Inferring concepts from noisy examples in Hopfield-like neural networks

Enzo Marinari[1]

[1]Sapienza Università Di Roma, Rome, Italy

We study a variant of the pseudo-inverse learning rule for Hopfield like neural networks, which allows the network to infer archetypal concepts on the basis of a limited number of examples. The mean-field replica theory for this model reveals how this generalization ability is mediated by a multitude of states, with diverse thermodynamic properties, coexisting with the standard Hopfield ones. They appear and vanish through smooth transitions or discontinuous jumps and, interestingly, show much stronger Replica Symmetry Breaking effects than the standard Hopfield model, as captured by our 1RSB analysis. Our results, in excellent agreement with numerical simulations, provide deeper insight into the interplay between memory storage and generalization in attractor neural networks.