When Continual Learning Diffuses: Phase Structures in Binary Neural Networks

A. Bedhotiya et al.[1],[2]

[1] PGI-15, Forschungszentrum Jülich [2] RWTH Aachen

We extend the diffusion framework of continual learning to incorporate generalized update rules. Drawing an analogy to statistical mechanics, we show that this framework exhibits emergent behavior, most notably the appearance of distinct learning and forgetting phases. These phases are characterized by different late time regimes of consolidation and plasticity.

We demonstrate that based on their hyperparameters, abroad class of update rules can be organized into such phase structures. Our approach builds on the diffusion formulation of continual learning, which relates learning and forgetting trajectories to synapse-level update dynamics via the probability distribution function (PDF) of latent weights. This PDF is obtained by promoting discrete task updates to continuous time, yielding a drift–diffusion partial differential equation.

The emergence of the phase structure follows from the temporal evolution of this PDF. While closed-form solutions are generally intractable for general update rules, analysis in the frequency domain provides insight into late time behavior. In particular, we find that the non-analytic structure of the PDF in frequency space uniquely determines long-term network dynamics, thereby defining distinct phases.

We illustrate these results for a two hyperparameter family of update rules, deriving a phase diagram Fig.1 that separates regimes of long-term consolidation and sustained plasticity. We further map commonly used update rules onto this diagram, including the straight-through estimator (STE), which appears as a critical point exhibiting scale invariant dynamics.

Our results reveal previously unexplored phase structures in continual learning, with practical implications for hyperparameter optimization, metaplasticity, and hardware aware learning rule design. We conclude by demonstrating one such hardware application [1].

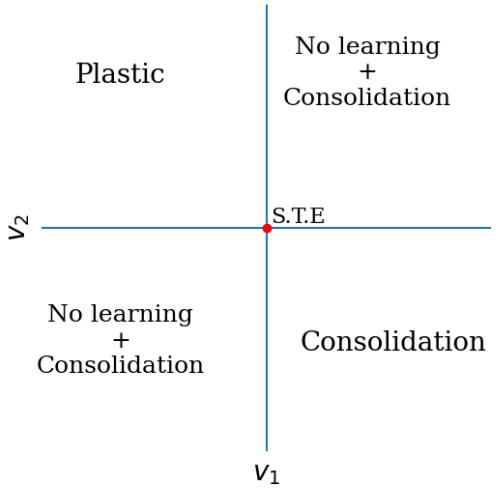

Figure 1: A phase diagram for two hyperparameter (v_1,v_2) family of update rules demonstrating distinct network dynamics in each quadrant. Approaching the origin brings the network closer to criticality and the axes represent phase transitions resulting in drastically different late time behavior.

[1] Bernard G. et al., Nano Lett. 2025, 25, 4, 1443–1450