

# Technical Details on BDT Analysis of Calorimeter Images

Matt Zhang, Wei Wei

## Features

We use boosted decision trees (BDTs) as a proxy for traditional feature-based analysis on calorimeter data. This provides a baseline against which to compare the results of neural net performance. All features used in the BDT are calculated using traditional methods. The complete list of features we use is as follows:

- total energy deposited in ECAL
- total number of hits registered in ECAL
- the ratio of energy deposited in ECAL first layer over energy deposited in second layer
- the ratio of energy deposited in ECAL first layer over all ECAL energy
- 1st through 6th moments in the x, y, and z dimensions for ECAL energy deposits
- all equivalent features for HCAL
- ratio of HCAL to ECAL energy
- ratio of number of hits in HCAL to ECAL
- n-subjettiness measures

We have included several n-subjettiness measures here mostly to help distinguish between (1) photons and (2) neutral pions which have split into two photons with a very small opening angle (below 0.01 radians). Specifically, we use the n-subjettiness algorithm [1] to calculate  $\tau_1$ ,  $\tau_2$ , and  $\tau_3$ , as well as their ratios  $\tau_2/\tau_1$  and  $\tau_3/\tau_2$ . These features measure how well energy depositions are described as belonging to either one, two, or three jets. These jets are calculated with a modified version of the exclusive anti-kT algorithm [2], based on treating individual calorimeter hits as small subjects originating from the interaction point, and subsequently combining them into larger jets. In this modified version of exclusive anti-kT, since we are ignoring the possibility of beam jets, we force the algorithm to continue combination until exactly N jets are left, removing the ability to terminate calculation early. Because n-subjettiness



Figure 1: ECAL for a single event with one, two, and three best anti-kT jets superimposed.

is computationally expensive to calculate, we use it only in photon vs. neutral pion classification, and not in electron vs. charged pion.

In Figure 1 we show the ECAL for a single event with the one, two, and three best anti-kT jets superimposed on top of the calorimeter hits. We see that two jets describe the event very well. In almost all cases, having more reconstructed jets will further improve the fit to data, but the ratio  $\tau_{N+1}/\tau_N$  will be close to one if the additional jet is unimportant. In this case,  $\tau_2/\tau_1$  is small while  $\tau_3/\tau_2$  is close to one.

## Hyperparameter Scan

In order to find the best training parameters, we now perform a hyperparameter scan over three variables for our BDT. First, we look at the maximum depth allowed for each tree, scanning from 2 to 5. Second, we scan over the number of trees, going from 400 to 1000 in steps of 200. Third, we examine the effect of learning rate, looking at rates of 0.01, 0.1, and 0.5. Final test accuracies after training on 360,000 events for a single epoch are shown in Figures 2 and 3. This hyperparameter scan was performed using photon vs. neutral pion samples, but it's assumed that the electron vs. charged pion task has the same hyperparameter performance.

## Photon vs. Neutral Pion

The ROC for the best hyperparameter point at `point4` is shown in Figure 4, and the photon vs. neutral pion discrimination at that point is shown in Figure 5. Finally, the features with highest discriminating power are shown in Figure 6.



Figure 2: BDT test accuracies for a hyperparameter scan over depth and nEstimators. The learning rate is fixed at 0.1.



Figure 3: BDT test accuracies for a hyperparameter scan over learning rate and nEstimators. The maximum depth is fixed at 5.



Figure 4: ROC for BDT trained at best hyperparameter point.



Figure 5: BDT response for electrons vs. charged pions. Test and training curves match up, showing no overtraining.



Figure 6: Discrimination power of BDT features, ranked in order of importance.

### **Electron vs. Charged Pion**

The ROC for the best hyperparameter point is shown in Figure 7, and the electron vs. charged pion discrimination at that point is shown in Figure 8. The features with highest discriminating power are shown in Figure 9. Note that n-subjettiness features are not included in this classification task.



Figure 7: ROC for BDT trained at best hyperparameter point.



Figure 8: BDT response for electrons vs. charged pions. Test and training curves match up, showing no overtraining.



Figure 9: Discrimination power of BDT features, ranked in order of importance.

## References

- [1] <https://arxiv.org/pdf/1011.2268.pdf>
- [2] <https://arxiv.org/pdf/hep-ph/9305266.pdf>