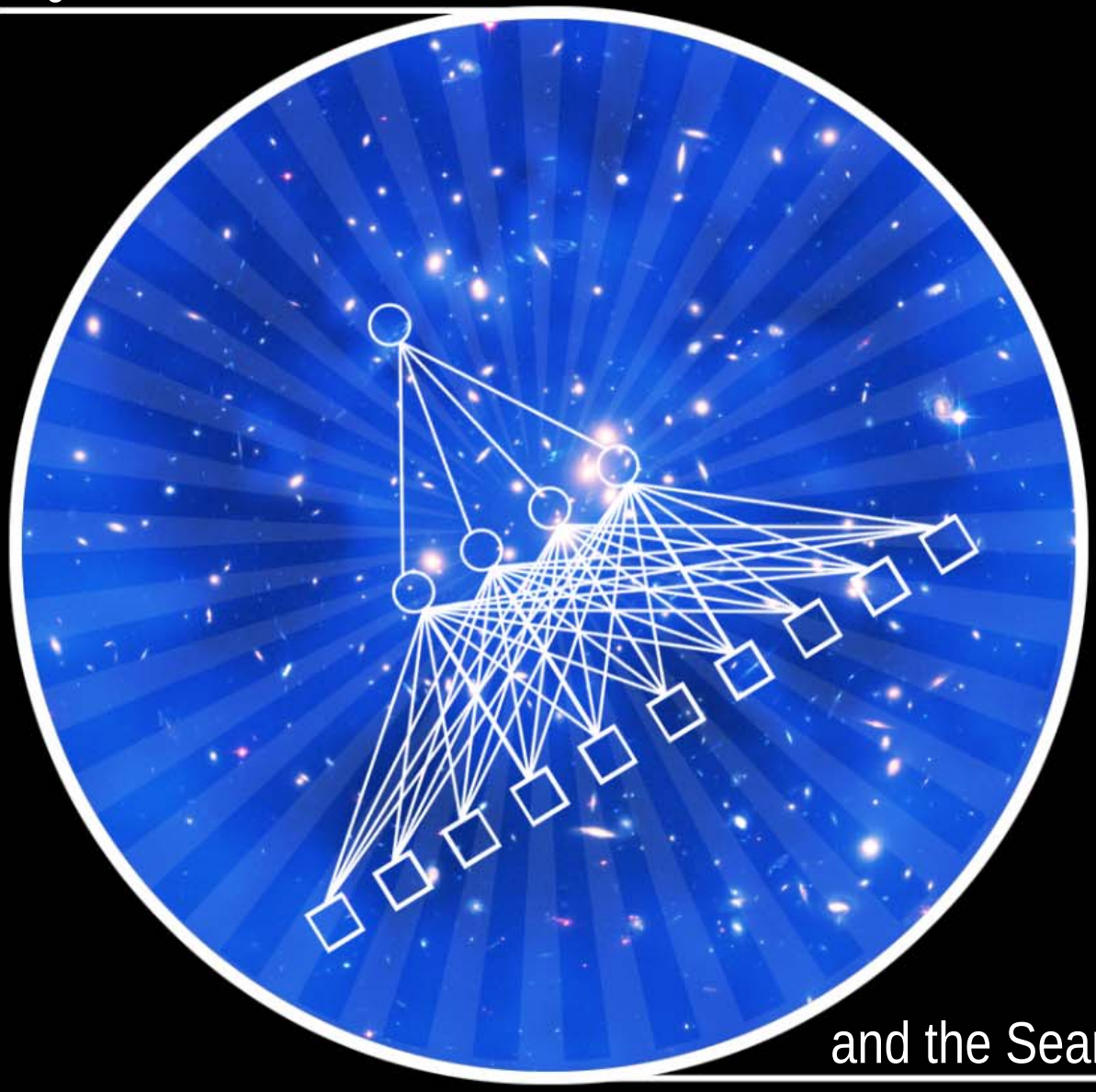
August 2, 2007



Neural Networks and the Search for Dark Matter

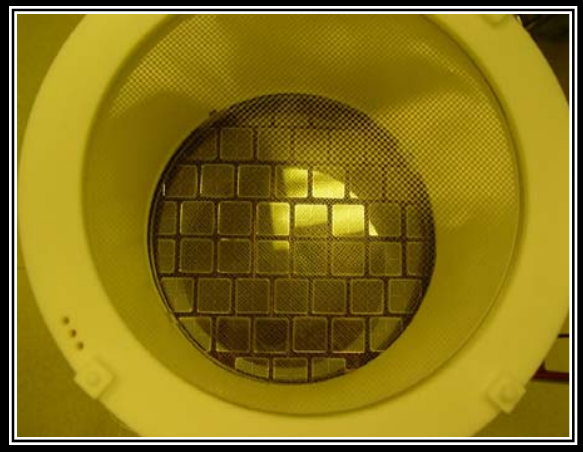
Marc Z. Miskin

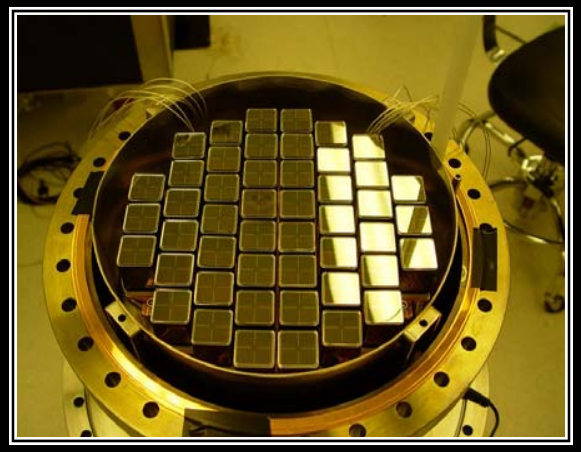
Introduction

- The XENON10 Detector
- Neural Network Overview
- The Double Hit Detection Network

The XENON10 Detector

- The XENON10 Detector uses two arrays of Hamamatsu PMT's to:
  - Convert light to charge and distinguish nuclear recoils from electron recoils
  - Reconstruct events in X and Y to facilitate fiducial volume cuts

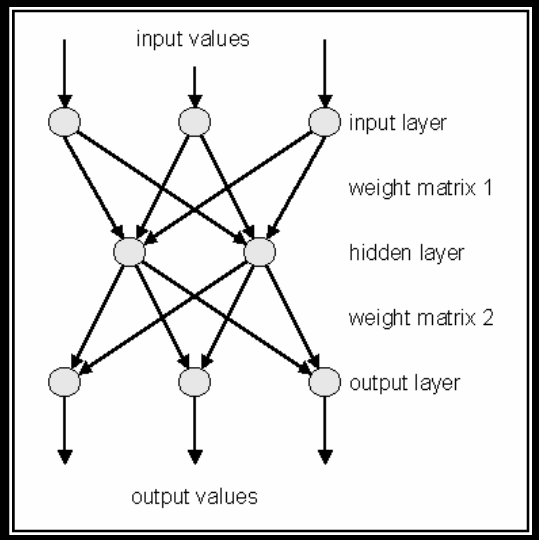




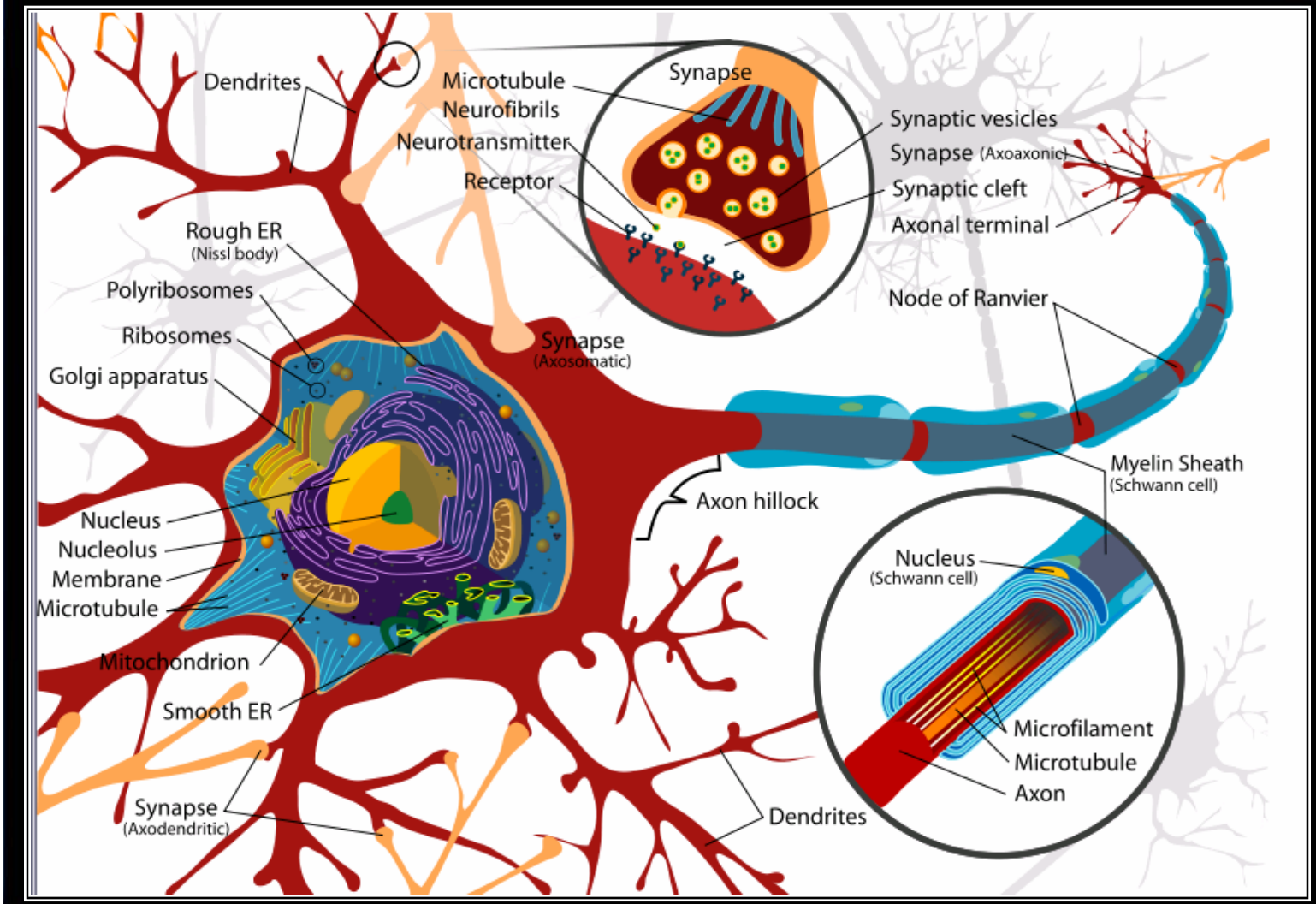
- One source of background which remains is a double hit event, that is a multiple scattering event being recognized as a single event
- Given information from the PMT's, a properly designed NN can recognize and remove such signals

Neural Network

- NN is a processing network modeled on the human brain
  - Processing elements are multiply connected both in and out. (10^{15}-10^{16} in humans)
  - The inputs to a processing element are weighted and summed (weights can be positive or negative)
  - A threshold is subtracted
  - The result is fed to an activation function (generally a sigmoid) to produce an output
  - Adjusting the weights and thresholds allows the network to produce different responses
  - Given a training set of data, a network can be adjusted automatically to produce proper outputs for certain inputs
  - A well trained network can at times interpolate results well outside its training range



Neural Network

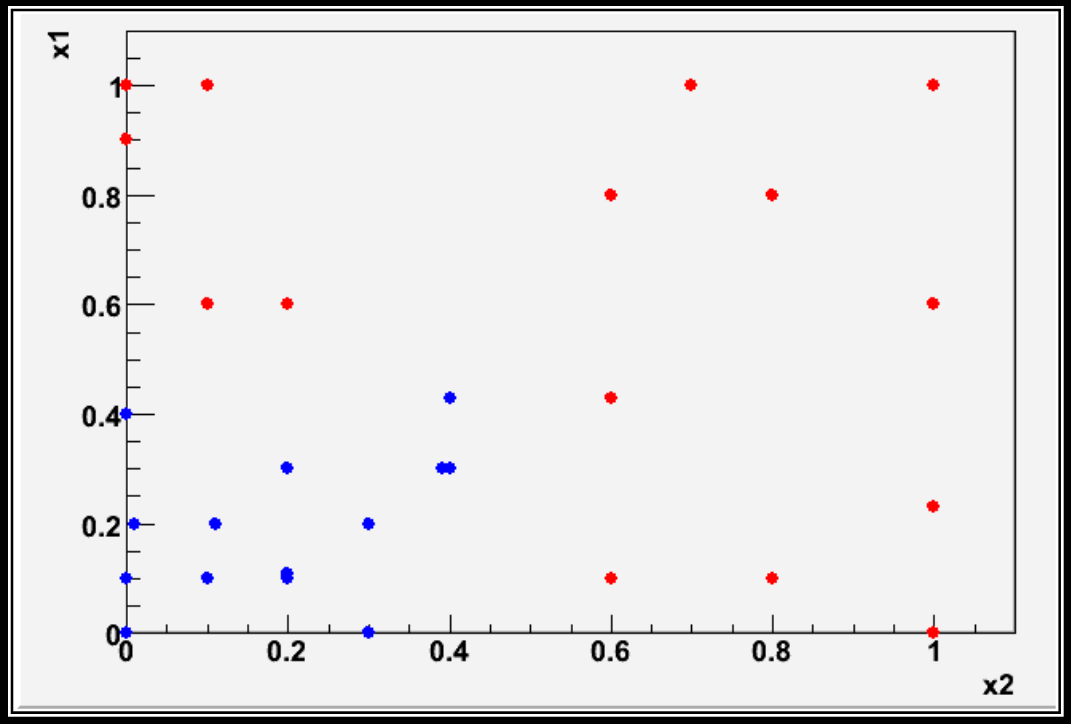


Neural Network

- Learning is analogous to fitting a curve
- There exists some error between the output of the network for a given set of training data and the desired response
- The network weights are adjusted to move across an “error space” and attempt to locate a minimum
- Movement governed by the learning rule
- Repeated for a set number of epochs
- The effectiveness of this mapping is dictated by the topology of the network and the appropriateness of the training set
    - Ex- Pattern complexity
    - Ex- Poor training data
    - Ex- Under fitting and Over fitting

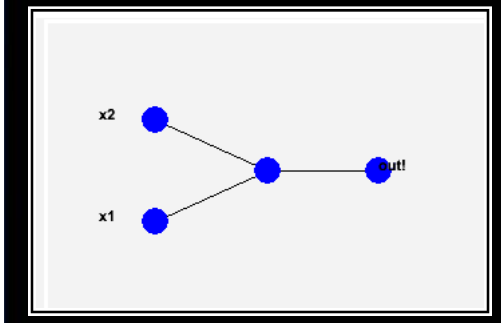
Neural Network

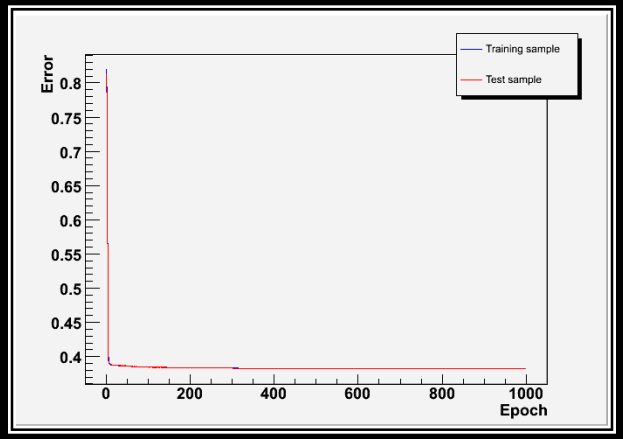
- A basic example: The OR Gate

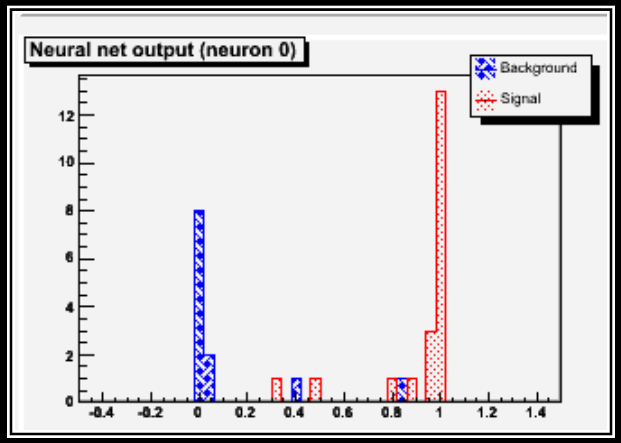


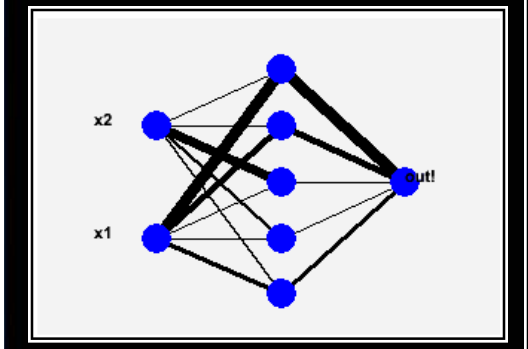
Neural Network

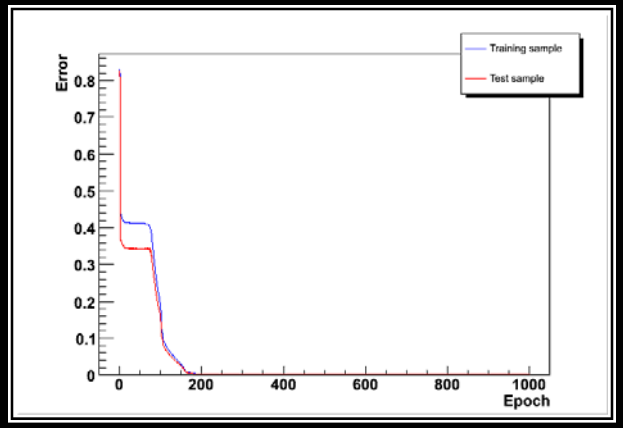
- Different network designs meet different success

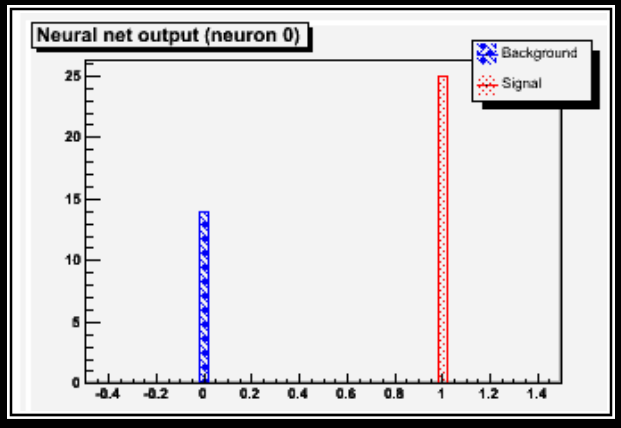










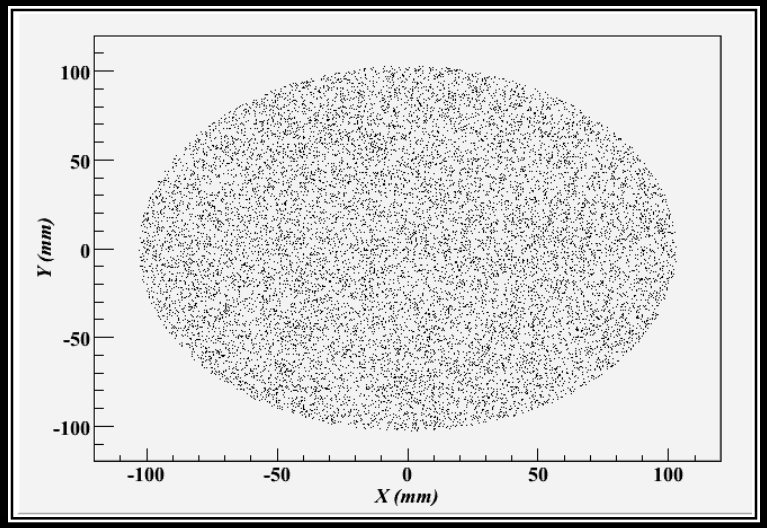


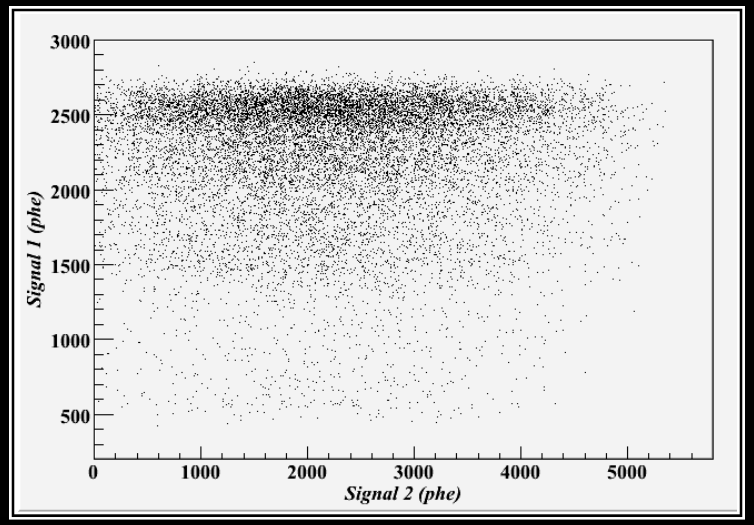
Neural Network

- Comments on the OR Gate:
  - What happens if a negative signal is presented?
  - Note the “saddle” in the training curve
  - What if the training set had been unbalanced?
  - What if the training data were uncorrelated to the modeled pattern?
- Minimizing the error in training is not equivalent to minimizing the actual error.
- Tests must be made to evaluate the true accuracy of the network

The Double Hit Neural Network

- Training data:
  - Uses each of the top 47 PMTs as input
  - The value at each input is the ratio of the amount of photoelectrons at that PMT to the total quantity of photoelectrons
  - 3×10^4 examples randomly distributed in XY
  - Equal number of single and double hit examples
  - Initial hits generally deposit 2500 phe, the second hit ranges from 0 to 5000 phe





The Double Hit Neural Network

- Network structure
  - Originally designed without a hidden layer, problem proved too complex
  - Changed to use one input layer with 47 neurons, one hidden layer with 47 neurons, and a single neuron output layer
  - Uses hyperbolic tangent activation functions on the input and hidden layers and a sigmoid output function
- Trained until error remained constant for 1000 Epochs

The Double Hit Neural Network

- Networks were tested with 3×10^4 examples randomly distributed in X and Y
- First signal of mostly 25000 phe, second signal in the range of 0 to 5000 phe
- Networks evaluated by MSE, Correlation Coefficient, AIC, MDL

$$MSE = \frac{\sum_{i=0}^{N_{trial}} (d_i - y_i)^2}{N_{trial}}$$

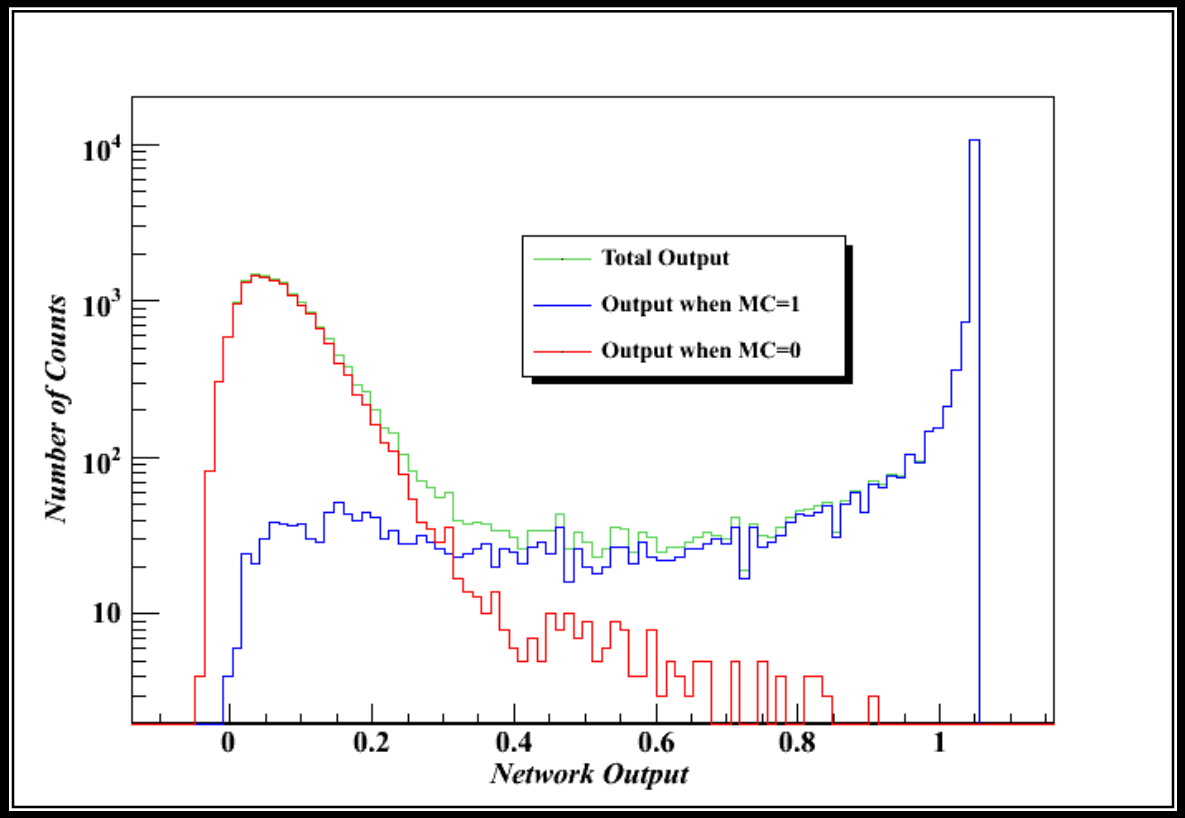
$$AIC = N_{training} \ln(MSE) + 2k$$

$$MDL = N_{training} \ln(MSE) + 0.5k \ln(N_{training})$$

$$r = \frac{\sum_{i=0}^{N_{trial}} (x_i - \bar{x})(d_i - \bar{d}) \Big/ N_{trial}}{\sqrt{\sum_{i=0}^{N_{trial}} (x_i - \bar{x})^2 \Big/ N_{trial}} \sqrt{\sum_{i=0}^{N_{trial}} (d_i - \bar{d})^2 \Big/ N_{trial}}}$$

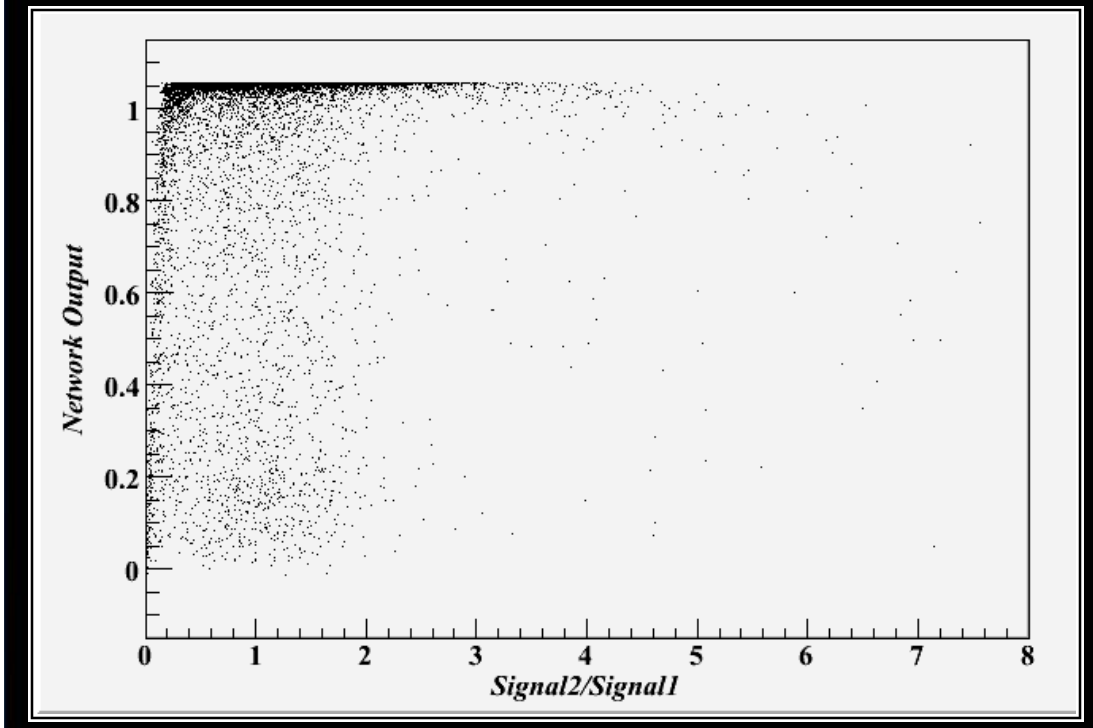
The Double Hit Neural Network

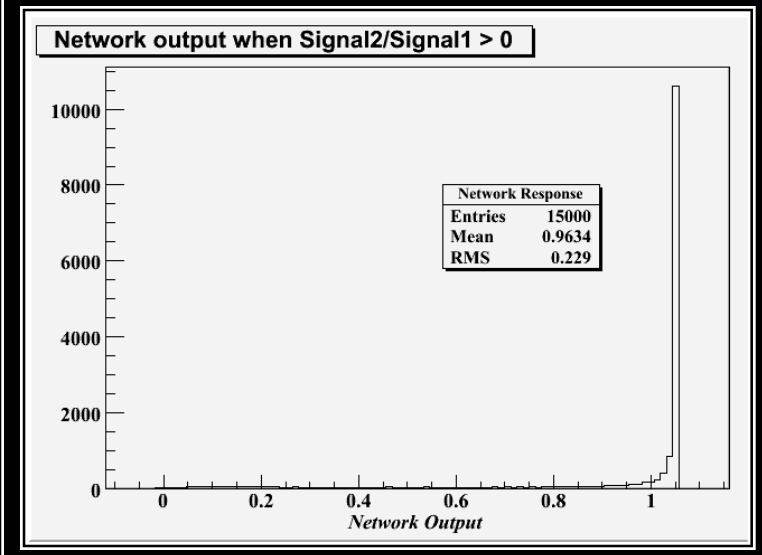
- Original training set used events over the full range of the detector
- MSE: 0.028027
- MDL: -95361.128462
- Correlation Coefficient: 0.929623
- Akaike Information Criterion: -1026929.041



The Double Hit Neural Network

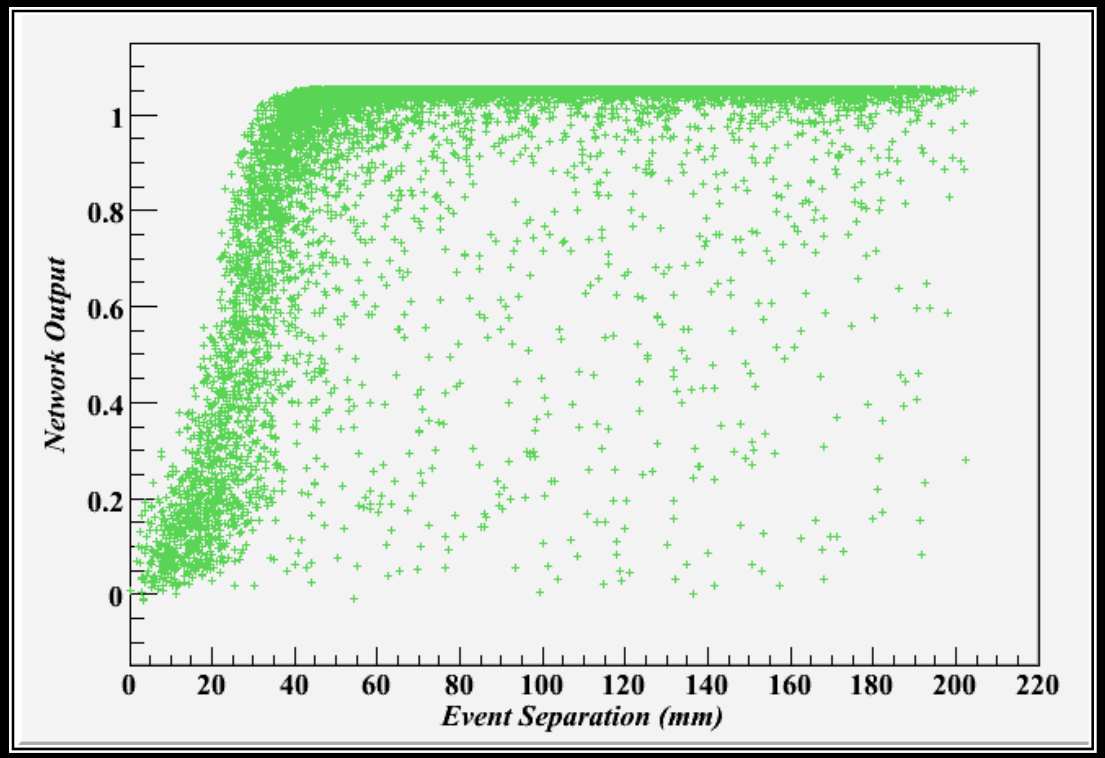
- Original training set used events over the full range of the detector

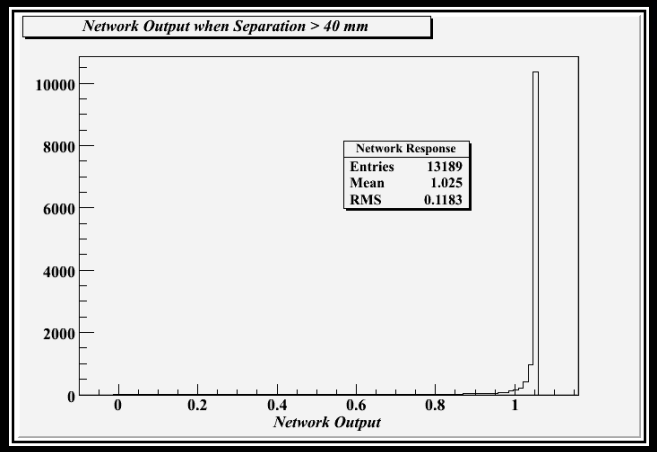


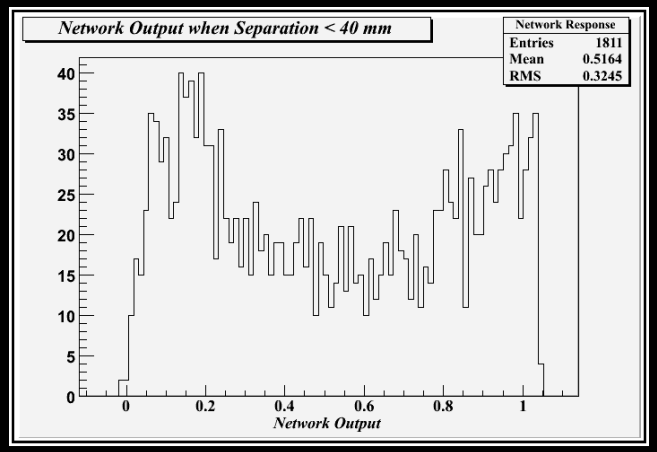


The Double Hit Neural Network

- Original training set used events over the full range of the detector

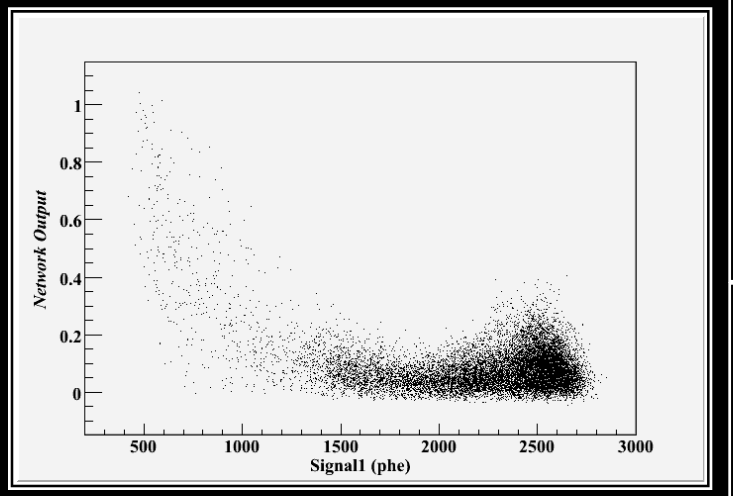


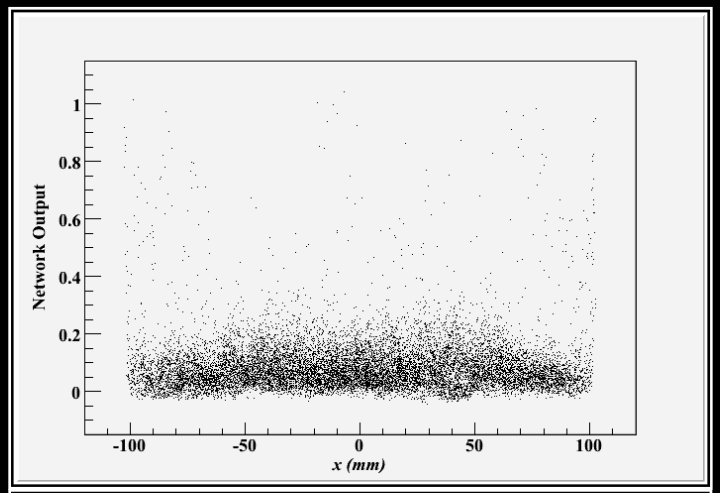


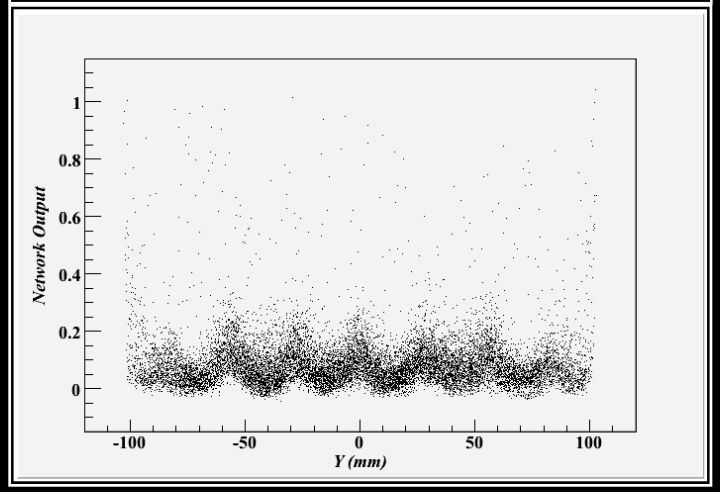


The Double Hit Neural Network

- Original training set used events over the full range of the detector

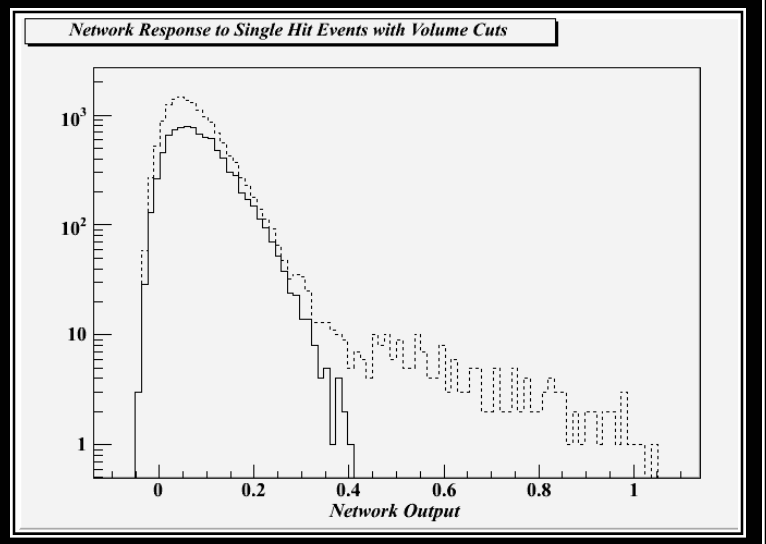


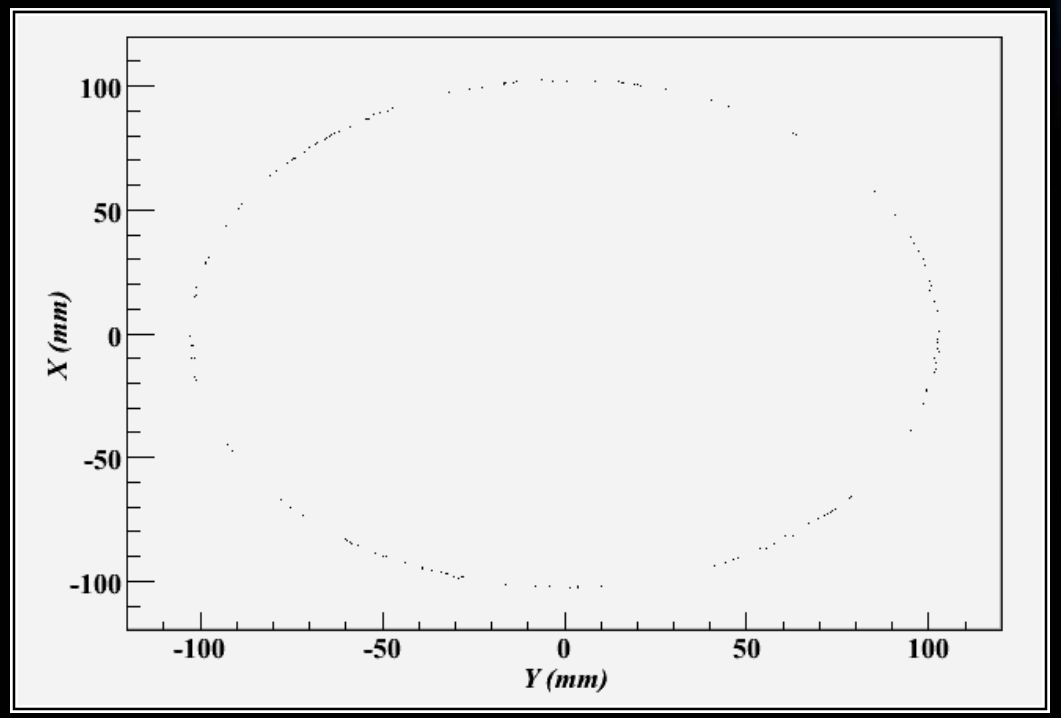




The Double Hit Neural Network

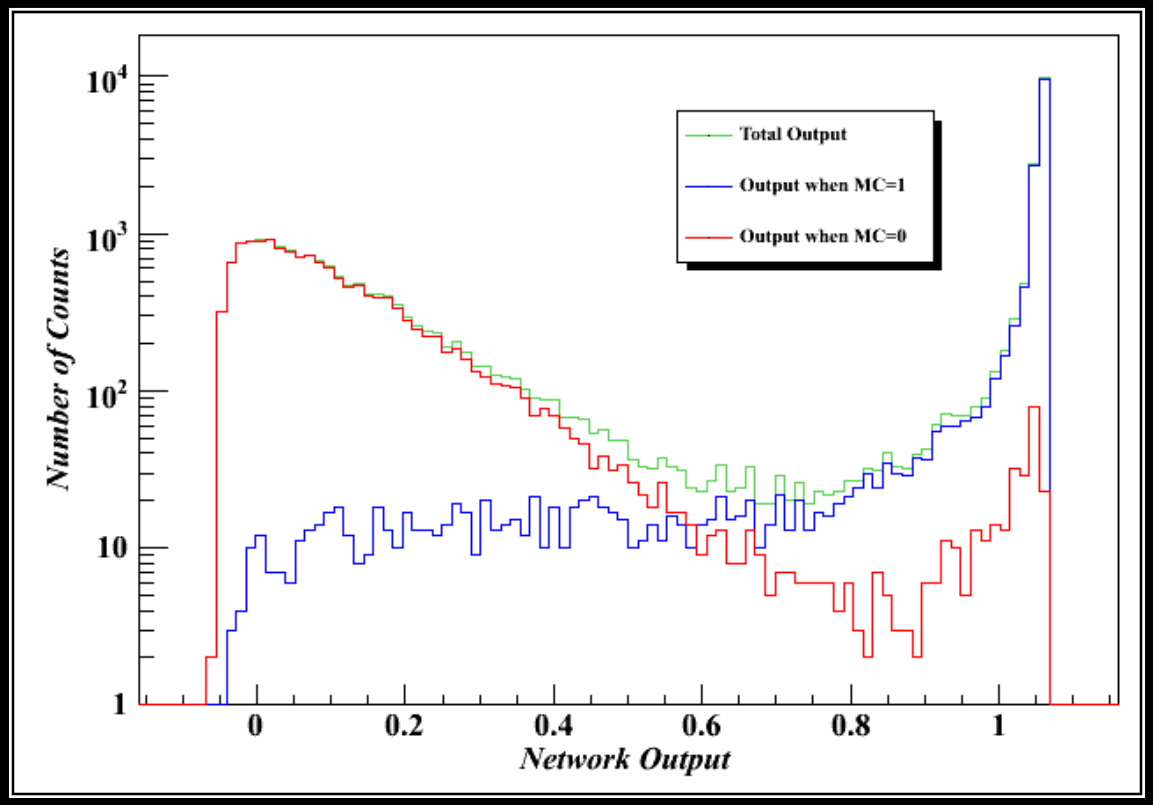
- Original training set used events over the full range of the detector





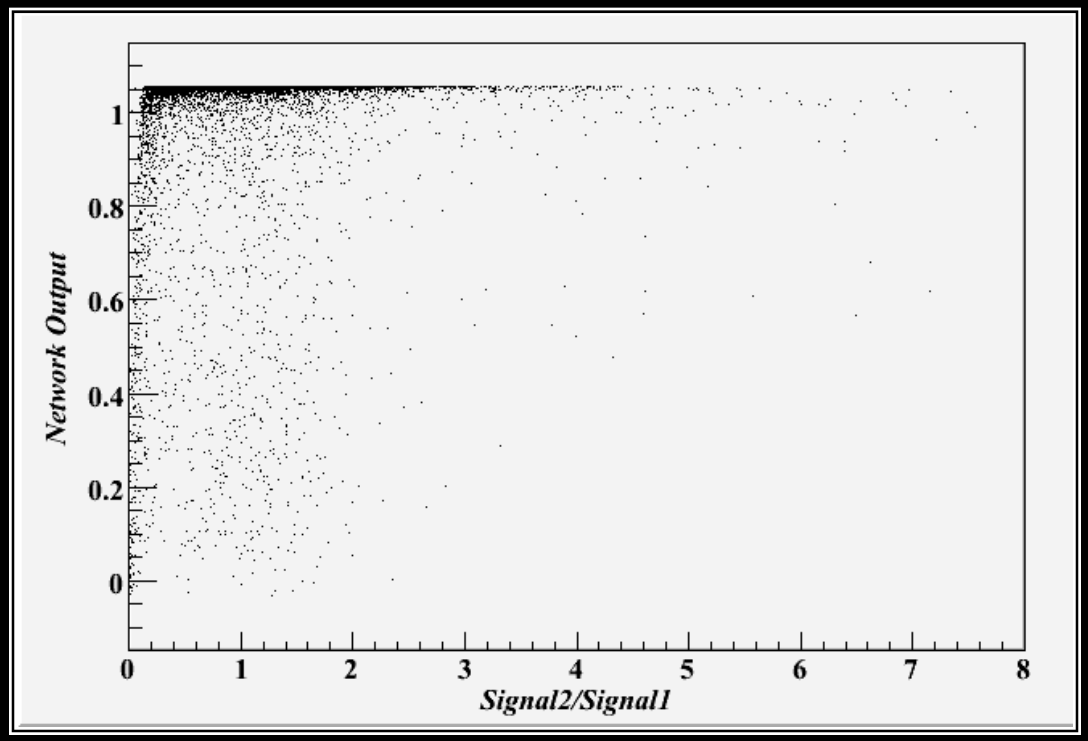
The Double Hit Neural Network

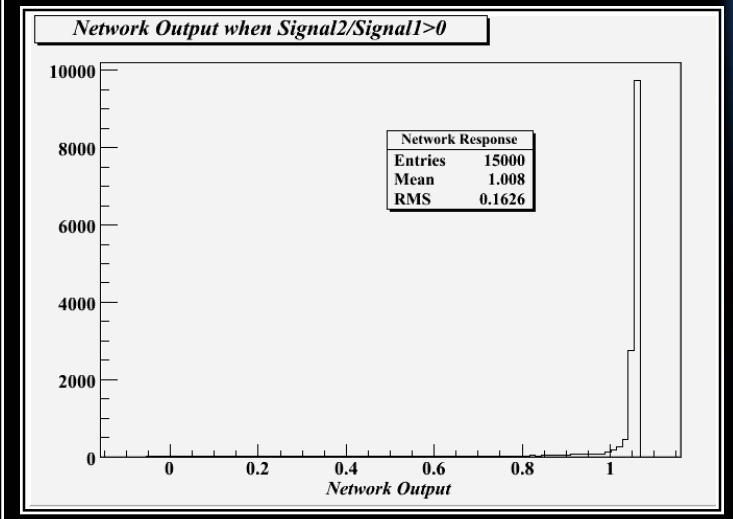
- Updated training set only used events separated by 5-40 mm
- MSE: 0.029985
- Akaike Information Criterion: -100603.3557
- Correlation Coefficient: 0.93275
- MDL: -93335.44232



The Double Hit Neural Network

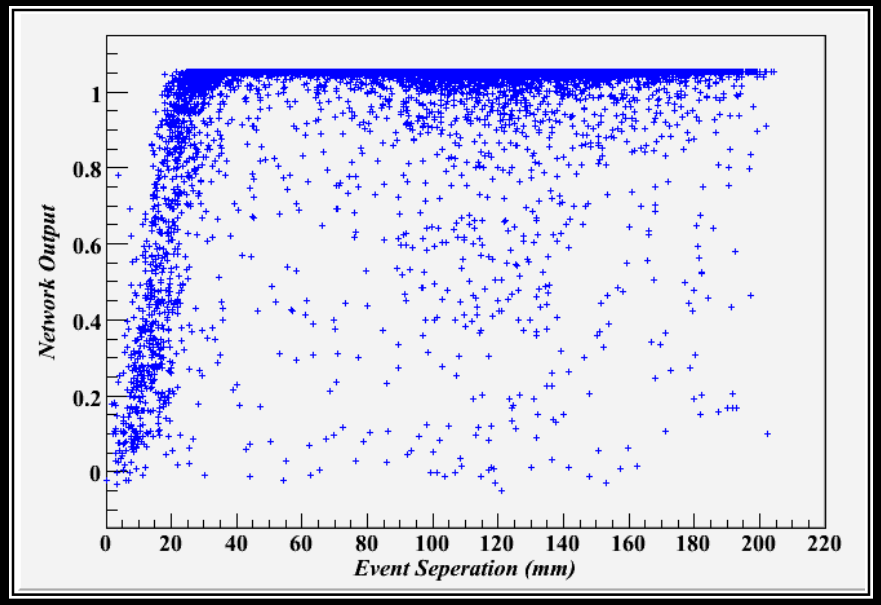
- Updated training set only used events separated by 5-40 mm

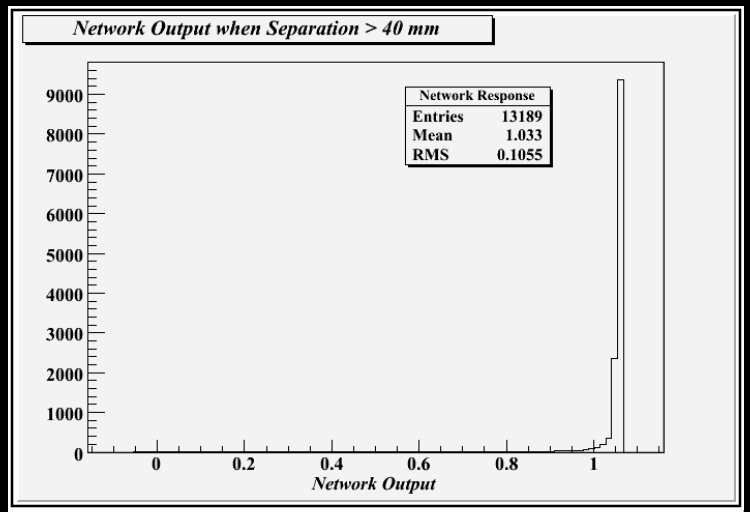


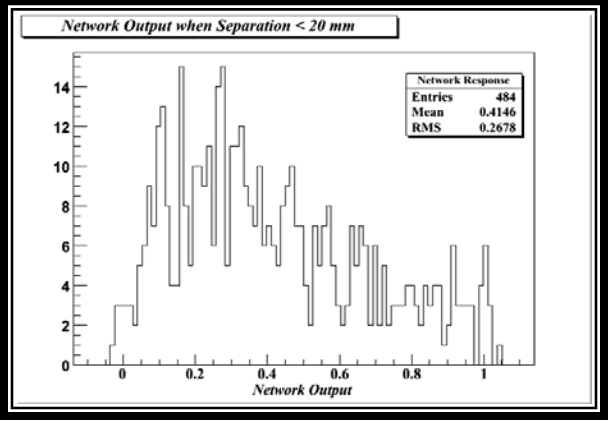


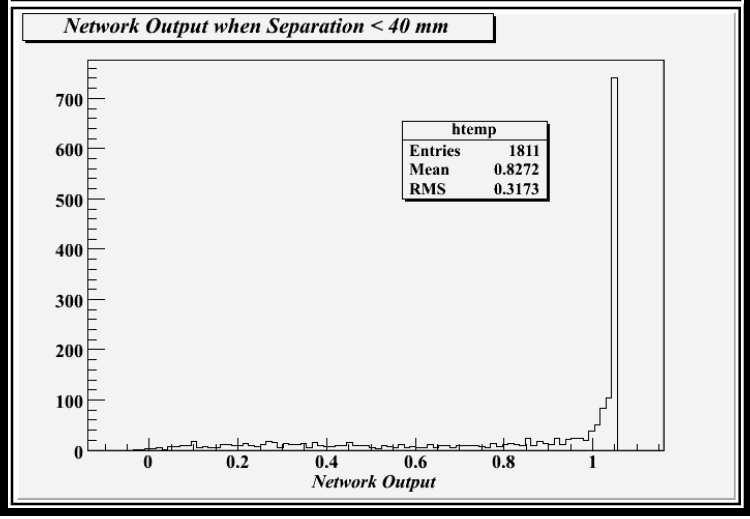
The Double Hit Neural Network

- Updated training set only used events separated by 5-40 mm



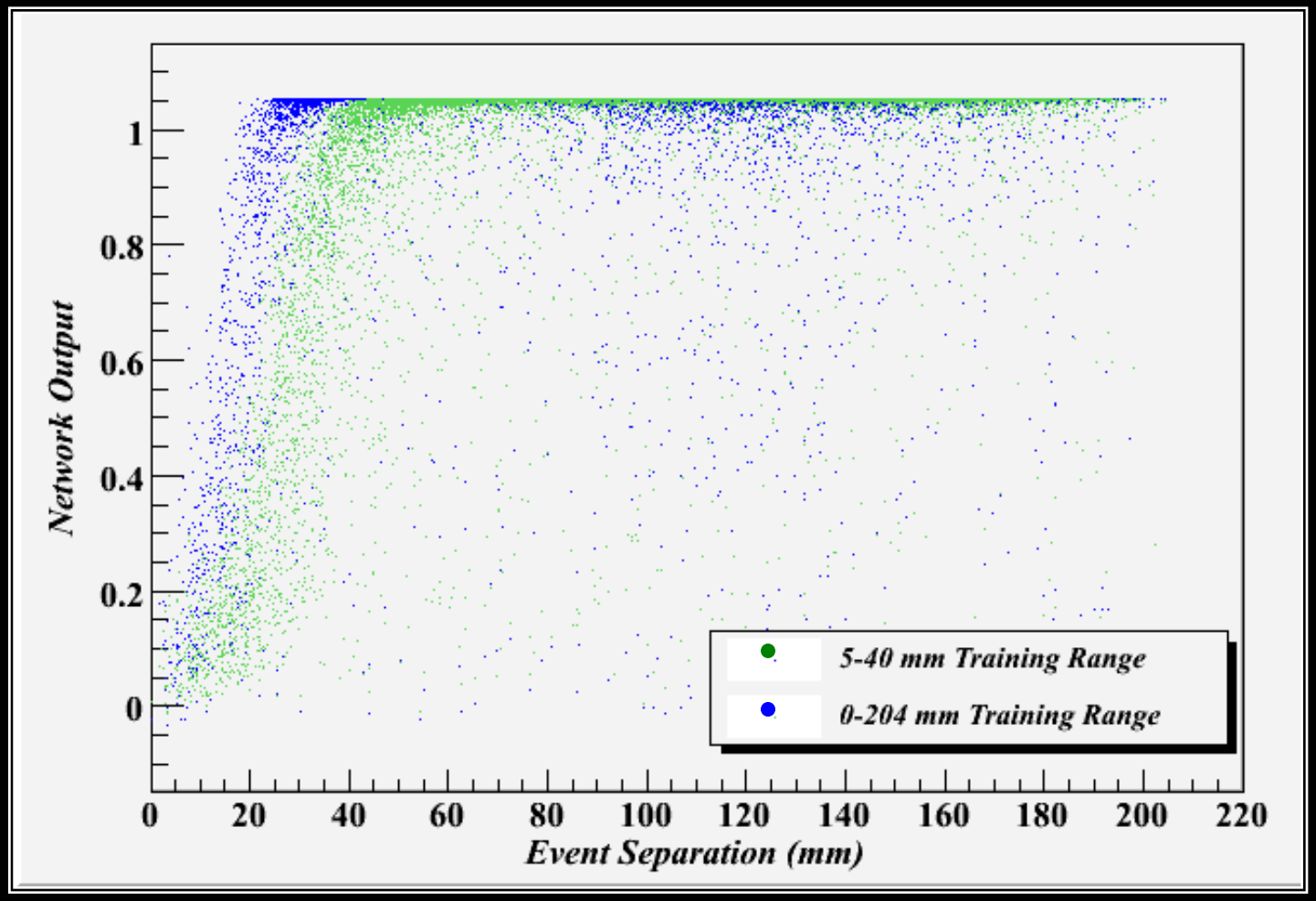






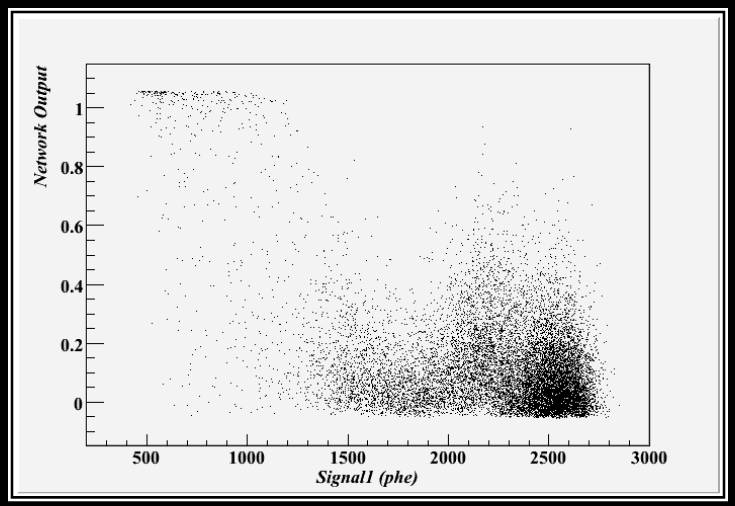
The Double Hit Neural Network

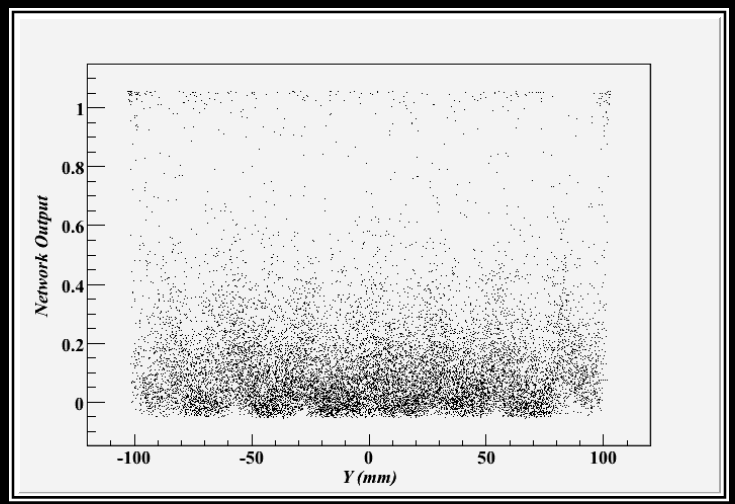
- Updated training set only used events separated by 5-40 mm

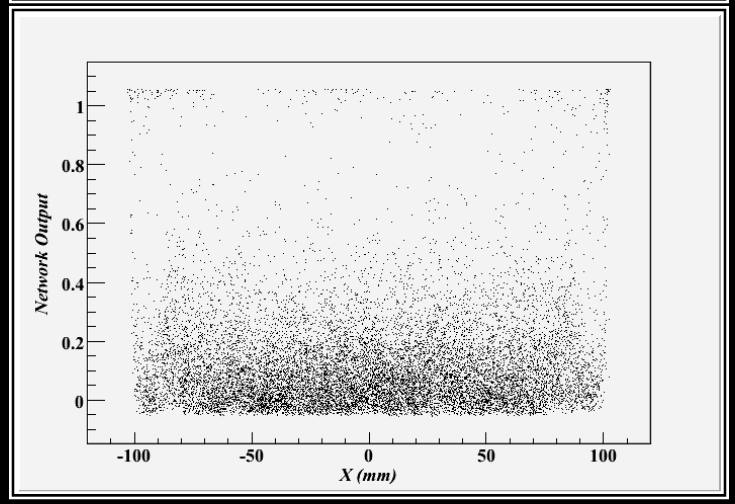


The Double Hit Neural Network

- Updated training set only used events separated by 5-40 mm

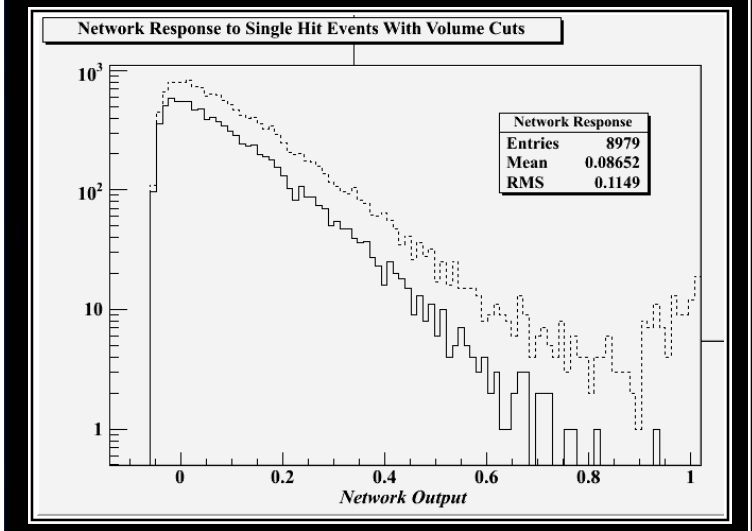


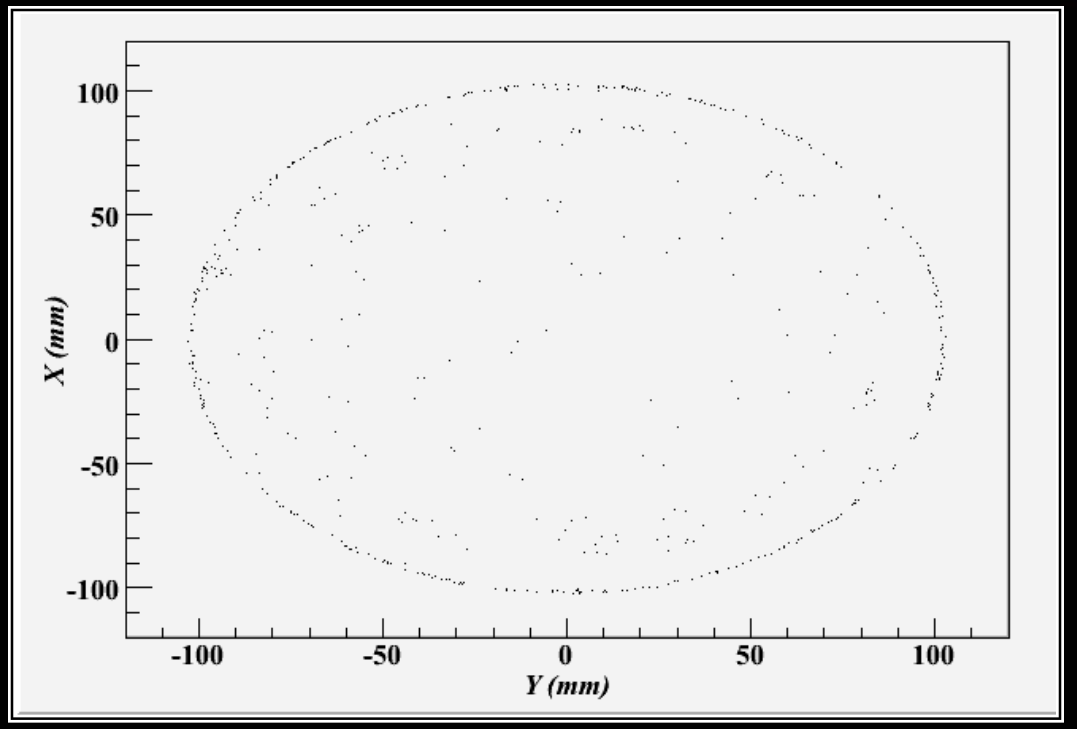




The Double Hit Neural Network

- Updated training set only used events separated by 5-40 mm

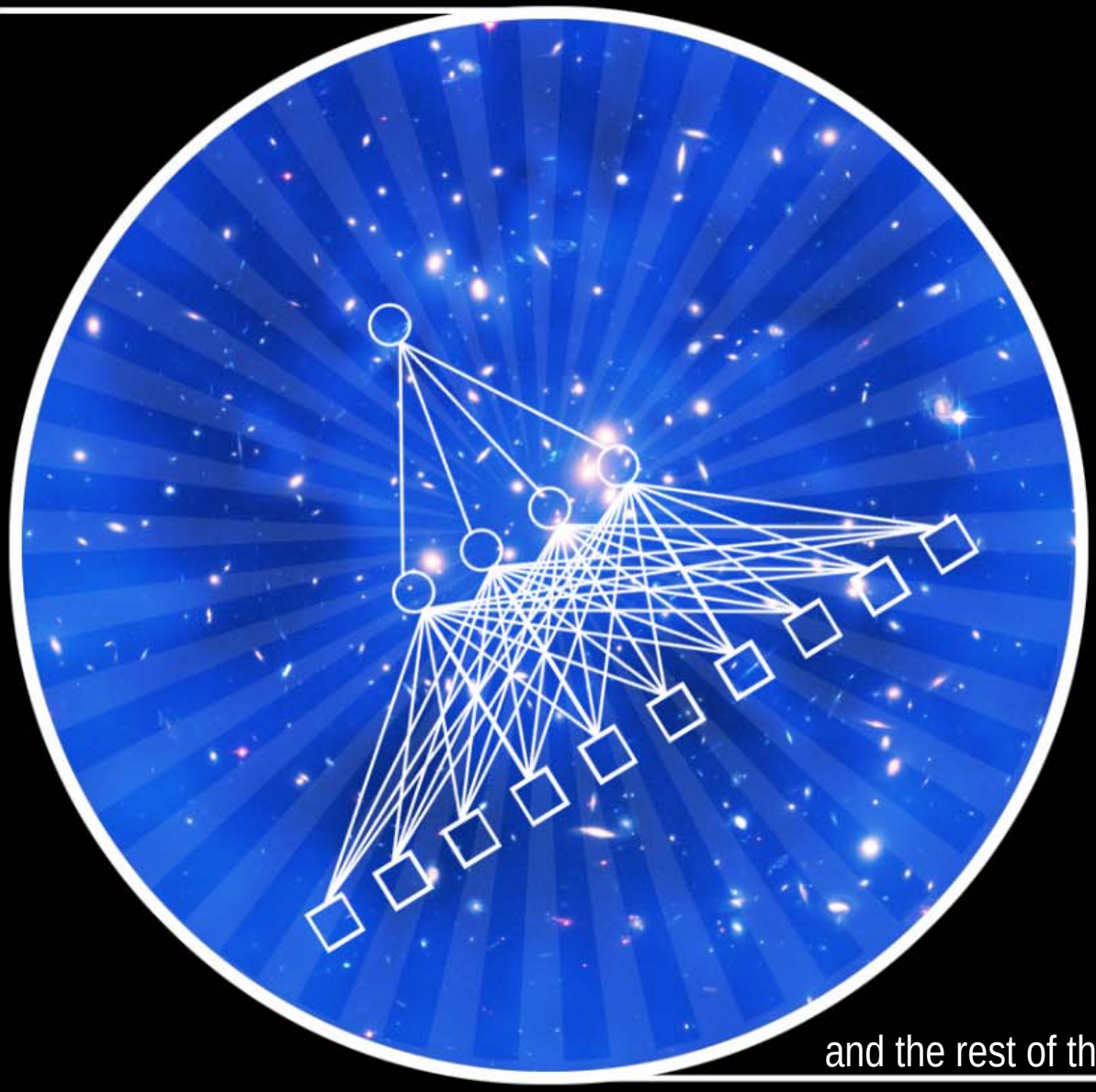




Conclusions

- Evidently the each training set provides statistically equivalent fits
  - The network trained on a smaller dataset is better at reconstructing closer events
  - The network trained on the full set less frequently interprets single hit events as double hits.
- One or both can be employed to improve background reduction
- Future of the Double Hit Network
  - Preprocessing script to change the data into a linearly separable problem or reduce the necessary number of neurons
    - Bright-space
    - Grouping PMT's
  - Reconstructing event locations of double hits

THE END



Special thanks to:
John Parsons
Elena Aprile
Roberto Santorelli
and the rest of the XENON Collaboration

THANK YOU