Search for high transverse momentum $H \to b\bar{b}$ with CMS

Final presentation

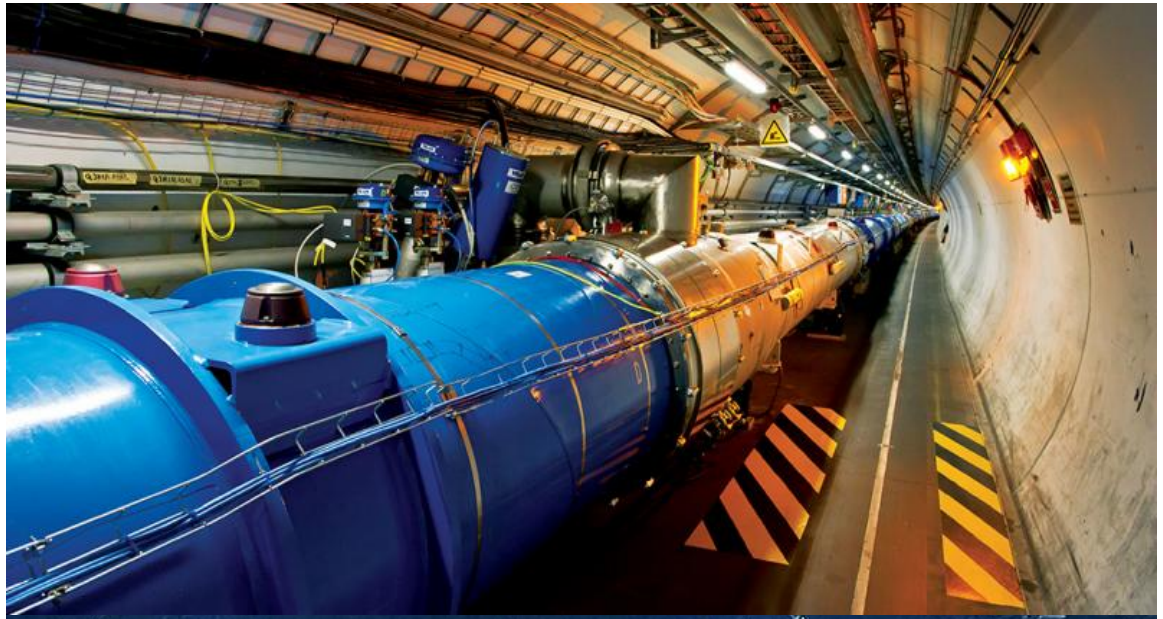
Eleonora Pasino

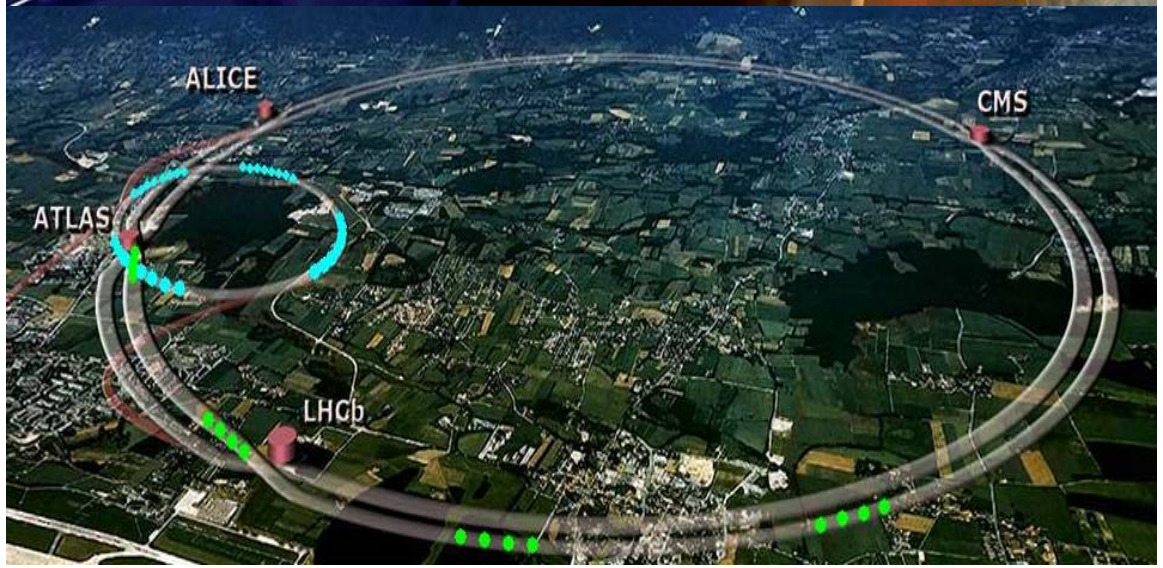
Supervisor: Caterina Vernieri

September 27th, 2017

LHC

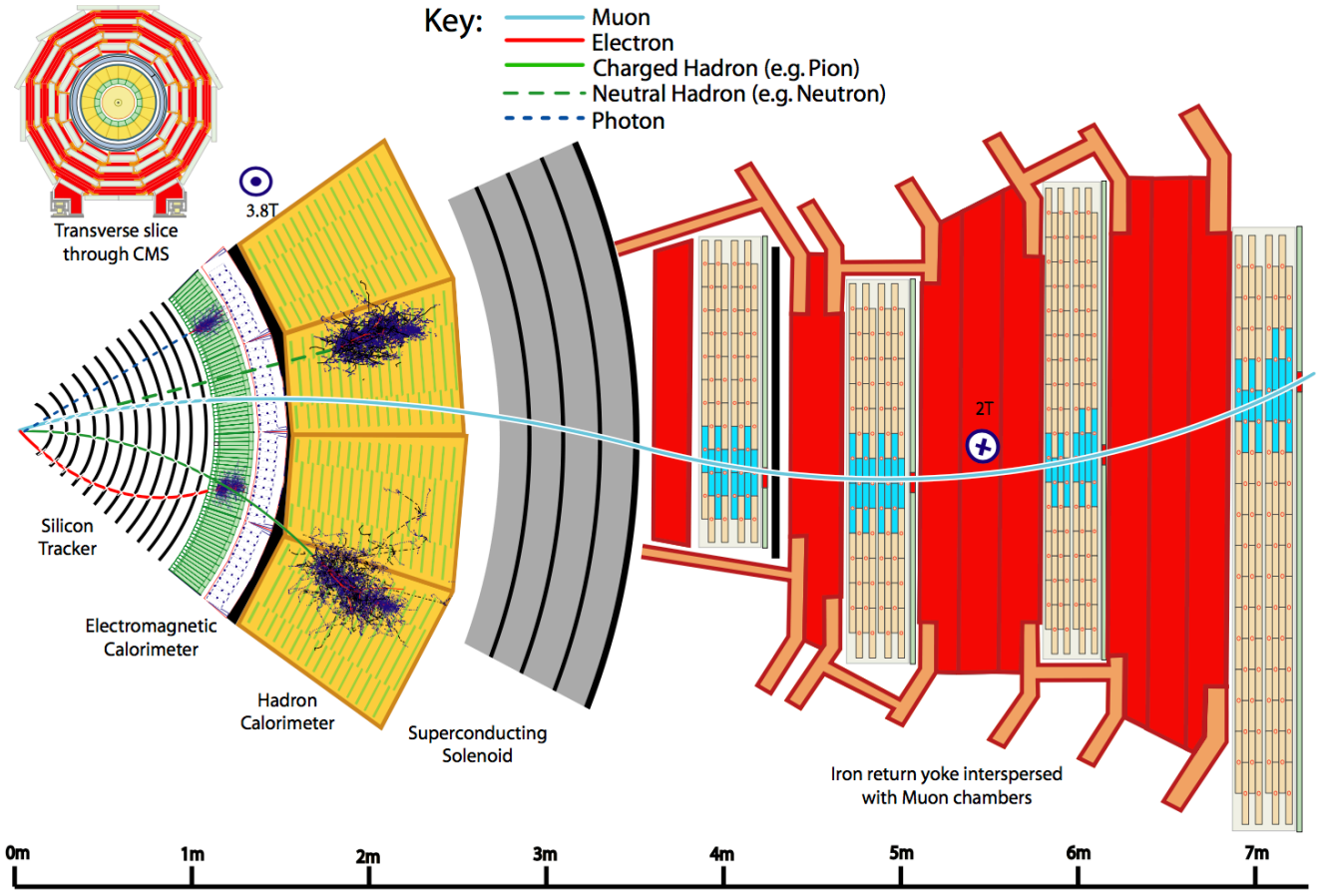
- Large Hadron Collider:
  - Proton-proton collisions
  - 2010-2012: luminosity 25fb^{-1}, Run I, 7/8 TeV
  - 2015-2016: luminosity 40 fb^{-1}, Run II, 13 TeV
- 2017: Run just restarted, 13 TeV
- Four big experiments: ATLAS, CMS, ALICE, LHCb





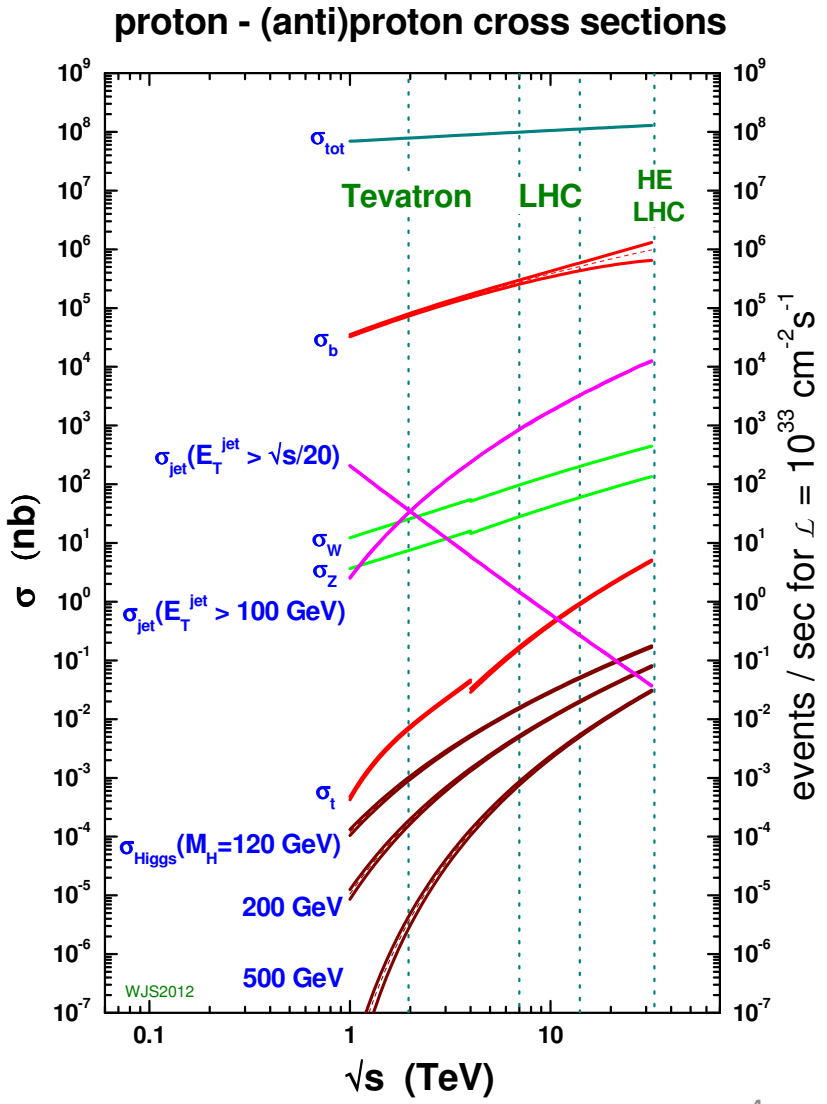
CMS

- General-purpose detector
  - Test Standard Model and search for beyond SM theories
- Huge solenoid magnet that generates a field of 4 tesla
- Several particles detectors arranged in concentric shells: tracker; electromagnetic calorimeter; hadron calorimeter; muon chambers

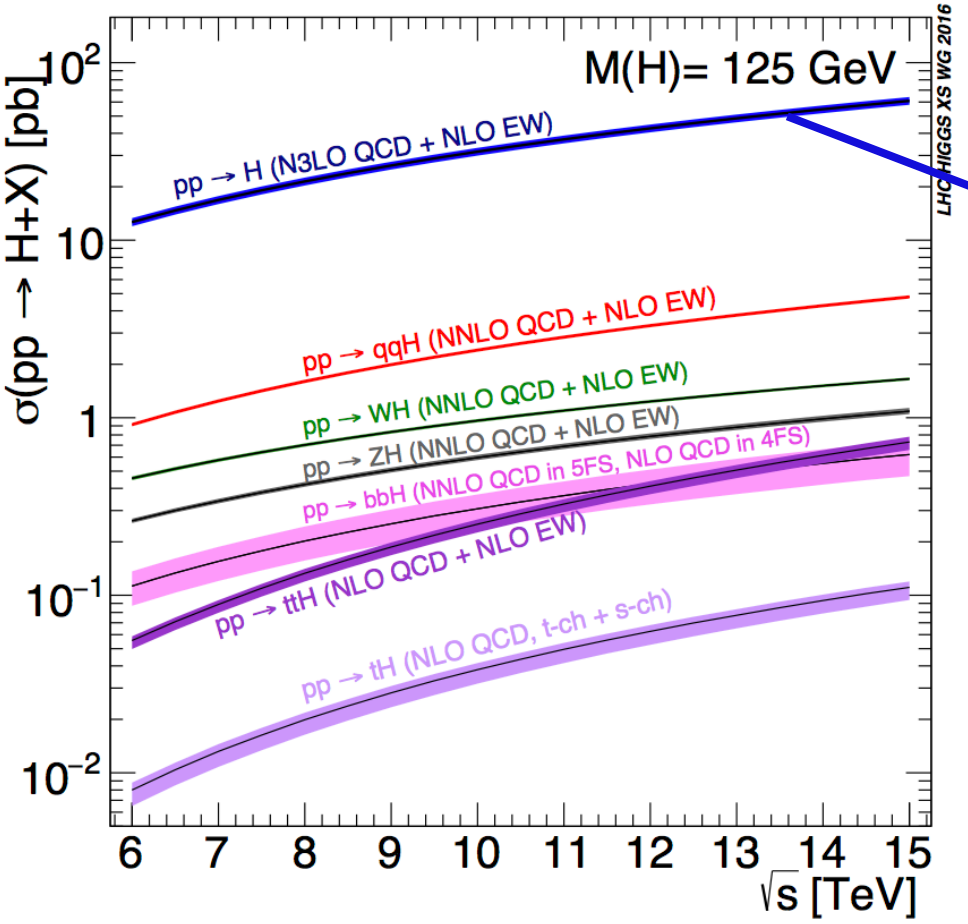


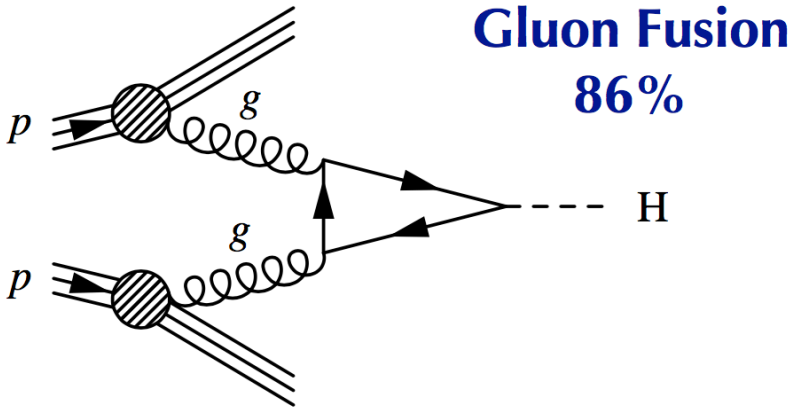
HIGGS BOSON

- Discovered in 2012 by CMS and ATLAS from indipendent observations in $\gamma\gamma$ and ZZ decay modes (best mass resolution)
- QCD cross section is order of magnitude larger than the Higgs production cross section
- Higgs boson decay in b quarks: $H \rightarrow b\bar{b}$
  - It has not been observed yet
  - It has the largest BR for SM Higgs (~57%)



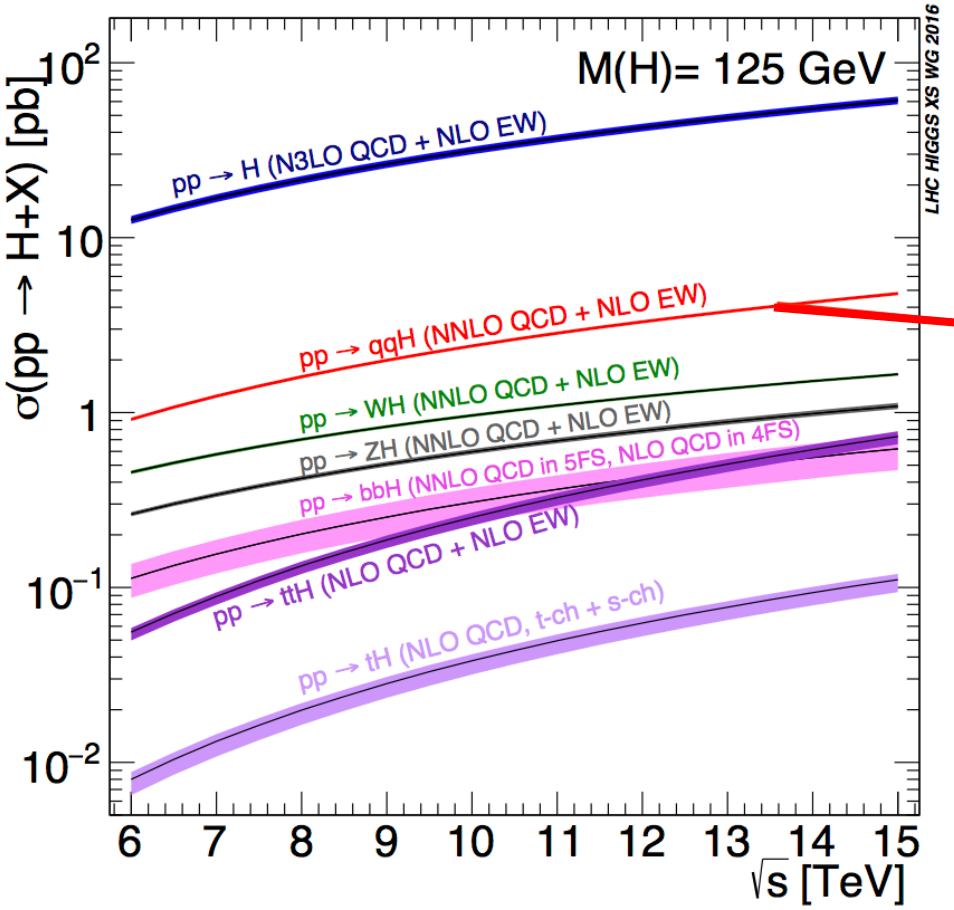
$H \rightarrow b\bar{b}$: MECHANISMS OF PRODUCTION

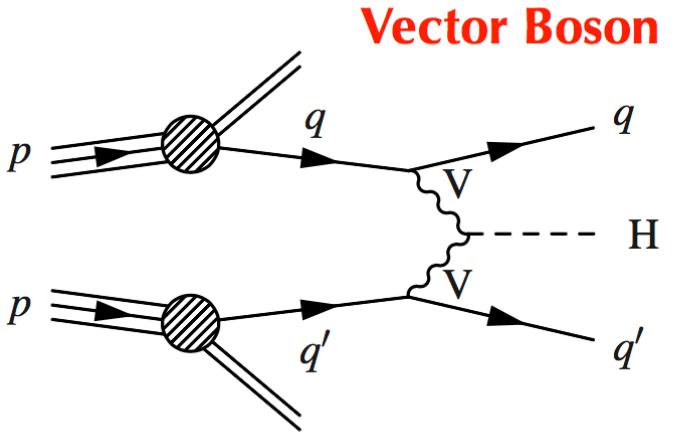




- Gluon-gluon fusion (GF) → 86% : most abundant channel. Large background from QCD production of b quarks.

$H \rightarrow b\bar{b}$: MECHANISMS OF PRODUCTION

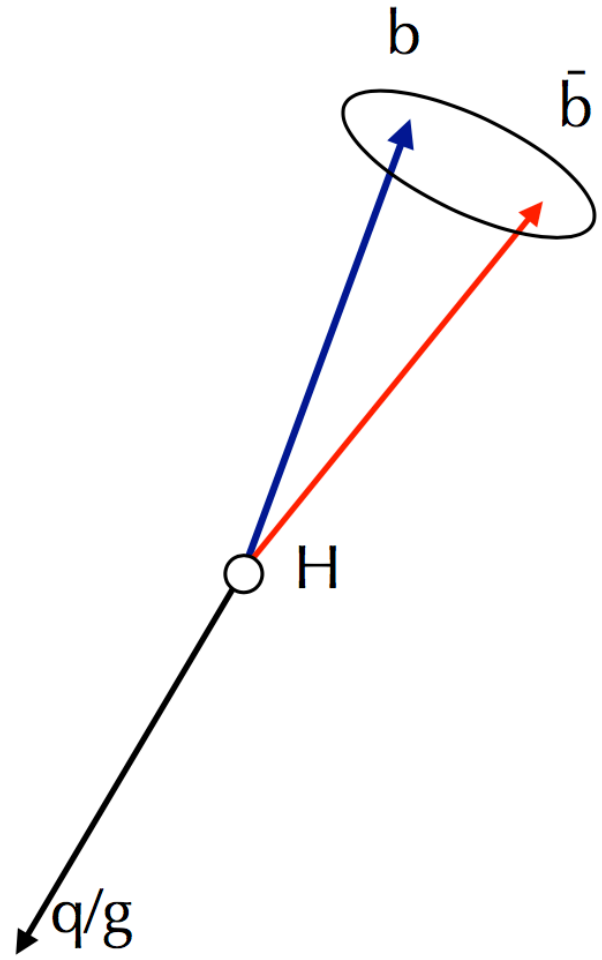




- VBF → 7% : very distinctive topology. Two quarks with large pseudorapidity.

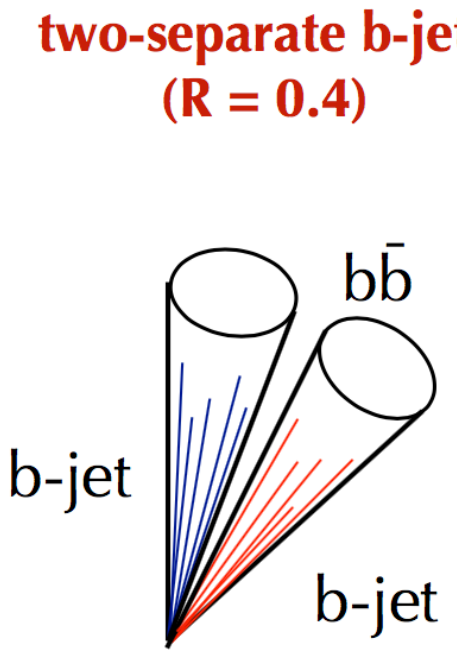
GLUON FUSION PROCESS

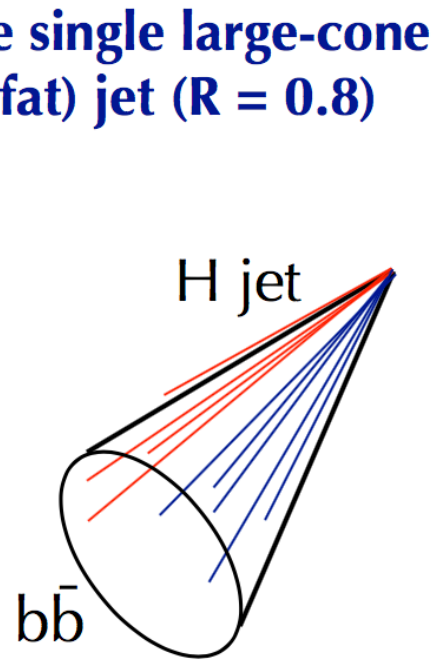
- It was considered impossible because of the large background but it has been proved to be possible by requiring high trasverse momentum Higgs (HIG-17-010)
- I focused on the gluon fusion production mode
- A result has been published based on 35.9 fb^{-1} of luminosity and 13 TeV center of mass energy
  - The sensitivity to the $H \rightarrow b\bar{b}$ is 1.5 σ
- I contributed to the optimization of the event selection analyzing the discrimination between quark and gluon jets



HIGGS-JET RECONSTRUCTION

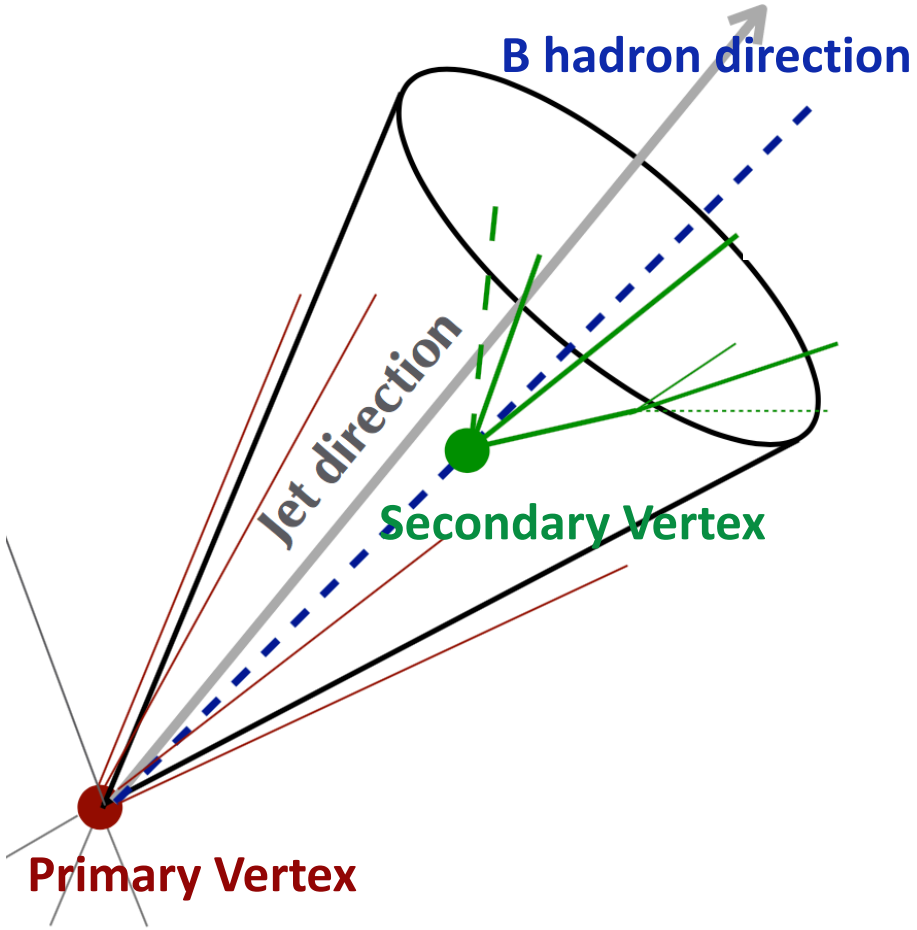
- A quark is reconstructed as a cluster of particles, jet, with the AK algorithm caractherized by the distance parameter R
- For high pt Higgs, the strategy is to consider one single large-cone (fat) jet with R=0.8 instead of taking two separate b-jets with R=0.4 for the two quarks
- Higgs jet: AK8 algorithm





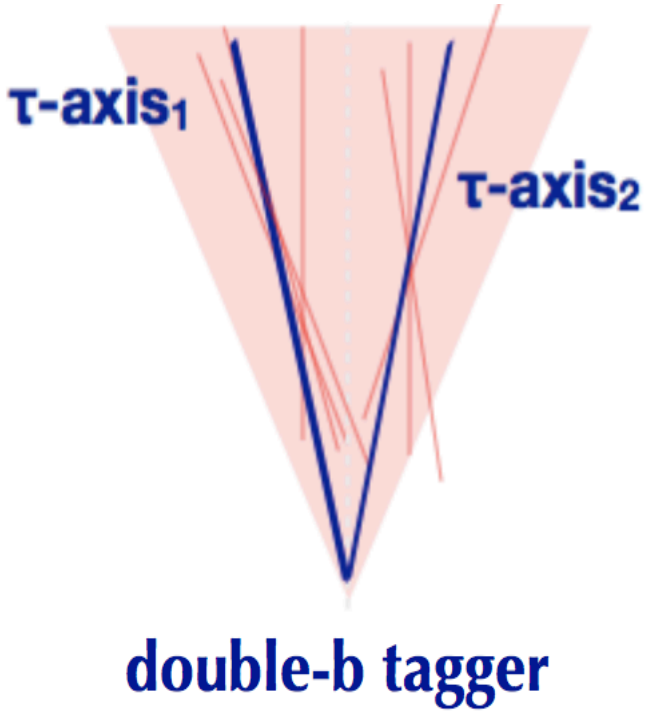
HIGGS-JET RECONSTRUCTION

- Tool for identifying $H \rightarrow b\bar{b}$ at high trasverse momentum: b-tagging
- b quark produced by the decay of the Higgs boson hadronizes in a B hadron (primary vertex)
- B hadron has a quite long lifetime, thus it can fly for some distance in the detector before decaying in a jet of other particles (secondary vertex)
- B hadron decay tracks are mostly produced in a cone in the B hadron flight direction which is considered the jet direction



HIGGS-JET RECONSTRUCTION

- New approach: DOUBLE-b tagging
- Identifies two B hadron decay chains from the two b quarks within the same fat jet
- Discrimination between b or non-b jet: combining tracking and vertexing information; presence of secondary vertex; flight distance and direction of B hadrons; multiplicy, mass and energy of the jet



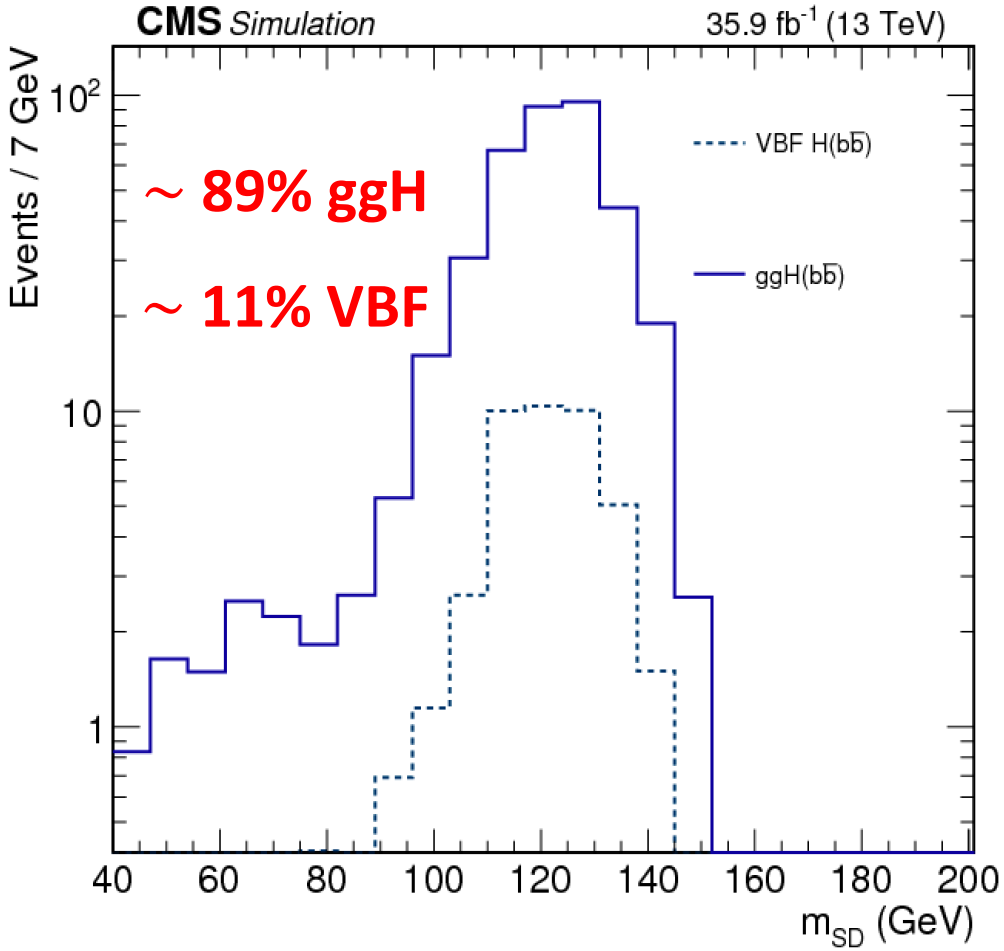
EVENT SELECTIONS

The selections used to make the analysis are the following; most of the cuts are made to improve the signal and reject the background:

- Pt of the Higgs jet > 350 GeV : the gluon fusion process is possible requiring a high pt Higgs
- Double-b tagger > 0.9 : reduce the QCD background
- Jet mass > 40 GeV : reduce the QCD background
- N2DDT < 0 : it's a variable that determines the probability that a jet contains 2 prongs; less than 0 because lower values mean it's more likely to have 2 prongs
- We are getting rid of events that have muon, electrons, and taus

EVENT SELECTIONS

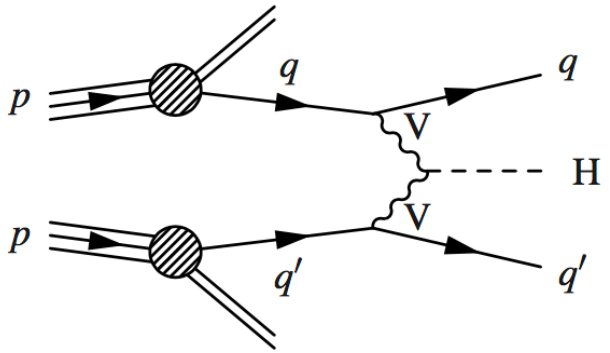
- Plot of the mass of the Higgs jet with all the selections
- Percentages: 89% for ggH and 11% for VBF
- Find some observables through which it is possible to better separate gluon fusion events from vector boson fusion events and improve the percentages



EVENT SELECTIONS

- Two quarks with large pseudorapidity
- AK algorithm with R = 0.4
- Quarks: AK4 algorithm
- Information of the 6 highest Pt AK4 jets
- Selections:
  - Pt > 30 GeV
  - $|\eta| < 2.5$
  - $dR = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

    dR (AK4 jet, Higgs jet) > 0.3

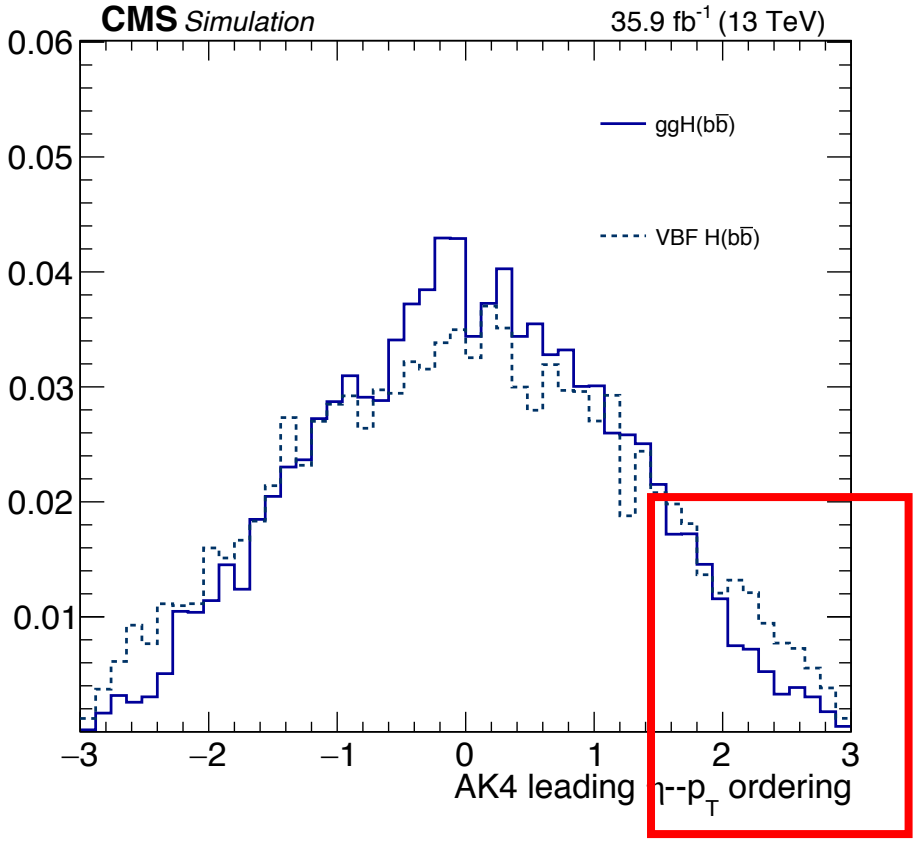


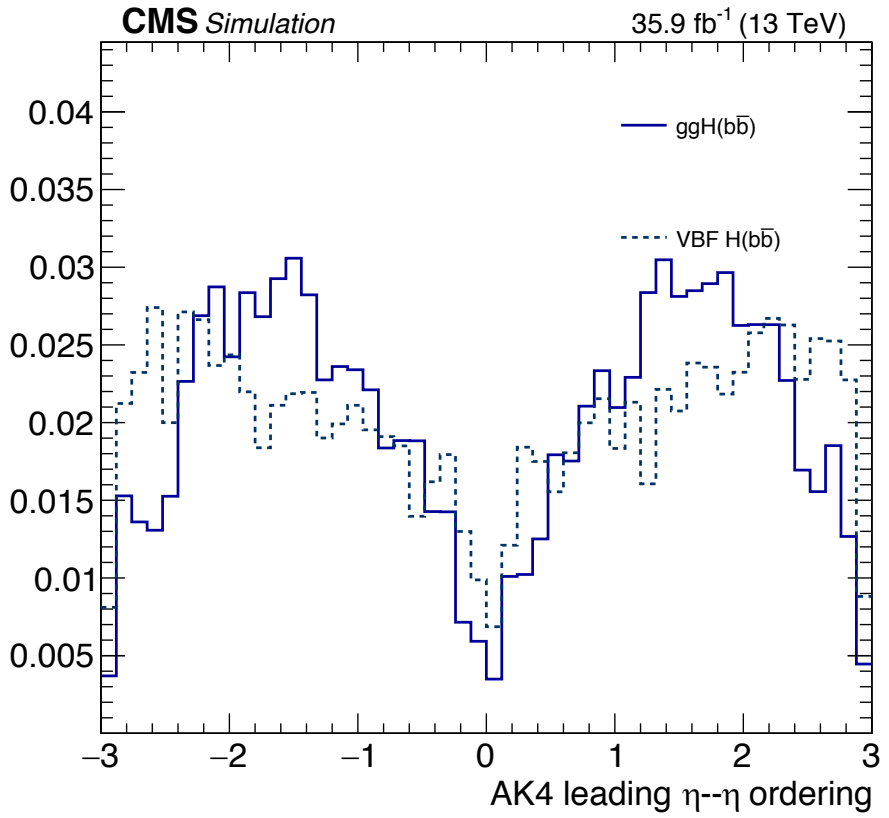
VBF
7%

- Sorting:
  - Pt-ordering
  - Eta-ordering

ANALYSIS

- Distributions of Eta for AK4 leading jet for both the sorting (linear scale)





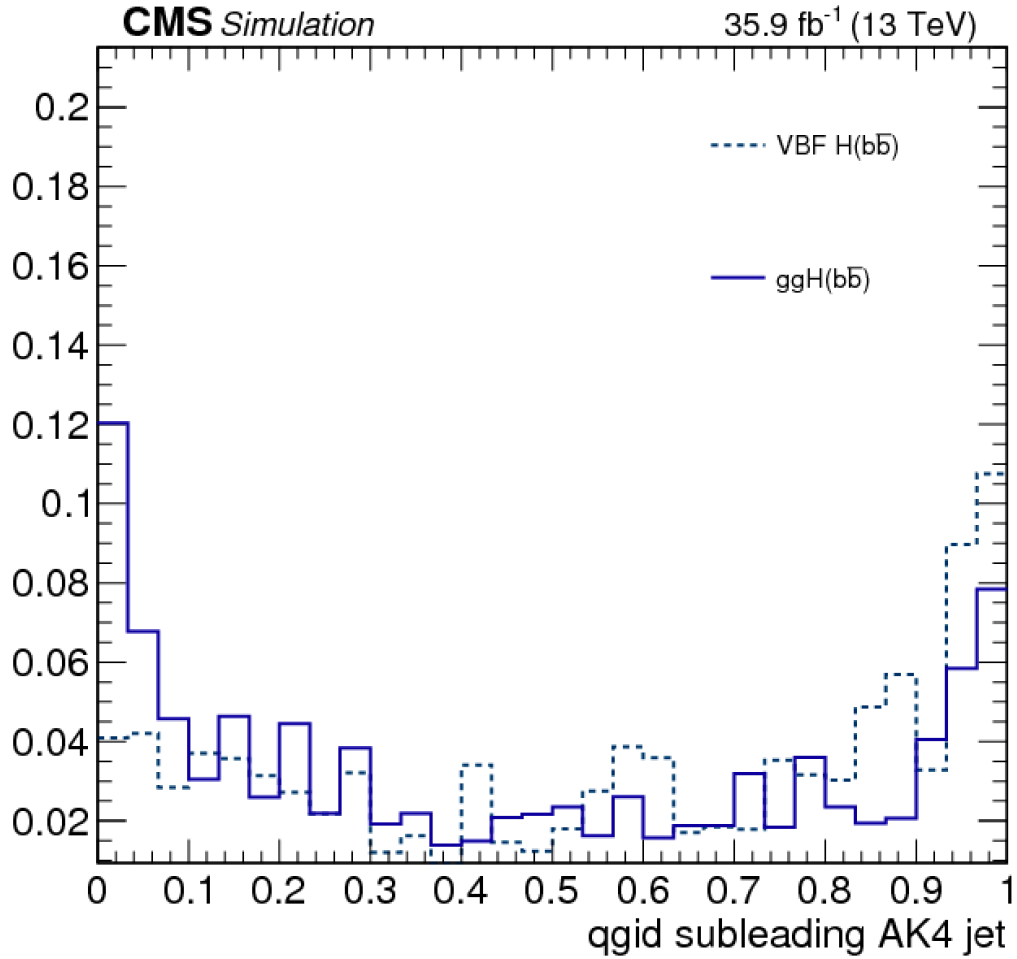
QUARK/GLUON DISCRIMINATION

- Quark/gluon discriminant: different variables considered
  - Multiplicity: total number of particles reconstructed within the jet
  - Jet energy sharing variable: it's 1 for a jet made of only one particle and 0 for a jet made of an infinite number of particles
  - Angular spread in the $\eta - \phi$ plane

- It defines a probability that the jet is coming from a quark or from a gluon: 1 if it's a quark jet and 0 if it's a gluon jet
  - Gluon jets: wider, higher multiplicity, uniform energy fragmentation
  - Quark jets: narrow, hard constituents, significant fraction of energy

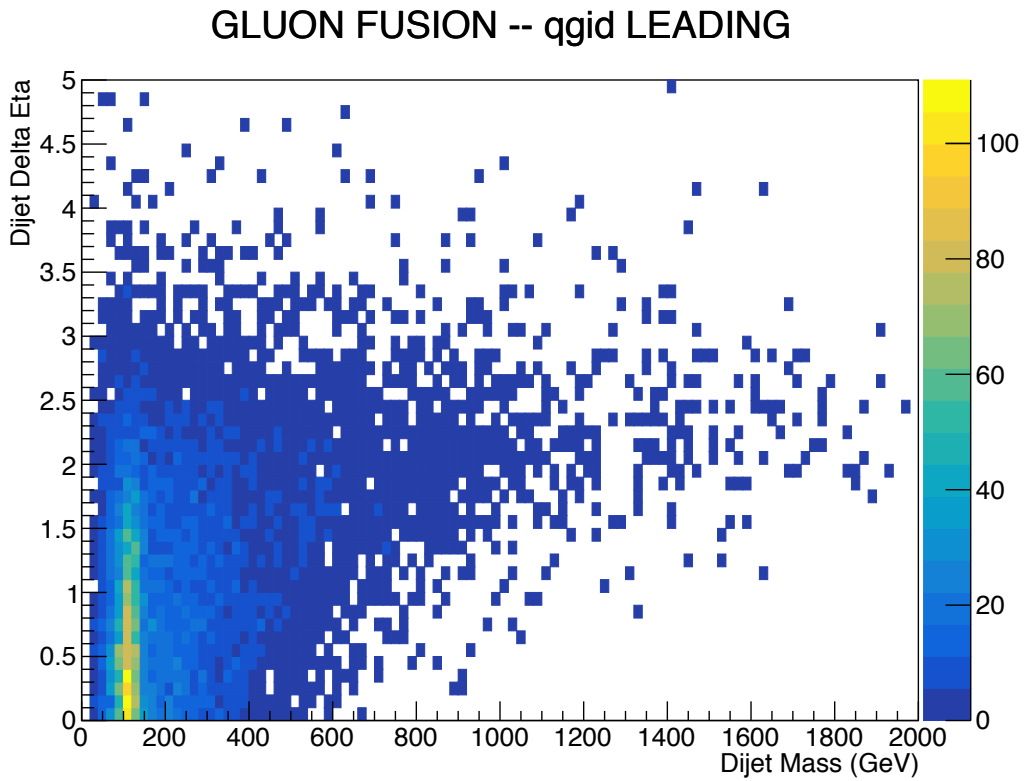
QUARK/GLUON DISCRIMINATION

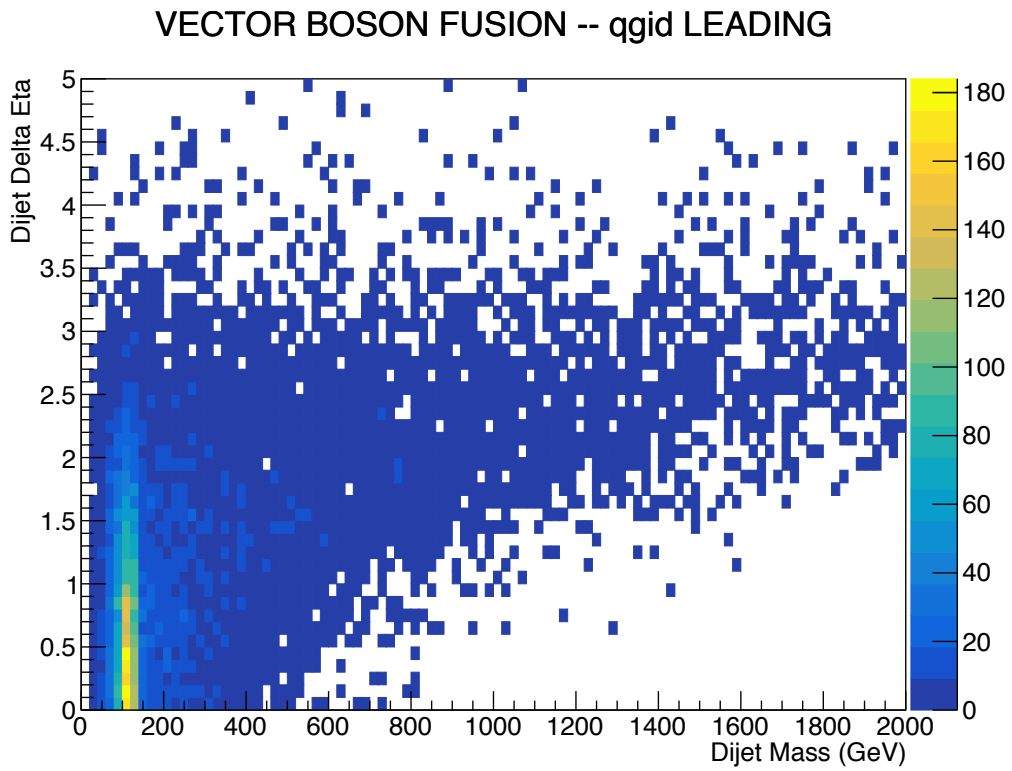
- Calculate the q/g discrinimant for all the jets
- Use this variable to sort the jets
- Distribution of quark/gluon discriminant for the subleading AK4 jet: tiny discrimination



QUARK/GLUON DISCRIMINATION

- Jets with the highest value of the quark/gluon discriminant parameter to calculate the Mass and Delta Eta





QUARK/GLUON LIKELIHOOD RATIO (QGLR)

- Ratio built as a combination of the quark/gluon discriminant of the single jets
- Range of values: between 0 (gluon jets) and 1 (quark jets)
- Probability that the whole event contains two quark jets instead of having only the probability that one single jet is coming from a quark or a gluon
- L(Qq,Gg) is the probability that Q jets are coming from Q quarks and G gluons

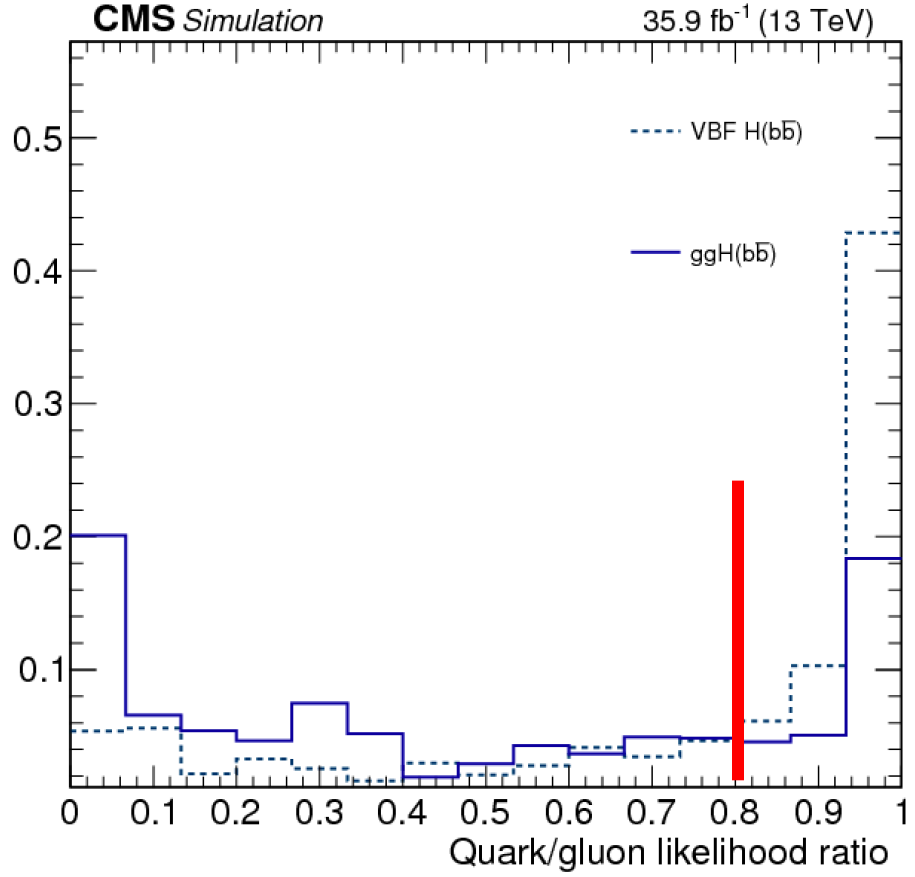
$$q_{LR(Q\mathrm{v}A)} = \frac{L(Qq, 0g)}{L(Qq, 0g) + L(Aq, (Q-A)g)}$$

- f are probability distribution functions
- ζ is the value of the quark/gluon discriminant of the single jet

$$L(Qq, Gg) = \sum_{i_1} \sum_{i_2 \neq i_1} \cdots \sum_{i_{Q+G} \neq i_1, \ldots, i_{Q+G-1}} \left\{ \prod_{k \in \{i_1, \ldots i_Q\}} f_{\mathrm{q}}(\zeta_k) \prod_{m \in \{i_{Q+1}, \ldots i_{Q+G}\}} f_{\mathrm{g}}(\zeta_m) \right\}$$

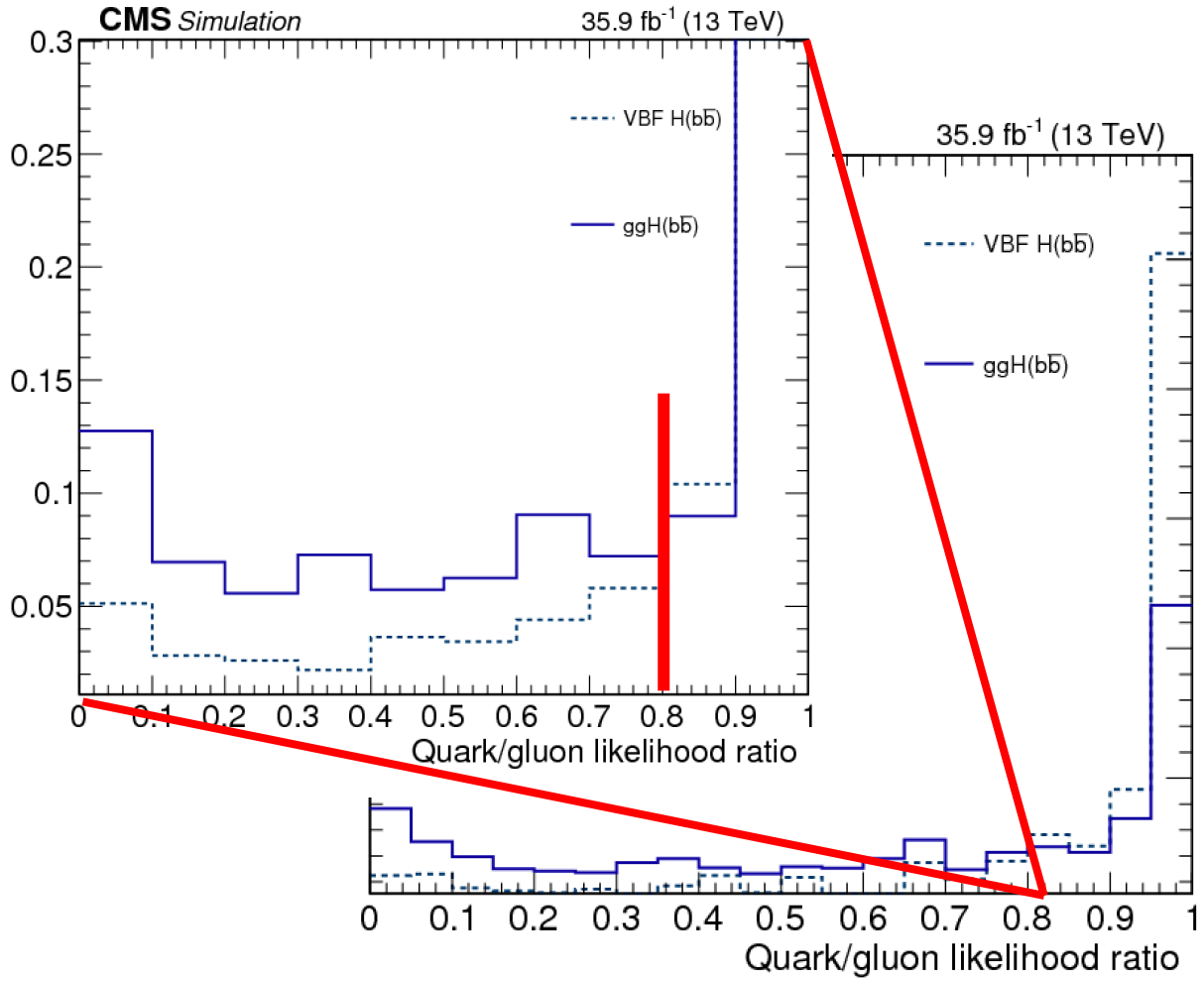
QUARK/GLUON LIKELIHOOD RATIO (QGLR)

- Quark/gluon likelihood ratio for each event
- Discrimination for QGLR < 0.8
- I improved the discrimination selecting a smaller value for the dR cut, tuning the number of AK4 jets considered matched to the Higgs jet
- dR > 0.3: best cut



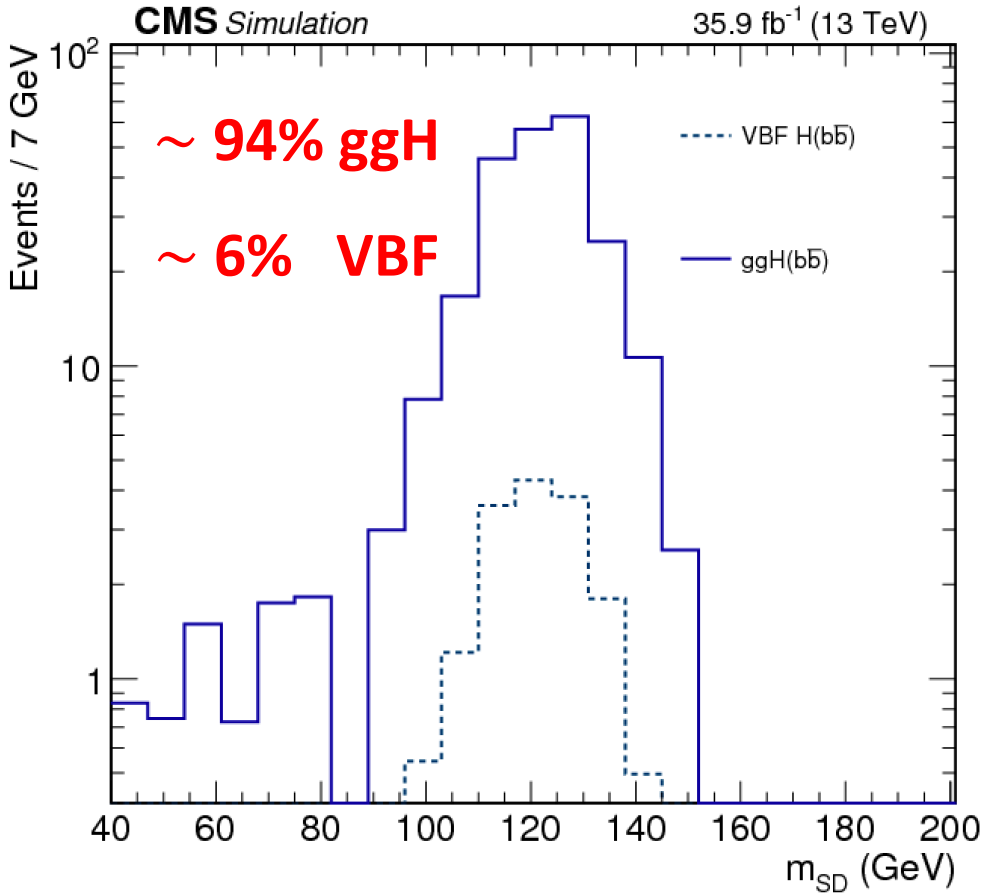
QUARK/GLUON LIKELIHOOD RATIO (QGLR)

- Plot of the quark/gluon likelihood ratio with dR > 0.3
- Good discrimination between the ggH and VBF processes
- Under the value 0.8: bigger contribution from ggH process



QUARK/GLUON LIKELIHOOD RATIO (QGLR)

- Cut on the quark/gluon likelihood ratio to calculate the Higgs jet mass distribution
- QGLR < 0.8 : the ggH process is dominant
- Plot of the mass distribution with all the cuts
- Percentages: 94% for ggH and 6% for VBF



CONCLUSIONS

- My work started with learning how to use PyROOT and then I wrote some parts of the code
- I learnt how to access the information stored in some analysis ntuples and analyze them using Python and ROOT
- I have familiarized with the phenomenology of the processes and event reconstruction in CMS
- I found that the quark/gloun likelihood ratio was a good variable to better separate gluon fusion events from vector boson fusion events
- I contributed to the event selection and improved the percentages of the contributions of the two processes

THANK YOU!