

Analysis of Higgs boson decay into four leptons at 8 TeV

Matteo Malucchi

September 2022

Abstract

This project contains an analysis of the decay $H \rightarrow ZZ \rightarrow 4l$ using reduced NanoAOD files created from [CMS Open Data](#). The analysis follows loosely the official CMS analysis published in 2012 [[Phys. Lett. B 716 \(2012\) 30](#)]. First, the dataset is "skimmed", i.e. all events which are not of interest for the reconstruction of Higgs bosons are removed, and the various observables necessary for the analysis are computed. The remaining variables are then plotted and the invariant mass of the four leptons is fitted in order to measure the Higgs mass. Subsequently, a Deep Neural Network is trained using as input the simulated signal and background Monte Carlo samples and as discriminant variables the invariant masses of the two reconstructed Z bosons and the five angles formed by the leptons in the final state as described in detail in the article [[Phys.Rev.D86:095031,2012](#)]. Then, the DNN is evaluated on the whole dataset in order to obtain a graph in which the distribution of the DNN Discriminant versus the invariant mass of the four leptons is plotted. This shows a clear separation between signal and background, hence a further cut on the data based on this discriminant can be applied in order to obtain a "cleaner" sample and better discriminate the signal from the background.

The majority of the project is written in Python and uses extensively PYROOT, in particular RDataFrame for the skimming and the TMVA library with Keras for the machine learning portion. In addition, many functions used in this analysis are written in C++ and are invoked by the Python modules thanks to the ROOT interpreter.