

## Enric Boix: The staircase property and the leap complexity

Which functions  $f : \{+1, -1\}^d \rightarrow \mathbb{R}$  can neural networks learn when trained with SGD? In this talk, we will consider functions that depend only on a low-dimensional projection of the input. We will study the dynamics of two-layer neural networks trained by SGD with  $O(d)$  samples and will show that a hierarchical property, the “merged-staircase property”, is both necessary and nearly sufficient for learning in this setting. For the more general setting of  $O(d^c)$  samples, we will define the “leap complexity” of the function, and argue that it controls whether the neural network will learn. Based on joint work with Emmanuel Abbe and Theodor Misiakiewicz.