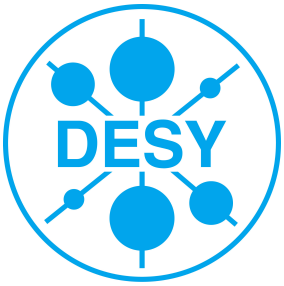

Study of the QCD background for MSSM Higgs Boson $\rightarrow b\bar{b}$ with LHC Run 2015 data

Connor Thorburn, Universita' degli Studi di Torino, Italy

September 8, 2016

Abstract

A study of the QCD background for neutral Higgs bosons decaying into a $b\bar{b}$ quark pair is presented. The analysis is based on 2015 LHC data, collected by the CMS experiment, with collisions at a center-of-mass of $13TeV$. Comparisons with different MC samples are shown.

Contents

1 MSSM Higgs bosons

The Higgs boson with a mass of approximately 125 GeV was discovered in 2012. Its measured properties are in agreement with standard model (SM) predictions, but the particle may also be the first visible one of an "extended Higgs sector". One of the candidate theories dealing with this problem is the minimal supersymmetric extension of SM (MSSM). It is based on two charged Higgs bosons ($H^{\pm}$) and three neutral ones (A, H, h). A is CP odd, while H and h are CP even. The properties of MSSM are well described by two parameters, the mass of the A Higgs boson (m_A) and the ratio of vacuum expectation values of the two Higgs doublets ($tan\beta = v_2/v_1$)[1]. According to MSSM we could associate the 125 GeV Higgs boson to one of the CP Higgs bosons.

2 $b\bar{b}$ associated production

This study focuses on $b\bar{b}$ associated H/A production. This decay mode is an interesting object of study for several reasons: cross section is enhanced by a factor of $\approx tan^2\beta$,it helps reducing background and it is sensitive to the higgsino mass parameter μ and to bottom Yukawa coupling. The signal is searched in processes with at least three b-jets (a fourth jet requirement is not necessary since it would be beyond acceptance). Some examples of processes are presented in Figure 1. The main background, which is the object of this work, originates from QCD multijet production, with at least two energetic b-jets.

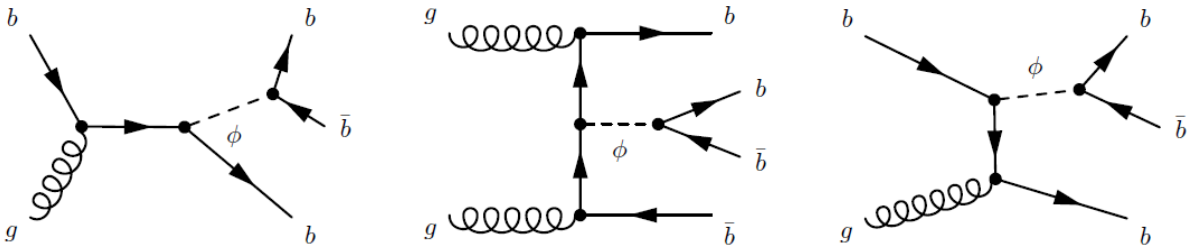

Figure 1: processes

3 The CMS detector

The CMS detector is based on a superconducting solenoid of 6m internal diameter, providing a magnetic field of 3.8T. Particle detection is performed by an inner tracker (silicon pixel and strip tracker), a crystal electromagnetic calorimeter, a brass and scintillator hadron calorimeter and muon gas-ionization detectors, displayed as shown in Figure 2.

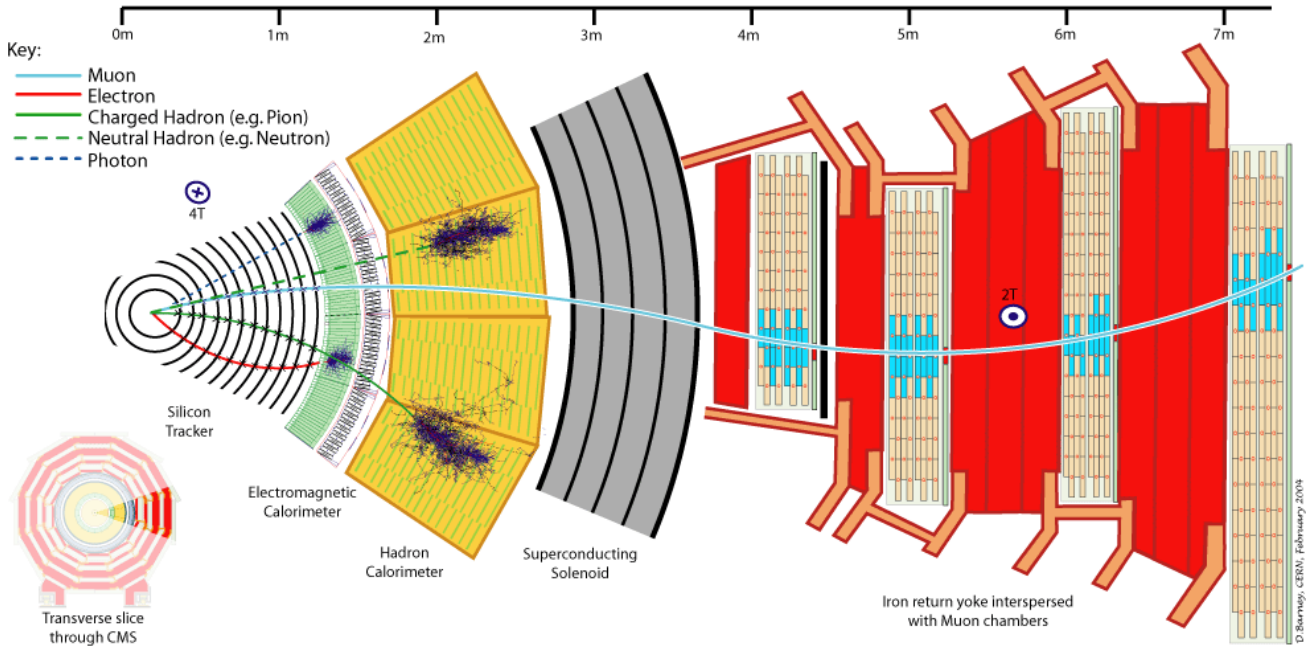

Figure 2: CMS detector

4 Data

LHC Run 2015 C and D data is studied. Integrated luminosity is $2.69 fb^{-1}$. Analysis is performed on BTagCSV MINIAOD ntuples.

5 Online selection

Signal selection at online level is performed by two dedicated triggers, one for the low mass (LM) and one for the high mass (HM) region. Trigger cuts are listed in Table 1 for both mass ranges.These triggers are aimed at reducing event rates and non $b\bar{b}$ background. This work focuses on LM region.

Table 1: Trigger cuts

LM trigger	HM trigger	purpose
- at least two jets with $p_T > 100GeV$ - $\|\Delta\eta\| < 1.6$	at least two jets with $p_T > 160GeV$	reduce rate
at least two jets with discr$_{CSV} > 0.9$	at least two jets with discr$_{CSV} > 0.85$	reduce non $b\bar{b}$ background

6 b-tagging

Jet b-tagging is performed by different types of algorithm. One example is the Combined Secondary Vertex algorithm (CSV) which is based on impact parameters and secondary vertex information. Figure 3 shows the relation between jet tagging efficiency and misidentification probability for some of such algorithms. There are three main working points (Loose, Medium, Tight) associated to different probabilities (see Table 2)[2].

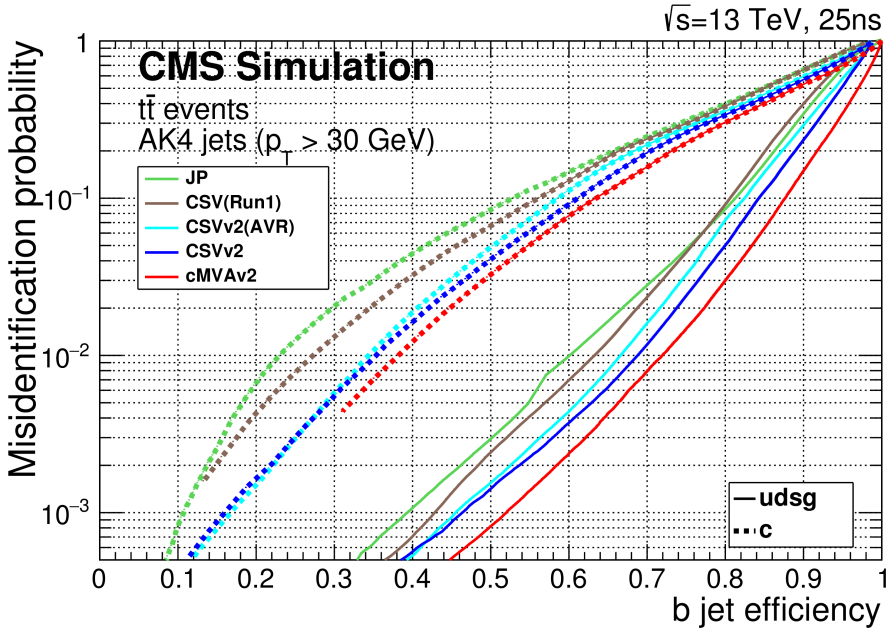

Figure 3: Misidentification probability curves

Table 2

working point	misidentification probability	CSVv2 discriminant
loose	10^{-1}	0.460
medium	10^{-2}	0.800
tight	10^{-3}	0.935

7 Offline selection

Offline double b-tag selection for LM region is performed. This selection requires only two of the three leading jets to pass tight CSV b tagging requirements. Applied leading and sub-leading jet cuts are listed in Table 3.

Table 3: Leading and Sub-Leading jet cuts

$p_T > 100GeV$
$\|\eta\| < 2.2$
$\|\Delta\eta\| < 1.6$
$\Delta R < 1$
$discr_{CSV} > 0.935$

8 Monte Carlo

QCD background is simulated with two different Monte Carlo samples: Madgraph5 and Pythia8. Basic information about MC generators is obtained from their data cards.

8.1 QCD bEnriched Madgraph5

Madgraph8 is a framework employed for simulation of processes to leading order (LO) and next to leading order (NLO) accuracy[3]. Input consists of simple processes, as shown below, with the possibility of covering all topologies.

$$p\,p \to j\,b\,\overline{b}$$
$$p\,p \to j\,j\,b\,\overline{b}$$
$$p\,p \to j\,j\,j\,b\,\overline{b}$$

The matrix elements of such processes are given as output. Then the parton shower part (generated with Pythia8) is added. In this study just LO processes are considered. The Madgraph5 sample is unweighted, so reweighting must be performed using either Ht or Jet multiplicity. Ht is chosen in this case. The procedure is the following: first of all Ht data and unweighted MC distributions are obtained, then the fuction $Data/MC$ is built and it is used for reweighting the MC sample. Ht distributions before and after reweighting are shown in Figure 4.

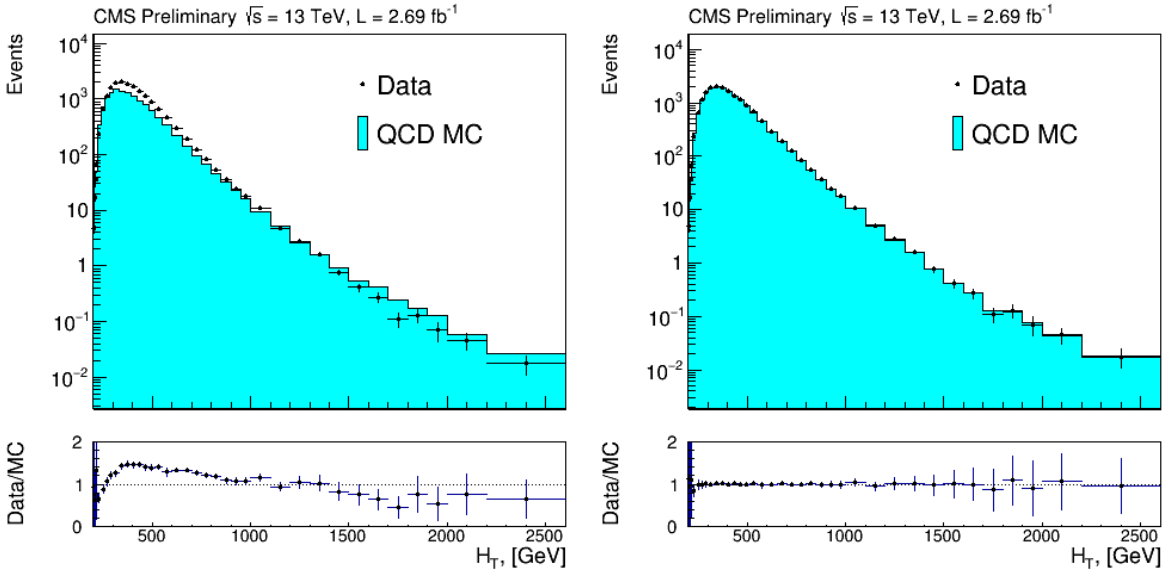

Figure 4: Ht distribution before (left) and after (right) reweighting

8.2 QCD Pythia8

Pythia8 is employed in the simulation of high energy Physics processes to LO accuracy. N-tuples are produced devided in p_T bins by using the HardQCD class, which contains the processes for QCD jet production above a minimum p_T threshold (30 GeV in this case). If such threshold is too low, too large cross sections are obtained[4]. Some of the processes included in the class are listed below:

$$
\begin{aligned}
g\ g &\to g\ g \\
g\ g &\to q\ \overline{q},\ q\ =\ (u,d,s) \\
q\ g &\to q\ g \\
\overline{q}\ g &\to \overline{q}\ g \\
q\ q' &\to q\ q' \\
q\ \overline{q'} &\to \overline{q}\ \overline{q'} \\
q\ \overline{q} &\to g\ g \\
q\ \overline{q} &\to q'\ \overline{q'}
\end{aligned}
$$

In this case both matrix elements and parton shower part are given as output.

9 Data and MC comparisons

Data and MC distributions for the following parameters are studied: $p_T, \eta, \Delta\eta$, $diJet\ object\ p_T$, $diJet\ object\ mass$.

9.1 p_T distribution

Data and MC distribution of parameter p_T is presented for leading (Figure 5) and subleading (Figure 6) jets. Data distribution trends like a parton fraction function. At high values of the parameter a power law trend can be recognised, followed by an exponential trend at the centre of the plot. There is an initial low entry (close to cut value) which is due to trigger efficiency and cut reasons.
Di-Jet object p_T distribution is presented in Figure 7.
The performance of both MC is good, though Madgraph5 performs slightly better.

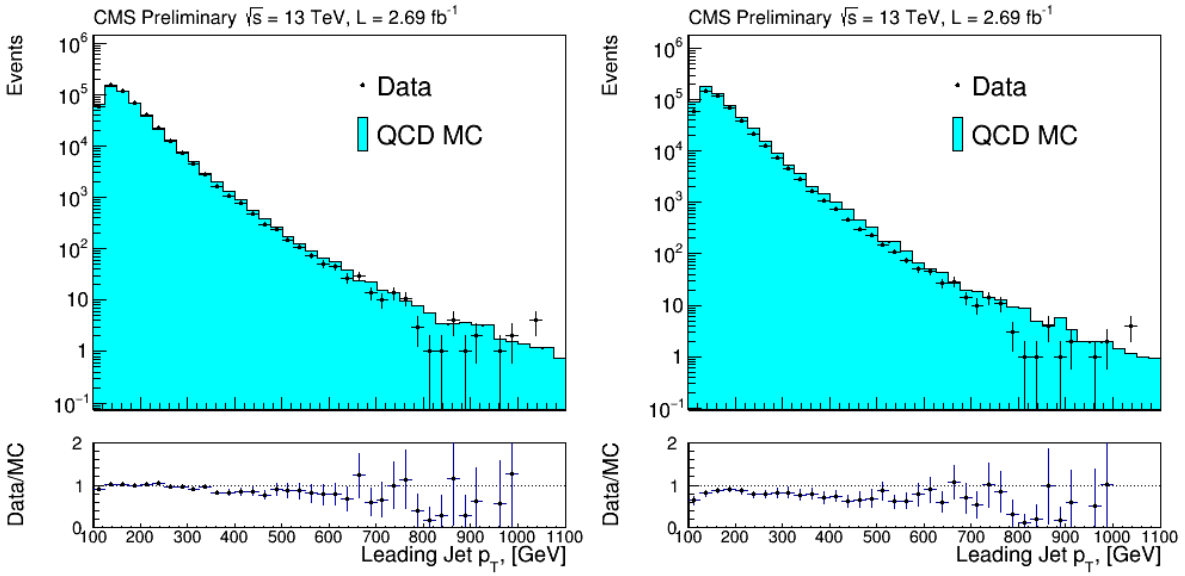

Figure 5: Leading jet p_T data and MC distributions (Madgraph5 left, Pythia8 right).

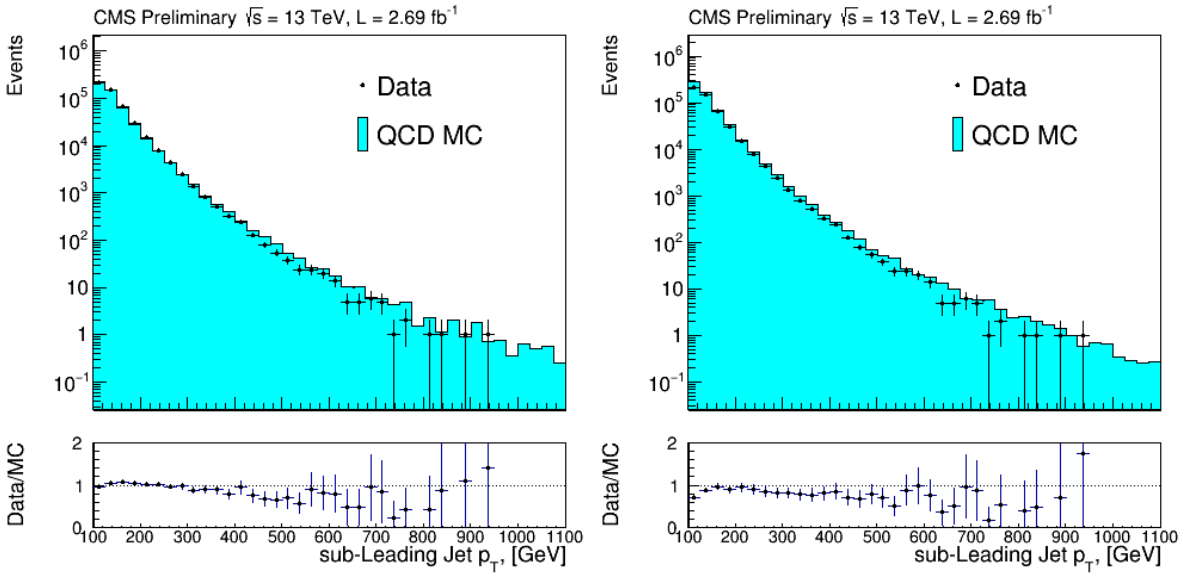

Figure 6: Sub-leading jet p_T data and MC distributions (Madgraph5 left, Pythia8 right).

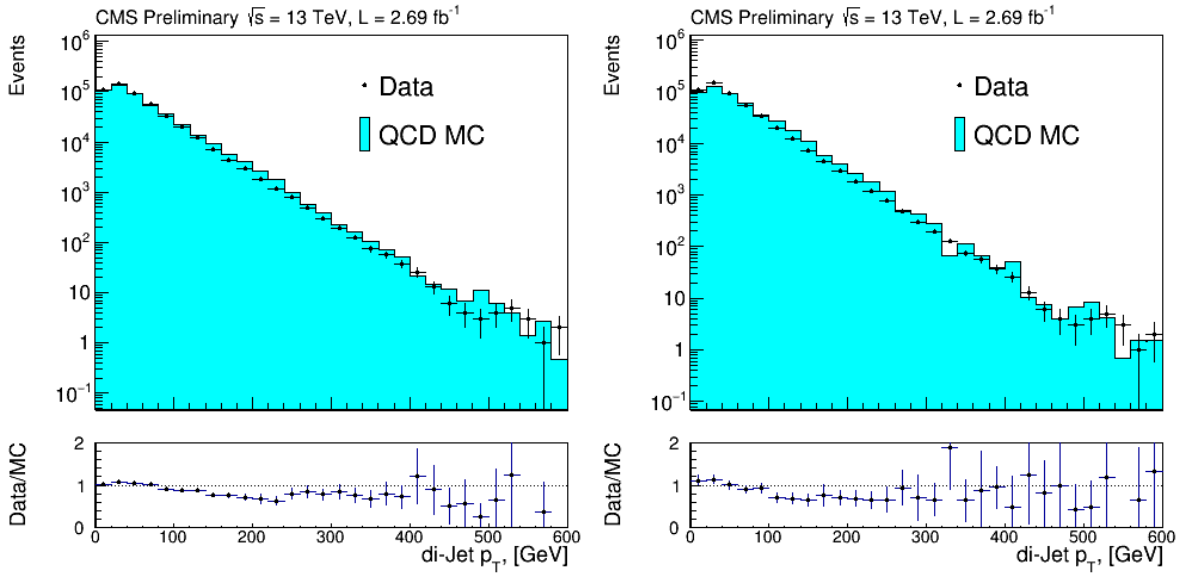

Figure 7: Di-Jet object p_T data and MC distributions (Madgraph5 left, Pythia8 right).

9.2 Pseudorapidity distribution

Pseudorapidity (η) distribution for leading (Figure 8) and sub-leading (Figure 9) jet is studied. The trend is symmetric in both cases. MC samples perform equally good, but Pythia8 is more dominated by statistical uncertainty. It is also intersting to study the pseudorapidity difference of the two jets ($\Delta\eta$). As can be seen in Figure 10, the distribution is dominated by a flat central trend.Regarding MC samples, some structures can be recognised at the corners of the Data/MC ratio plot in the case of Madgraph5 but are absent in Pythia.

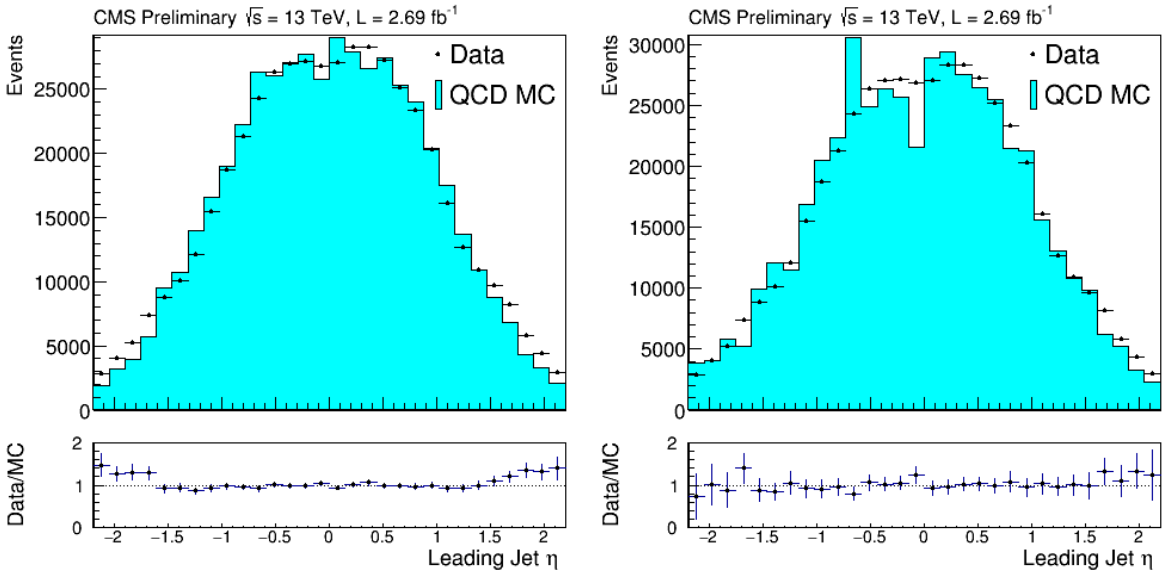

Figure 8: Leading jet η data and MC distributions (Madgraph5 left, Pythia8 right).

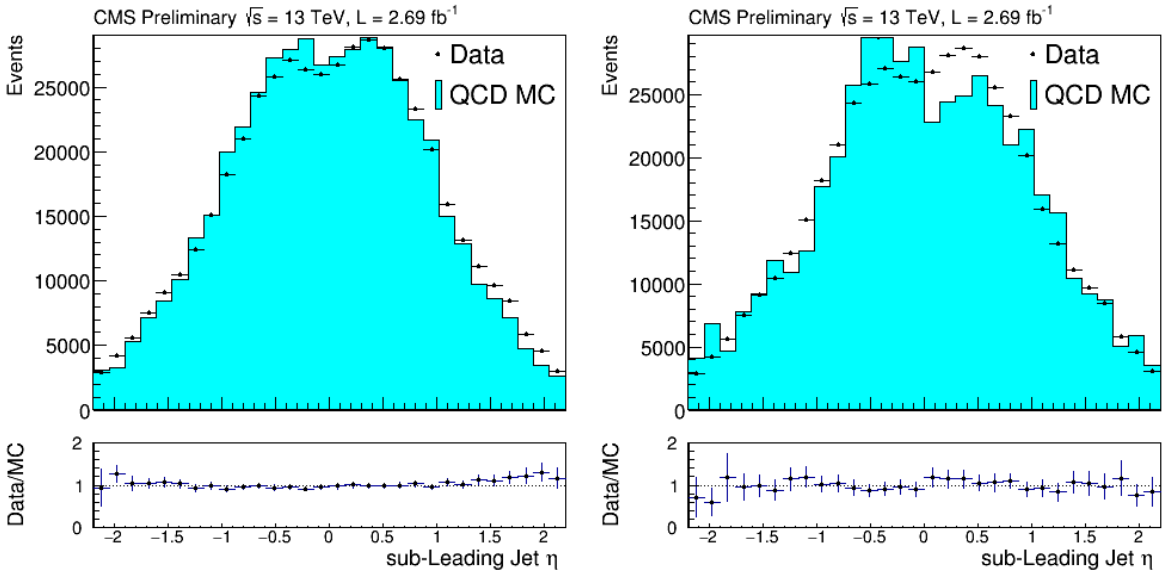

Figure 9: Sub-leading jet η data and MC distributions (Madgraph5 left, Pythia8 right).

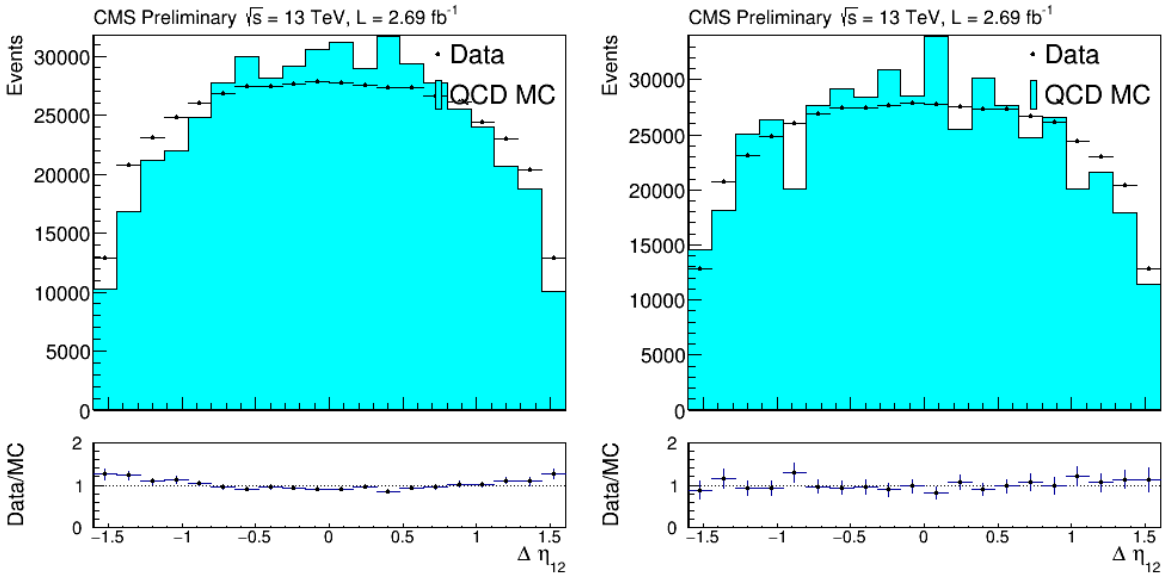

Figure 10: Jet $\Delta\eta$ data and MC distributions (Madgraph5 left, Pythia8 right).

9.3 Di-Jet object mass distribution

Di-Jet object mass distribution is presented in Figure 11. Prior knowledge of background shape to MSSM $H \to b\bar{b}$ associated production can be obtained from this distribution. Also in this case good data MC agreement was achieved, but no new resonances can be observed (no peak at 750 GeV).

10 Conclusions

- The aim of this work was to study the QCD background for MSSM H $\to b\bar{b}$ associated production and to make comparisons between data and MC for key variables.
- As can be seen there is a general good agreement between different MC simulations and data, though Madgraph5 seems to perforom slightly better than Pythia8 and has more statistics. Regarding online selection,a large high-p_T sample was selected by dedicated triggers.
- Prior knowledge of background shape to MSSM $H \to b\bar{b}$ associated production is obtained. What's more the confirmation of the validity of the QCD models studied can be used in further steps of the analysis, for example for signal selection optimisation and background fit.

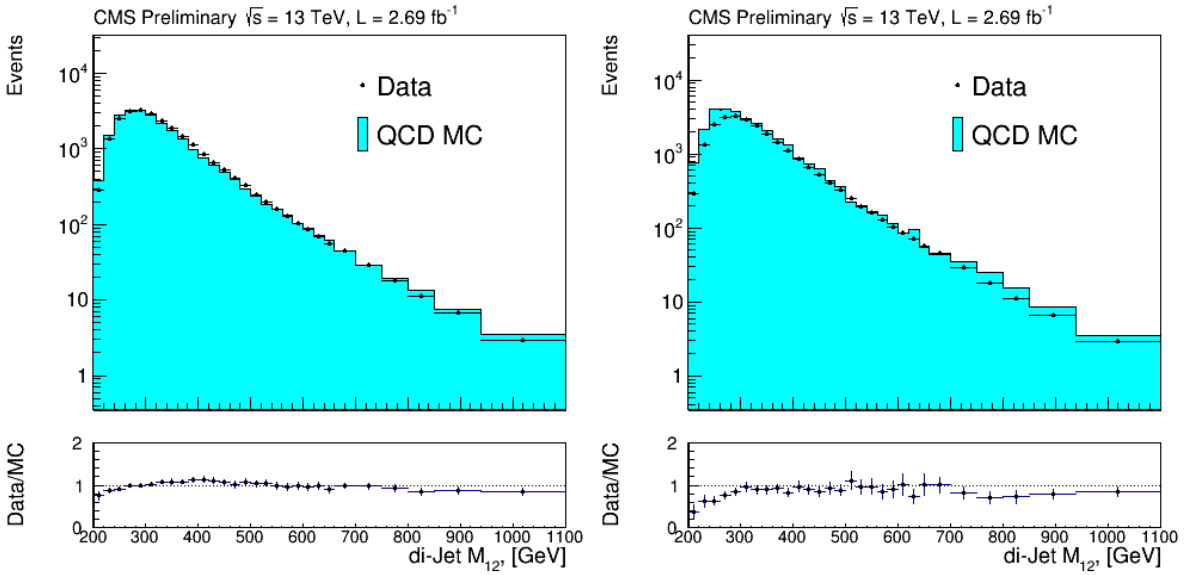

Figure 11: Di-Jet object *mass* data and MC distributions (Madgraph5 left, Pythia8 right).

References

[1] CMS Collaboration, "Search for Neutral MSSM Higgs bosons decaying into a pair of bottom quarks", *arXiv:1506.083229v2 [hep-ex] 24 Jan 2016.*

[2] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", *CMS PAS BTV-15-001.*

[3] MadGraph Web Page, *http://madgraph.physics.illinois.edu/*

[4] Pythia Web Page, *http://home.thep.lu.se/ torbjorn/Pythia.html*