Emergence of giant component in percolated artificial neural networks

Despite the fact that major recent artificial intelligence breakthroughs rely on the introduction of new architectures [1], from a theoretical perspective, the structure-function relationship of such networks remains largely unexplored. In this work, we explore how sparsity affects functional robustness, particularly in the context of pruning techniques designed to reduce parameter count. Recent studies have demonstrated that training deep learning models can induce heavy-tailed degree distributions [2, 3], which are known to significantly influence network dynamics. These distributions are revealed by diverse pruning strategies.

We explore the implications of such an observation through the framework of percolation theory applied to the MNIST dataset (**Fig. 1a**). Our results show that the formation of a giant connected component systematically precedes the recovery of network performance. Specifically, percolating edges in descending order of weight accelerates the emergence of a giant component compared to random Bernoulli percolation. This phenomenon correlates with an early recovery of network performance (**Fig. 1b,c**).

Notably, by adapting the theoretical percolation threshold for random networks, we derived an accurate analytical prediction of the percolation threshold in layered networks. We show that in networks composed of successive same-sized layers, the threshold stays independent of the number of layers in the limit of large width compared to the number of layers. Thus, the threshold only depends on the width of the network and is given by $\rho = 1/2n$ where n is the number of nodes of each layer (**Fig. 1c**, top panel). Our findings suggest that learning inherently promotes sparse and robust structures, offering insights into the implicit architectural biases of trained networks.

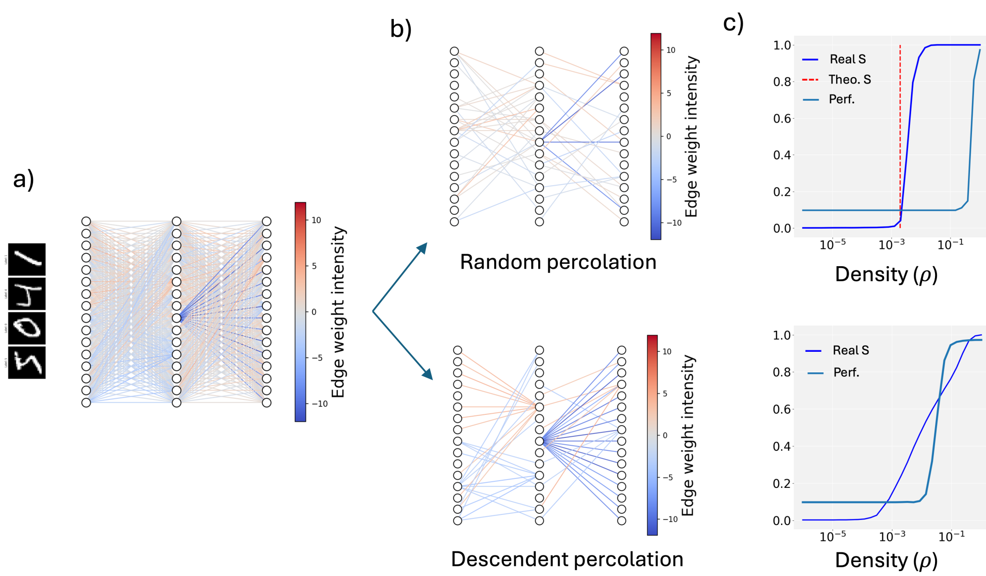

Figure 1: a) Multi-layer perceptrons trained to solve MNIST datasets. b) We propose two percolation frameworks. Random percolation in which we do not account for edge weights magnitude, and Descendent percolation which consists in adding progressively edges in descending order of their weights magnitude. c) Giant component size S and performance against connection density.

References

[1] Ashish Vaswani et al., 2023.

[2] Mocanu et al. *Nat. Commun.*, 9(1):2383, 2018.

[3] Chris Kang et al. *Neural Networks*, 187:107308, 2025.