DOUBLE PEAK DETECTION IN XY THROUGH NEURAL NETWORKS FOR XENON10

Marc Z. Miskin
August 3, 2007

ABSTRACT:
To reduce the background stemming from multiple scattering events, a neural network (NN) was developed to distinguish single peak light patterns from multiple peak light patterns

1. INTRODUCTION:

The XENON10 experiment utilizes a time projection chamber to search for potential dark matter candidates, such as the WIMP. A 14 kg liquid xenon mass interacts with a variety of energetic particles such that scintillation light is produced by particle collisions. This light, when detected, can be analyzed to describe both the type of collision and its location.

In spite of this sophistication, the XENON10 experiment is susceptible to misinformation from a double scattering event. Should a particle electron recoil in the xenon multiple times, the resulting signal can be misinterpreted as a nuclear recoil event. Remedying this inappropriate detector response would yield more accurate measurements in the XENON10 experiment and an improved dark matter search.

The NN provides an ideal algorithmic solution: NN have continually excelled at pattern recognition and their ease of development makes them particularly desirable. One simply needs to construct an appropriate training set, teach the network to recognize the double hit light patterns, and finally evaluate the efficiency at double peak detection. A well designed, trained network can then be employed to further strengthen experimental accuracy.

2. METHOD:

2.1 Neural Network Training Set

One of the critical elements to developing a successful NN is the training set. Obviously, a network cannot properly identify the desired pattern if it is developed with faulty training data. To be sure that the network truly recognizes single and double peak events in X and Y, careful precautions were taken so that signal brightness and event location were not influential in network output.

The full training set consisted of 3×10^4 events randomly distributed in the X and Y plane. In figure 1, one may see a scatter plot of the events in X and Y. To ensure balanced data and prevent favoritism towards single or double hit response, the data set consisted of 15000 single hit examples and 15000 double hit examples.

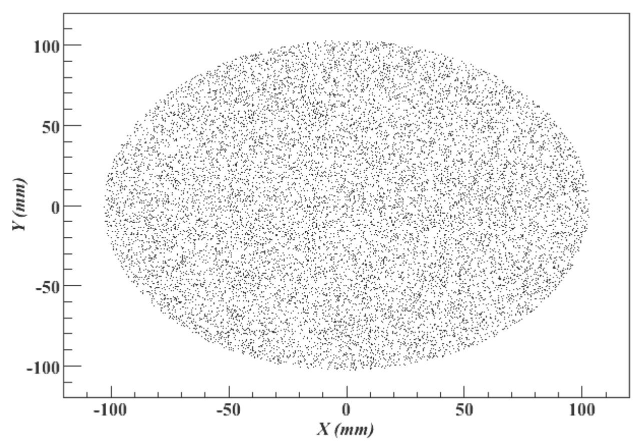


Figure 1. A scatter plot of the X and Y positions of events in the NN training set.

To prevent light intensity from influencing network decision, the initial and single hit events were created such that the collisions would result in imparting approximately 2500 photoelectrons (phe) to the photomultiplier tubes (PMTs) of the XENON10 detector, whereas secondary hits were allowed to span in magnitude from 0 to 5000 imparted phe. This provided a rich range of ratios between first and second scintillation intensity and reduced the probability that the relationship between brightness of light produced by collisions would influence the decision making process. In figure 2, one may see a scatter plot of initial and secondary signal intensities.

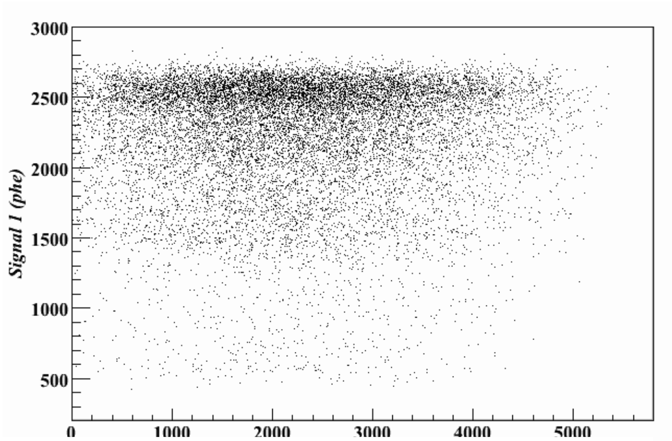


Figure 2. A scatter plot relating the initial and secondary signal intensities in the NN training set.

2.2 Neural Network Structure

Because the effects of single and double hits would be most pronounced in PMT data, all 47 top PMTs were chosen as inputs to the network. Specifically, each input takes the ratio of the total photoelectrons observed at that PMT to the total photoelectrons observed on all top PMTs. This ratio was employed to prevent signal brightness from influencing network decisions.

The remainder of the network is a multilayer perceptron neural network topology with one hidden layer of 47 neurons and a single output neuron. The hidden layers, as well as the input layer, use hyperbolic tangent activation functions to ensure robust modeling. The output neuron uses a sigmoid activation function to classify events between single hits (a response of zero) and double hits (a response of one).

2.3 Neural Network Testing

In order to protect against over learning and to better assess the network's understanding of double hit patterns, a testing set of events, independent of the training set, was developed. This set also consisted of $3x10^4$ events randomly distributed in X and Y. Moreover, the initial collision consistently produced 2500 phe whereas the second varied from 0 to 5000 phe.

2.3 Variations on the Network

Slight variations were made to both network topology and training data in attempts to develop more efficient programs. In an effort to develop better resolution for closer double hit events, one network of identical topology was trained using strictly double hit events separated by 5 mm to 40 mm. Additionally, in an attempt to reduce network complexity, another network was developed without a hidden layer. This network would not train to a sufficient accuracy and henceforth is not discussed in the remaining sections. Presumably, the problem of double peak pattern recognition is too complex to be solved by such a simplistic design.

3. DATA AND ANALYSIS:

3.1 Performance Evaluation

The success of a neural network can be evaluated by an array of statistical tools; the most straightforward of which are the mean square error (MSE), Akaike information criterion (AIC), minimum description length (MDL), and the correlation coefficient.

Given that both the network response and the desired response are known, deviation can be assessed in the fashion of mean squared error (MSE) by defining the MSE as

$$MSE = \frac{\sum_{i=0}^{N_{trial}} (d_i - y_i)^2}{N_{trial}} \qquad [1]$$

where N_{trial} is the number of events in a given trial, d_i is the output of the network for a given event, and y_i is the desired response. In principle, this value, analogous to χ^2, allows one to test how well a network output matches the desired data.

However, the MSE test has a shortcoming in that it provides no assessment of how network complexity influences accuracy. A more complex network can always provide a more accurate response at the cost of processing efficiency and at the risk of over modeling. The AIC and MDL analyze the network error in light of network complexity. The respective formulas are:

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

and

$$MDL = N_{training} \ln(MSE) + 0.5k \ln(N_{training}) \qquad [3]$$

where MSE is the mean square error, k is the number of network weights, and $N_{training}$ is the number of examples presented in the training set. Objectively, minimizing these values implies heightened efficiency.

The correlation coefficient provides a means of asserting that network response is in fact correlated to desired response. The correlation coefficient is defined as

$$r = \frac{\sum_{i=0}^{N_{trial}} (x_i - \bar{x})(d_i - \bar{d}) \Big/ N_{trial}}{\sqrt{\sum_{i=0}^{N_{trial}} (x_i - \bar{x})^2 \Big/ N_{trial}} \sqrt{\sum_{i=0}^{N_{trial}} (d_i - \bar{d})^2 \Big/ N_{trial}}} \qquad [4]$$

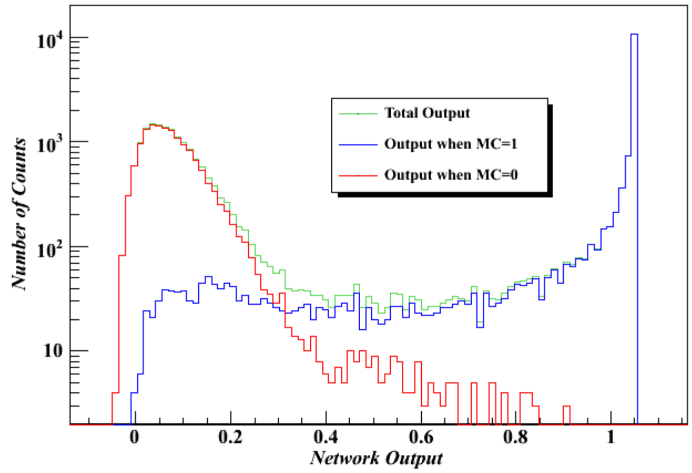
where x_i is the network output for a given trial, N_{trial} is the number of events in a given set, and d_i is the desired network response. The closer a correlation coefficient is to unity, the higher the probability that the network has appropriately modeled the data.

Analyzing the NN trained on the full detector range and the NN trained on the limited range in the scope of these criterion shows nearly equivalent network performance. The exact values may be seen in table 1.

	Full Training Set NN	Limited Training Set NN
MSE	0.028027529	0.02998538
AIC	-1026929.041	-100603.3557
MDL	-95361.128462	-93335.44232
Cor. Coef.	0.929623	0.93275

Table 1. MSE, AIC, MDL, and Correlation Coefficient for the NN trained on the limited range and the NN trained on the full range

Finally, for a more direct understanding of error in network response, figure 3 displays each network's output under the conditions of various expected results.



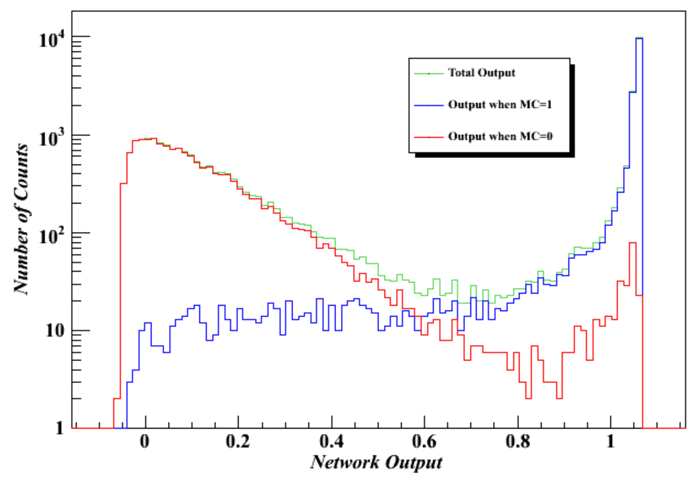


Figure 3. Histograms displaying network response for different expected outputs from the full training set NN (top) and the limited training set NN (bottom).

3.2 Network Output Related to Physical Parameters

One can note from figure 3 that the network trained on the full set has superior single peak resolution whereas the network trained on the limited set has superior double peak resolution. Careful analysis reveals the causes for these mistakes are related to the physical conditions of the events.

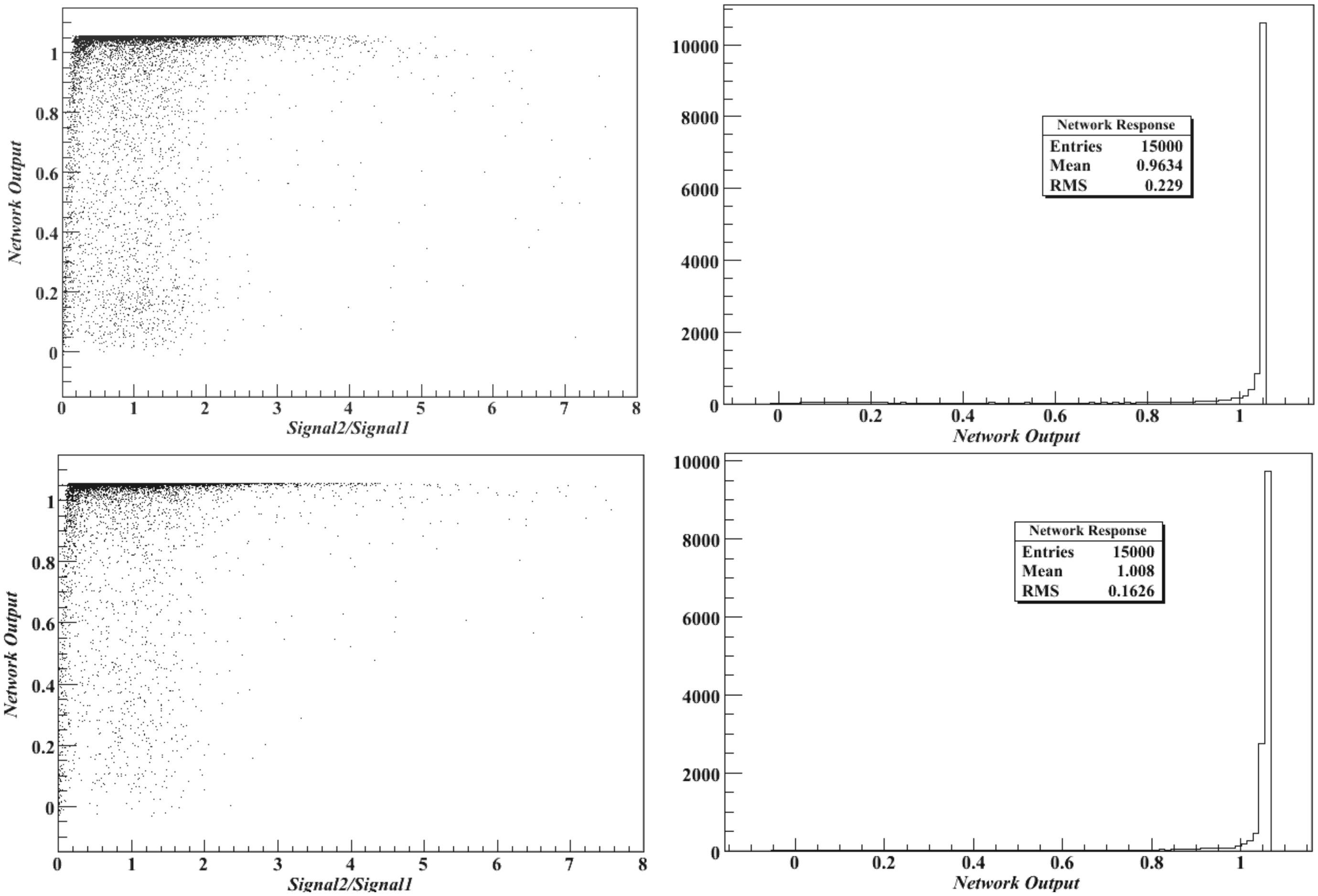


Figure 4: A scatter plots relating signal intensity ratio and network output for the NN trained on the full set (top left) and the NN trained on the limited set (bottom left) and histograms of output from the network trained over the full range (top right) and the limited range (bottom right) when the intensity ratio is greater than zero

The primary event properties which could influence double hit recognition are event separation and the relative intensity between the first and second hit. However, a well trained network should have its light pattern recognition ability independent of the relative intensity. In figure 4, one can note scatter plots comparing the ratio of signal intensities and network response as well as histograms of network response when the number of secondary phe divided by the number of primary phe is greater than zero. One can verify a majority of the signal is correctly recognized as double hits. Moreover, the scatter plots show variation in the ratio of phe yields little effect on network output for either NN suggesting both have been well trained.

Unlike the ratio of initial to secondary phe, event separation should produce a discernable effect on both networks. As event separation goes towards zero, the qualities of the light pattern observed by the PMTs should bear more resemblance to a single hit. Figure 5 provides a scatter plot comparing event separation and network response for both network designs. One should note that while both networks generally perform well, each possesses a particular breakdown region. For the network trained over the full span of the detector, events separated by less than 40 mm are difficult to distinguish whereas the NN trained over the set limited from 5-40 mm only struggles to identify double hits when the two collisions are 20 mm or closer. Histograms are also included in figure 6 for verification.

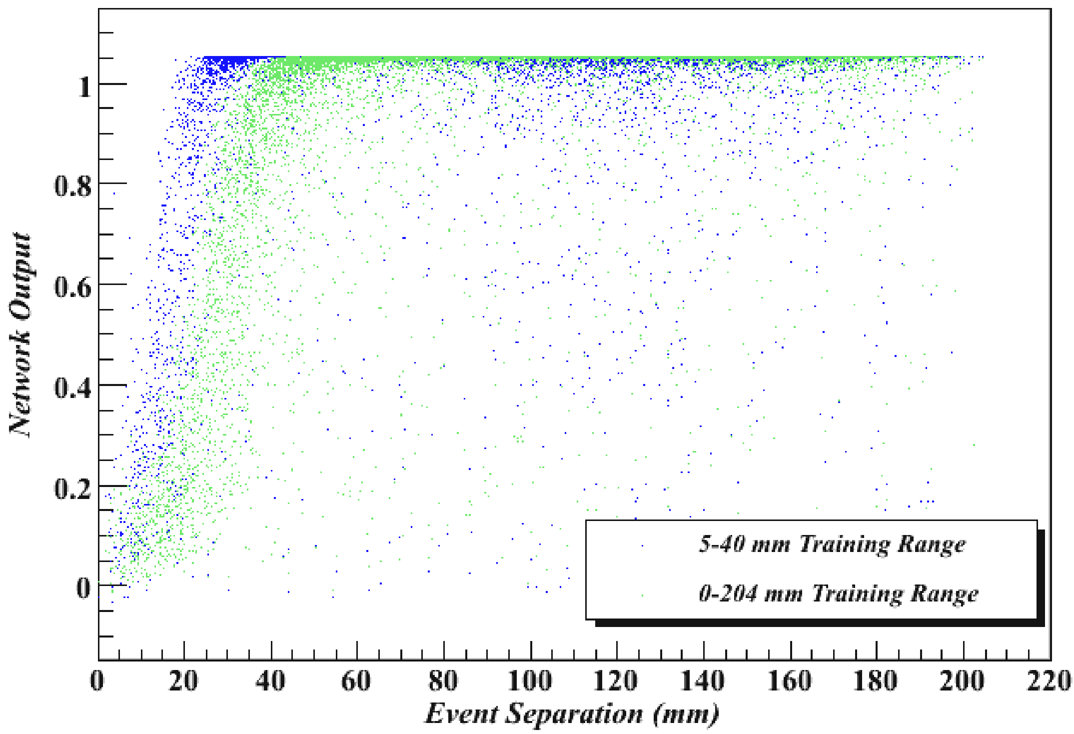


Figure 5. Network output vs. event separation for the full training set (green) and the limited training set (blue)

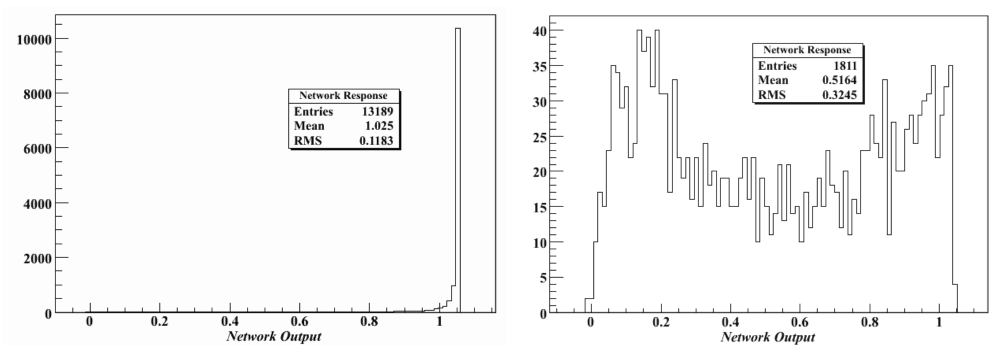
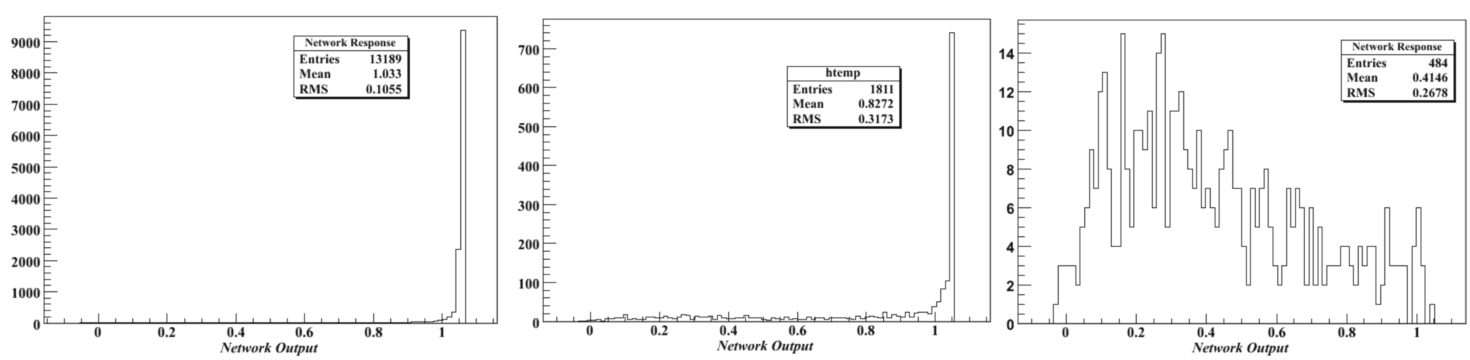
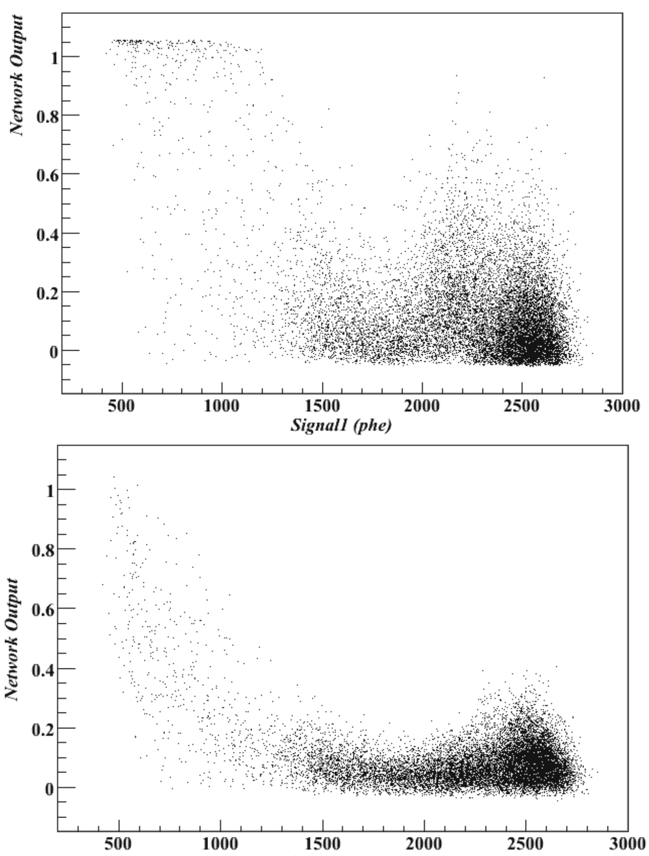


Figure 6. Full set network response for event separation greater than 40 (top left) and less than 40 (top right). Limited set network response for event separation greater than 40 (bottom left), less than 40 (bottom middle), and less than 20 (bottom right)



In contrast to double hit detection, single hit event recognition should not be influenced by brightness or event location. Figure 7 seem to assert that, in both networks, X, Y, and imparted phe have little effect on determining network output individually. However, one can note that should the single hit events incorrectly interpreted as double hit events (that is, a network response greater than 0.5) be plotted in the XY plane, a majority of them fall on the outer edge of the detector. One can see such plots in figure 8. This is fortuitous as such indecisions will be removed later by the fiducial volume cuts. In figure 9 one can see the network response to single hits with events outside of the fiducial volume removed and note the decreased error, particularly in the case of the network trained on the full range of events.



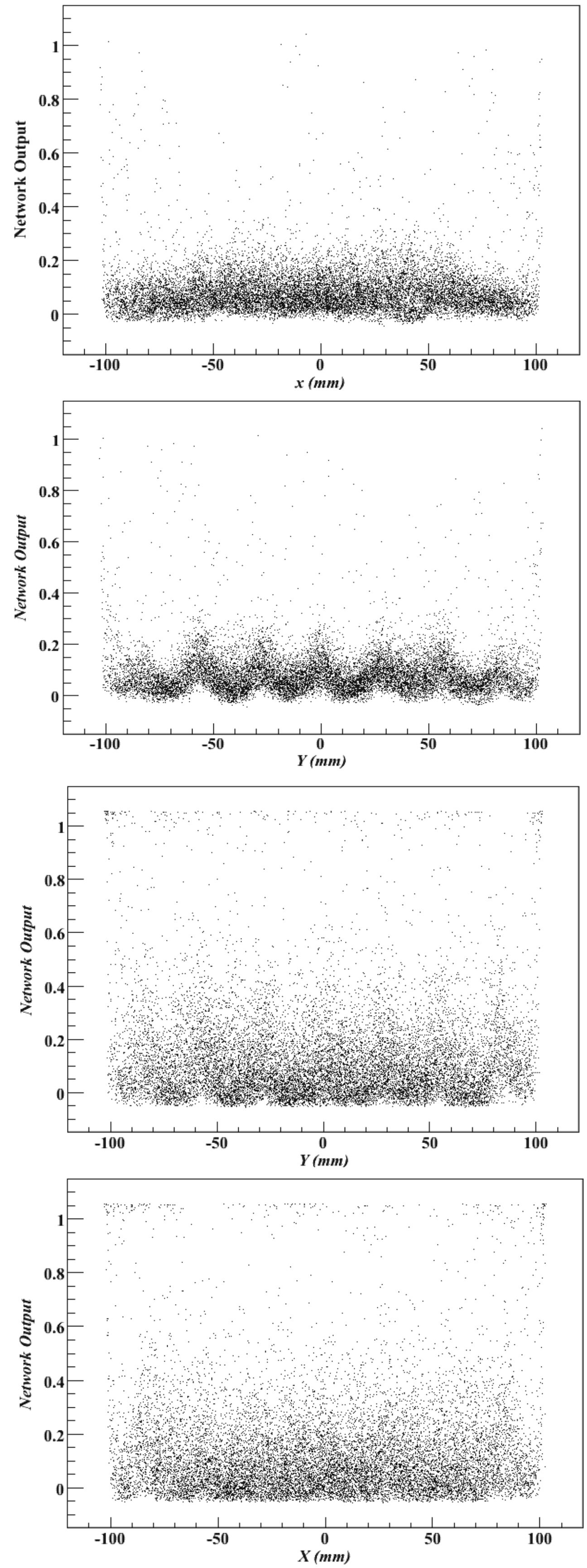


Figure 7. Single hit intensity vs. network output for the full set NN (top left) and limited set NN (bottom left). Single hit position in x vs. network output for the full set NN (top right) and the limited set NN (bottom right). Single hit position in y vs. network output in the full set NN (second down, right) and the limited set NN (third down, right)

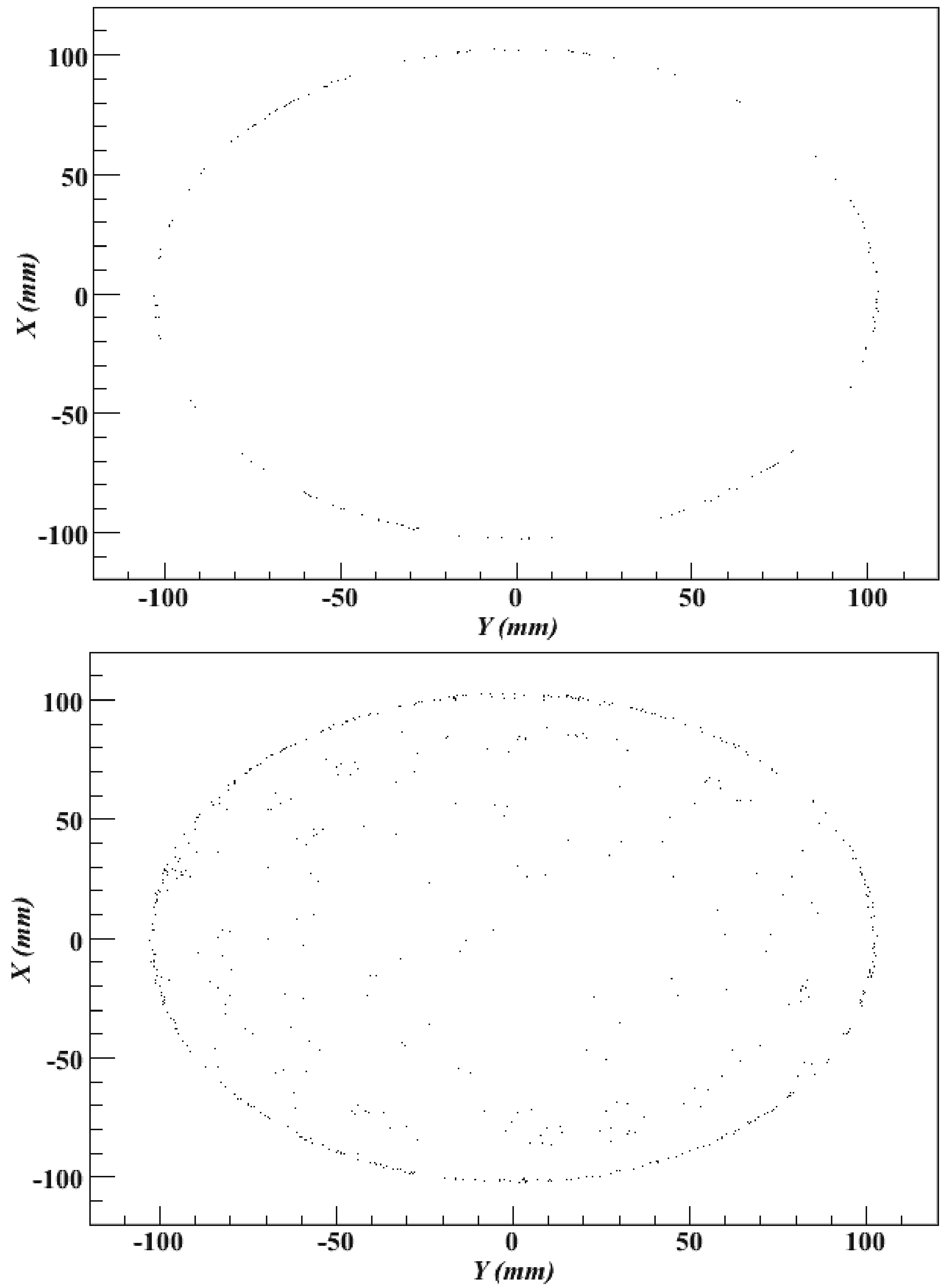
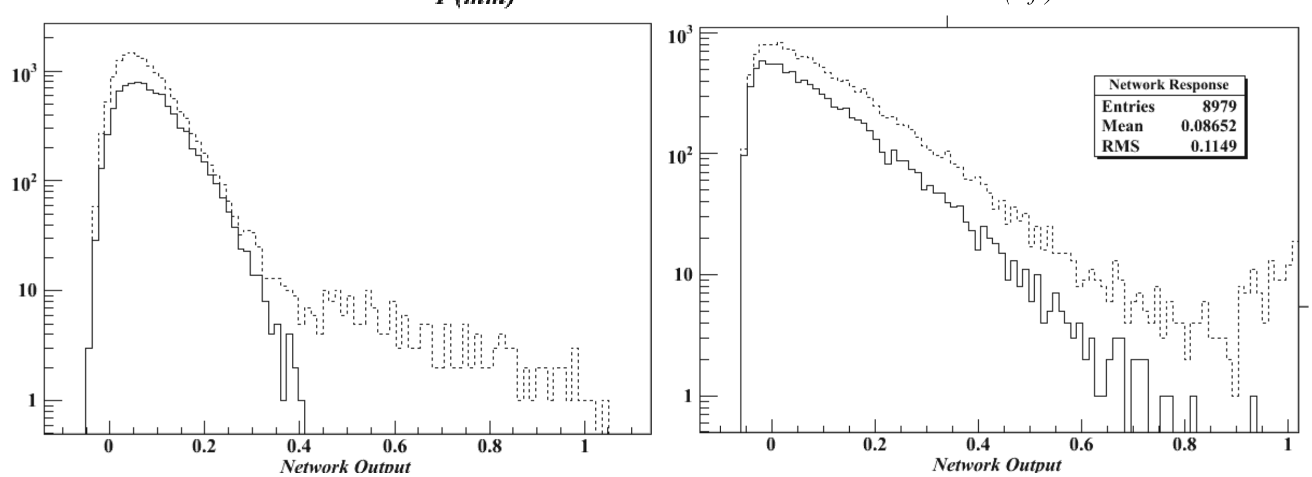


Figure 8. (left) Locations of single hits misinterpreted by the limited set network (bottom) and the full set network (top).

Figure 9. (bellow) Network output for single hit events before fiducial volume cuts (dashed) and after (solid) in the limited training set NN (right) and the full training set NN (left)



4. Conclusions:

While statically the two NN created for double peak recognition are equally effective, assessment in light of physical parameters shows the network trained on the full span dataset excels at recognizing single collision events while the network trained on the limited set is more effective at recognizing double peak events. It should be noted that the mistaken single hits of the limited training set NN will be reduced by fiducial volume cuts whereas the mistakes of the full set network will predominately be unchanged.

Ultimately, the nature of implementation for these networks is left open: one network or both may be employed to reduce background in the XENON10 experiment.

Future improvements for the multiple peak detection NN might include a manipulation of the input data into a less complex pattern space. A preprocessing script could be employed to transform the signal into limited terms of brightness or to group the PMTs. This would reduce the number of inputs and could greatly improve the algorithm efficiency. Thus far, several attempts have been made to make such a reduction, but all of which have yielded little success. Hence, while the possibility to simplify remains open, future attempts should be made cautiously.

5. Acknowledgements:

I have a great amount of appreciation for Dr. John Parsons and Dr. Elena Aprile for facilitating this REU program. Moreover, Dr. Roberto Santorelli helped this project significantly with his guidance on NN and is deserving of credit. Finally, I have a great amount of admiration and gratitude for the entire XENON collaboration for their professionalism, dedication to the advancement of scientific knowledge, and willingness to teach. Certainly everyone involved in the project is worthy of an appreciable amount of thanks.