



Workshop EPFL-Inria
November 23 and 24, 2023, Lyon

Geovani Rizk, EPFL

Title: « Robust Distributed Learning: Tight Error Bounds and Breakdown Point Under Data Heterogeneity »

Abstract:

The theory underlying robust distributed learning algorithms, designed to resist adversarial machines, matches empirical observations when data is homogeneous. Under data heterogeneity however, which is the norm in practical scenarios, established lower bounds on the learning error are essentially vacuous and greatly mismatch empirical observations. This is because the heterogeneity model considered is too restrictive and does not cover basic learning tasks such as least-squares regression. We consider in this paper a more realistic heterogeneity model, namely (G, B) -gradient dissimilarity, and show that it covers a larger class of learning problems than existing theory. Notably, we show that the breakdown point under heterogeneity is lower than the classical fraction $1/2$. We also prove a new lower bound on the learning error of any distributed learning algorithm. We derive a matching upper bound for a robust variant of distributed gradient descent, and empirically show that our analysis reduces the gap between theory and practice.