

# More neurons or more complex neurons?

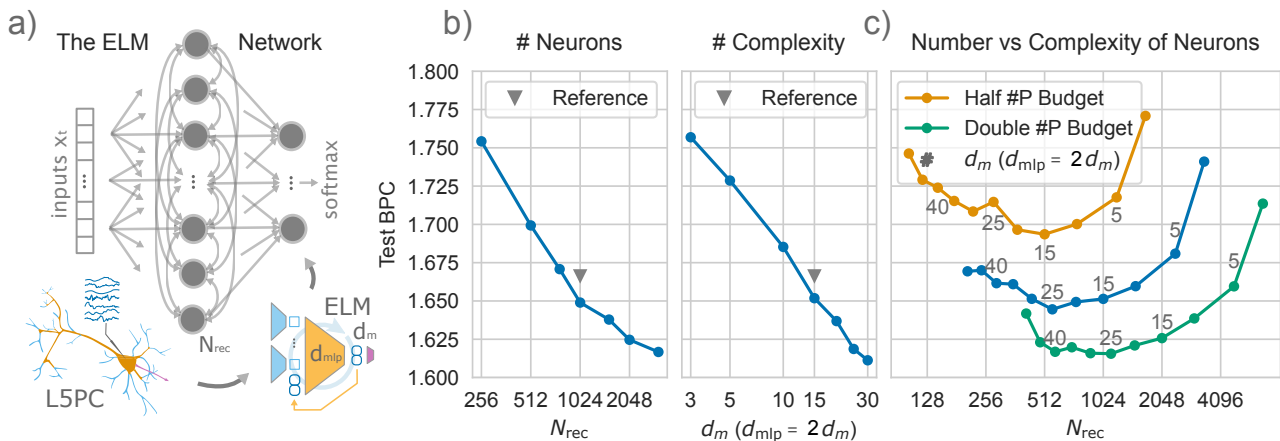
Aaron Spieler<sup>(1,2)</sup>, Georg Martius<sup>(1,3)</sup> and Anna Levina<sup>(1,2)</sup>

<sup>(1)</sup> University of Tübingen, Germany, <sup>(2)</sup> Max Planck Institute for Biological Cybernetics, Tübingen, Germany,

<sup>(3)</sup> Max Planck Institute for Intelligent Systems, Tübingen, Germany

Biological cortical neurons are sophisticated computational units, integrating thousands of synaptic inputs through intricate dendritic trees across multiple timescales. Yet most artificial neural architectures employ simple, memory-less neurons, raising a fundamental question: given a fixed resource budget, should one build networks from many simple neurons or fewer, more sophisticated ones?

To investigate this trade-off, we construct recurrent networks from Expressive Leaky Memory (ELM) neurons [1]—biologically inspired recurrent cells whose complexity can be independently controlled along several dimensions: the number of leaky memory units  $d_m$ , the sophistication of nonlinear input integration ( $l_{mlp}$ ,  $d_{mlp}$ ), and the number of connections per neuron  $d_s$ . In the ELM Network,  $N_{rec}$  such independently parametrized neurons are recurrently wired as a hidden layer, with a subsequent smaller feedforward readout layer. In the following, we investigate how these components individually, or jointly under resource constraint, shape performance on a language modeling benchmark (enwik8).



**Figure 1: Performance scales with both neuron count and per-neuron complexity, but a fixed parameter budget reveals an optimal intermediate complexity.** **a)** The ELM Neuron and Network architecture. **b)** Scaling neuron count (left) or per-neuron complexity (right) independently improves performance (lower BPC is better). **c)** At fixed parameter budget the optimal per-neuron resource allocation is non-trivial and shifts toward *more and more complex* neurons as budget increases.

Our scaling experiments demonstrate that while scaling individual network components might give monotonic performance improvements, under a fixed total parameter budget they start competing for optimal allocation. For neuron count vs per-neuron expressivity this results in a U-shaped optimum that shifts toward more neurons of higher complexity as budget grows; additional resources are best invested into both dimensions simultaneously, never saturating just one. Interestingly, an information-theoretic modeling approach of the layer’s channel capacity can qualitatively reproduce this findings. These results establish that neuron-level complexity is a computational resource worth investing in, not merely a biological curiosity. The existence of a budget-dependent optimal complexity echoes findings on compute-optimal scaling in large language models [2], and suggests that architectures with expressive units, whether biological or neuromorphic, may achieve superior resource efficiency.

[1] A. Spieler et al., Proc. ICLR, 2024.

[2] J. Kaplan et al., arXiv:2001.08361, 2020.