This is the errata for the paper "**Optimal Complexity in Decentralized Training**" at proceedings: `https://proceedings.mlr.press/v139/lu21a.html`. The paper is then corrected at Arxiv (`https://arxiv.org/abs/2006.08085`). We apologize for the confusion that may have caused. Please feel free to contact Yucheng Lu (`yl2967@cornell.edu`) if you have any further questions regarding this file or the paper.

Corollary 1. The description of the original Corollary 1 is:

"For every $\Delta > 0$, $L > 0$, $n \in \{2, 3, 4, \cdots\}$, $\lambda \in [0, \cos(\pi/n)]$, $\sigma > 0$, and $B \in \mathbb{N}^+$, there exists a loss function $f \in \mathcal{F}_{\Delta,L}$, a set of underlying oracles $O \in \mathcal{O}_{\sigma^2}$, a gossip matrix $W \in \mathcal{W}_n$ with second largest eigenvalue being λ, and a graph $G \in \mathcal{G}_{n,D}$, such that no matter what $A \in \mathcal{A}_{B,W}$ is used, $T_\epsilon(A, f, O, G)$ will always be lower bounded by

$$\Omega\left(\frac{\Delta L \sigma^2}{nB\epsilon^4} + \frac{\Delta L}{\epsilon^2\sqrt{1-\lambda}}\right).$$

"

where it should be

"For every $\Delta > 0$, $L > 0$, $n \in \{2, 3, 4, \cdots\}$, $\sigma > 0$, and $B \in \mathbb{N}^+$, there exists a loss function $f \in \mathcal{F}_{\Delta,L}$, a set of underlying oracles $O \in \mathcal{O}_{\sigma^2}$, a gossip matrix $W \in \mathcal{W}_n$ with second largest eigenvalue being $\lambda = \cos(\pi/n)$, and a graph $G \in \mathcal{G}_{n,D}$, such that no matter what $A \in \mathcal{A}_{B,W}$ is used, $T_\epsilon(A, f, O, G)$ will always be lower bounded by

$$\Omega\left(\frac{\Delta L \sigma^2}{nB\epsilon^4} + \frac{\Delta L}{\epsilon^2\sqrt{1-\lambda}}\right).$$

"

A missing K in the proof. In Equation (83), the last term should be $\frac{24\alpha^3\tilde{\sigma}^2L^2K}{n}$, not $\frac{24\alpha^3\tilde{\sigma}^2L^2}{n}$. Note that this missing K propagates for a few steps, but does not change the final bound. The missing Ks are put back in the latest Arxiv version.