δ Multiplexed Gradient Descent: Perturbative Learning with Astrocytes

Nick Skuda[1], Ryan O'Loughlin[12], Bakhrom Oripov[1], Noah Chongsiriwatan[3], Ian Whitehouse[3], Wolfgang Losert[3], Bradley Hayes[2], Adam McCaughan[1], Sonia Buckley[1]

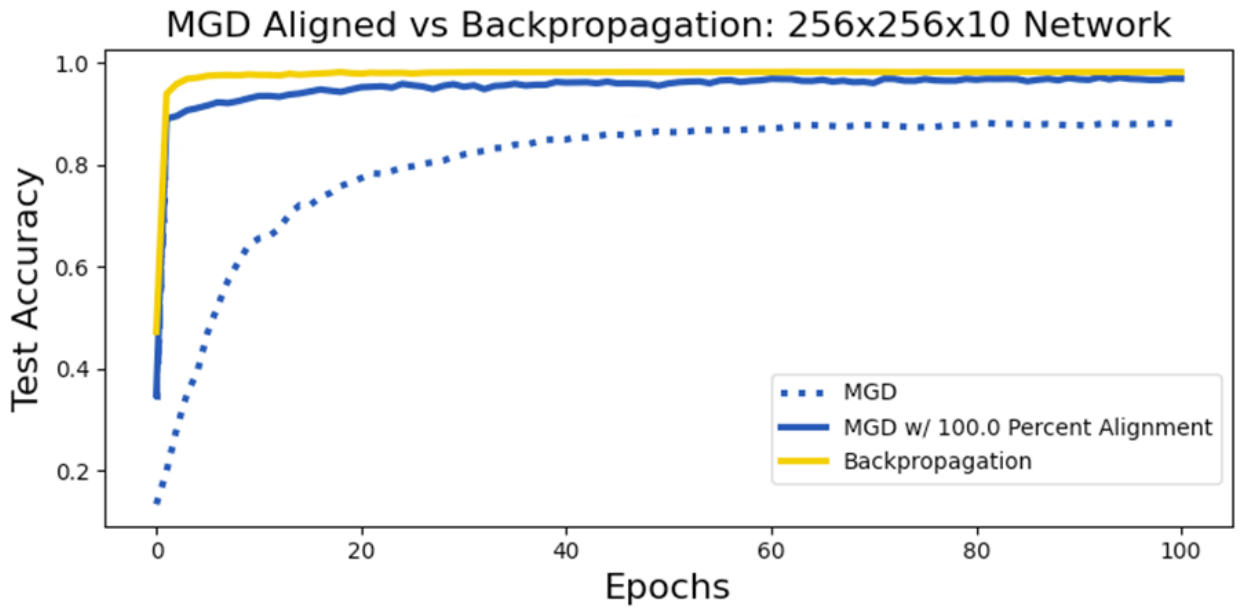
1 National Institute of Standards and Technology

2 University of Colorado Boulder

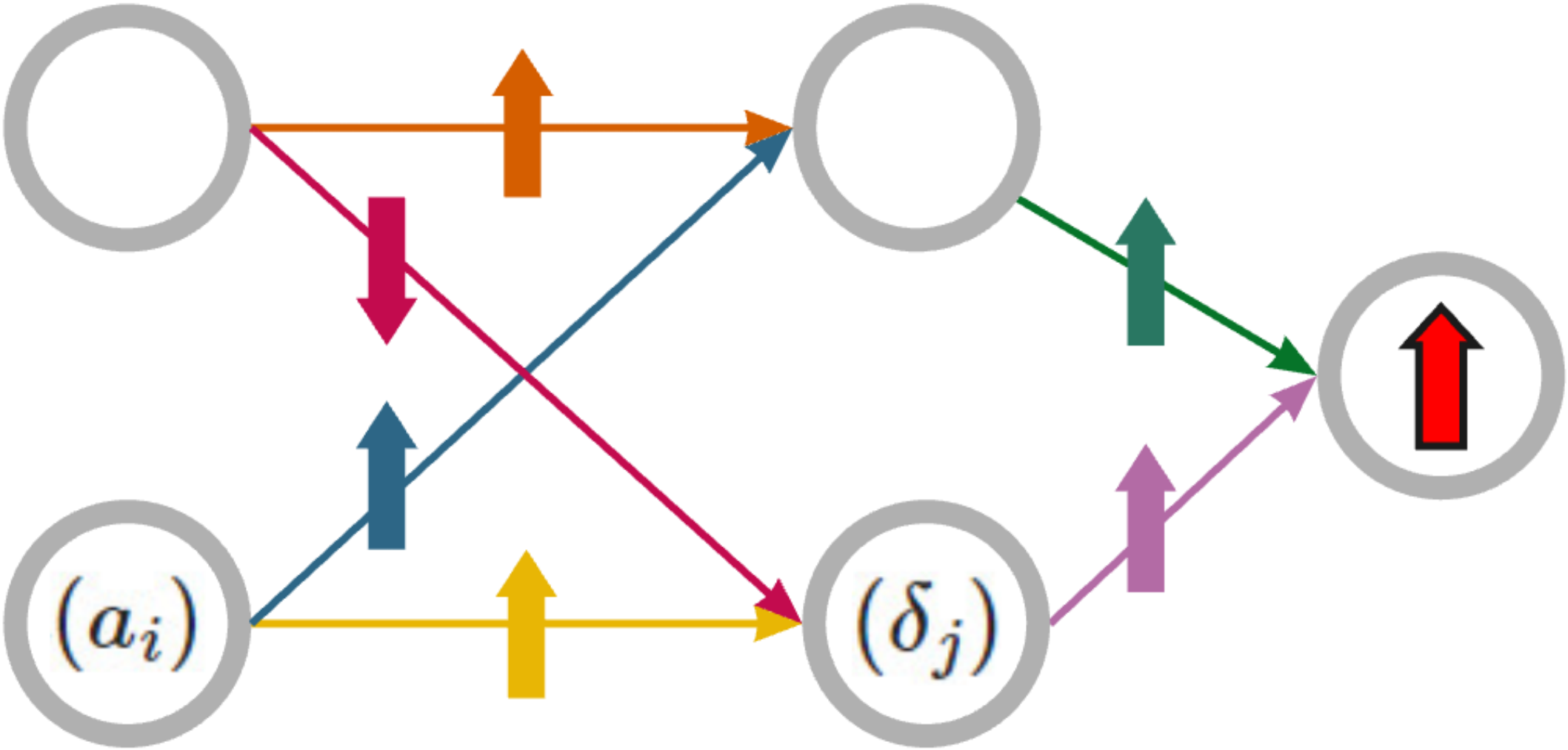
3 University of Maryland

Improving Hardware Friendly Training

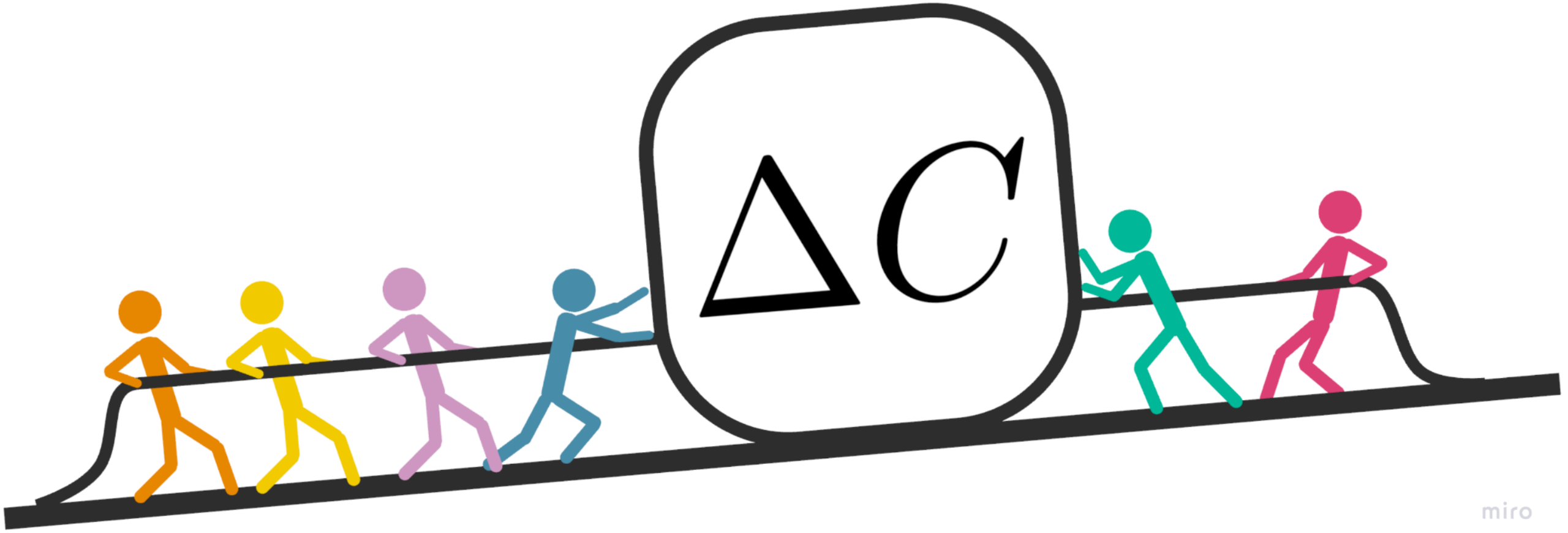
- Neuromorphic hardware is important for better efficiency
- Hardware based training is challenging
- Perturbative algorithms like SPSA or MGD are hardware friendly
  - Itd be good if faster (iterations)
- 60x Speedup in iterations



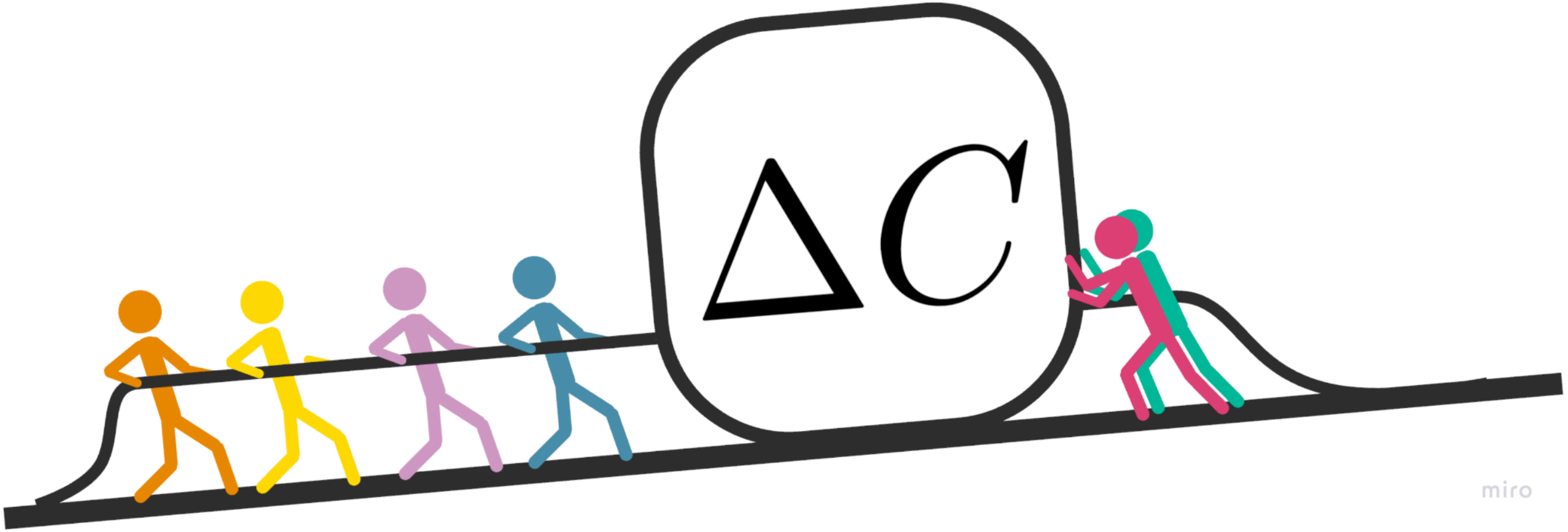
Multiplexed Gradient Descent



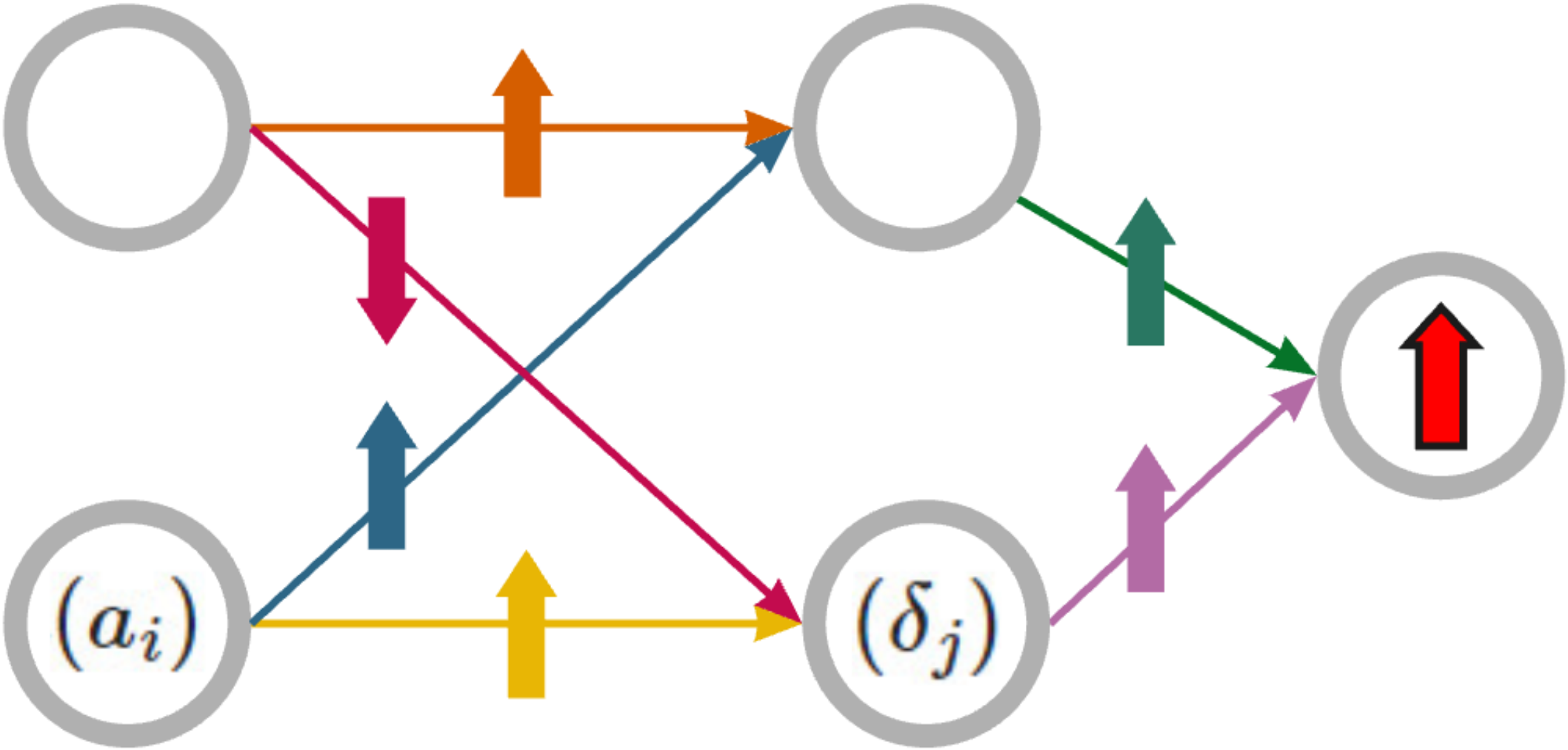
MGD w/ Random Perturbations



MGD w/ Aligned Direction Perturbations

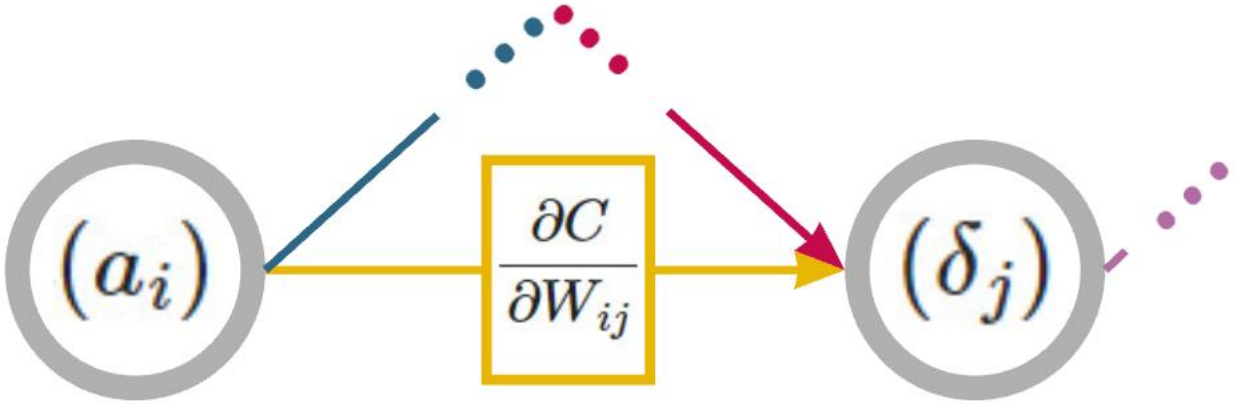


How to get correct signs?

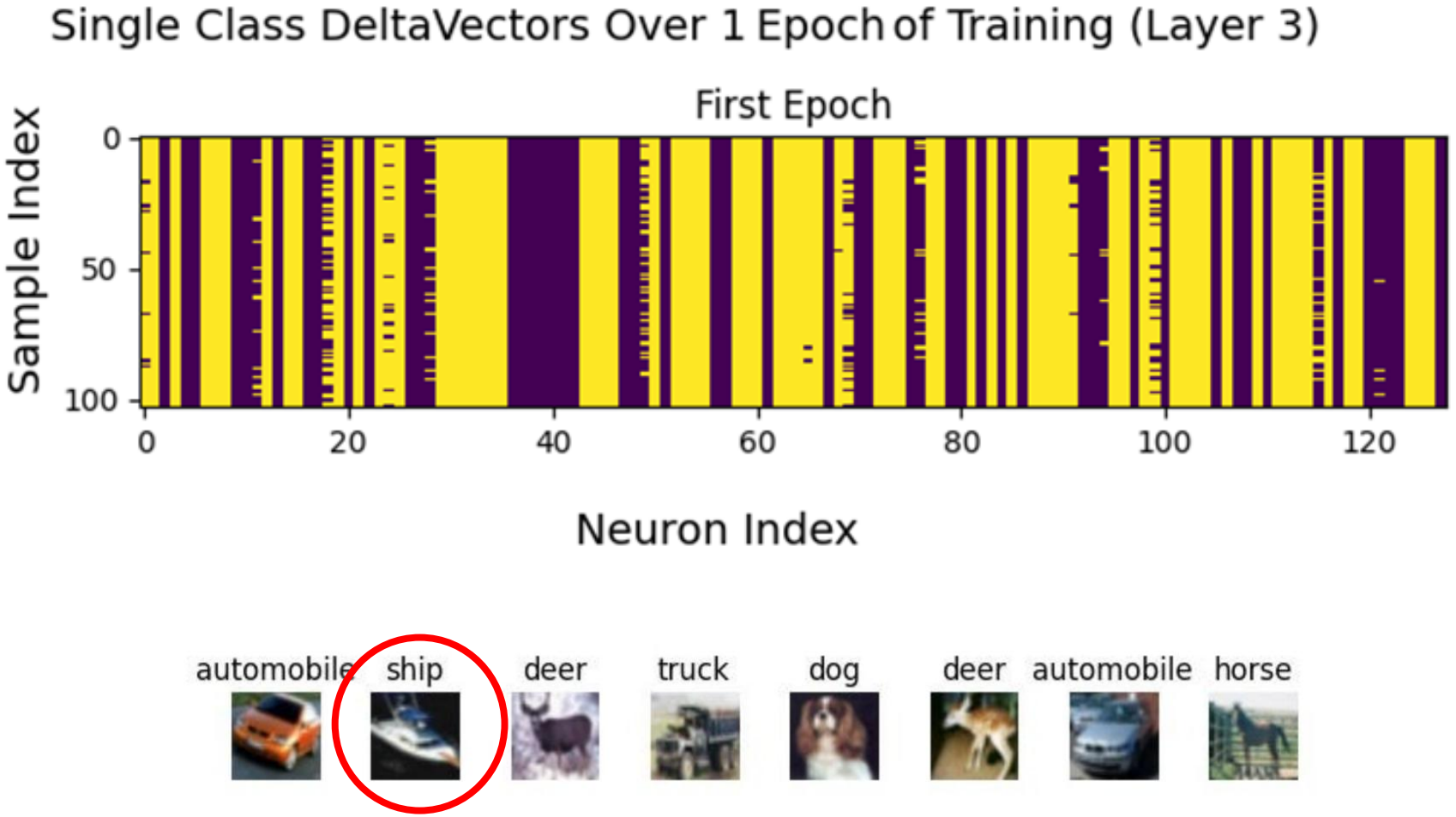


Neuron Bias Gradient Estimation

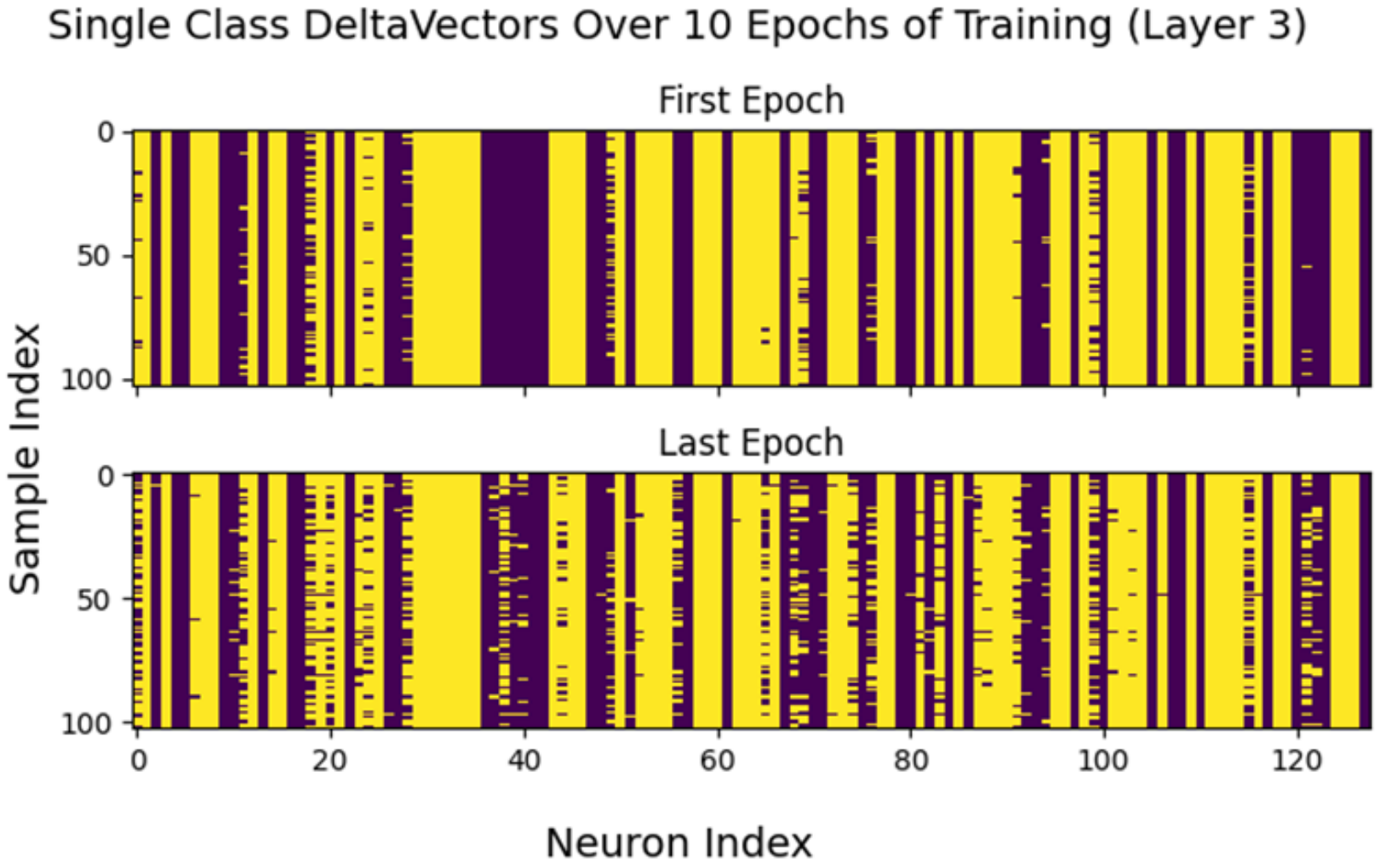
- Would like correct signs
  - If know activations from one layer
  - and bias gradient from next layer
  - can get a good guess of direction
- Know activations locally for free
- Without backprop how to get those bias gradients?
  - And ideally not estimate constantly



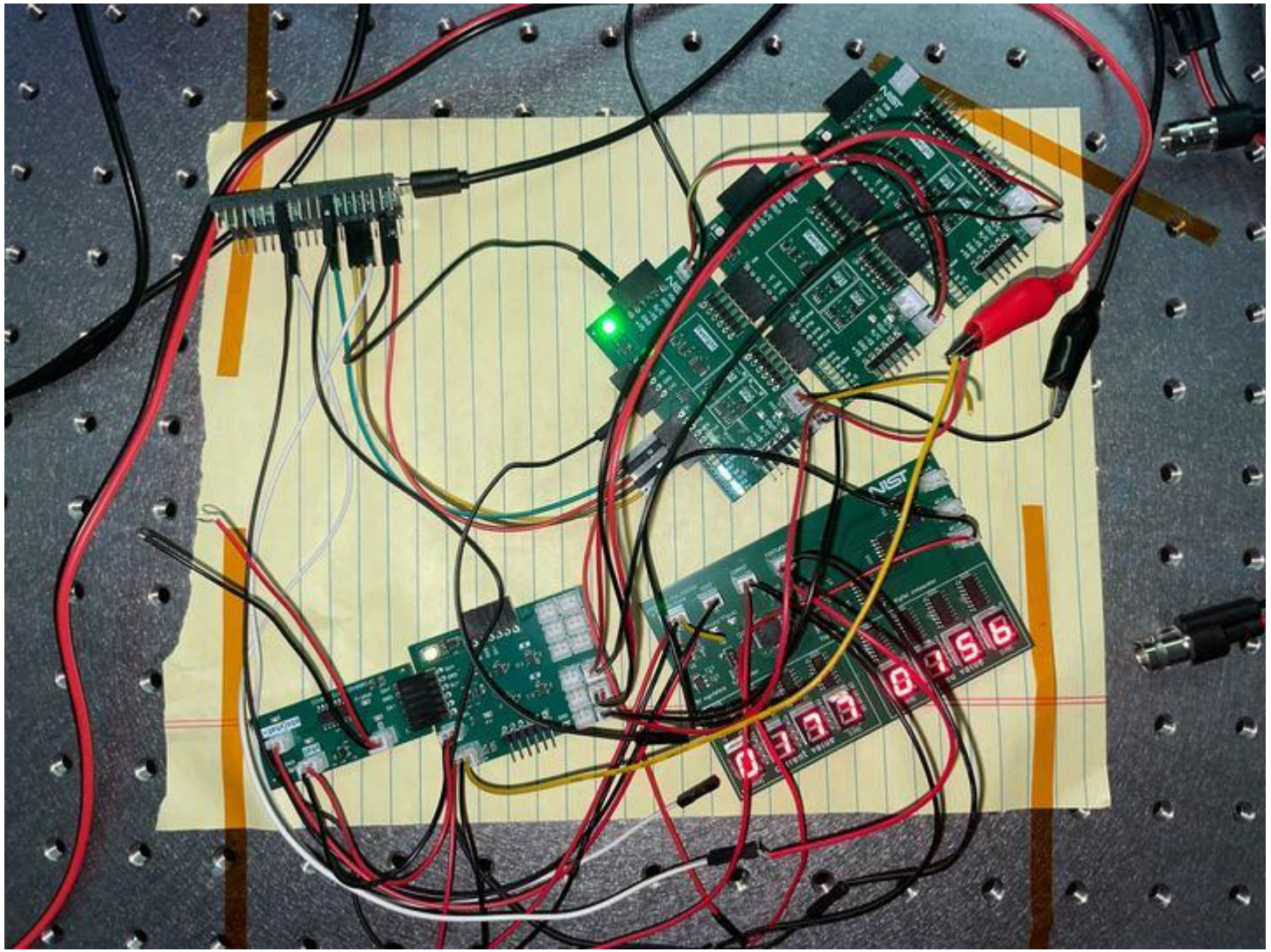
How to get correct signs? Estimate from class



How to keep correct signs? Rare updates

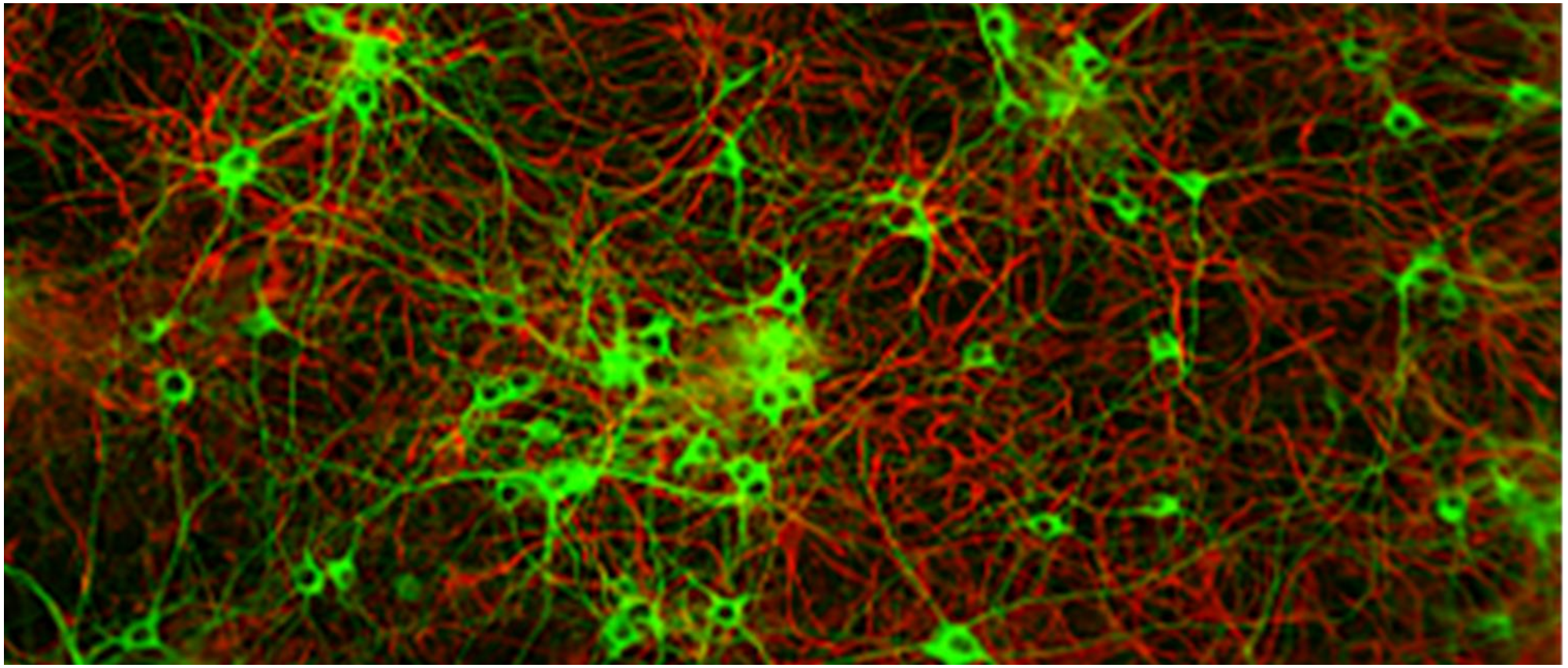


How can we do this in hardware?



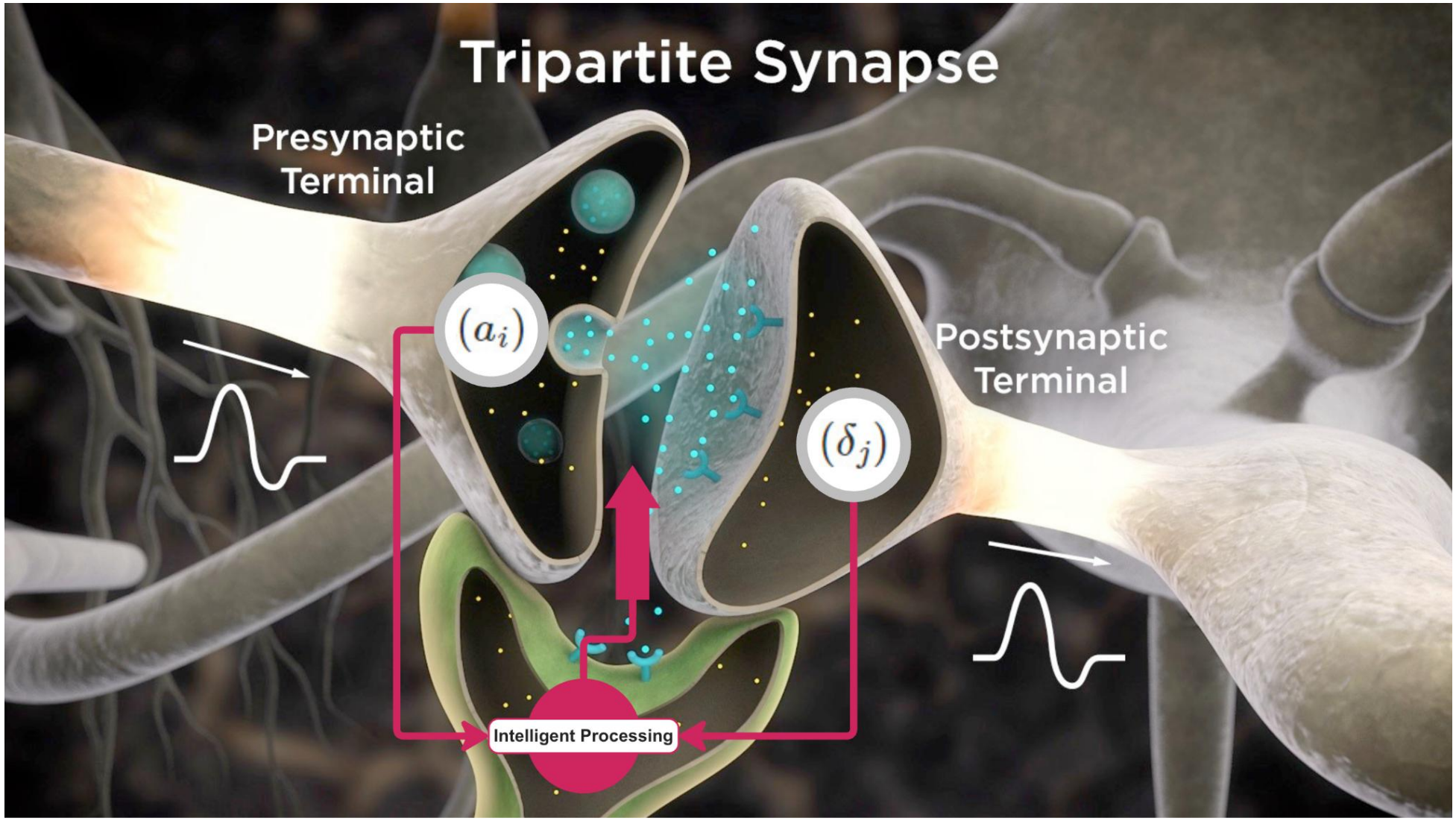
How can we do this in hardware?

Brain Inspiration!

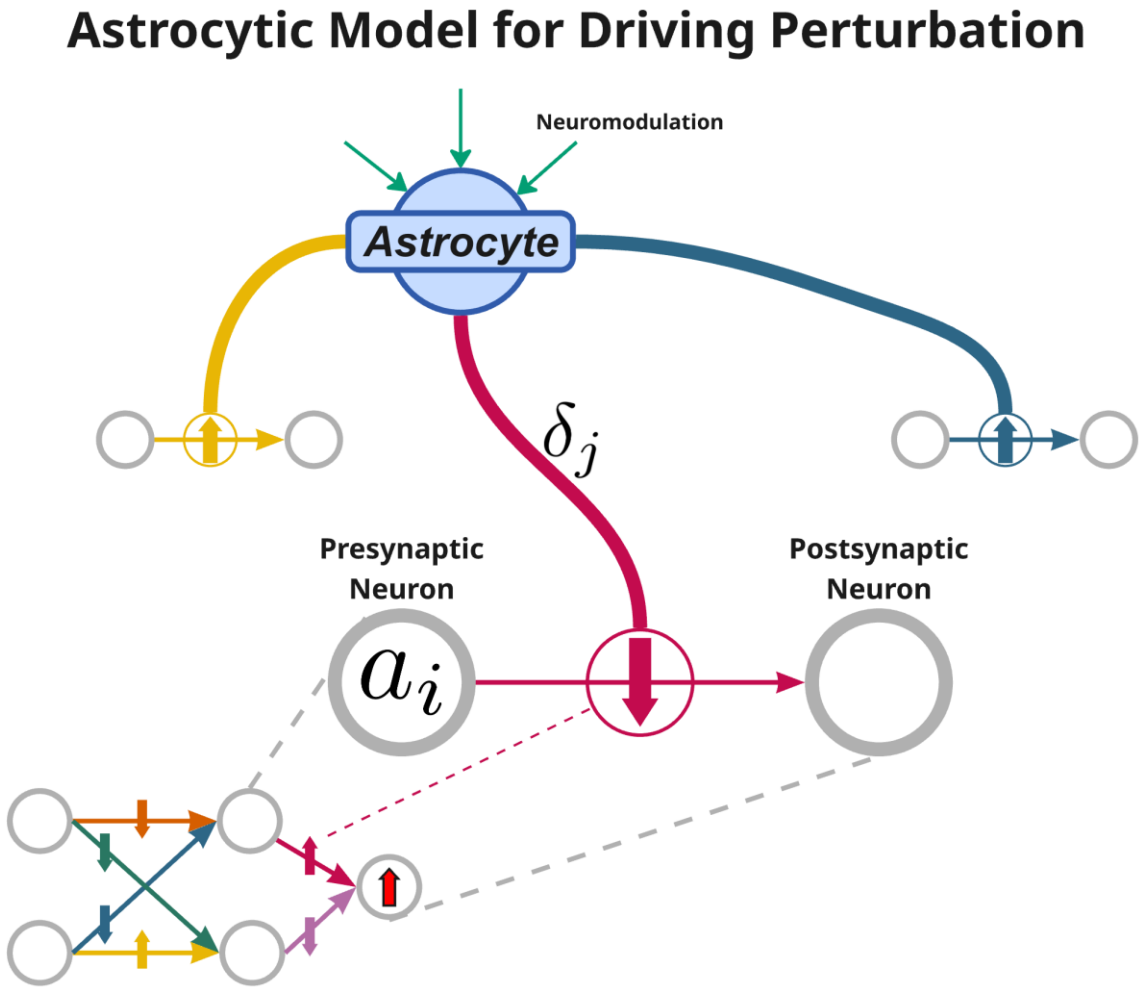


Red astrocytes, green neurons. Credit: Losert Lab Biocomputing Team

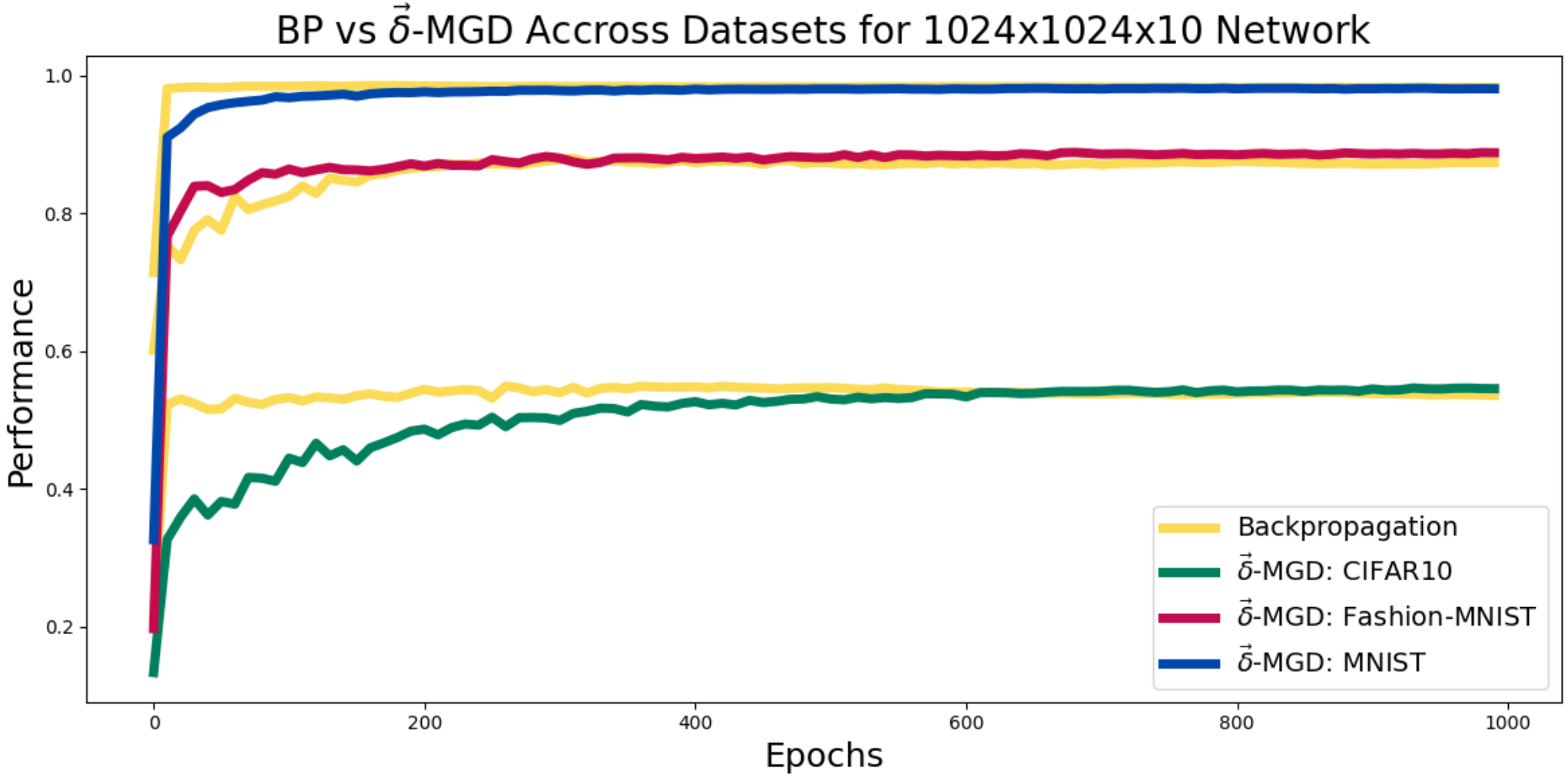
Astrocytes


Tripartite Synapse
Presynaptic Terminal
(a_i)
(δ_j)
Postsynaptic Terminal
Intelligent Processing

Astro-inspired δ Subnetworks



DNN Results


BP vs $\vec{\delta}$-MGD Accross Datasets for 1024x1024x10 Network
Performance
1.0
0.8
0.6
0.4
0.2
0
200
400
600
800
1000
Epochs
Backpropagation
$\vec{\delta}$-MGD: CIFAR10
$\vec{\delta}$-MGD: Fashion-MNIST
$\vec{\delta}$-MGD: MNIST

LLM Results

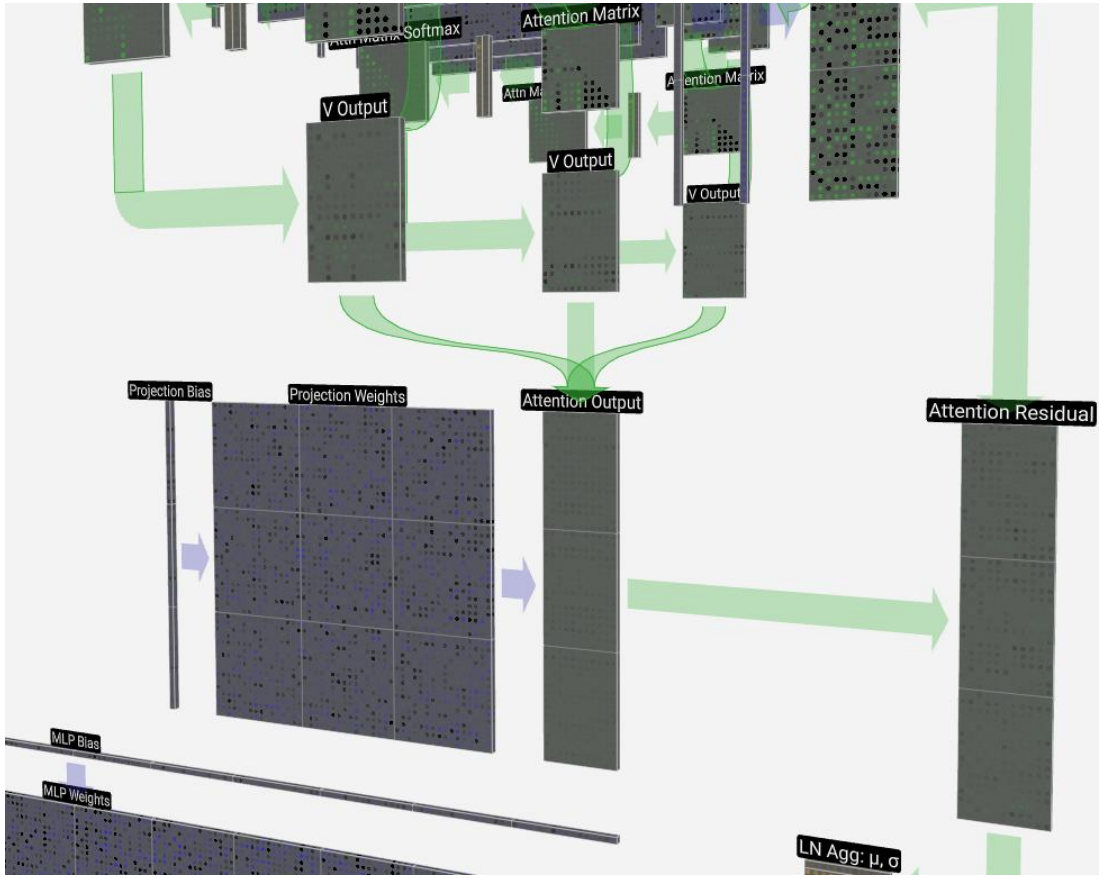
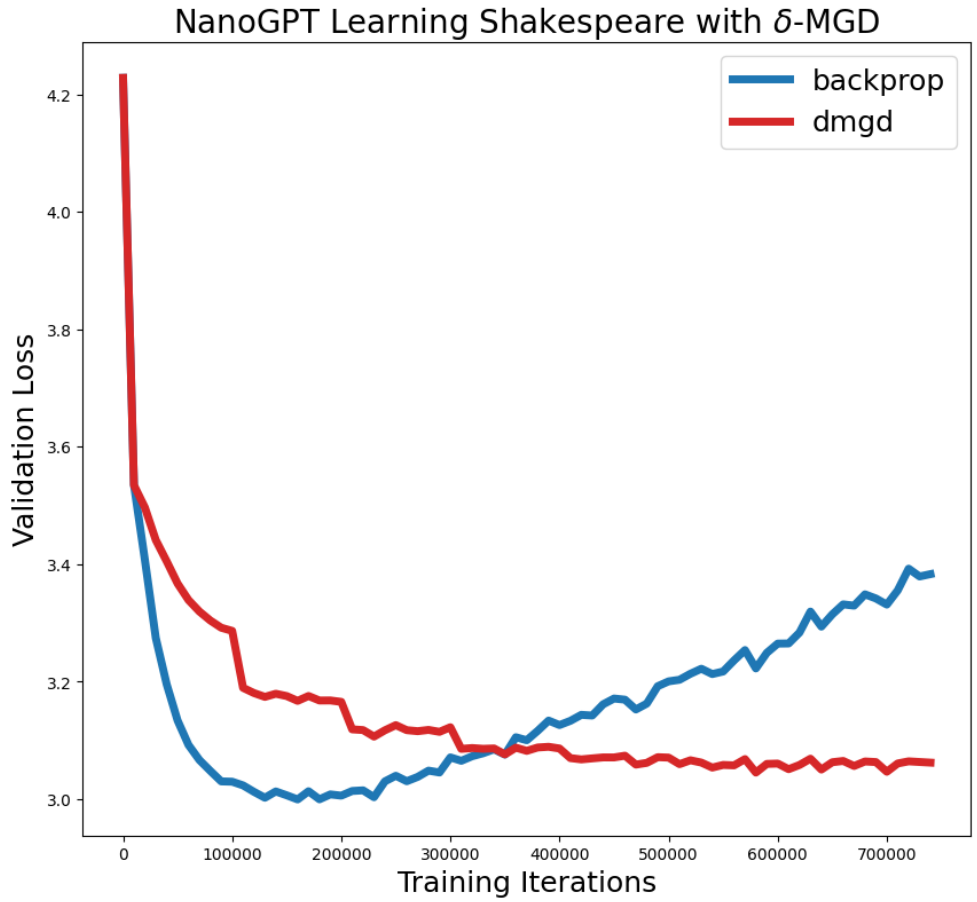


Image from Brendan Bycroft's LLM visualizer of the Projection Layers in NanoGPT between the self-attention layers and the MLP
https://bbycroft.net/llm



Questions?

δ Multiplexed Gradient Descent: Perturbative Learning with Astrocytes

Nick Skuda[1], Ryan O'Loughlin[12], Bakhrom Oripov[1], Noah Chongsiriwatan[3], Ian Whitehouse[3], Wolfgang Losert[3], Bradley Hayes[2], Adam McCaughan[1], Sonia Buckley[1]

1 National Institute of Standards and Technology

2 University of Colorado Boulder

3 University of Maryland