bstract template for Youth in High-Dimensions

. Giorlandino , S. Goldt , and . Maillard[2]

International School of dvanced Studies (SISS), Trieste, Italy
[2] *INRI Paris & DI ENS, PSL University, Paris, France*

Large language models demonstrate remarkable ability in factual recall, but the fundamental limits of memorizing and retrieving many input–output associations remain unclear. We study these limits in a minimal setting: a linear associative memory that must map p input embeddings in ^{d} to their corresponding -dimensional targets via a single linear map, while keeping each mapped input well separated from all other targets. Unlike in supervised classification, this strict separation condition induces p constraints per association and produces correlations between constraints through the shared outputs. ere, we characterise the storage *capacity* p) of a linear associative memory, i.e. the maximum number of input-output patterns it can store reliably, in the following ways. We introduce a simpler "decoupled" capacity problem in which, for each input, the full set of competing output patterns is independently re-sampled. We find numerically that the original and the decoupled problem exhibit a striking similarity. We then characterise the capacity p) of this decoupled model using tools from the statistical physics of disordered systems. Our results clarify the fundamental scaling law governing linear associative memories, and provide a starting point for the analysis of more complex models.