

Training dynamics of ReLU Networks: a Path-lifting Perspective

Rémi Gribonval

Can we hope to decipher the role of the well-known rescaling symmetries of ReLU networks parameterizations in their training dynamics? The talk will explore recent advances in this direction that exploit the path-lifting, a rescaling-invariant polynomial representation of the parameters of general ReLU networks. Despite its combinatorial dimension, the path-lifting turns out to be not only a convenient mathematical analysis tool: it also gives rise to a computational toolbox to reveal useful properties of the function corresponding to a ReLU network, from Lipschitz regularity to convexity. As we will see, the path-lifting viewpoint also leads to simple modifications of gradient descent that accelerate network training.