla f_{i,\mu_{i,k}}(x_{i,k})$ with $\Delta_{i,k} := \nabla f_{i,\mu_{i,k}}(x_{i,k}) - g_{i,k}$,

$$f_{i,\mu_{i,k}}(x_{i,k}) - f_{i,\mu_{i,k}}(x^*) \leq \langle g_{i,k}, z_{i,k} - x^* \rangle + \langle \Delta_{i,k}, x_{i,k} - x^* \rangle + \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle.$$

Multiplying by $\bar{r}_k/n$ and summing over i and k gives

$$\begin{aligned} \sum_{k=0}^{t-1} \bar{r}_k (f(\tilde{x}_t) - f(x^*)) &\leq \underbrace{\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, z_{i,k} - x^* \rangle}_{\text{weighted regret bound}} + \underbrace{\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle \Delta_{i,k}, x_{i,k} - x^* \rangle}_{\text{gradient noise error}} \\ &\quad + \underbrace{\sum_{k=0}^{t-1} \frac{2L\bar{r}_k}{n} \sum_{i=1}^n \mu_{i,k}}_{\text{bias error}} + \underbrace{\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle + \sum_{k=0}^{t-1} \frac{L\bar{r}_k}{n} \sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|}_{\text{consensus error}}, \end{aligned} \quad (2)$$

We begin with the projected descent component inherited from the centralized POEM recursion.

Lemma 1 (Weighted regret bound). *Assume $\|g_{i,k}\| \leq Ld$ for all i, k . Then, for every $t \geq 1$,*

$$\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, z_{i,k} - x^* \rangle \leq 3D_{\mathcal{X}} \bar{r}_t \sqrt{\hat{G}_{t-1}} + \frac{\sqrt{nd} D_{\mathcal{X}} L \bar{r}_t}{1 - \sigma},$$

where $\hat{G}_{t-1} = \max_i \{G_{i,t}\}$.

The next term is stochastic: the algorithm uses the zeroth-order estimator $g_{i,k}$ instead of the smoothed gradient $\nabla f_{i,\mu_{i,k}}(x_{i,k})$. The following lemma controls the resulting martingale difference sequence.

Lemma 2 (Estimator noise). *Assume $\|g_{i,k}\| \leq Ld$ almost surely for all $i \in [n]$ and $k \geq 0$. Then, for every $\delta \in (0, 1)$, there exists an event $\Omega_{T,\delta} \subseteq \Omega$ with $\mathbb{P}(\Omega_{T,\delta}) \geq 1 - \delta$ such that, simultaneously for all $1 \leq t \leq T$,*

$$\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle \Delta_{i,k}, x_{i,k} - x^* \rangle \leq 8D_{\mathcal{X}} \bar{r}_{t-1} \sqrt{\theta_{t,\delta} \hat{G}_{t-1} + 4L^2 d^2 \theta_{t,\delta}^2}.$$

where $\theta_{t,\delta} := \log(60 \log(6t/\delta))$.

Since D-POEM works with smoothed local objectives, we next bound the bias introduced by the smoothing radii $\mu_{i,k}$.

Lemma 3 (Smoothing bias). *For every $t \geq 1$,*

$$\sum_{k=0}^{t-1} \frac{2L\bar{r}_k}{n} \sum_{i=1}^n \mu_{i,k} \leq 4D_{\mathcal{X}} L\bar{r}_t \sqrt{dt}.$$

The remaining term is caused by the mismatch between local iterates and their network average. We handle this in two steps. First, we estimate the average size of the disagreement process. Second, we substitute that estimate into the consensus-error sum in (2).

Lemma 4 (Time-averaged consensus error). *Under Assumptions 1–3 and common initialization $x_{1,0} = \dots = x_{n,0}$, for any horizon $K \geq 1$, define*

$$e_k := \left(\sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|^2 \right)^{1/2},$$

then

$$\frac{1}{K} \sum_{k=0}^{K-1} e_k \leq \frac{D_{\mathcal{X}}}{1-\sigma} \sqrt{\frac{1}{K} \sum_{i=1}^n \log\left(\frac{G_{i,K-1}}{G_{i,0}}\right)}.$$

The previous lemma quantifies how much disagreement can accumulate on average along the trajectory. The next bound feeds that control back into the decomposition and turns it into an explicit estimate for the disagreement penalty.

Lemma 5 (Consensus error). *For every $t \geq 1$, define*

$$e_k := \left(\sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|^2 \right)^{1/2},$$

then

$$\begin{aligned} & \sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle + \sum_{k=0}^{t-1} \frac{L\bar{r}_k}{n} \sum_{i=1}^n \|x_{i,k} - \bar{x}_k\| \\ & \leq \frac{2}{n} \sum_{k=0}^{t-1} \bar{r}_k \left(\sum_{i=1}^n \|g_{i,k}\|^2 \right)^{1/2} e_k + \frac{L}{\sqrt{n}} \sum_{k=0}^{t-1} \bar{r}_k e_k. \end{aligned}$$

Combining the primary bounds with the decomposition (2) yields the main finite-time guarantee. The result keeps the network term explicit, which is useful before passing to the expected-rate statement.

Theorem 1 (Function gap). *Under Assumptions 1–3 and the hypotheses of Lemma 2, on the event $\Omega_{T,\delta}$ from Lemma 2, for every $1 \leq t \leq T$,*

$$\begin{aligned} f(\tilde{x}_t) - f(x^*) &\leq \frac{16D\chi\theta_{T,\delta}}{R_t/\bar{r}_t} (\sqrt{\hat{G}_T} + Ld + L\sqrt{dT}) + \frac{\sqrt{nd}D\chi L}{R_t/\bar{r}_t(1-\sigma)} \\ &\quad + \frac{1}{R_t} \left[\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle + \sum_{k=0}^{t-1} \frac{L\bar{r}_k}{n} \sum_{i=1}^n \|x_{i,k} - \bar{x}_k\| \right]. \end{aligned} \quad (3)$$

The final step is to show that, after choosing the most favorable weighted output time and taking conditional expectation, the remaining disagreement term contributes only the standard network factor $1/(1-\sigma)$.

Theorem 2 (Conditional convergence rate). *Under the assumptions of Theorem 1 and common initialization $x_{1,0} = \dots = x_{n,0}$, the output $\bar{x}_{\text{out}}$ defined below satisfies*

$$\mathbb{E}[f(\bar{x}_{\text{out}}) - f(x^*) \mid \Omega_{T,\delta}] = \tilde{\mathcal{O}}\left(\frac{LD\chi\sqrt{d}}{\sqrt{T}} + \frac{LD\chi\sqrt{d}}{(1-\sigma)\sqrt{T}} + \frac{\sqrt{nd}D\chi L}{T(1-\sigma)}\right), \quad (4)$$

where

$$\tau_T := \arg \max_{1 \leq t \leq T} \frac{R_t}{\bar{r}_t}, \quad \bar{x}_{\text{out}} := \tilde{x}_{\tau_T},$$

and $\mathbb{P}(\Omega_{T,\delta}) \geq 1 - \delta$.

For this choice of τ_T , the monotone-sequence bound used in the POEM analysis yields

$$\frac{R_{\tau_T}}{\bar{r}_{\tau_T}} \geq \frac{T}{e(1 + \log(D\chi/r_\epsilon))}.$$

Corollary 1 (Stochastic zeroth-order and communication complexity). *Under the assumptions of Theorem 2 and $T \geq n$, to guarantee $\mathbb{E}[f(\bar{x}_{\text{out}}) - f(x^*) \mid \Omega_{T,\delta}] \leq \varepsilon$, it suffices to take*

$$T_\varepsilon = \tilde{\mathcal{O}}\left(\frac{L^2 D_\chi^2 d}{\varepsilon^2} \left(1 + \frac{1}{(1-\sigma)^2}\right)\right). \quad (5)$$

Since each iteration of Algorithm 1 query zeroth order oracle two times and two communication rounds, the per-agent stochastic zeroth-order query complexity and the communication complexity are

$$N_{\text{comm}}(T_\varepsilon) = N_{\text{ZO}}^{(i)}(T_\varepsilon) = \tilde{\mathcal{O}}\left(\frac{L^2 D_\chi^2 d}{\varepsilon^2} \left(1 + \frac{1}{(1-\sigma)^2}\right)\right)$$

Remark 1. *The condition $T \geq n$ guarantees that the error term caused by the imperfect consensus of the local radii can be absorbed into the consensus error associated with x . This requirement is mild in typical decentralized optimization regimes, where the number of agents n is usually much smaller than the iteration horizon T . If in the exceptional case $n > T$, then performing $\sqrt{\frac{n}{T}}$ communication rounds for the consensus of the local radii is sufficient to reduce this error term, this communication cost is negligible because local radii is a scalar.*

5 Numerical Experiments

We evaluate D-POEM on decentralized stochastic hinge-loss binary classification:

$$\min_{x \in \mathcal{X}} f(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{(a,b)} [f_i(x; a, b)],$$

where $f_i(x; a, b) = \max\{0, 1 - b\langle a, x \rangle\}$, with $(a, b) \in \mathbb{R}^d \times \{\pm 1\}$ sampled uniformly from the local dataset of agent i . We conduct experiments on the **mushrooms** ($d = 112$, $N = 8124$), **a9a** ($d = 123$, $N = 32,561$), and **w8a** ($d = 300$, $N = 49,749$) datasets. The data are partitioned i.i.d. across the 20 agents. The communication network is modeled as an Erdős–Rényi graph with edge probability 0.25, and the feasible set $\mathcal{X}$ is the Euclidean ball of radius 1. In our experiments D-POEM compares against two variants of the distributed subgradient-free algorithm (DSF) [13]: the tuned version **DSF-T** and the default-setting version **DSF-D**, and two variants of MAZOPA [12]: the tuned version **MAZOPA-T** and the default-setting version **MAZOPA-D**. We use the subgradient gradient descent method to approximate the optimal solution x^* .

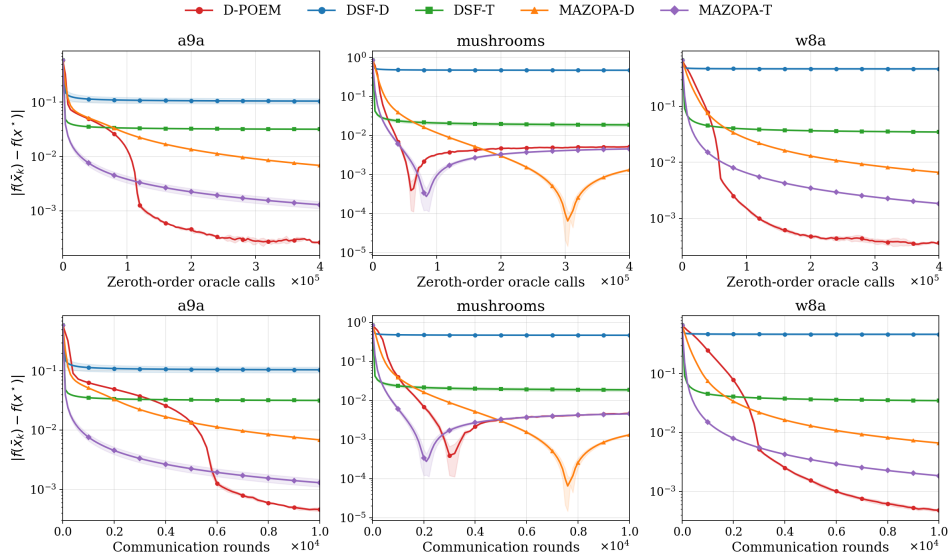


FIG. 1. Twenty-agent benchmark on **mushrooms**, **a9a**, and **w8a**. The top row plots objective at the network average versus total network-wide zeroth-order oracle calls; the bottom row plots the same objective versus communication rounds.

Figure 1 shows D-POEM is better than MAZOPA and DSF family method regardless of tuned or default, and the performance in D-POEM in the mushrooms is similar with MAZOPA family method. These results indicate that D-POEM is a robust method that performs well without parameter tuning.

TABLE. 1. Absolute difference $|f(\bar{x}_k) - f(x^*)|$ after 10^5 iterations for different values of r_ϵ . All reported entries are scaled by 10^{-4} . The starred column corresponds to $r_\epsilon = D_{\mathcal{X}} = 10^0$.

Dataset	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^0*	10^1
mushrooms	1.42	1.51	1.59	1.65	1.68	50.0
a9a	0.59	0.67	0.76	0.84	0.87	10.1
w8a	1.12	1.31	1.51	1.69	1.77	16.2

Table 1 makes the robustness claim more explicit. For all three datasets, the final objectives for $r_\epsilon \in [10^{-4}, 10^0]$ are nearly indistinguishable at the last iterate, while the degradation becomes visible only for much larger initial values such as 10^1 . In particular, the starred choice $r_\epsilon = 10^0 = D_{\mathcal{X}}$, which is the diameter of the feasible domain in our experiments, remains very close to the best final value on every dataset. This indicates that even choosing r_ϵ at the natural scale of the domain still achieves competitive performance. This is the sense in which different initializations of r_ϵ reach similar last function values.

To examine the connectivity of graph effect on our method, we use 5 graph topologies: complete, star, path, chain and random graph. The plots show our results is still robust to different graph with 20 nodes.

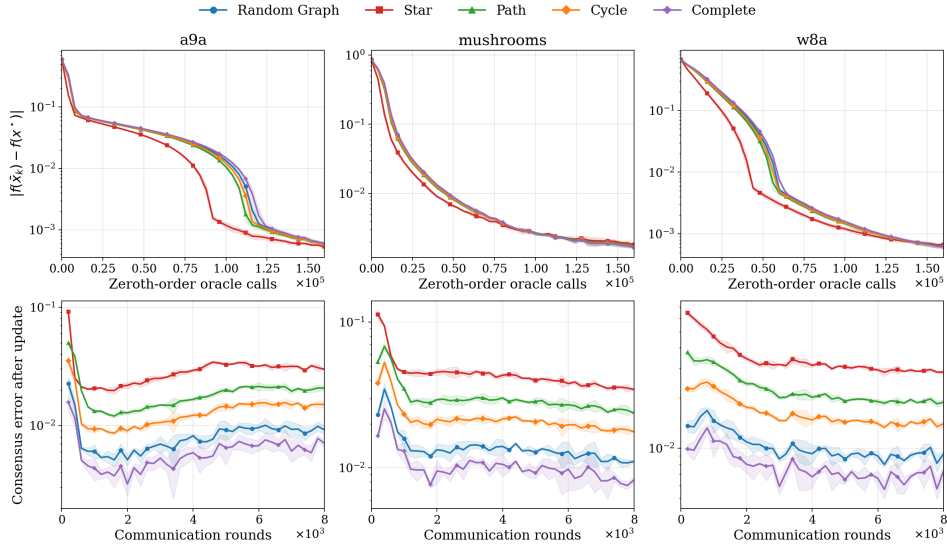


FIG. 2. performance of D-POEM under five 20-agent communication topologies.

6 Conclusion and Discussion

We studied decentralized stochastic convex optimization with zeroth-order oracle over a connected network and proposed D-POEM, a parameter-free distributed method that combines symmetric two-point gradient estimation with adaptive smoothing radii and stepsizes. We established a finite-time bound that separates the optimization term from the network disagreement term. The errors from optimization part almost match the convergence rate in [7]. The experiments on decentralized classification problems also indicate that D-POEM is robust without manual tuning and becomes competitive with tuned distributed zeroth-order baselines after a moderate number of iterations. From a practical viewpoint, the main benefit of D-POEM is the reduction of tuning effort in decentralized zeroth-order settings. In many black-box multi-agent problems, suitable stepsizes and smoothing radii are difficult to choose in advance, especially when problem scales vary across datasets or network topologies. D-POEM removes this requirement while remaining competitive with tuned baselines, making it attractive in settings where gradient information is unavailable and repeated hyperparameter search is costly.

Discussion. Several questions remain open. On the theoretical side, lower bounds for decentralized zeroth-order convex optimization are still lacking, and establishing them would help assess the efficiency of our method in terms of both communication and oracle complexity. It would also be valuable to sharpen the dependence on the network connectivity factor $1 - \sigma$, and derive fully high-probability guarantees for the final output.

References

- [1] D.S. Zalkind, E. Dall’Anese, L.Y. Pao, *Automatic controller tuning using a zeroth-order optimization algorithm*, Wind Energy Science, **5**:4 (2020), 1579–1600.
- [2] K.K. Patel, A. Saha, L. Wang, N. Srebro, *Distributed online and bandit convex optimization*, OPT 2022: Optimization for Machine Learning (NeurIPS 2022 Workshop), 2022.
- [3] C. Gratton, N.K.D. Venkatesh, R. Arablouei, S. Werner, *Privacy-preserving distributed zeroth-order optimization*, arXiv preprint arXiv:2008.13468, 2020.
- [4] Y. Nesterov, V. Spokoiny, *Random gradient-free minimization of convex functions*, Foundations of Computational Mathematics, **17**:2 (2017), 527–566.
- [5] J.C. Duchi, M.I. Jordan, M.J. Wainwright, A. Wibisono, *Optimal rates for zero-order convex optimization: The power of two function evaluations*, IEEE Transactions on Information Theory, **61**:5 (2015), 2788–2806.
- [6] O. Shamir, *An optimal algorithm for bandit and zero-order convex optimization with two-point feedback*, Journal of Machine Learning Research, **18**:52 (2017), 1–11.
- [7] K. Ren, L. Luo, *A Parameter-Free and Near-Optimal Zeroth-Order Algorithm for Stochastic Convex Optimization*, Forty-second International Conference on Machine Learning, 2025.
- [8] M. Ivgi, O. Hinder, Y. Carmon, *DoG is SGD’s best friend: A parameter-free dynamic step size schedule*, International conference on machine learning, 14465–14499, PMLR, 2023.

- [9] X. Chen, I. Kuruzov, G. Scutari, A. Gasnikov, *A parameter-free decentralized algorithm for composite convex optimization*, 2025 IEEE 64th Conference on Decision and Control (CDC), 1544–1551, IEEE, 2025.
- [10] I. Kuruzov, G. Scutari, A. Gasnikov, *Achieving linear convergence with parameter-free algorithms in decentralized optimization*, Advances in Neural Information Processing Systems, **37** (2024), 96011–96044.
- [11] J. Li, X. Chen, S. Ma, M. Hong, *Problem-parameter-free decentralized nonconvex stochastic optimization*, arXiv preprint arXiv:2402.08821, 2024.
- [12] D. Yuan, L. Wang, A. Proutiere, G. Shi, *Distributed zeroth-order optimization: Convergence rates that match centralized counterpart*, Automatica, **159** (2024), 111328.
- [13] Y. Wang, W. Zhao, Y. Hong, M. Zamani, *Distributed subgradient-free stochastic optimization algorithm for nonsmooth convex functions over time-varying networks*, SIAM Journal on Control and Optimization, **57**:4 (2019), 2821–2842.
- [14] W. Li, M. Assaad, *Distributed zeroth-order stochastic optimization in time-varying networks*, arXiv preprint arXiv:2105.12597, 2021.
- [15] E. Mhanna, M. Assaad, *Single point-based distributed zeroth-order optimization with a non-convex stochastic objective function*, International Conference on Machine Learning, 24701–24719, PMLR, 2023.
- [16] Z. Lin, J. Xia, Q. Deng, L. Luo, *Decentralized gradient-free methods for stochastic non-smooth non-convex optimization*, Proceedings of the AAAI Conference on Artificial Intelligence, **38**, 17477–17486, 2024.
- [17] J. Hou, X. Zeng, C. Chen, *Distributed gradient-free and projection-free algorithm for stochastic constrained optimization*, Autonomous Intelligent Systems, **4**:1 (2024), 6.
- [18] H. Mu, Y. Tang, Z. Li, *Variance-reduced gradient estimator for nonconvex zeroth-order distributed optimization*, arXiv preprint arXiv:2409.19567, 2024.
- [19] Z. Li, B. Ying, Z. Liu, C. Dong, H. Yang, *Achieving dimension-free communication in federated learning via zeroth-order optimization*, arXiv preprint arXiv:2405.15861, 2024.

A Proofs

A.1. Auxiliary Lemma.

Lemma 6 ([6], Lemma 10). *Under Assumption 2, each D-POEM estimator satisfies*

$$\|g_{i,k}\| \leq Ld \quad \text{and} \quad \mathbb{E}[\|g_{i,k}\|^2] \leq cL^2d,$$

where $c > 0$ is a numerical constant.

A.2. Primary bounds.

Proof of Lemma 1. Define

$$d_{i,k} := \|x_{i,k} - x^*\|, \quad \hat{d}_{i,t} := \max_{0 \leq s \leq t} d_{i,s}.$$

By nonexpansiveness of projection and $x^* \in \mathcal{X}$,

$$\|x_{i,k+1} - x^*\|^2 = \|\Pi_{\mathcal{X}}(z_{i,k} - \eta_{i,k}g_{i,k}) - \Pi_{\mathcal{X}}(x^*)\|^2 \leq \|z_{i,k} - \eta_{i,k}g_{i,k} - x^*\|^2.$$

Expanding and rearranging gives

$$\langle g_{i,k}, z_{i,k} - x^* \rangle \leq \frac{\|z_{i,k} - x^*\|^2 - \|x_{i,k+1} - x^*\|^2}{2\eta_{i,k}} + \frac{\eta_{i,k}}{2} \|g_{i,k}\|^2. \quad (6)$$

Multiplying by $\bar{r}_k$, inserting $\eta_{i,k} = \bar{r}_{i,k}/\sqrt{G_{i,k}}$, and setting

$$\mathcal{M}_{i,k} := \frac{\bar{r}_k - \bar{r}_{i,k}}{2\eta_{i,k}} (\|z_{i,k} - x^*\|^2 - \|x_{i,k+1} - x^*\|^2).$$

Since $\hat{G}_k = \max_i G_{i,k}$ and $\sum_i \|z_{i,k} - x^*\|^2 \leq \sum_i d_{i,k}^2$, summing over agents and time yields

$$\begin{aligned} \sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, z_{i,k} - x^* \rangle &\leq \frac{1}{2n} \sum_{k=0}^{t-1} \sum_{i=1}^n \sqrt{\hat{G}_k} (d_{i,k}^2 - d_{i,k+1}^2) \\ &\quad + \frac{1}{2n} \sum_{k=0}^{t-1} \sum_{i=1}^n \frac{\bar{r}_k \bar{r}_{i,k}}{\sqrt{G_{i,k}}} \|g_{i,k}\|^2 + \frac{1}{n} \sum_{k=0}^{t-1} \sum_{i=1}^n \mathcal{M}_{i,k}. \end{aligned}$$

The DoG telescoping bound [8, Lemma 1] and the standard adaptive-sum bound give

$$\begin{aligned} \frac{1}{2n} \sum_{k=0}^{t-1} \sum_{i=1}^n \sqrt{\hat{G}_k} (d_{i,k}^2 - d_{i,k+1}^2) &\leq 2D_{\mathcal{X}} \bar{r}_t \sqrt{\hat{G}_{t-1}}, \\ \frac{1}{2n} \sum_{k=0}^{t-1} \sum_{i=1}^n \frac{\bar{r}_k \bar{r}_{i,k}}{\sqrt{G_{i,k}}} \|g_{i,k}\|^2 &\leq D_{\mathcal{X}} \bar{r}_t \sqrt{\hat{G}_{t-1}}. \end{aligned}$$

Lemma 7 controls the remaining mismatch term, and combining these three estimates proves the claim. $\square$

Lemma 7 (Radius mismatch). *For every $t \geq 1$, define*

$$A_{i,k} := \|z_{i,k} - x^*\|^2 - \|x_{i,k+1} - x^*\|^2, \quad \mathcal{M}_{i,k} := \frac{\bar{r}_k - \bar{r}_{i,k}}{2\eta_{i,k}} A_{i,k}.$$

If $\mathbb{E}\|g_{i,k}^2\| \leq cL^2d$ for all i, k , then

$$\frac{1}{n} \sum_{k=0}^{t-1} \sum_{i=1}^n \mathcal{M}_{i,k} \leq \frac{\sqrt{nd} D_{\mathcal{X}} L \bar{r}_t}{1 - \sigma}.$$

Proof. Since W is nonnegative and row stochastic, $z_{i,k} \in \mathcal{X}$. Projection nonexpansiveness gives $\|z_{i,k} - x_{i,k+1}\| \leq \eta_{i,k} \|g_{i,k}\|$. Hence

$$A_{i,k} \leq 2D_{\mathcal{X}} \eta_{i,k} \|g_{i,k}\|, \quad \mathcal{M}_{i,k} \leq D_{\mathcal{X}} L \sqrt{d} |\bar{r}_k - \bar{r}_{i,k}|.$$

Let $\rho_k = (\bar{r}_{1,k}, \dots, \bar{r}_{n,k})^\top$, $\hat{\rho}_k = (\hat{r}_{1,k}, \dots, \hat{r}_{n,k})^\top$, $e_k = (I - J)\rho_k$, and $s_k = \hat{\rho}_k - \rho_{k-1}$, with $\rho_{-1} = \hat{\rho}_{-1} = r_\epsilon \mathbf{1}$. Since $\hat{r}_{i,k} = \max\{\bar{r}_{i,k-1}, \|x_{i,k} - x_{i,0}\|\}$, the coordinates of s_k are nonnegative. Also,

$$e_k = (W - J)e_{k-1} + (W - J)s_k, \quad \|e_k\| \leq \sigma \|e_{k-1}\| + \sigma \|s_k\|.$$

Therefore

$$\sum_{k=0}^{t-1} \|e_k\| \leq \frac{\sigma}{1 - \sigma} \sum_{k=0}^{t-1} \|s_k\|.$$

Since W is doubly stochastic, $\mathbf{1}^\top \rho_k = \mathbf{1}^\top \hat{\rho}_k$. Thus, using $s_k \geq 0$ coordinate-wise,

$$\sum_{k=0}^{t-1} \|s_k\| \leq \sum_{k=0}^{t-1} |s_k| = \mathbf{1}^\top \hat{\rho}_{t-1} - nr_\epsilon = n\bar{r}_{t-1} - nr_\epsilon \leq n\bar{r}_t.$$

Finally,

$$\frac{1}{n} \sum_{k=0}^{t-1} \sum_{i=1}^n \mathcal{M}_{i,k} \leq \frac{D_{\mathcal{X}} L \sqrt{d}}{n} \sum_{k=0}^{t-1} \|e_k\| \leq \frac{\sqrt{nd} D_{\mathcal{X}} L \bar{r}_t}{1 - \sigma}.$$

□

Proof of Lemma 2. Let $\mathcal{F}_{k-1}$ contain all iterates, radii, and oracle randomness available before the random direction and sample at iteration k are drawn. Define $d_{i,k} := \|x_{i,k} - x^*\|$, $\hat{d}_k := \max_{j \leq k} \max_{i \leq n} d_{i,j}$. Apply the POEM martingale argument [7, Lemma 5] to

$$X_k := \frac{1}{\hat{d}_k} \cdot \frac{1}{n} \sum_{i=1}^n \langle \Delta_{i,k}, x_{i,k} - x^* \rangle, \quad \hat{X}_k := \frac{1}{\hat{d}_k} \cdot \frac{1}{n} \sum_{i=1}^n \langle \nabla f_{i,\mu_{i,k}}(x_{i,k}), x_{i,k} - x^* \rangle.$$

Then X_k is a martingale difference, while Lemma 6 implies $|X_k|, |\hat{X}_k| \leq 2Ld$. The variance proxy is controlled by

$$\sum_{k=0}^{t-1} (X_k - \hat{X}_k)^2 \leq \frac{1}{n} \sum_{i=1}^n \sum_{k=0}^{t-1} \|g_{i,k}\|^2 \leq \hat{G}_{t-1}.$$

Using Lemma 5 of [7] with weights $Y_k = \bar{r}_k \hat{d}_k$ gives, on an event of probability at least $1 - \delta$, simultaneously for all $1 \leq t \leq T$,

$$\sum_{k=0}^{t-1} \bar{r}_k \cdot \frac{1}{n} \sum_{i=1}^n \langle \Delta_{i,k}, x_{i,k} - x^* \rangle \leq 8\bar{r}_{t-1} \hat{d}_{t-1} \sqrt{\theta_{t,\delta} \hat{G}_{t-1} + 4L^2 d^2 \theta_{t,\delta}^2}.$$

Using $\hat{d}_{t-1} \leq D_{\mathcal{X}}$ proves the lemma. □

Proof of Lemma 3. Using $\mu_{i,k} = \bar{r}_{i,k} \sqrt{d/(k+1)}$, $\bar{r}_{i,k} \leq D_{\mathcal{X}}$, and monotonicity of $\bar{r}_k$,

$$\sum_{k=0}^{t-1} \frac{2L\bar{r}_k}{n} \sum_{i=1}^n \mu_{i,k} \leq 2L\sqrt{d} \sum_{k=0}^{t-1} \frac{\bar{r}_k^2}{\sqrt{k+1}} \leq 2LD_{\mathcal{X}}\bar{r}_t\sqrt{d} \sum_{k=0}^{t-1} \frac{1}{\sqrt{k+1}} \leq 4D_{\mathcal{X}}L\bar{r}_t\sqrt{dt}.$$

□

A.3. Consensus error.

Proof of Lemma 4. Write the disagreement and forcing terms as

$$e_k := \left(\sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|^2 \right)^{1/2}, \quad u_k := \left(\sum_{i=1}^n \eta_{i,k}^2 \|g_{i,k}\|^2 \right)^{1/2}.$$

The empirical mean minimizes squared distances, projection is nonexpansive, and W is doubly stochastic. Hence $\bar{z}_k = \bar{x}_k$ and

$$e_{k+1} \leq \left(\sum_{i=1}^n \|z_{i,k} - \bar{x}_k\|^2 \right)^{1/2} + u_k \leq \sigma e_k + u_k,$$

where the last inequality is the contraction $\|W - J\|_2 = \sigma$ applied to the zero-mean vectors $\{x_{i,k} - \bar{x}_k\}_{i=1}^n$. Summing the recursion and using common initialization, $e_0 = 0$, gives

$$(1 - \sigma) \sum_{k=0}^{K-1} e_k \leq \sum_{k=0}^{K-1} u_k.$$

The forcing satisfies, by the adaptive stepsize and $\bar{r}_{i,k} \leq D_{\mathcal{X}}$,

$$u_k^2 = \sum_{i=1}^n \bar{r}_{i,k}^2 \frac{\|g_{i,k}\|^2}{G_{i,k}} \leq D_{\mathcal{X}}^2 \sum_{i=1}^n \frac{\|g_{i,k}\|^2}{G_{i,k}}.$$

Using $\sum_{k=0}^{K-1} a_k/A_k \leq \log(A_{K-1}/A_0)$ for $A_k = A_{k-1} + a_k$, followed by Cauchy–Schwarz, yields

$$\sum_{k=0}^{K-1} u_k \leq D_{\mathcal{X}} \sqrt{K} \left(\sum_{i=1}^n \log \left(\frac{G_{i,K-1}}{G_{i,0}} \right) \right)^{1/2}.$$

Substituting this bound above and dividing by $K(1-\sigma)$ proves the claim. $\square$

Proof of Lemma 5. Write

$$e_k := \left(\sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|^2 \right)^{1/2}.$$

By Cauchy–Schwarz, Jensen’s inequality, and the fact that W is doubly stochastic,

$$\sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle \leq 2 \left(\sum_{i=1}^n \|g_{i,k}\|^2 \right)^{1/2} e_k.$$

Indeed, $(\sum_i \|x_{i,k} - z_{i,k}\|^2)^{1/2} \leq 2e_k$, because $x_{i,k} - z_{i,k} = \sum_j w_{ij}(x_{i,k} - x_{j,k})$. Therefore

$$\sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle \leq \frac{2}{n} \sum_{k=0}^{t-1} \bar{r}_k \left(\sum_{i=1}^n \|g_{i,k}\|^2 \right)^{1/2} e_k.$$

The remaining term is immediate from Cauchy–Schwarz:

$$\sum_{k=0}^{t-1} \frac{L\bar{r}_k}{n} \sum_{i=1}^n \|x_{i,k} - \bar{x}_k\| \leq \frac{L}{\sqrt{n}} \sum_{k=0}^{t-1} \bar{r}_k e_k.$$

Adding the two estimates proves the lemma. $\square$

A.4. Proof of the main theorems.

Proof of Theorem 1. Combine the decomposition (2) with Lemmas 1, 2, and 3. On $\Omega_{T,\delta}$, monotonicity of $\bar{r}_t$, $\hat{G}_t$, and $\theta_{t,\delta}$, together with $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, gives

$$\begin{aligned} R_t(f(\tilde{x}_t) - f(x^*)) &\leq 16D_{\mathcal{X}}\theta_{T,\delta}\bar{r}_t(\sqrt{\hat{G}_T} + Ld + L\sqrt{dT}) + \frac{D_{\mathcal{X}}Ld\bar{r}_t}{1-\sigma} \\ &\quad + \sum_{k=0}^{t-1} \frac{\bar{r}_k}{n} \sum_{i=1}^n \langle g_{i,k}, x_{i,k} - z_{i,k} \rangle + \sum_{k=0}^{t-1} \frac{L\bar{r}_k}{n} \sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|, \end{aligned}$$

where the numerical constant absorbs the absolute constants in the preceding lemmas. Dividing by $R_t = \sum_{k=0}^{t-1} \bar{r}_k$ gives (3). $\square$

A.5. Expected-rate proof.

Proposition 1 (Explicit conditional expected rate). *Under the assumptions of Theorem 2,*

$$\mathbb{E}[f(\bar{x}_{\text{out}}) - f(x^*) \mid \Omega_{T,\delta}] \leq \frac{D_{\mathcal{X}}Ld}{T(1-\sigma)}e(1 + \log(D_{\mathcal{X}}/r_\epsilon)) \quad (7)$$

$$\begin{aligned} &+ \frac{16e(\log(\frac{D_{\mathcal{X}}}{r_\epsilon})_+)}{T} D_{\mathcal{X}}\theta_{T,\delta} \left(\frac{\sqrt{r_\epsilon^2 + TcL^2d}}{\sqrt{1-\delta}} + L(d + \sqrt{dT}) \right) \\ &+ \frac{e(\log(\frac{D_{\mathcal{X}}}{r_\epsilon})_+)(2\sqrt{cd} + 1)LD_{\mathcal{X}}}{(1-\delta)(1-\sigma)\sqrt{T}} \sqrt{\log\left(1 + \frac{TcL^2d}{\min_i G_{i,0}}\right)}. \end{aligned} \quad (8)$$

Proof. Let $\Lambda_T := \frac{e(1+\log(D_{\mathcal{X}}/r_\epsilon))}{T}$. The monotone-sequence bound for the adaptive radii gives $\bar{r}_{\tau_T}/R_{\tau_T} \leq \Lambda_T$ and therefore Theorem 1, applied at $t = \tau_T$, gives (3). The non-consensus part in that bound is at most

$$16D_{\mathcal{X}}\theta_{T,\delta}\Lambda_T(\sqrt{\hat{G}_T} + Ld + L\sqrt{dT}).$$

The accumulator term satisfies $\hat{G}_T \leq r_\epsilon^2 + \sum_{k=0}^{T-1} \max_{i \in [n]} \|g_{i,k}\|^2$, so Lemma 6 gives $\mathbb{E}[\hat{G}_T] \leq r_\epsilon^2 + TcL^2d$. By Jensen's inequality and $\mathbf{P}(\Omega_{T,\delta}) \geq 1 - \delta$,

$$\mathbb{E}[\sqrt{\hat{G}_T} \mid \Omega_{T,\delta}] \leq \sqrt{\mathbb{E}[\hat{G}_T \mid \Omega_{T,\delta}]} \leq \frac{\sqrt{r_\epsilon^2 + TcL^2d}}{\sqrt{1-\delta}}.$$

Then

$$\frac{D_{\mathcal{X}}Ld\bar{r}_{\tau_T}}{R_{\tau_T}(1-\sigma)} \leq \frac{D_{\mathcal{X}}Ld}{T(1-\sigma)}e(1 + \log(D_{\mathcal{X}}/r_\epsilon)).$$

For the bracketed consensus term in (3) at $t = \tau_T$, Lemma 5 gives the nonnegative upper bound, with $e_k := (\sum_{i=1}^n \|x_{i,k} - \bar{x}_k\|^2)^{1/2}$,

$$\frac{2}{nR_{\tau_T}} \sum_{k=0}^{\tau_T-1} \bar{r}_k \left(\sum_{i=1}^n \|g_{i,k}\|^2 \right)^{1/2} e_k + \frac{L}{\sqrt{n}R_{\tau_T}} \sum_{k=0}^{\tau_T-1} \bar{r}_k e_k.$$

Since this upper bound is nonnegative, conditioning on $\Omega_{T,\delta}$ costs a factor at most $(1-\delta)^{-1}$. Using $\bar{r}_k/R_{\tau_T} \leq \Lambda_T$ for $k \leq \tau_T - 1$, and $\mathbb{E}[(\sum_{i=1}^n \|g_{i,k}\|^2)^{1/2} | \mathcal{F}_{k-1}] \leq L\sqrt{ncd}$, its conditional expectation is at most

$$\frac{(2\sqrt{cd} + 1)L\Lambda_T}{(1-\delta)\sqrt{n}} \sum_{k=0}^{T-1} \mathbb{E}[e_k].$$

Lemma 4, Jensen's inequality, and $\mathbb{E}[G_{i,T-1}] \leq G_{i,0} + TcL^2d$ yield the cumulative disagreement bound

$$\sum_{k=0}^{T-1} \mathbb{E}[e_k] \leq \frac{D_{\mathcal{X}}\sqrt{nT}}{1-\sigma} \sqrt{\log\left(1 + \frac{TcL^2d}{\min_i G_{i,0}}\right)}.$$

Consequently, the conditional consensus contribution is at most

$$\frac{e(1 + \log(D_{\mathcal{X}}/r_{\epsilon}))(2\sqrt{cd} + 1)LD_{\mathcal{X}}}{(1-\delta)(1-\sigma)\sqrt{T}} \sqrt{\log\left(1 + \frac{TcL^2d}{\min_i G_{i,0}}\right)}.$$

Substituting the accumulator and consensus estimates into the conditional finite-time bound proves (7). $\square$

Proof of Theorem 2. Proposition 1 is an explicit version of the rate. Its first line is $\tilde{\mathcal{O}}(LD_{\mathcal{X}}\sqrt{d}/\sqrt{T})$, while its network line is $\tilde{\mathcal{O}}(LD_{\mathcal{X}}\sqrt{d}/((1-\sigma)\sqrt{T}))$. The remaining logarithmic factors and lower-order T^{-1} terms are hidden by $\tilde{\mathcal{O}}(\cdot)$, giving (4). $\square$

JIAWEI CHEN
MOSCOW INDEPENDENT RESEARCH INSTITUTE OF ARTIFICIAL INTELLIGENCE,
9 INSTITUTSKIY PER.
MOSCOW, 141701, DOLGOPRUDNY RUSSIA
APPLIED AI INSTITUTE
BOLSHOY BOULEVARD, 30, BLD. 1
MOSCOW 121205, RUSSIA
Email address: jiawei.chen@phystech.edu

ALEXANDER VIKTOROVICH ROGOZIN
MOSCOW INDEPENDENT RESEARCH INSTITUTE OF ARTIFICIAL INTELLIGENCE,
9 INSTITUTSKIY PER.
MOSCOW, 141701, DOLGOPRUDNY RUSSIA
Email address: aleksandr.rogozin@phystech.edu