PNN studies update

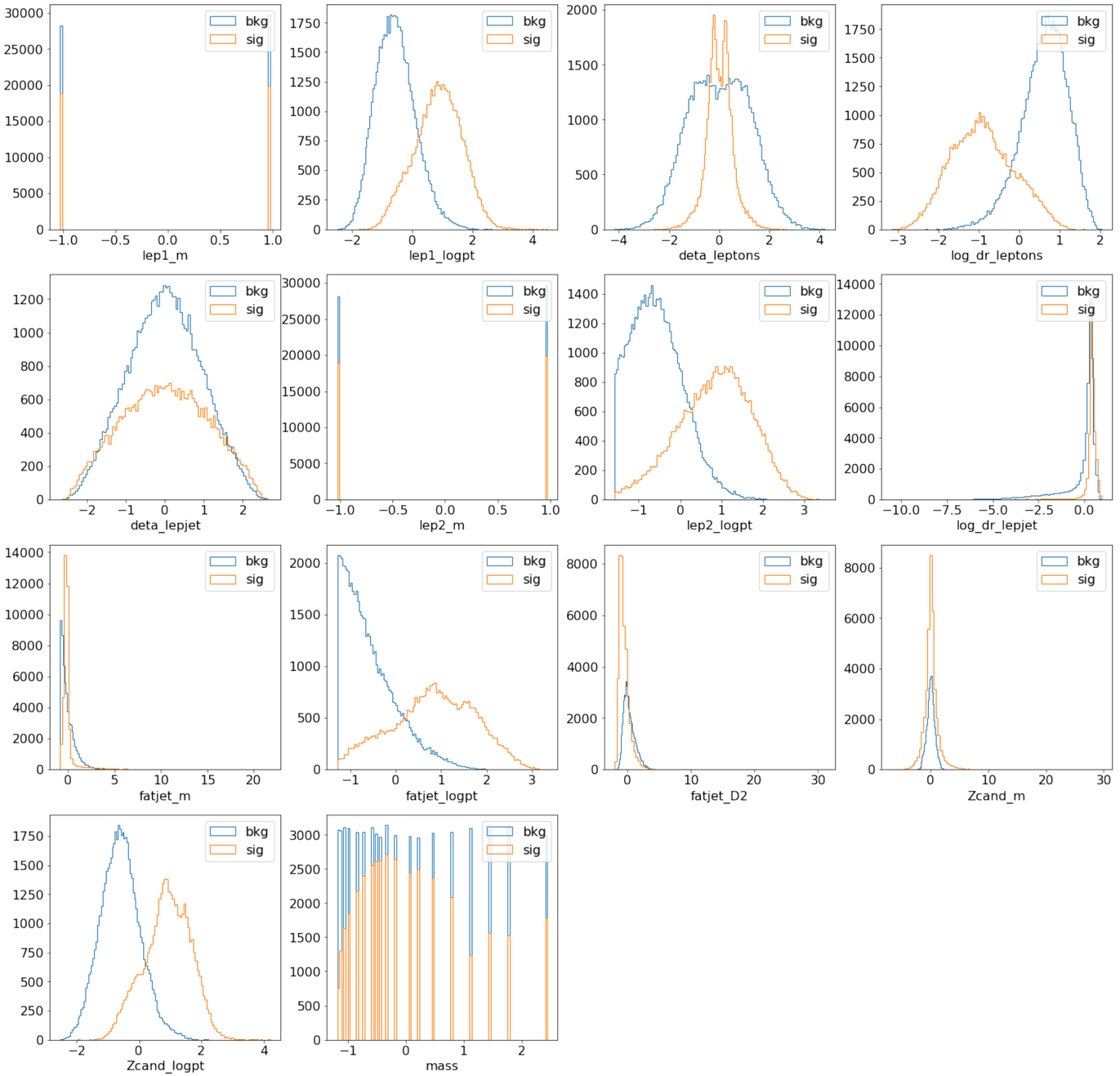
Martino Centonze, 26/04/2022

A (slightly) different approach to PNN

- No negative MC weights in training set —> signal/bkg equalization + signal masses equalization with definite positive weights
- MC weights in test set (as usual)
- New combination of variables in input
- Simplified architecture of the PNN (less layers and nodes)
- Testing the model:
  - Loss performance on test set
  - Optimized significance

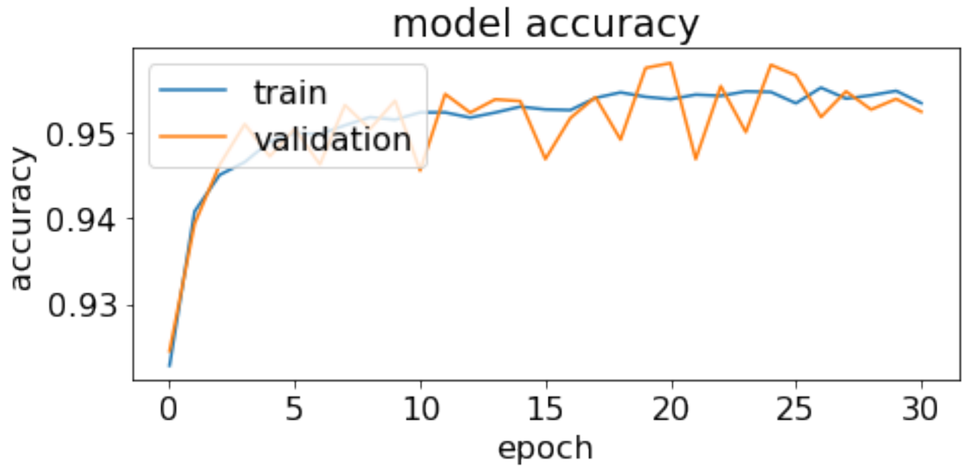
Input variables

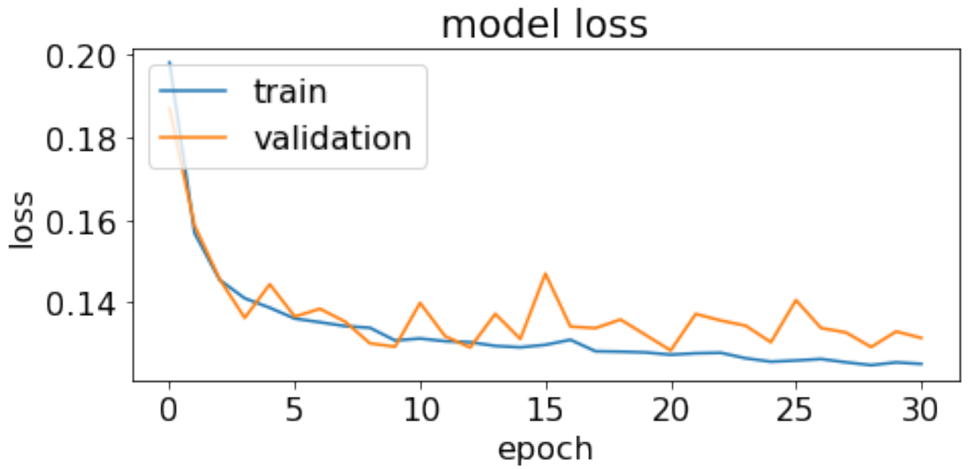
VBF merged Radion



The model - PNN

- 2 hidden layers;
- 32 nodes each;
- Output: softmax;
- Loss: categorical_crossentropy;
- LR: 0.01 with Adam optimizer;
- No dropout.





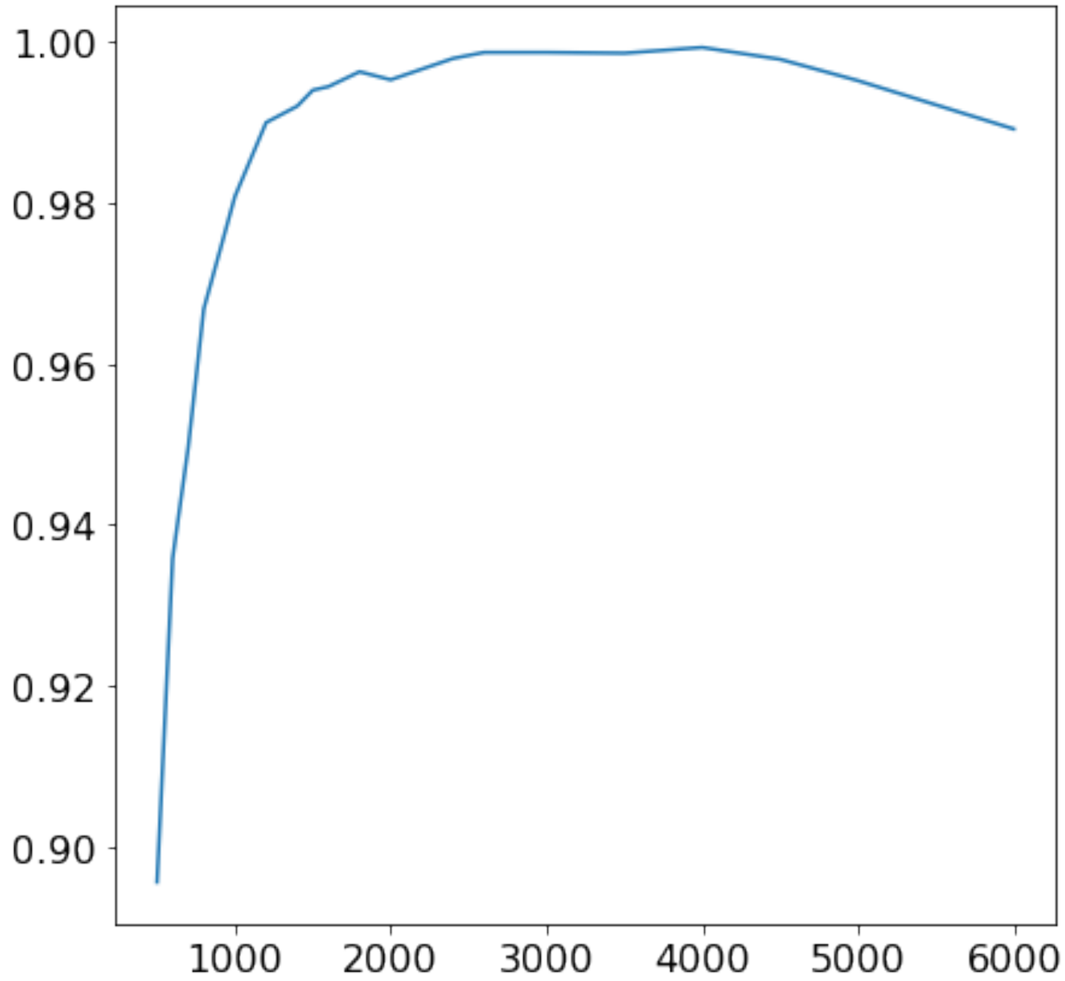
Performance on test set (with MC weights)
Loss: 0.096
Accuracy: 0.957

VBF merged Radion

The model - PNN

- 2 hidden layers;
- 32 nodes each;
- Output: softmax;
- Loss: categorical_crossentropy;
- LR: 0.01 with Adam optimizer;
- No dropout.

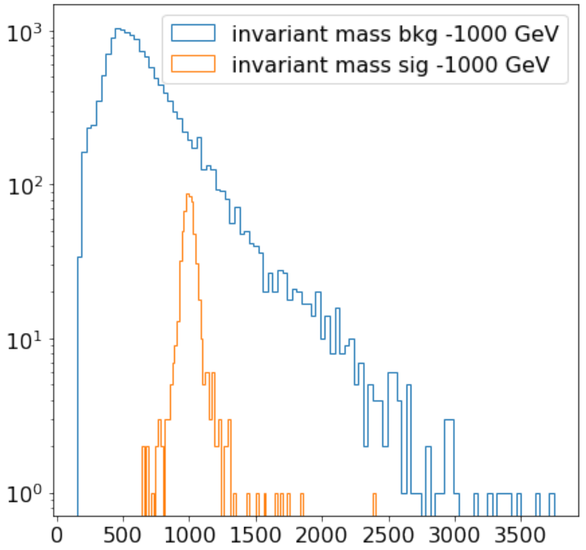
AUC as a function of the resonance mass

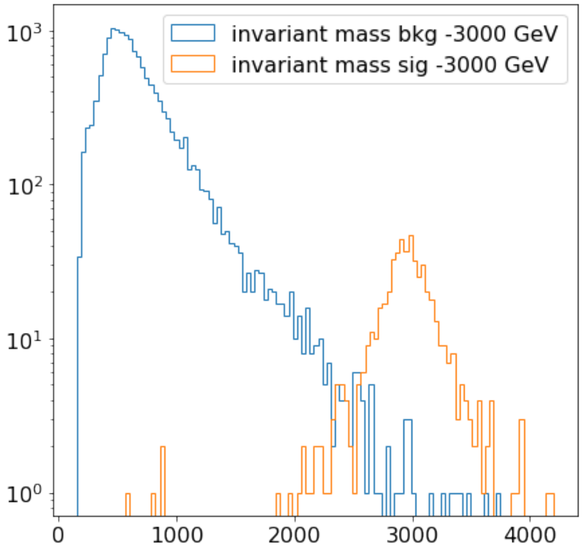


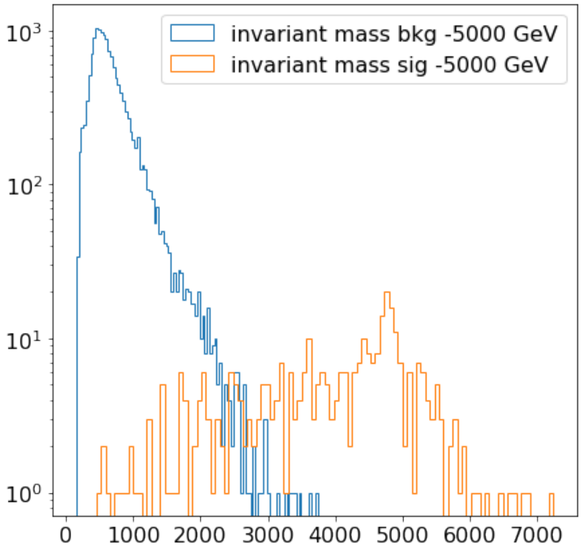
VBF merged Radion

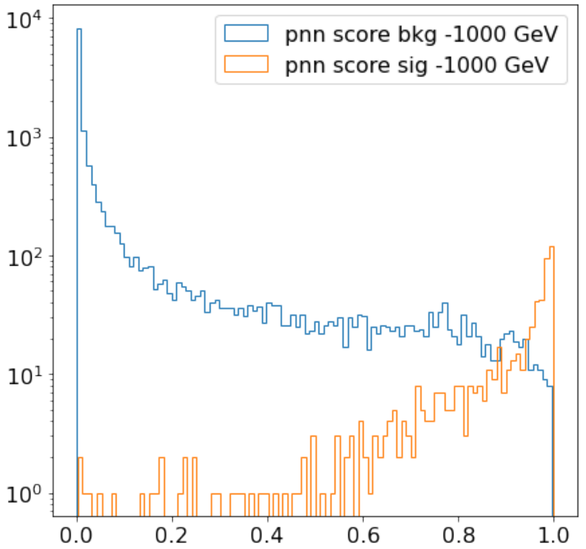
Some distributions of the pnn score - X_boosted_m

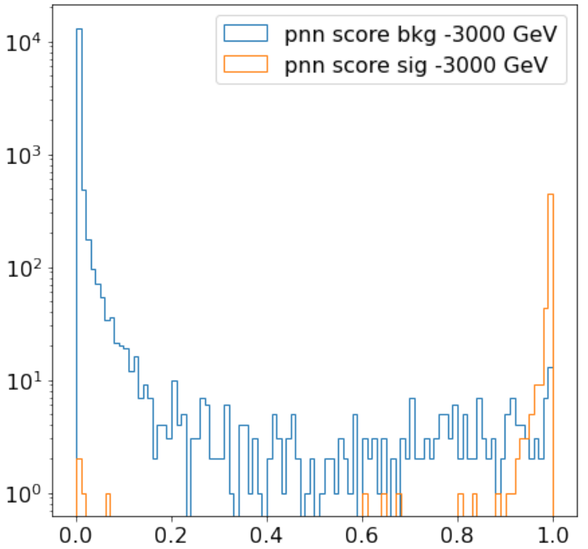
VBF merged Radion

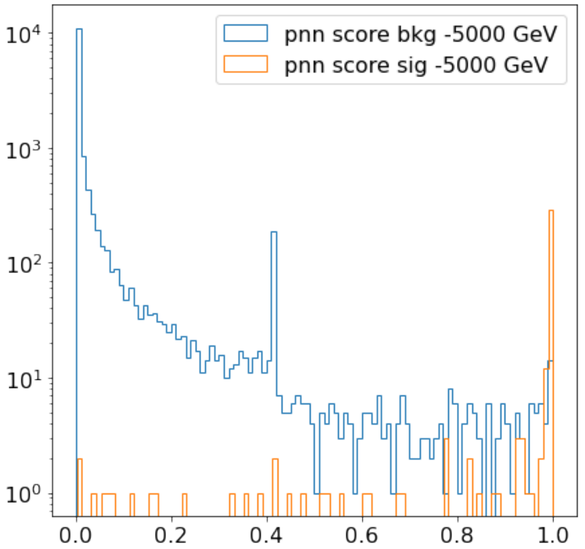








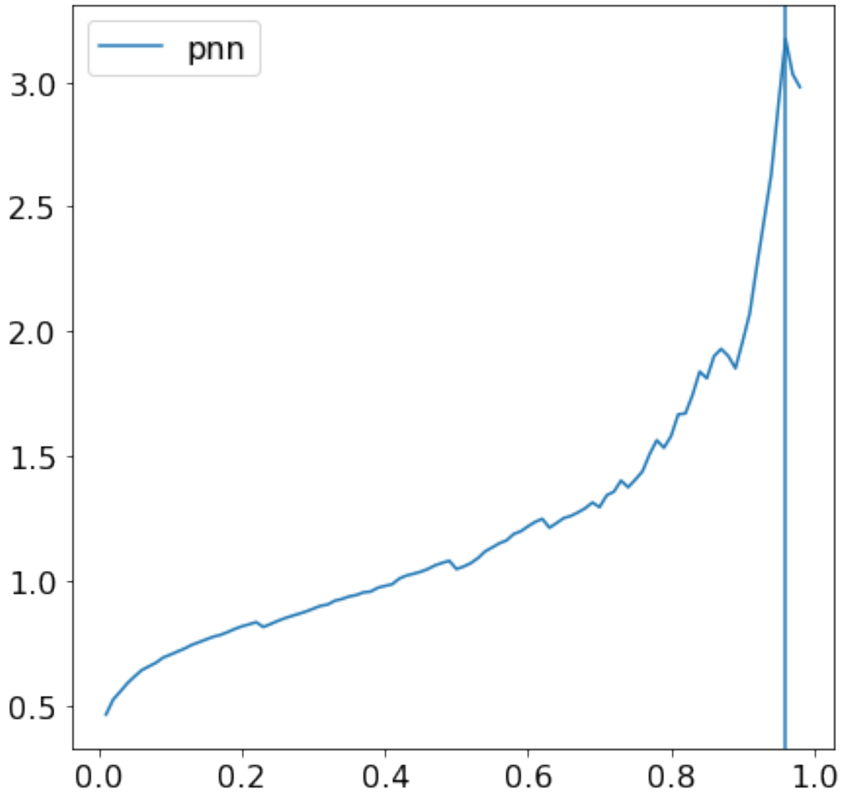


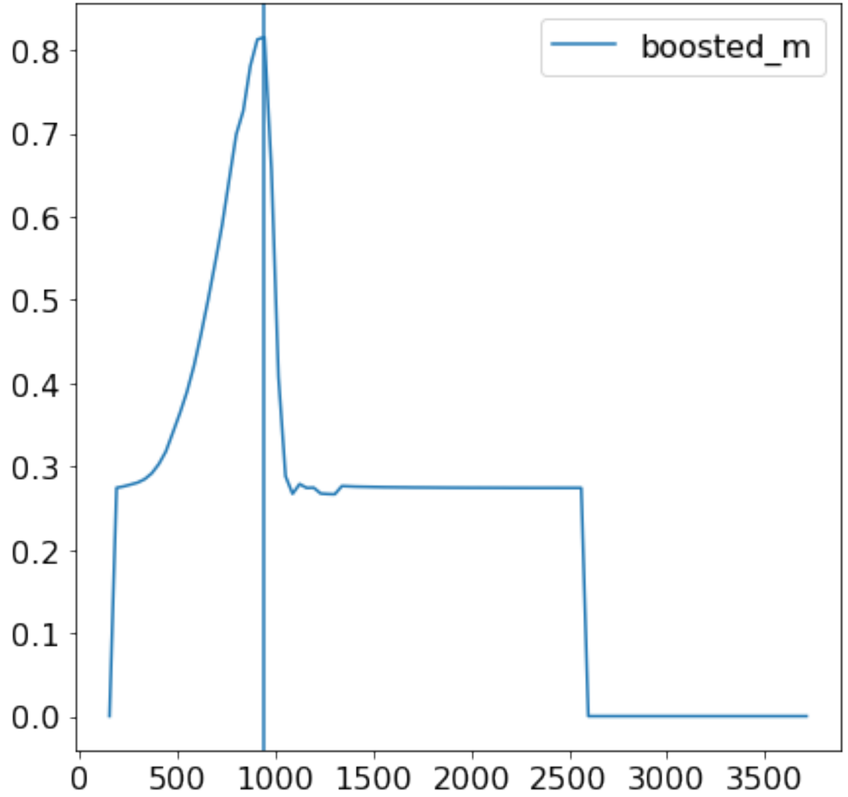


Significance

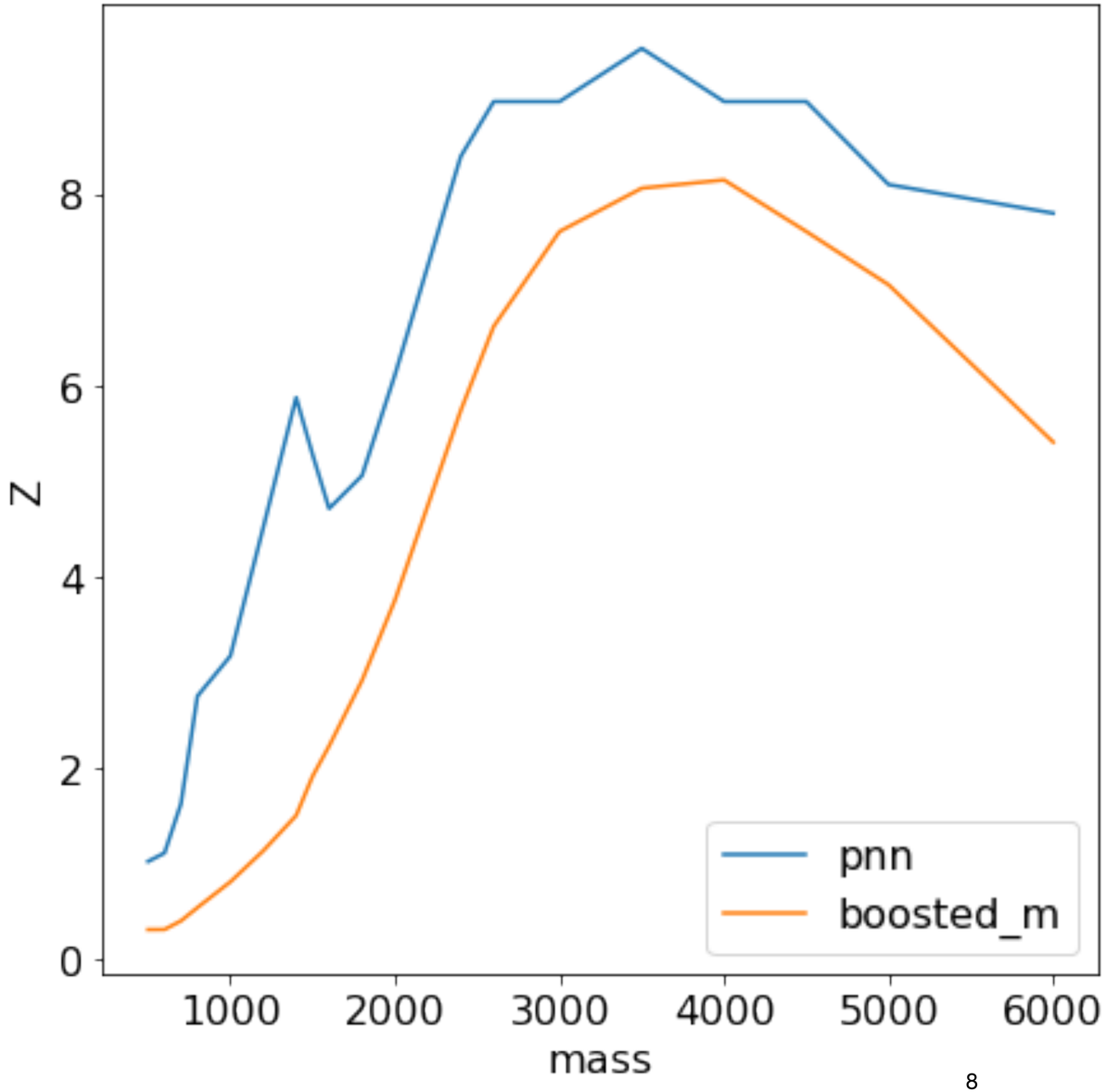
VBF merged Radion

- Signal scaled (significance formula well defined for a b/s > 1e-1)
- Double binning optimization (Luigi's method).

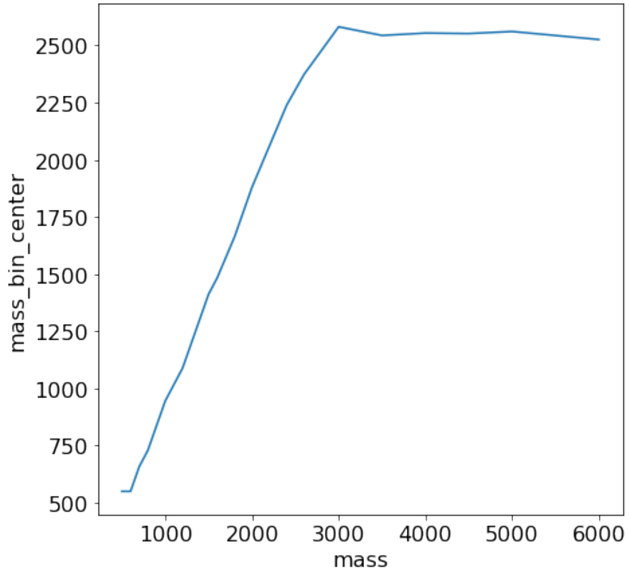


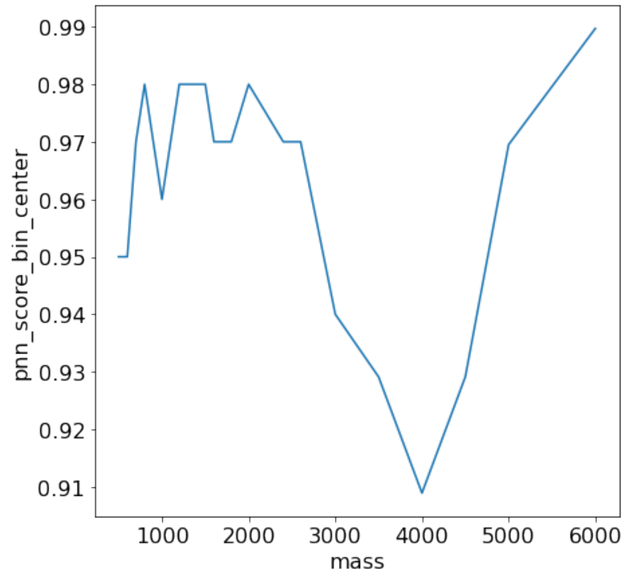


Significance



VBF merged Radion





Conclusions

- To do:
  - Generalization to other signals/channels;
  - Compute significance on selected region (e.g. MergHP ZZ)
- Results:
  - Promising results (higher pnn significance over X_boosted_m)
  - Better performance that previous model;
  - The model can be further optimized (idea: use pca for deriving its correct size)