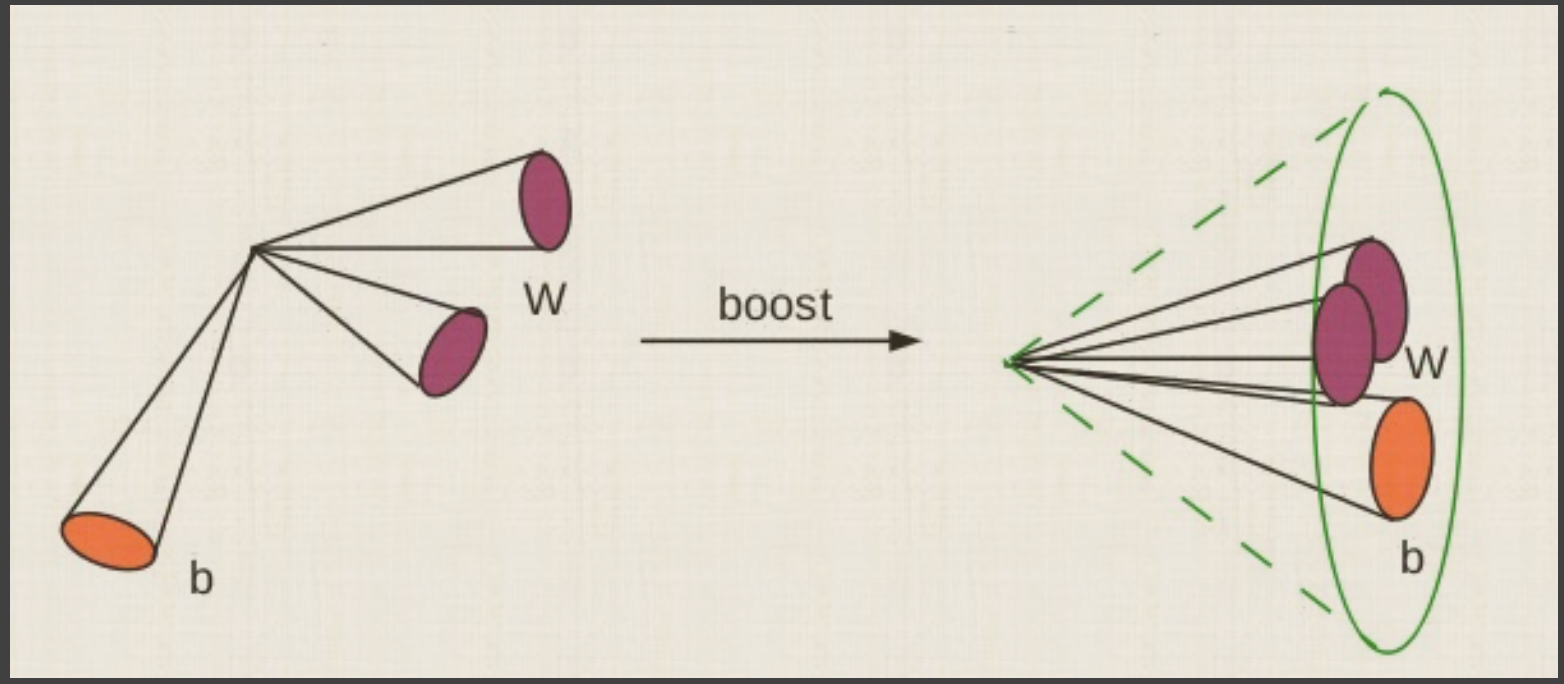
Multivariate Jet Calibration Using Neural Networks

-Dayton Grogan

-Duke University

Large-Radius Jets



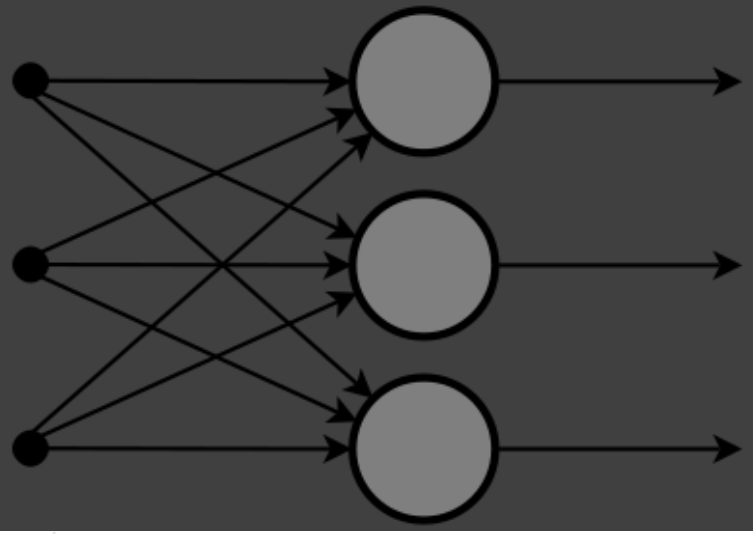
Why do we care?

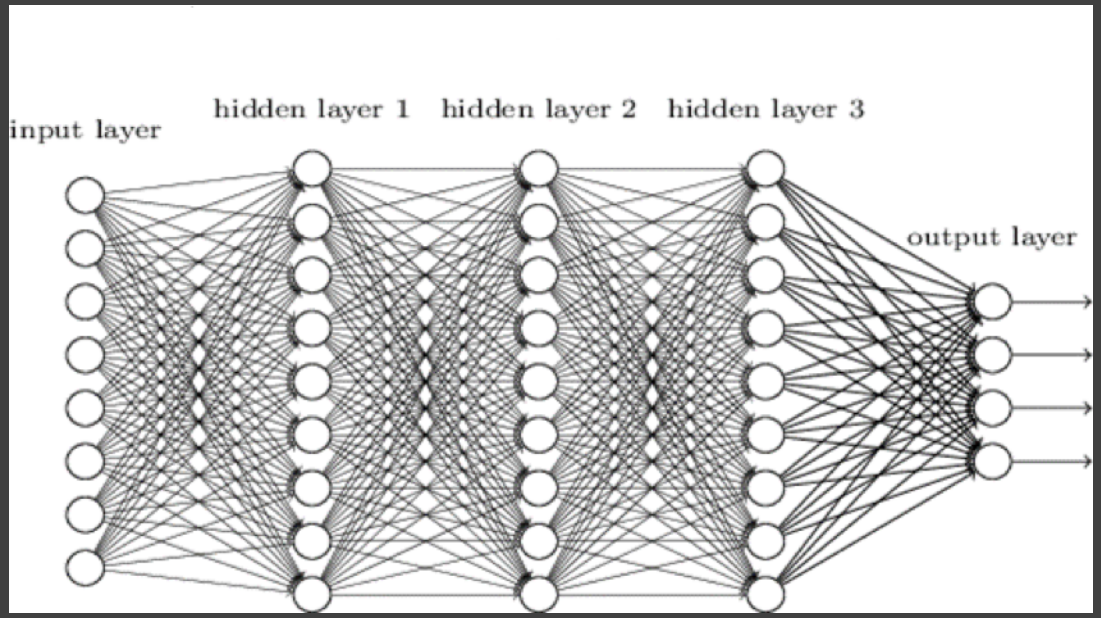
- Provides simplified event reconstruction
  - Jet **energy** → mass scale of the process
  - Jet **mass** → identity of particle
- Understanding scale and resolution is important!

Energy Calibration

- Calibration function, $R \approx \frac{True\ energy}{Reconstructed\ energy}$
- $true\ pT \approx R(Jet\ pT, \eta) \times reconstructed\ pT$
- Does more information give more precision?
- Limitations
  - not very practical
- → Use Neural Networks!

Neural Networks





Why Neural Nets?

- Accommodates many variables
- Easily updatable

Network Training

- Input: simulated data
- Output: calibration factor R
- loss function: $L2 = (\frac{true\ pT}{reco\ pT}$ - R$)^2$

Simulated data sample:

- simulated di-jet events
- $p_T > 200$ GeV, $|\eta| < 2.0$
- matched geometrically to true jet

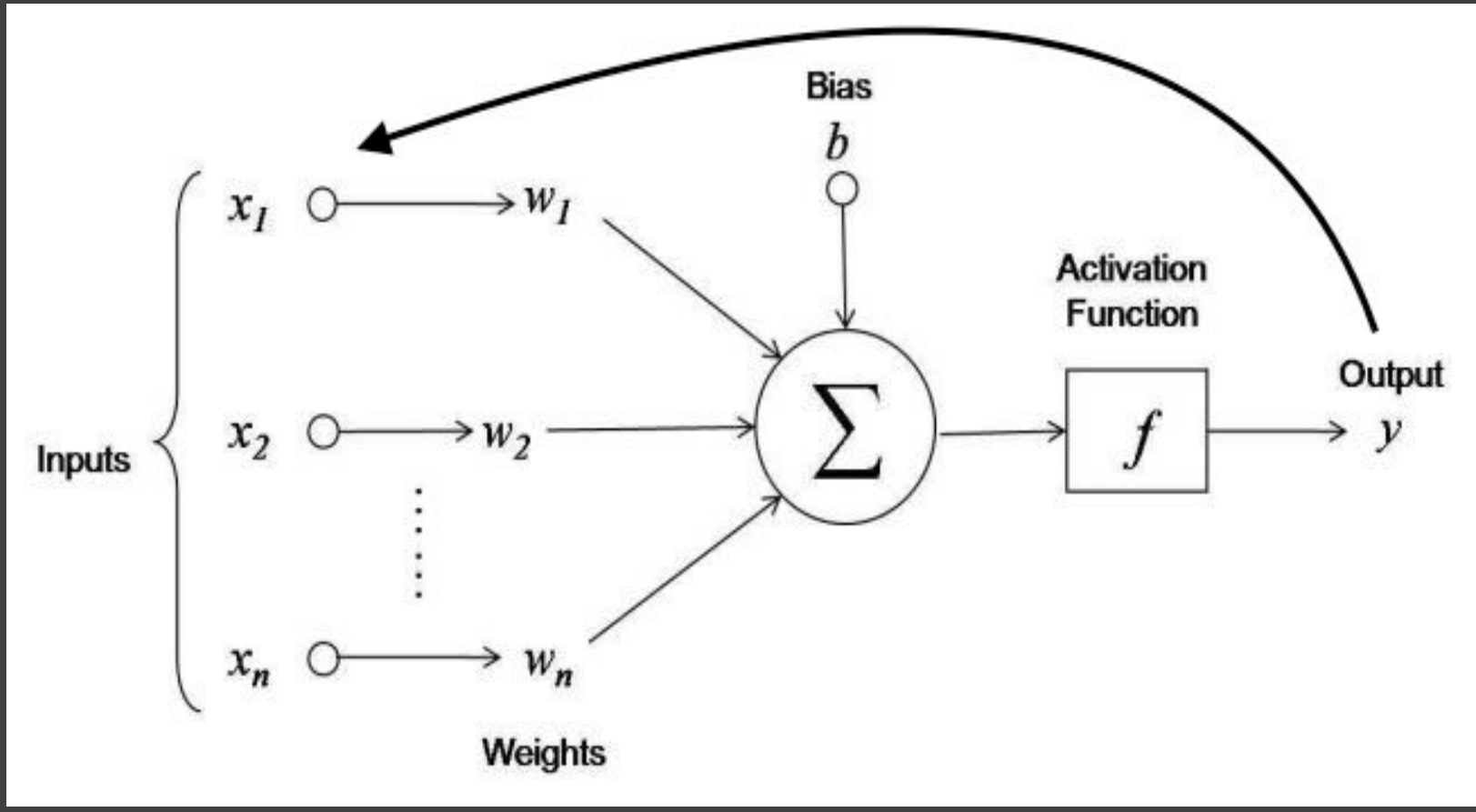
Network Configuration

Many options!

- Number of layers
- Number of nodes (neurons)
- activation functions
- Propagation algorithms

Does it matter?

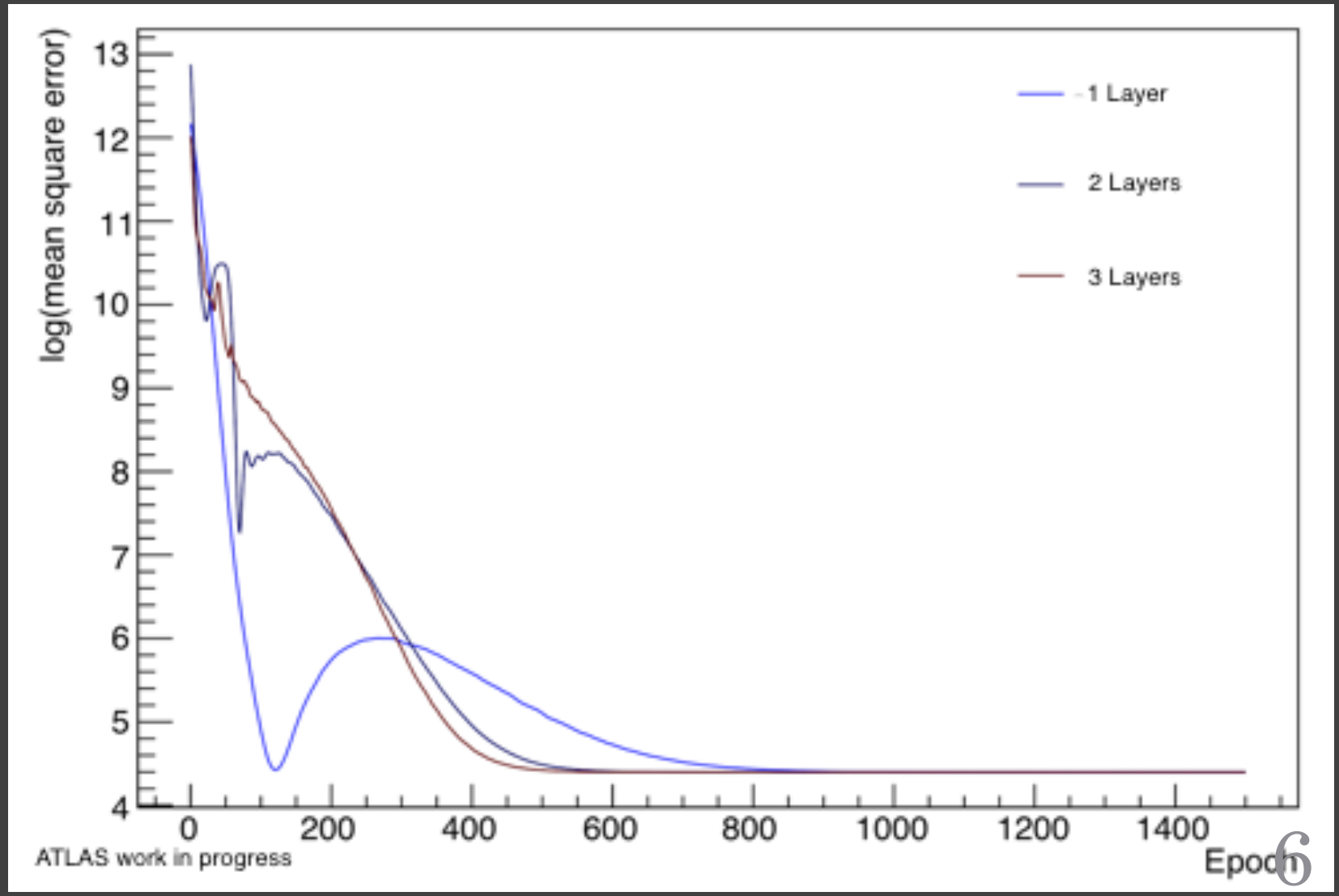
- Yes!
- Effects on **runtime** and **convergence**



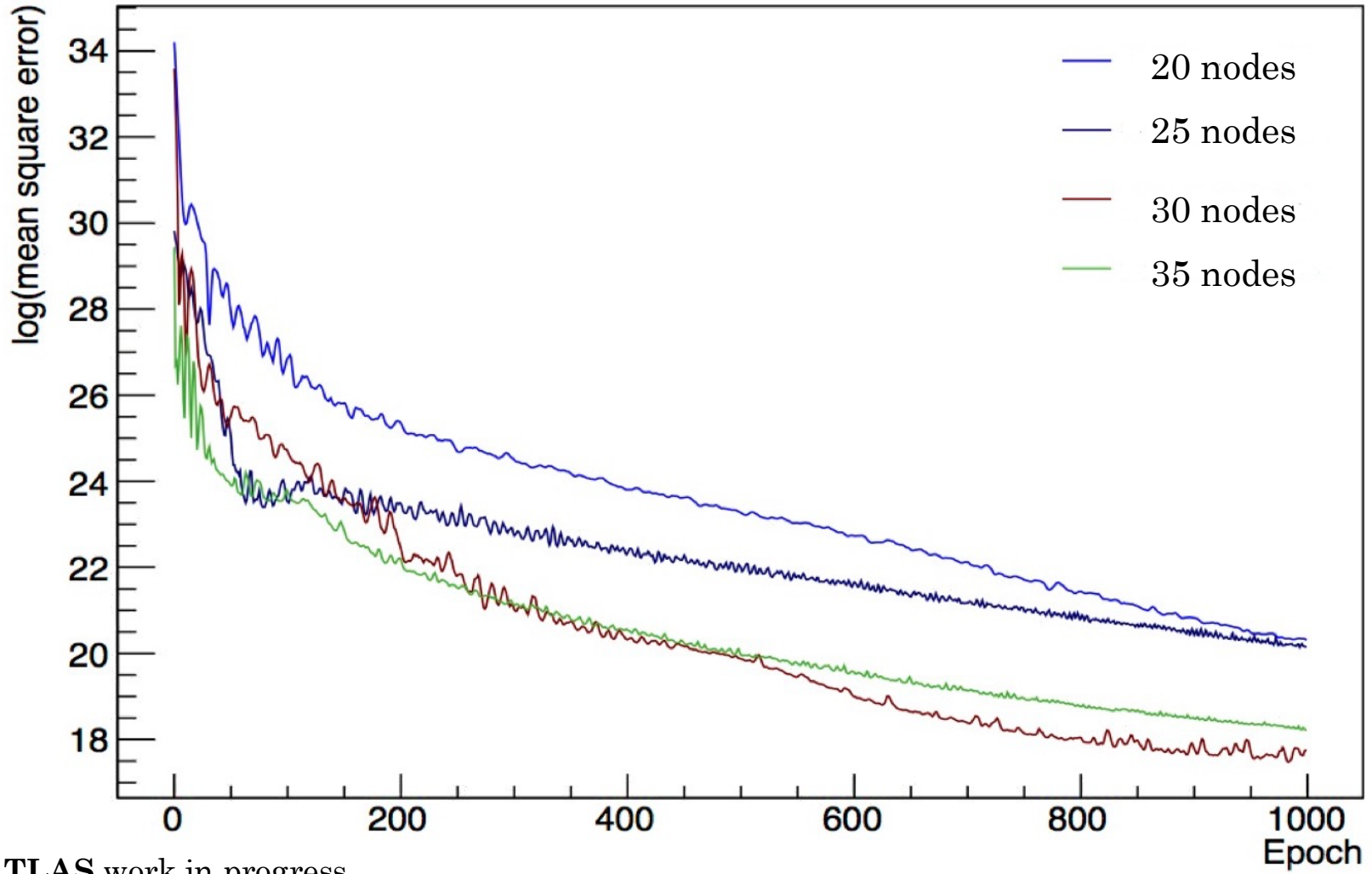
Epoch: one forward pass and one backward pass of *all* training examples

Impact of Structure: Layers

- Multiple layers may improve convergence
- Runtime increased by factor of 2.5/epoch
- Still effective to use a single layer



Impact of Structure: Nodes



ATLAS work in progress

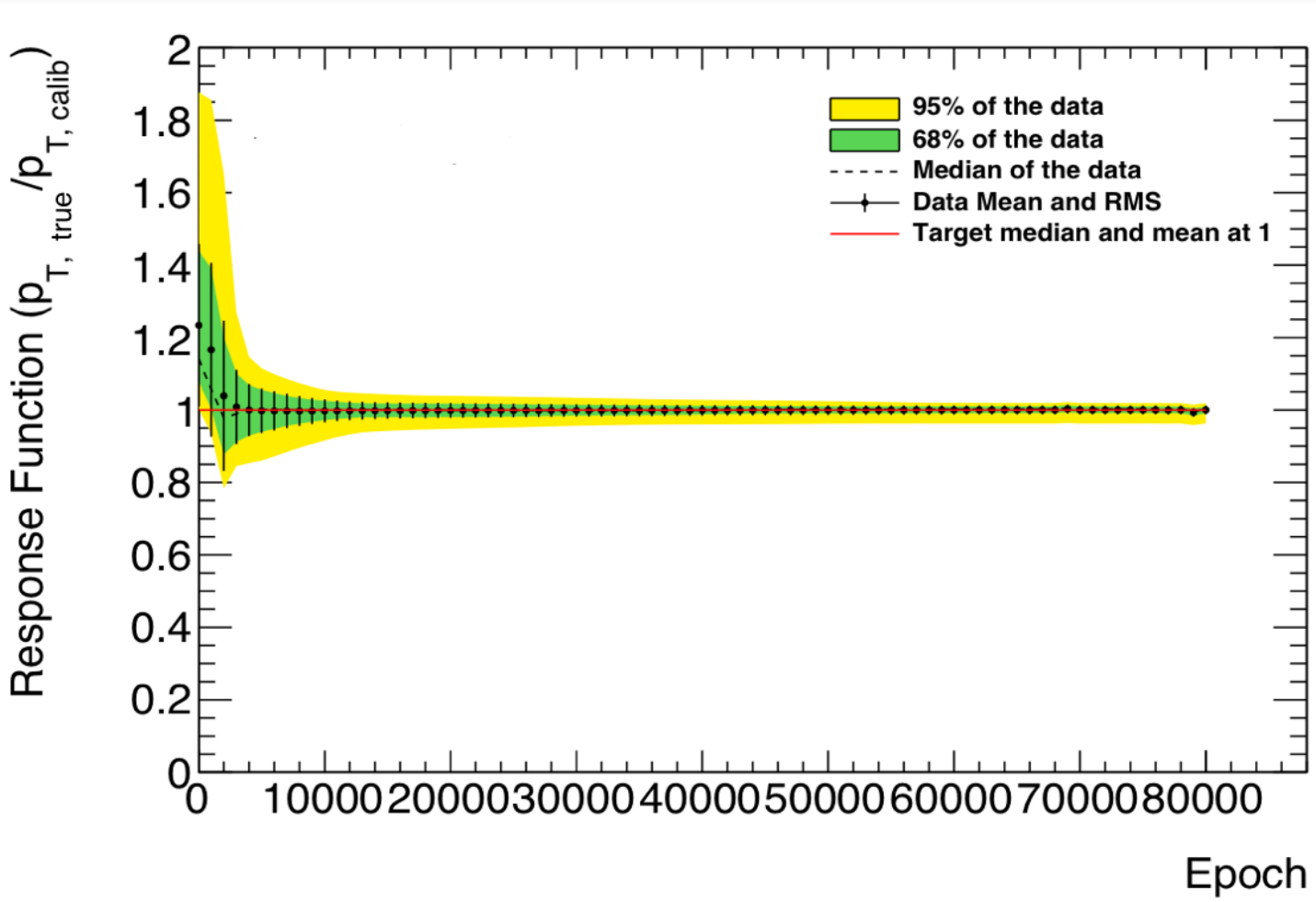
- Performance not correlated with nodal number
- More input variables requires more nodes

Final Network Structure

Network Parameters

- Number of layers: 1
- Number of nodes per layer: 20-50
- Activation function: Tanh
- Propagation algorithm: Adam

Preliminary Results



ATLAS work in progress

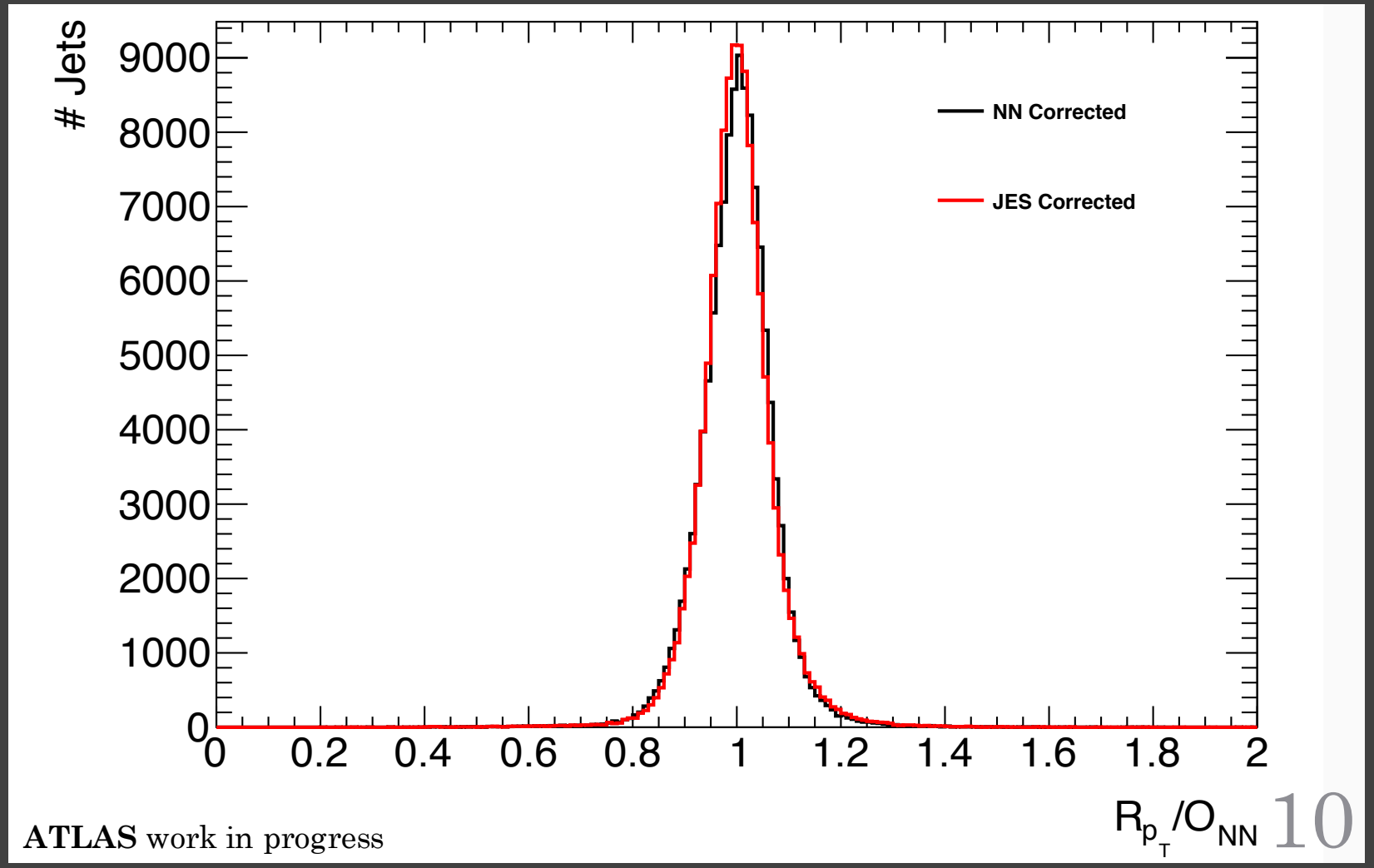
So, how are we doing?

- Successfully calibrated!
- convergence occurs within 10,000 epochs and takes roughly 2 hours.

Preliminary Results

How do we compare?

- NN looks very similar to by-hand JES calibration



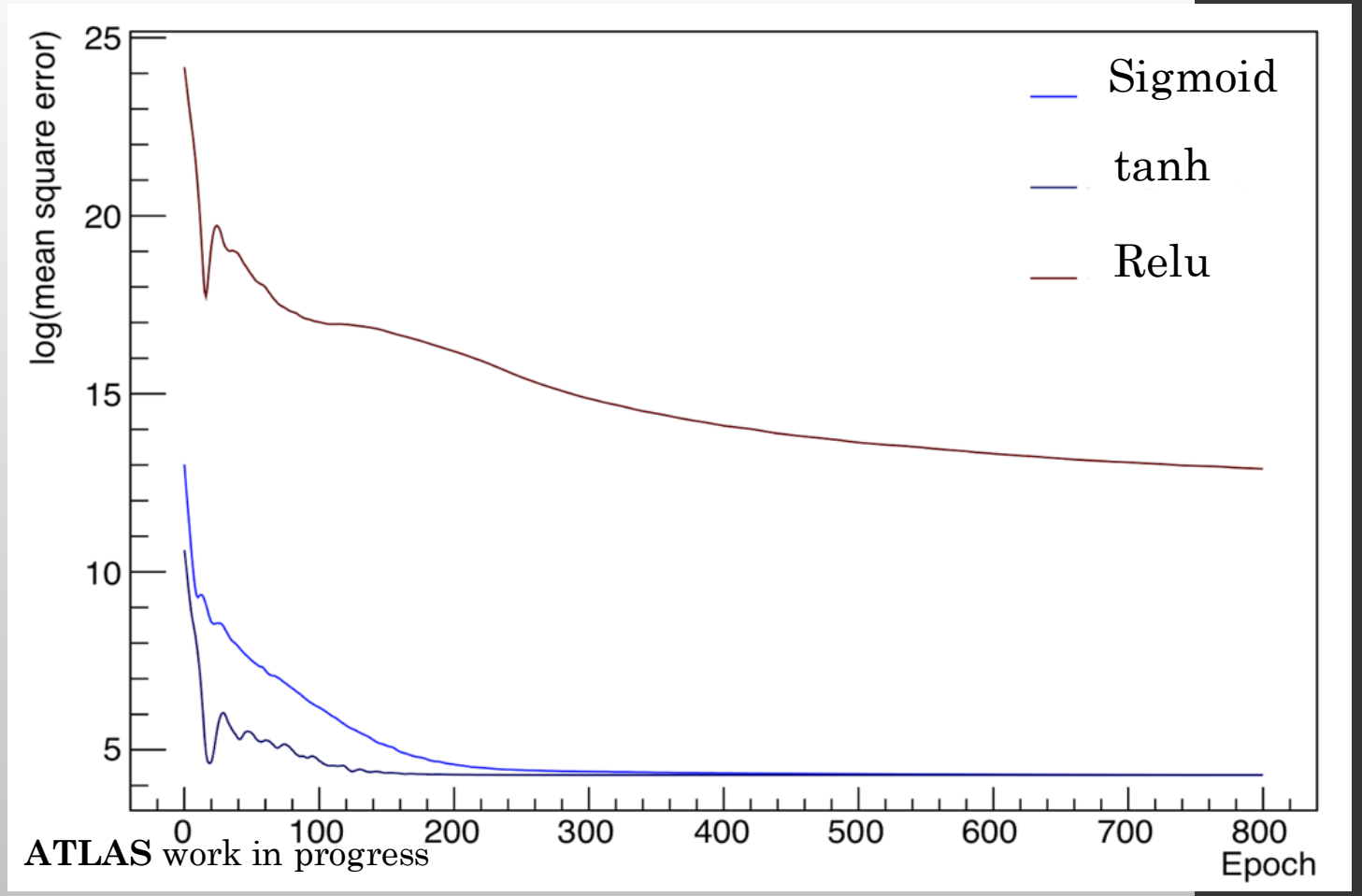
Future Directions

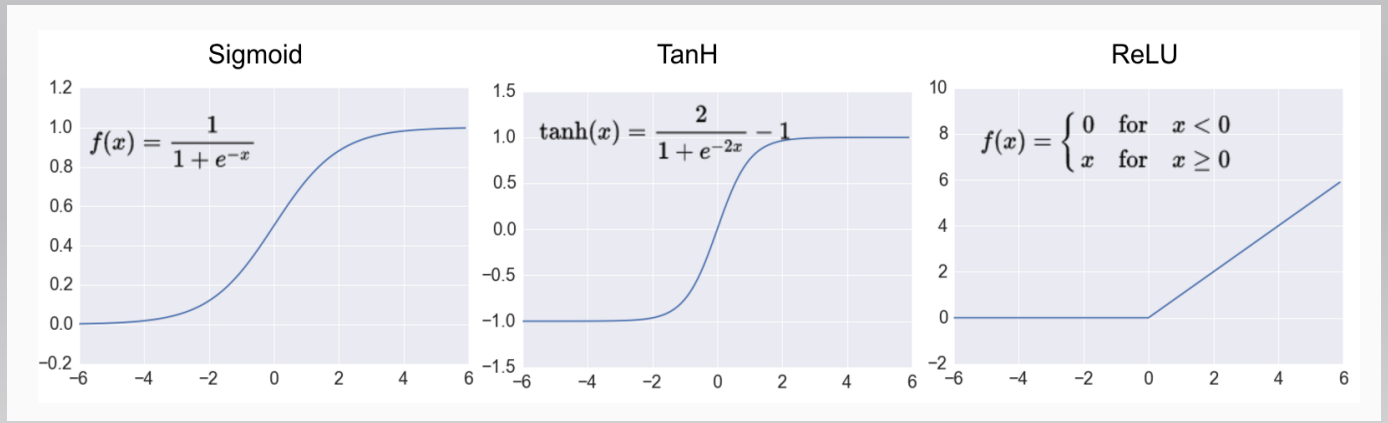
- Formalize optimization strategy and find optimal p_T calibration
- Repeat studies with new (jet substructure) input variables and less generic jets
- Expand strategy to calibrate other observables
    - particularly jet mass!

Backup Slides

Impact of Structure: Activation Function

- Relu (rectified linear unit)
- Sigmoid and tanh are more comparable





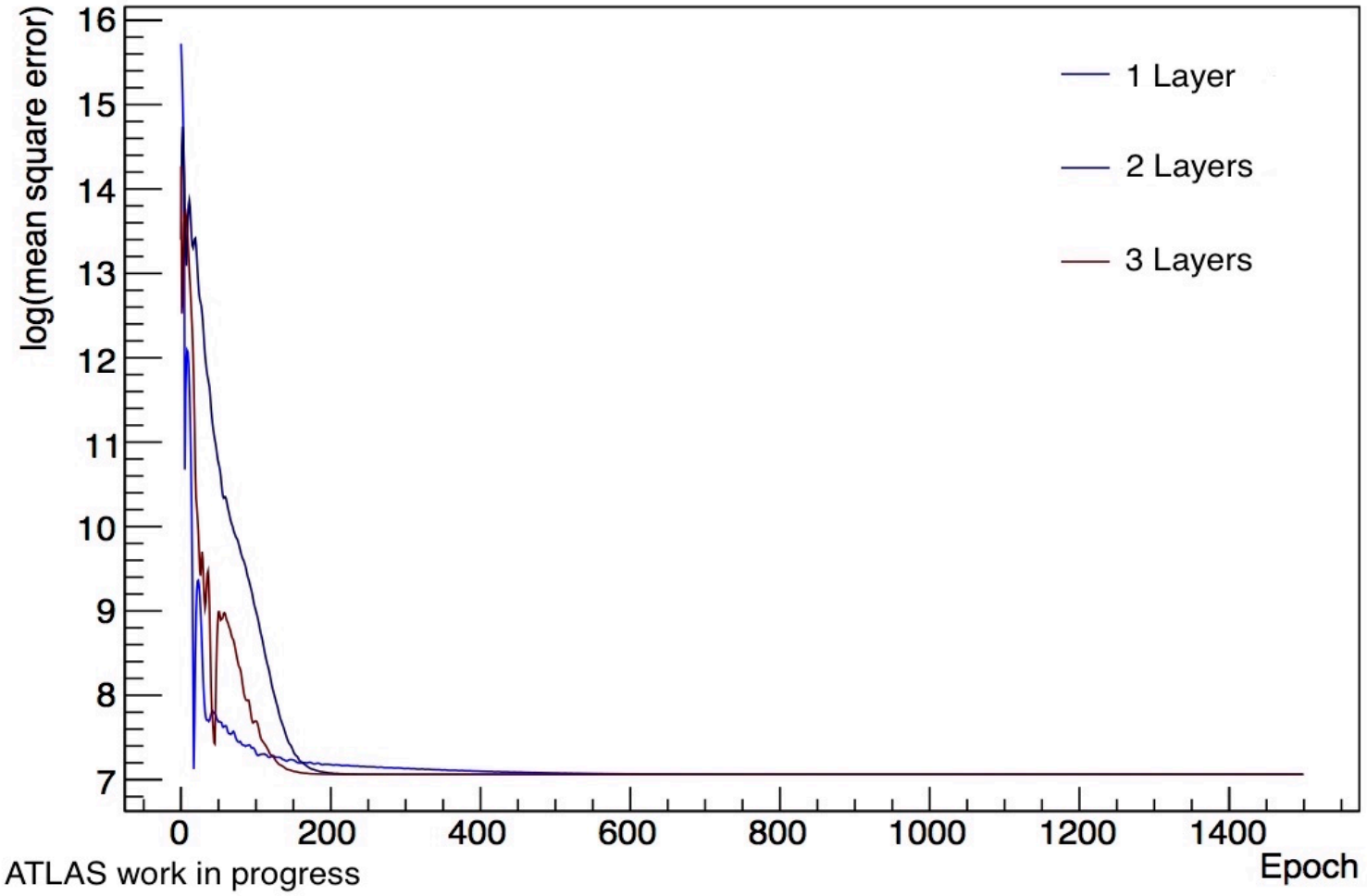
Back up Slides: Layers

Layers	Final Training MSE	Final Validation MSE	Training time
1	378.272	67.8496	107 sec
2	362.789	83.2652	189 sec
3	360.262	85.7885	250 sec
4	360.221	85.7614	290 sec

One data file (JZW7) and 20 nodes

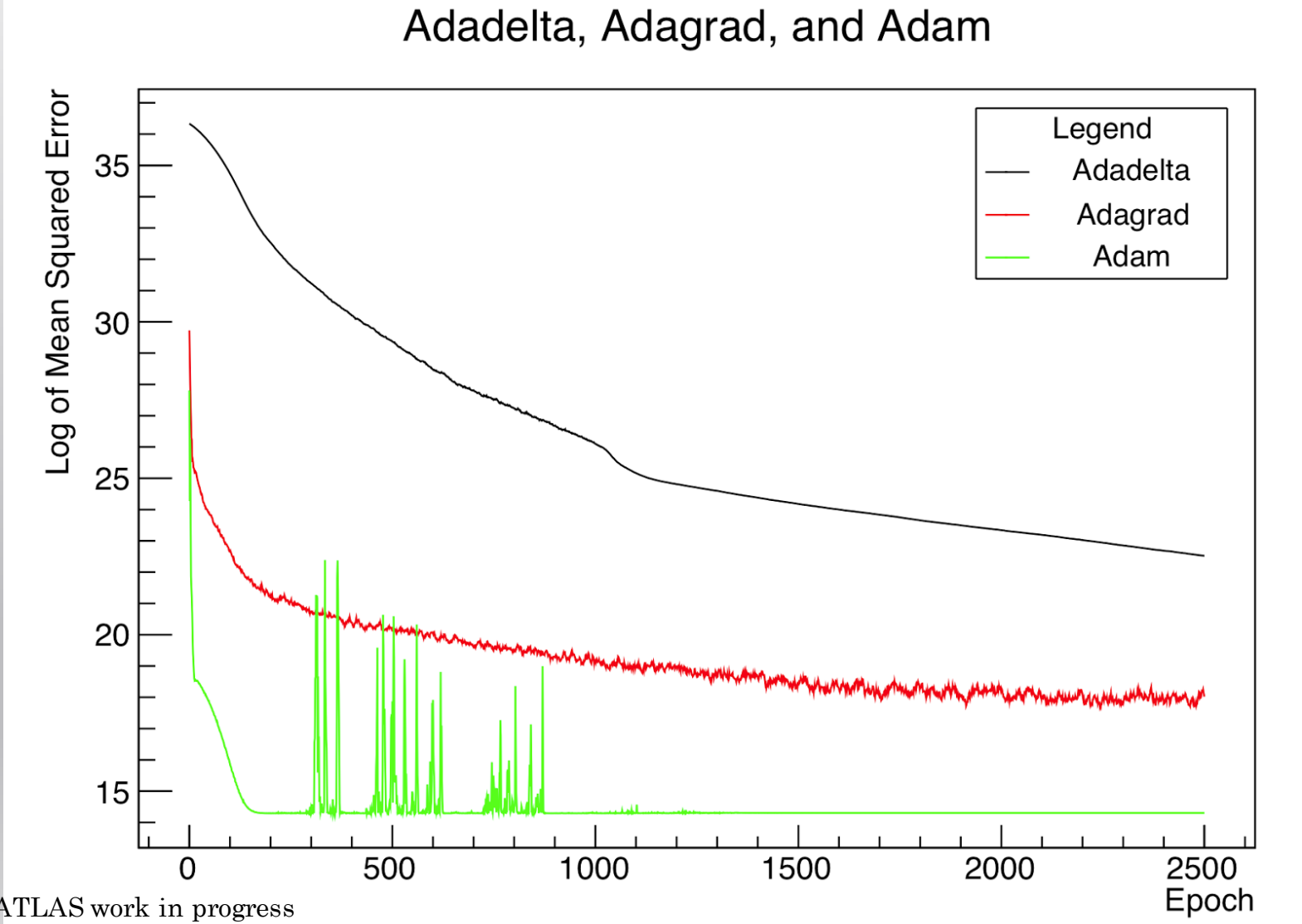
ATLAS work in progress

Backup Slides: Layers



- Convergence is comparable among the different layers
- Only see difference when lower numbers of nodes are used

Backup Slides: Propagation algorithm



ATLAS work in progress

- although there are strong early fluctuations away from best calibration, Adam converges fastest

- Other algorithms are smoother, but may be susceptible to getting stuck in a local minimum.

Backup Slides: Calibration Comparison