Distributed Momentum for Byzantine-resilient Stochastic Gradient Descent

El-Mahdi El-Mhamdi [2]   Rachid Guerraoui [1]   Sébastien Rouault [1]

[1]École Polytechnique Fédérale de Lausanne (EPFL)

[2]École polytechnique

May 6, 2021

A Simple Guessing Game

A Simple Guessing Game

$$X_1 \ \ldots \ X_8 \underset{\textit{i.i.d.}}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$$

A Simple Guessing Game

$$X_1 \;\ldots\; X_8 \underset{\textit{i.i.d.}}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$$

🎯 Tell whether μ positive

A Simple Guessing Game — 1st round

$$(x_1 \ldots x_8) = (8, 7, 3, 5, 9, 8, 7, 5)$$

A Simple Guessing Game — 1st round

$$(x_1 \ldots x_8) = (8, 7, 3, 5, 9, 8, 7, 5)$$

“μ very probably positive”

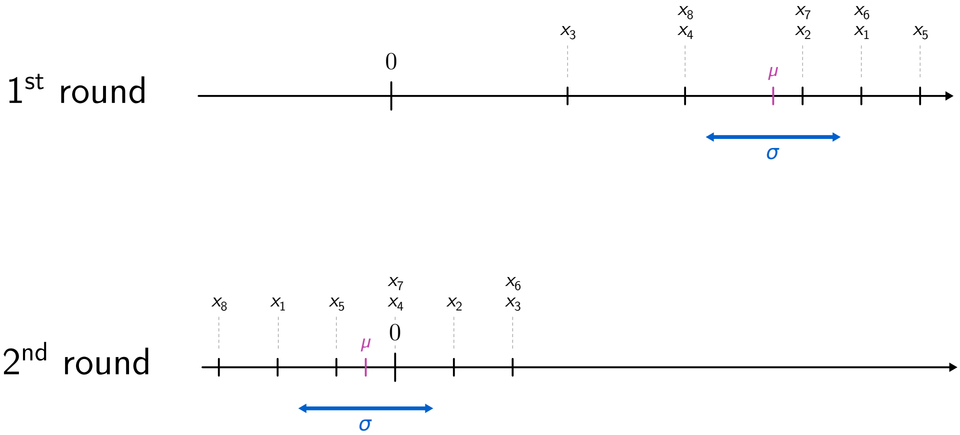
A Simple Guessing Game — 2nd round

$$(x_1 \ldots x_8) = (-2, 1, 2, 0, -1, 2, 0, -3)$$

A Simple Guessing Game — 2nd round

$$(x_1 \dots x_8) = (-2, 1, 2, 0, -1, 2, 0, -3)$$

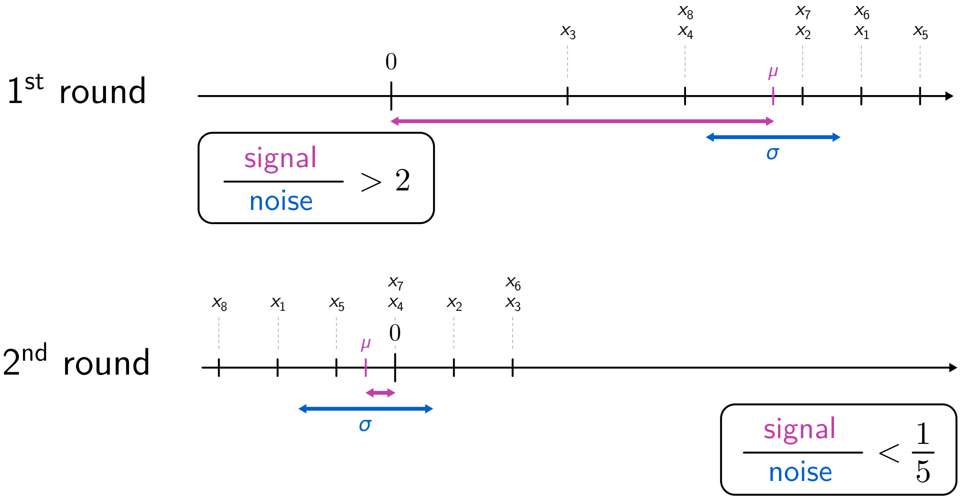
Use Signal-to-Noise Ratio



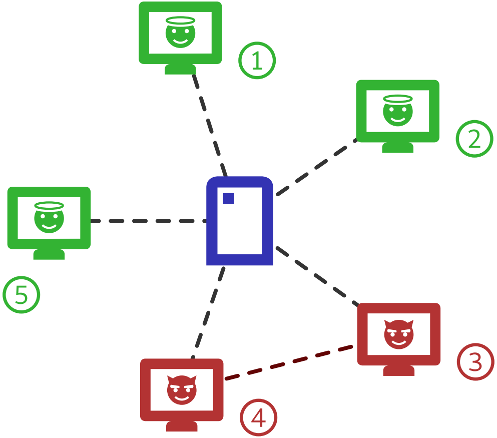
1st round

2nd round

Use Signal-to-Noise Ratio



The Problem of Byzantine-resilience

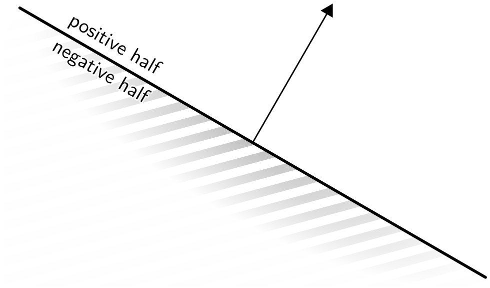


The Problem of Byzantine-resilience

Defenses based on **statistical robustness**

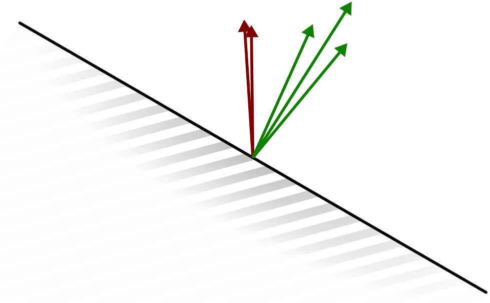
The Problem of Byzantine-resilience

Defenses based on **statistical robustness**



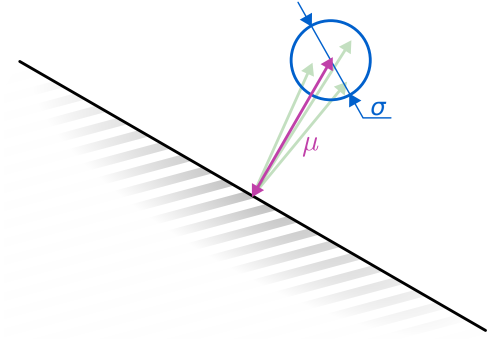
The Problem of Byzantine-resilience

Defenses based on **statistical robustness**



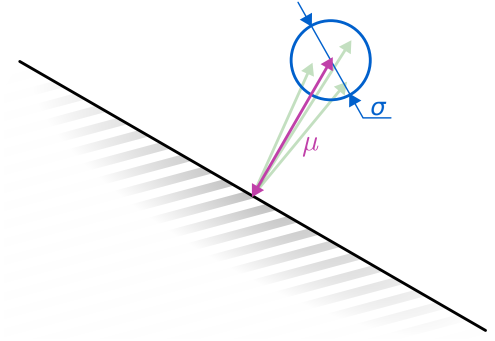
The Problem with Byzantine-resilience

$$\frac{\text{signal}}{\text{noise}} > \textit{some constant}$$



The Problem with Byzantine-resilience

$$\frac{\text{signal}}{\text{noise}} \overset{⚠}{<} \textit{some constant}$$



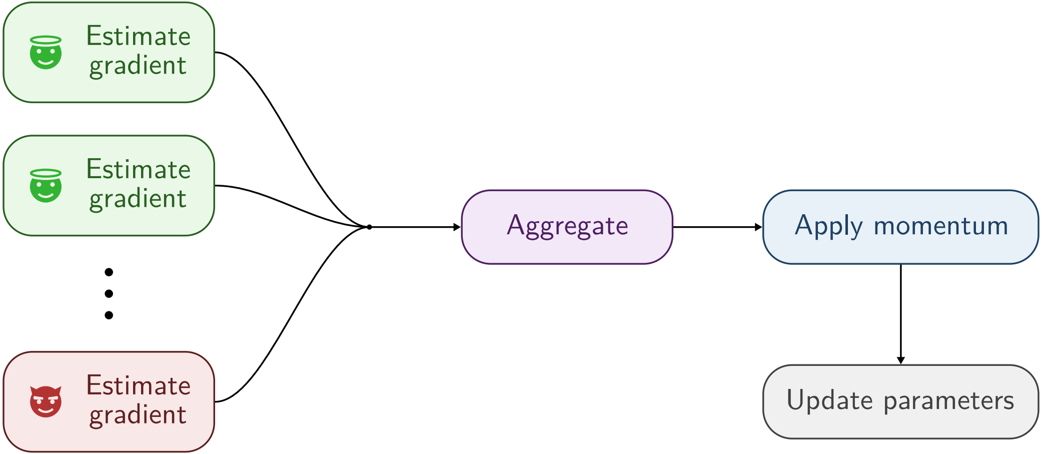
Momentum at the Workers

Expected gradient changes **slowly**

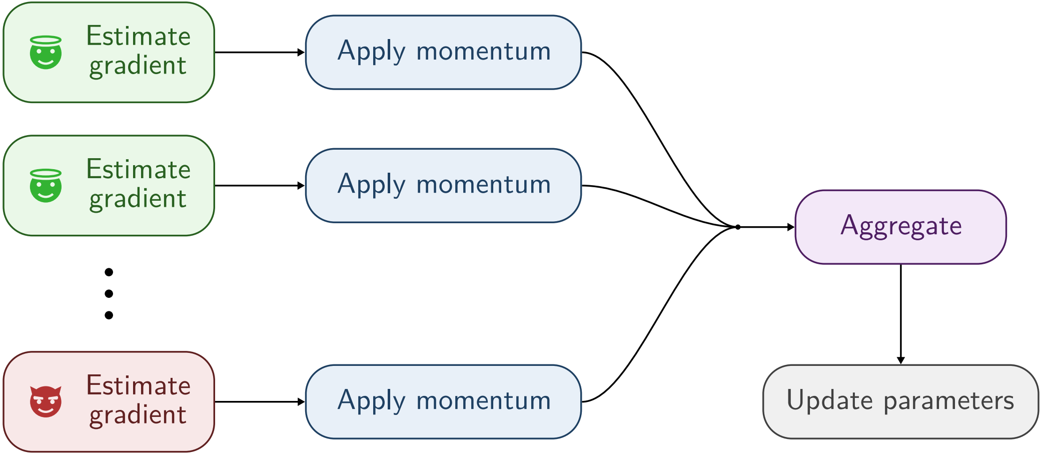
$\Downarrow$

$\sum$ of stochastic gradients
could **cancel noise**

Momentum at the Workers

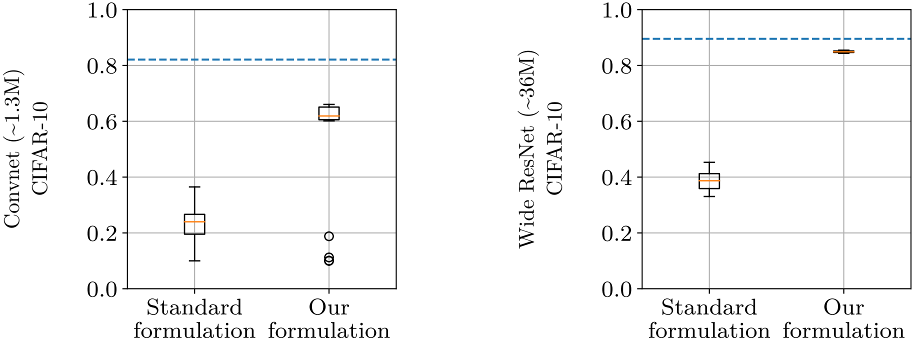


Momentum at the Workers



Small change, Large effects

Maximum top-1 cross-accuracy observed
over the duration of the training (under attack)



More in the paper...

Many more results

(758 different setups tested)

aggregated in the paper

`https://openreview.net/forum?id=H8UHdhWG6A3`

Fully reproducible

(experiments + graphs)

in one command

`https://github.com/LPD-EPFL/ByzantineMomentum`

Credits

- Chess logo by Jakub Steiner (`https://jimmac.eu/`) licensed under CC (BY-SA).
- Bullseye logo by Dragonartz (`https://dragonartz.net/`) licensed under CC (BY-NC-SA).
  From the original, I have modified the shadow of the arrow.
- Two-buttons meme by Jake Clark (`https://jake-clark.tumblr.com`).
  Of course, this drawing being a meme, the text was not in the original.