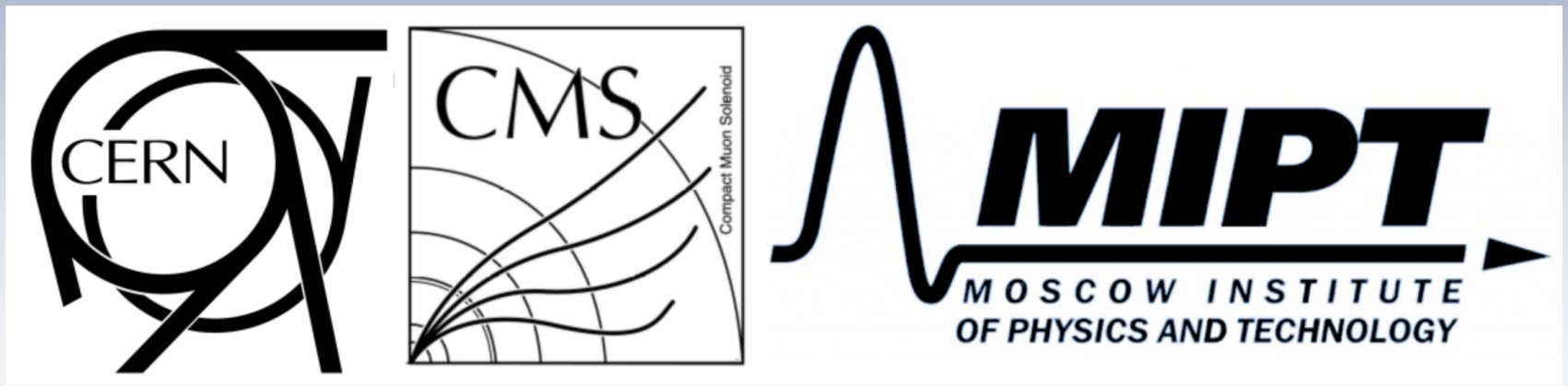
ggH rejection for double Higgs search at LHC

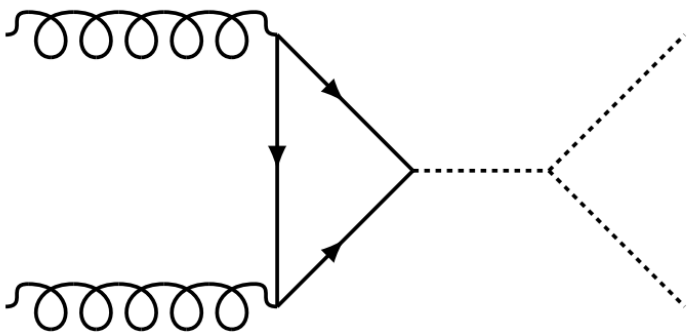
Anna Ershova
supervisors: Maxime Gouzevitch, Thong Nguyen

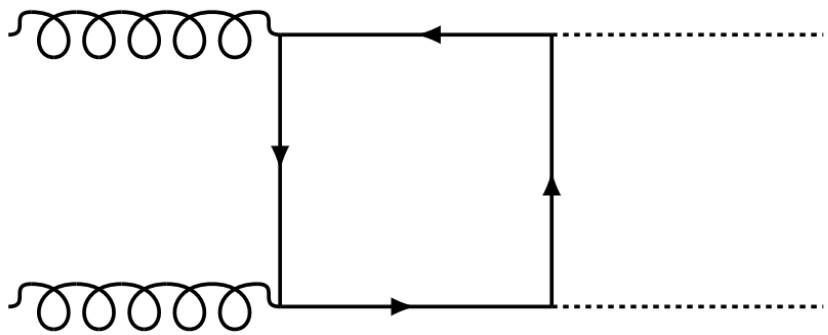
EP-CMG Department

5 Aug 2019

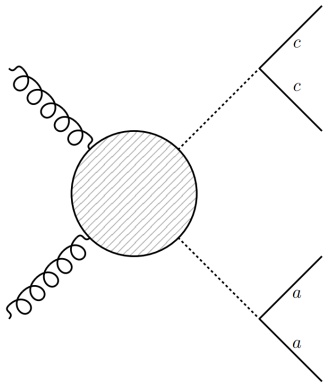


Motivation





HH prodiction



background

- ttH
- ggH

leading HH$\rightarrow \gamma\gamma\bar{\mathrm{b}}\mathrm{b}$ signal

Prerequisites

Strategy

- find variables which are different for signal and background samples
- try TMVA BDT to understand perspective improvements
- try more advanced DNN algorithm
- use bayessian optimization to tune DNN parameters

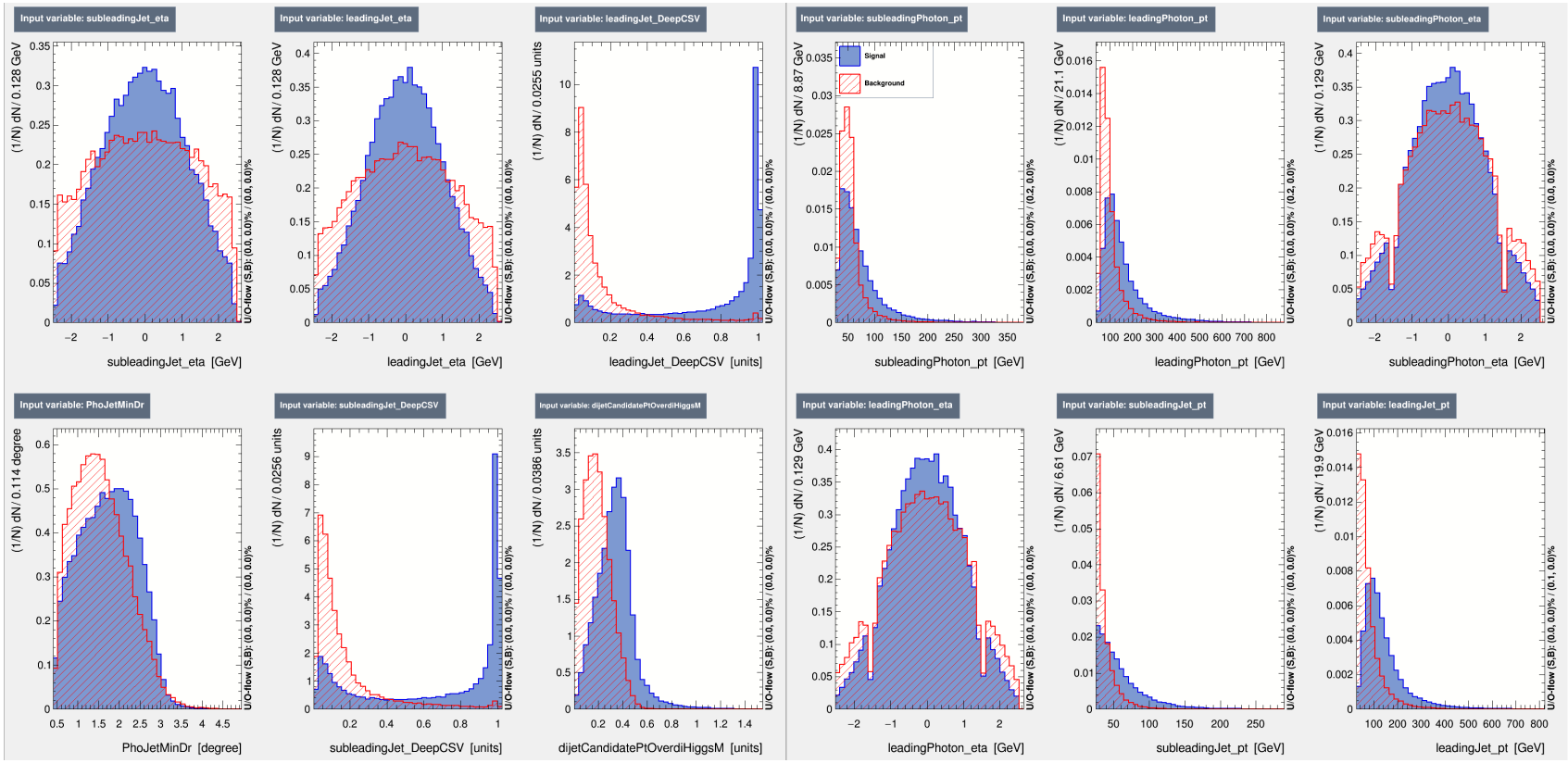
Used tools

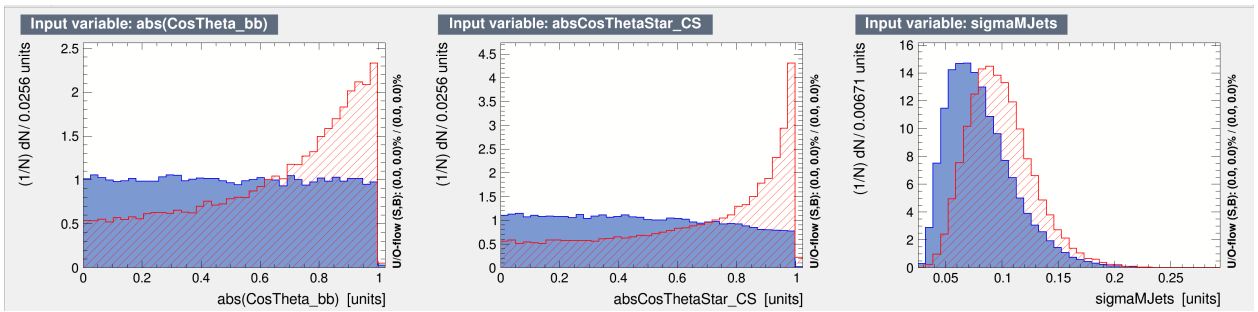
Variables used for training

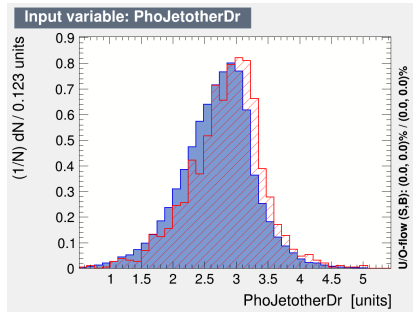
- subleadingPhoton_pt
- leadingPhoton_pt
- subleadingPhoton_eta
- leadingPhoton_eta
- subleadingJet_pt
- leadingJet_pt
- subleadingJet_eta
- leadingJet_eta
- leadingJet_DeepCSV
- PhoJetMinDr
- subleadingJet_DeepCSV
- dijetCandidatePtOverdiHiggsM
- absCosTheta_bb
- HHbbggMVA
- sigmaMJets
- PhoJetotherDr

Where HHbbggMVA is probability of the signal event, PhoJetMinDr – distance between closest jets, PhoJetotherDr – distance between another pair of jets, (sub)leadingJet_DeepCSV – btagging score of that jet

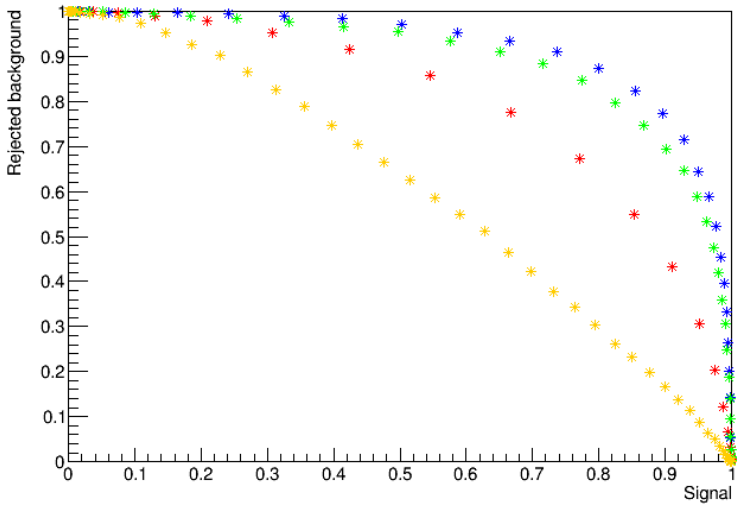
Signal and background variables





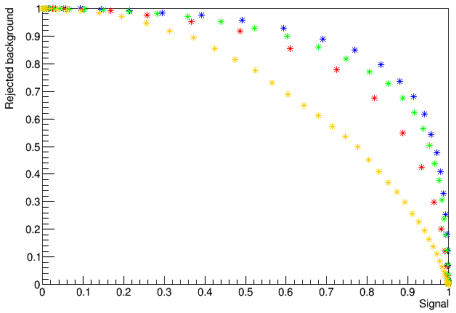


TMVA BDT performance

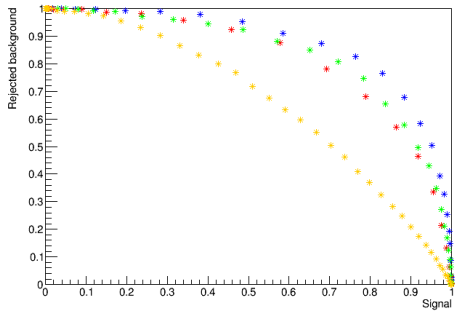


	MVA cut	No MVA cut
Btagging		
No Btagging		

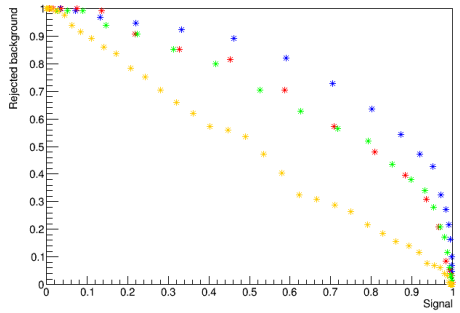
MVA>0.25



0.25<MVA<0.50

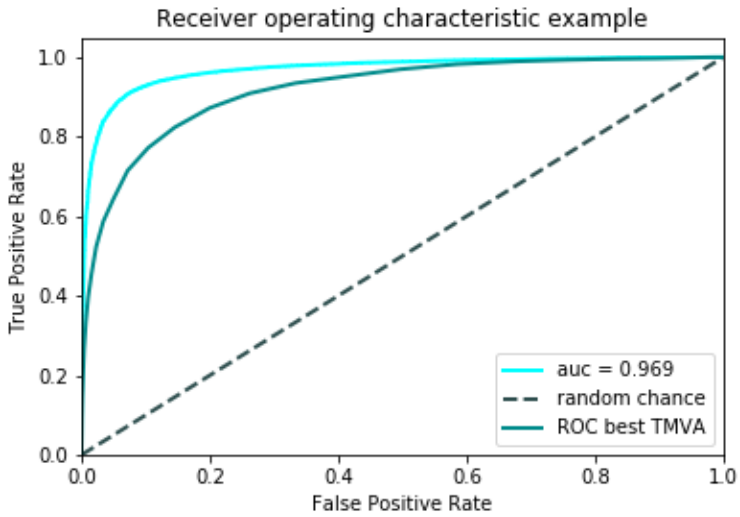


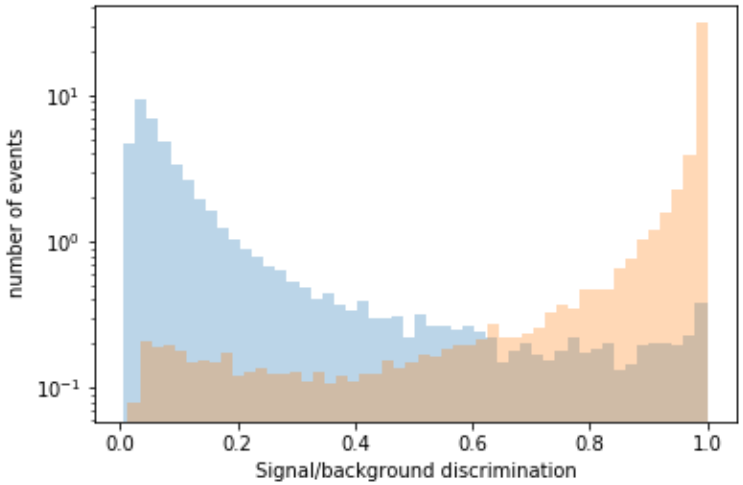
0.50<MVA<0.75



MVA>0.75

DNN perfomance





Neural network performance

Threshold	Signal Efficiency	Background Contamination
0.1320	0.9802	0.3512
0.3566	0.9502	0.1543
0.4743	0.9352	0.1126
0.6567	0.9002	0.0673
0.9559	0.7001	0.0131
0.9926	0.5001	0.0042
0.9984	0.3001	0.0011

Bayessian optimisation results

- best hidden layers = 2
- best initial nodes = 88
- best dropout = 0.730104980246572
- best batch size = 2184
- best learning rate = 1.9989431770225823e-05

Summary

- variables distinguishing signal and background events were selected
- Deep Neuron Network was trained
- Parameters of the DNN were optimized using Bayessian optimisation