Learning nonlinearities for identifying regular structure in the brain's inference algorithm

KiJung Yoon and Xaq Pitkow

Neurons in the brain are not simply linear filters followed by a half-wave rectification, and exhibit properties like divisive normalization, coincidence detection, and history dependence. Instead of fixed canonical nonlinear activation functions such as sigmoid, tanh, and relu, other nonlinearities may be both more realistic and more useful. We are particularly interested in multivariate (e.g. two- or three-argument) nonlinearities like $f(\mathbf{w}_1 \cdot \mathbf{x}, \mathbf{w}_2 \cdot \mathbf{x}, ...)$ which could allow inputs that arise from multiple distinct pathways such as feedforward, lateral, or feedback connections, or different dendritic compartments. Such multi-argument nonlinearities could allow one feature to modulate the processing of others. Many single-argument nonlinearities permit universal computation, but the *right* nonlinearity could allow faster generalization during inference, both for the brain and for artificial networks.

To address this, we parameterize the nonlinear input-output transformation flexibly by an "inner" neural network, which becomes a 'subroutine' called from the conventional "outer" network. These parameters are shared across all layers and all nodes of a given cell type. We evaluate fully-connected feedforward networks on image classification tasks given a diverse set of random initial conditions. We focus especially on the two-argument nonlinearities learned from MNIST and CIFAR-10 datasets.

We demonstrate that learned two-argument nonlinearities are reliably shaped roughly like quadratic functions, possibly with a linear transformation on the inputs such as a shift and/or rotation. We therefore separate the training and testing phases by a phase in which we fit an algebraic functional form to the learned inner-network nonlinearities. The algebraic nonlinearity does indeed perform as well as the more richly parameterized nonlinearity in the tasks.

In general, these nonlinearities are particularly well-suited for contextual gating of information, and an integral part of the message-passing inference in the brain, because they allow us flexible methods to learn canonical messages as they transform parameters of a population code.

Additional Detail

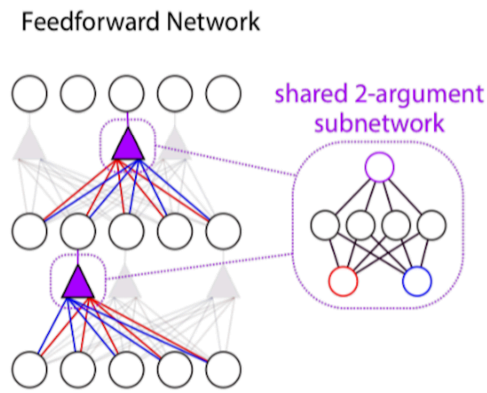

Figure 1. Architecture of feedforward network using a 2-argument nonlinearity described by a subnetwork.

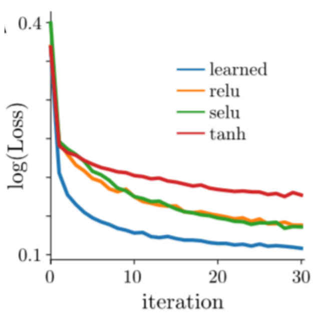

Figure 2. We also use the best dropout rate for the model and the controls, which turns out to be different. The preliminary result shows that the learning curve preserves the faster learning of our network than the point-wise nonlinearities, which suggests a possible advantage of multivariate nonlinearities.

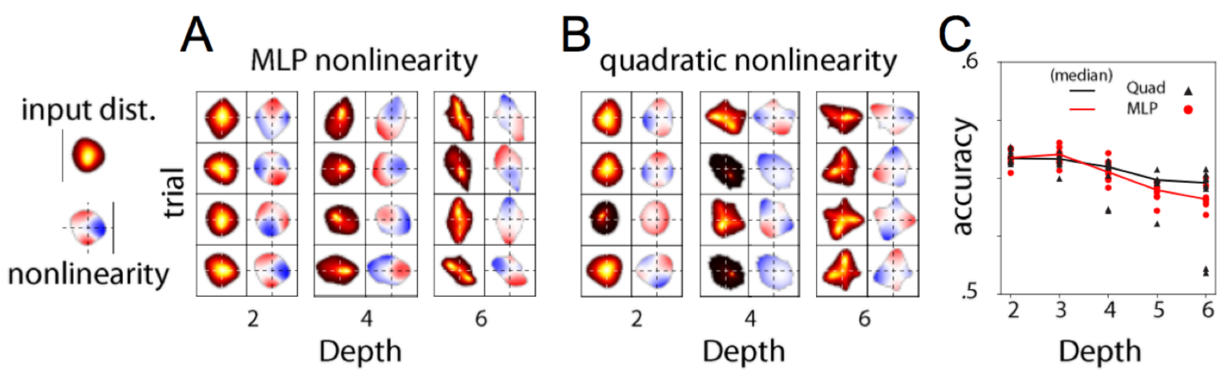

Figure 3. **A**: Nonlinearities (output in red/white/blue) and the resultant distribution of inputs to those nonlinearities (density in black/red/yellow), each learned by a multilayer perceptron (MLP) on CIFAR-10. Functions are masked to keep the most probable 90% of inputs. Functions here were fit over four independent trials, varying the depth from 2 to 6 layers. **B**: 2D quadratic fit to these MLP nonlinearities. **C**: Similar model performance for MLP (red) and the quadratic fits to inner network nonlinearities (black).