Projection Study for the Measurement of the $b\bar{b}H$ production at the region of High Luminosity of the LHC

Maria Carolina Kherlakian

Supervisor: Adinda de Wit

September 5, 2018

Abstract

This project aims to determine the possibilty of measurement of the production of the SM Higgs boson in association with a pair of b-quarks in the $\tau\bar{\tau}$ final state in the region of High Luminosity (HL) of the LHC. First it is described the method to separate the signal from its major background, then a normalization of the 2016 CMS MSSM Higgs $\rightarrow \tau\tau$ analysis to the region of High Luminosity of the LHC is presented together with the future expected limits for this channel.

Contents

1 Introduction

It is expected that several processes of production and decay of the Higgs boson will be able to be identified with the high amount of data provided by the High Luminosity of the LHC. In 2037 the integrated luminosity of the LHC will be increased to 3000 fb^{-1}, with a center of mass energy of 14TeV. In this report the possibility of detecting the production of the SM Higgs boson in association with a pair of b-quarks (Fig. 1) in the region of HL is analyzed. Due to the small cross section of this channel one expects to have a higher possibility to detect a signal with a larger data set. To do that we need a good way to separate the signal in the best way from its major background, the $t\bar{t}$ decay (Fig. 4). For this purpose, simulated samples are analyzed and then it was chosen the best variable to accomplish the task. We proceed to the normalization of the CMS MSSM Higgs $\rightarrow \tau\tau$ analysis to the region of HL, from which we could extract the limits of the channel of interest for integrated luminosities reaching 3000 fb^{-1}.

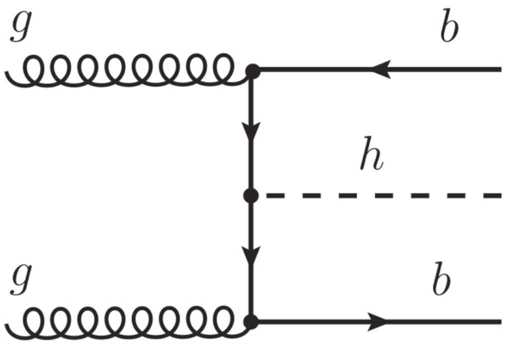

Figure 1: Diagram for the production of the Higgs boson in association with a pair of b-quarks.

2 Theory

The cross section as a function of the center of mass energy of colliding beams for several channels of production of the Higgs boson can be seen in Fig. 2. The pink line is the cross section expected for the channel of interest, and it is evident that this channel is not so easy to be found, since it has one of the lowest cross sections between all the other channels presented. This first challenge in the measurement of the signal of the $b\bar{b}H$ can be overcome with a larger data set (a higher luminosity), as described in the previous section.

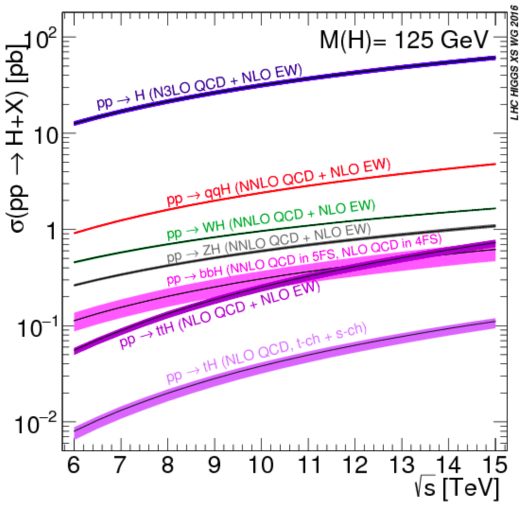

Figure 2: Cross section for different modes of production of the Higgs boson.

The second challenge to be overcome is the background coming from the $t\bar{t}$ decay, which has the same signature as our channel of interest. In Fig. 3 it is possible to see that the analyzed final state of the Higgs boson is a pair of two leptons, which is also present in the background (Fig. 4). In this report we will focus on the muon plus hadronic τ final state of the Higgs. Other common backgrounds to this channel are the Z $\rightarrow \tau\tau$, jet $\rightarrow$ hadronic τ fakes and w+jet production, however the one that plays the major role is the $t\bar{t}$ background.

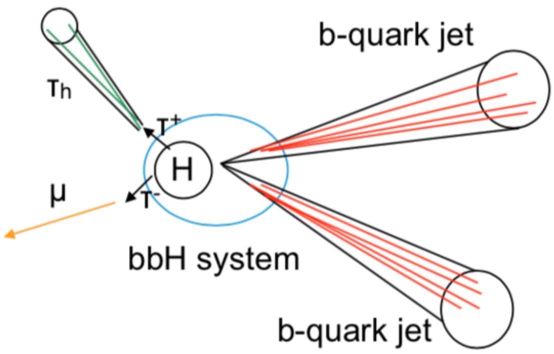

Figure 3: Scheme of the $b\bar{b}H$ production with a final state of two τ leptons.

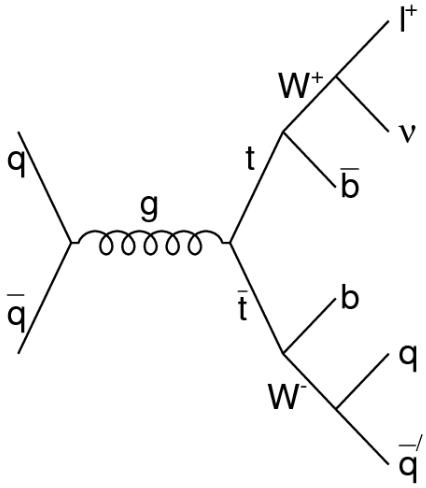

Figure 4: Main background for the channel of interest.

One of the ways to have a good separation between signal and background is searching for a variable that has different shapes for these two cases. One can look, for instance, to the total transverse mass (Eq. 1), but many others can be used, for example the transverse momentum, or the mass of the final state of the Higgs boson. In Eq. 1, MET is the missing transverse energy.

$$M_T^{TOT} = \sqrt{M_T^2(MET,\mu) + M_T^2(MET,\tau_h) + M_T^2(\mu,\tau_h)} \tag{1}$$

$$M_T(a,b) = \sqrt{2\,P_t^a\,P_t^b \cdot (1 - cos\Delta\phi_{ab})} \tag{2}$$

3 The CMS detector

The CMS (Compact Muon Solenoid Detector) (Fig. 5) is located in one of the four collision points of the LHC. It is a general purpose detector, designed to be able to identify any new physical phenomena revealed by the LHC. It has a cylindrical form with 15 meters high and 21 meters long and its shape was designed to detect muons with very accurately. The CMS detector is composed of a silicon tracking system, a lead tungstate crystal electromagnetic homogeneous calorimeter (ECAL), a brass and scintillator sampling hadron calorimeter (HCAL) and muon chambers made of ionizing gas outside of the solenoid, which provides a magnetic field of 3.8T. The CMS detector has to be prepared for the region of High Luminosity of the LHC in order to improve the efficiency of the detector for this condition. This will be achieved in two Phases. In the Phase-I Upgrade it is planned to do modifications in the Pixel Tracker, HCAL and Level-1 Trigger systems [1]. This phase has already been completed in the start of 2017.

The Phase II Upgrade will be installed from the beginning of 2024 and is expected to be largely completed in 2026.

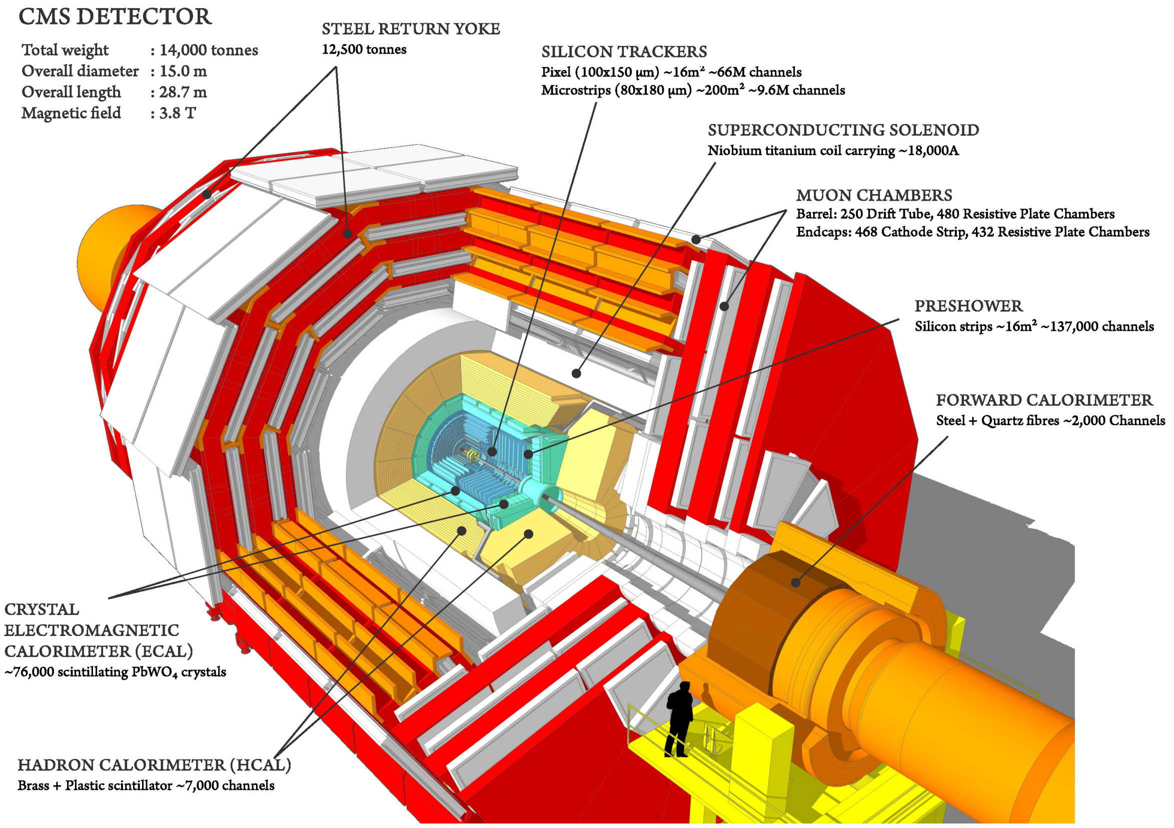

Figure 5: The CMS detector.

4 Methods

4.1 Production of data samples

To select events of the $b\bar{b}H$ with final state $\mu + \tau$ of the Higgs, and produce the data samples, we can use the following criteria:
General criteria plus criteria to select the final state $\tau\tau$ of the Higgs:

- At least one b-tagged jet
- Charge of Muon = - Charge of Tau
- Selection of the distance in the $\eta \times \phi$ plane between $\mu + \tau_h$: $\Delta R = \sqrt{\eta^2 + \phi^2}$: ΔR > 0.5
- No additional muons

- No electrons

Muon criteria:

- Transverse momentum > 23GeV
- $|\eta| < 2.1$
- Relative Isolation < 0.15
- medium ID

Tau criteria:

- Transverse momentum > 30GeV
- $|\eta| < 2.3$
- Tight working point of MVA isolation
- pass Tau ID decay mode

With the data samples I could search for a good variable to separate signal and background.

4.2 Separating signal and background

Using the simulated samples I could look into several variables and decide which one would separate in the best way the signal from the background. In the histograms below, three different scenarios were analyzed: the final state of the Higgs ($\mu + \tau$) (Fig. 6), the first object of the decay (μ) (Fig. 7) and the second object of the decay (τ) (Fig. 8). In each case, four different variables were analyzed: the mass (M), the total transverse momentum (P_t), the pseudorapidity (η) and the phi angle (ϕ).

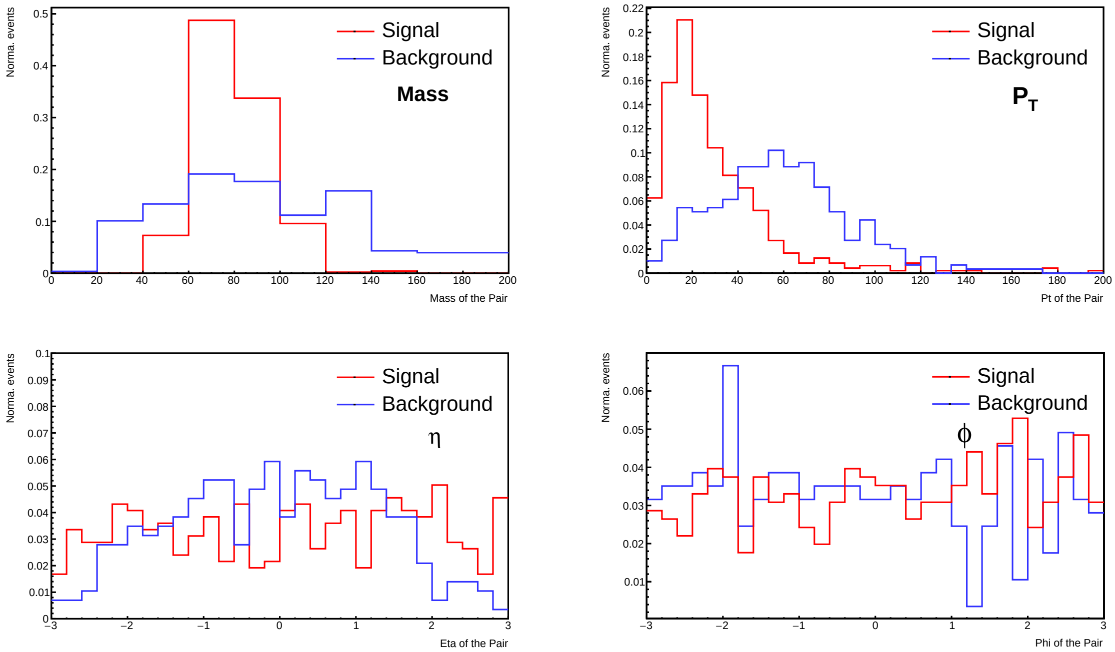

Figure 6: Variables for the combination of the final state.

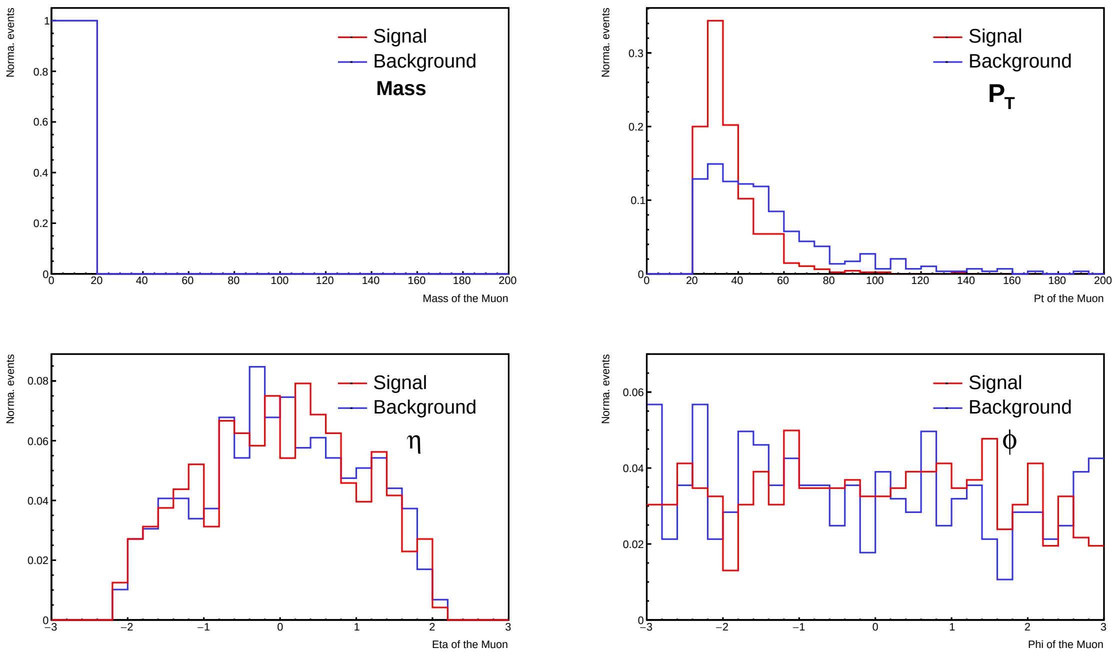

Figure 7: Variables for the μ.

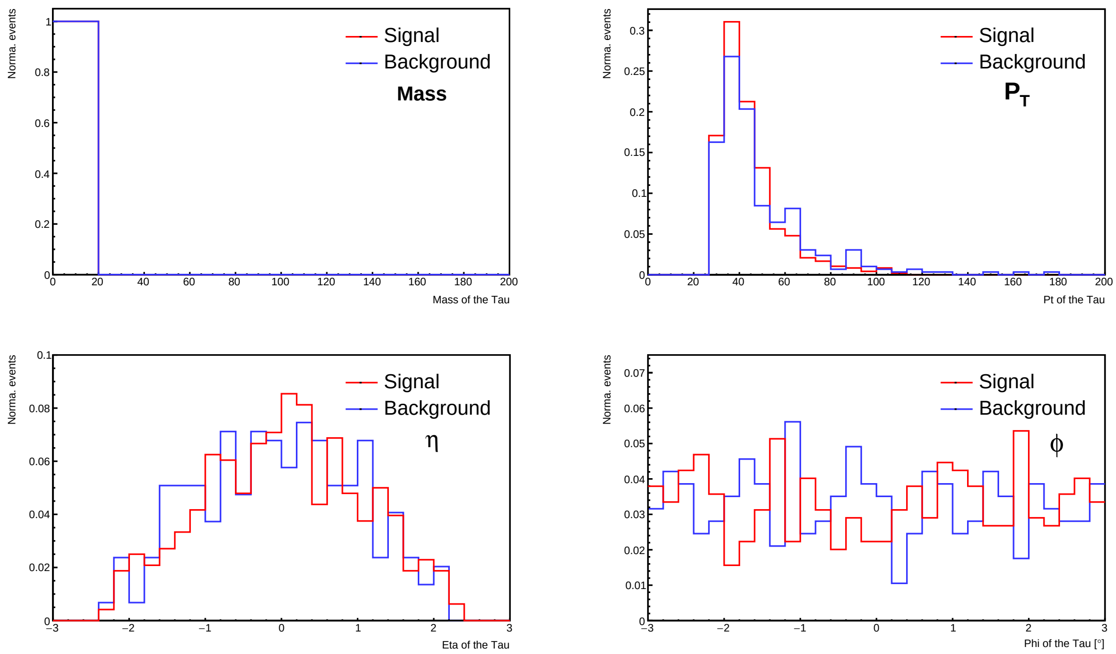

Figure 8: Variables for the τ.

It is possible to see that the best variables that could better separate signal and background are the mass of the pair and the total transverse momentum of the pair. All the other variables have similar shapes for signal and background. Although, this is not a precise way of choosing the best variable, since we are using only the judgment of our vision to do that. A more precise way to do it is to calculate the area of intersection for each of these variables and proceed the analysis with the one that has the lowest percentage of intersection between signal and background. The percentage of intersection for each of the 12 variables are presented in Table 1.

Table 1: Percentage of intersection between signal and background for the pair, the first object and the second object.

Variable	Pair %	Muon %	Tau %
Mass	51.4	100.0	81.9
P_t	51.1	67.8	90.2
η	85.0	85.8	80.1
ϕ	77.0	96.1	96.1

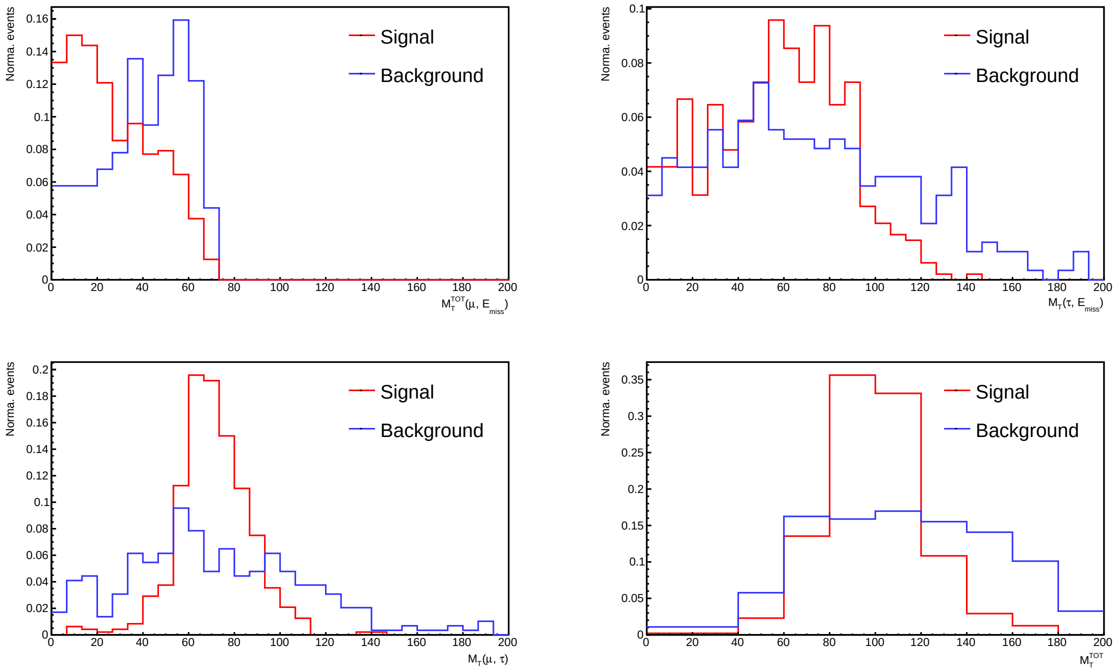

Figure 9: Variables for the transverse mass.

We can also look to the transverse mass between the Muon plus the missing transverse energy (MET), the Tau plus the missing energy, the Muon plus the Tau and to the total transverse mass in Fig. 9, where missing energy corresponds to the energy related to particles that cannot be directly detected by the CMS, for example neutrinos.

Table 2: Percentage of intersection between signal and background for different combinations of transverse mass.

variable	%
$\mu + MET$	68.5
$\tau + MET$	77.8
$\mu + \tau$	62.3
$Total$	64.1

The variables that have the lowest percentage of intersection between signal and background are the mass of the pair and the transverse momentum of the pair. Since the values of intersection are similar for these two variables I decided to proceed the analysis using the mass of the pair as the variable with the best separation signal/background.

4.3 Normalization of the 2016 CMS MSSM Higgs $\rightarrow \tau\tau$ analysis to the region of High Luminosity of the LHC

The MSSM H$\rightarrow \tau\tau$ analysis [2] searches for a beyond-the-standard model Higgs boson produced in association with 2 b-quarks. This signature is the same as our process of interest, and because the analysis also considered masses around the standard model Higgs boson mass, we can use this analysis to scale the signal- and background predictions to higher luminosities and extract limits on the process of interest.

The variable chosen to separate signal and background by this analysis was the total transverse mass (Fig. 10). To improve the separation of signal and background, we scale the shapes used for limit setting in the analysis we project using the division of the mass of the pair and the total transverse mass (Fig. 11), since the first variable shows a better separation in the region of HL.

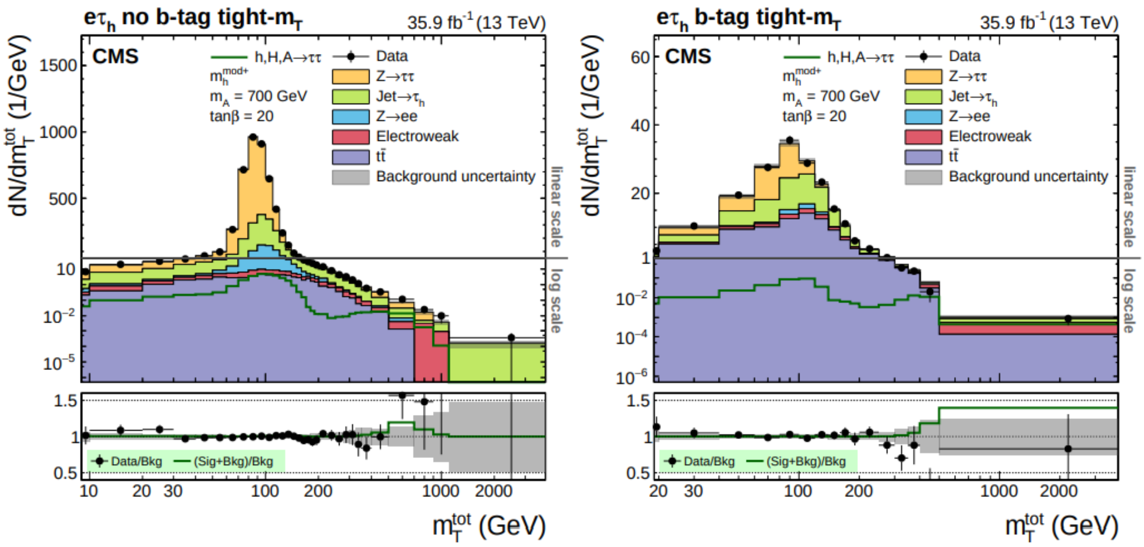

Figure 10: Total transverse mass in two categories of the analysis described above

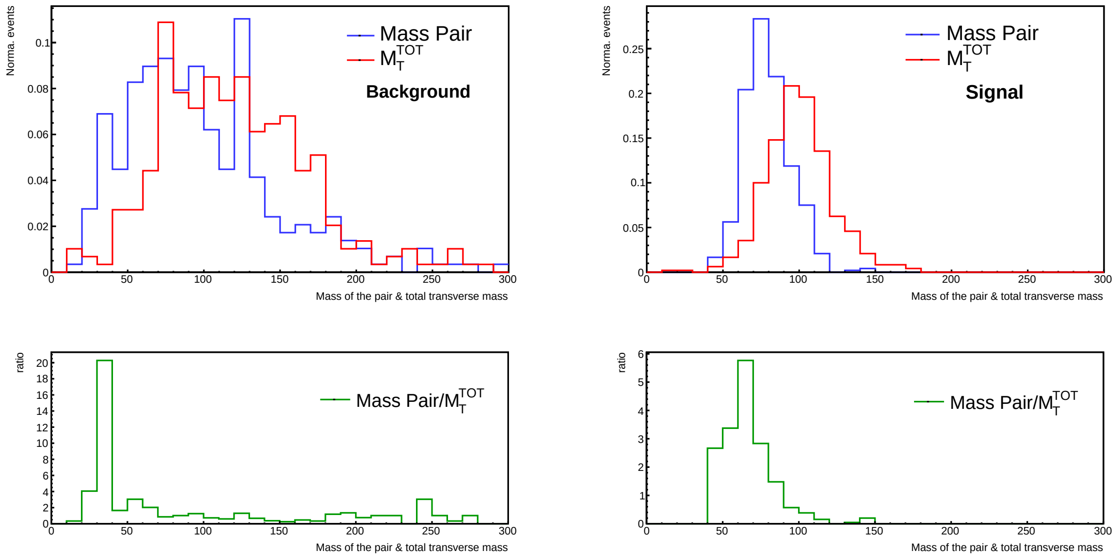

Figure 11: Division of the mass of the pair and the total transverse mass.

In Fig. 11 we have the division of the mass of the pair by the total transverse mass for two possible cases: using the signal data samples or the background data samples. Another possible situation would be maintaining the same shapes from the previous analysis. Then we have a total of three different cases to do the projection:

1. **Background:** $b\bar{b}H$ signal scaled using signal ratio on Fig. 11, all other processes scaled using background ratio on Fig. 11.
2. **Signal:** $t\bar{t}$ background scaled using background ratio on Fig. 11, all other processes scaled using signal ratio on Fig. 11.
3. **Nominal:** same shapes as 2016 analysis.

After choosing the option for the scaling of the shapes of the 2016 analysis, we can also choose one of the following two scenarios for the uncertainties that are going to be used in the limit setting:

1. **Scenario S1:** no changes to uncertainties from the 2016 analysis, however, uncertainties due to the finite size of the simulated samples are ignored.
2. **Scenario S2:**
   - theoretical uncertainties = 50% of their initial value,

- all other experimental uncertainties, for example on the Muon identification efficiency, are scaled down to $1/\sqrt{L}$ until a lower limit is reached.

The Scenario S1 for the uncertainties is bit more pessimistic, while the second one is a bit more optimistic. That said we expect the limits for the channel of interest to be lower in the second case.

5 Results

Given all the possible combinations of scaling for the 2016 analysis and for the uncertainties, we can search for the combination that is going to produce the lowest limit to the channel of interest in the region of HL. The opposite can also be done, i.e. search for the combination that provides the worst scenario for detecting the signal.

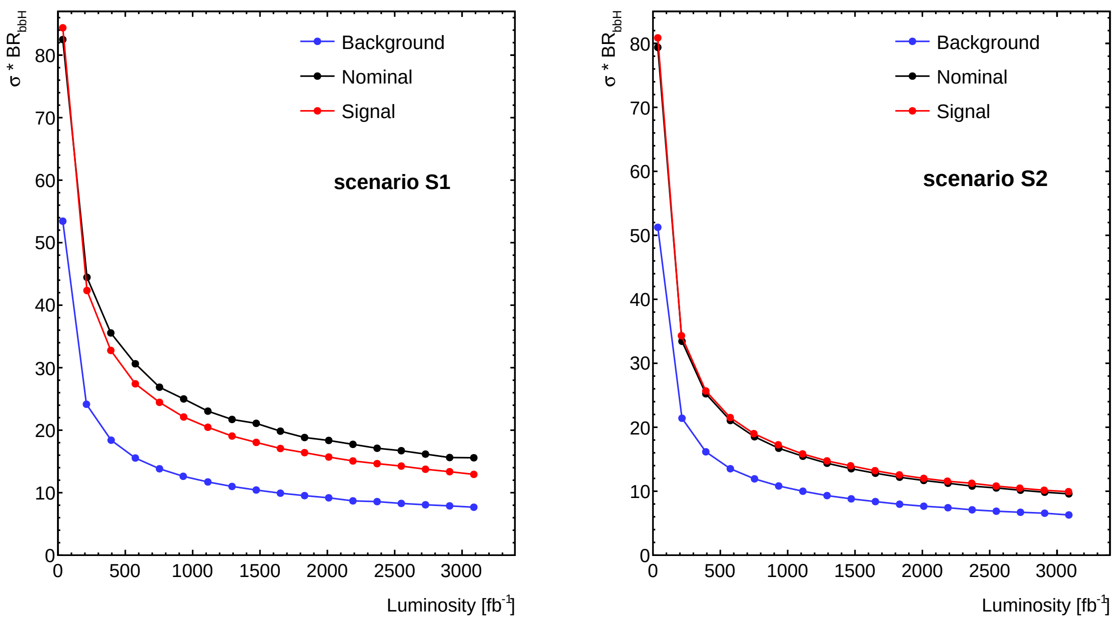

Figure 12: Comparison between three cases for scaling the 2016 analysis and the two cases of uncertainty scaling.

In Fig. 12 we can see that for both options analyzed for the uncertainties scaling the background is the one with the lowest limit. The nominal case and the signal case are the ones with highest limits for the scenario S1 and S2, respectively. In the best case (Background+S2) the limit is 6.5 times the Standard Model prediction for this channel, which is not yet enough to measure a signal.

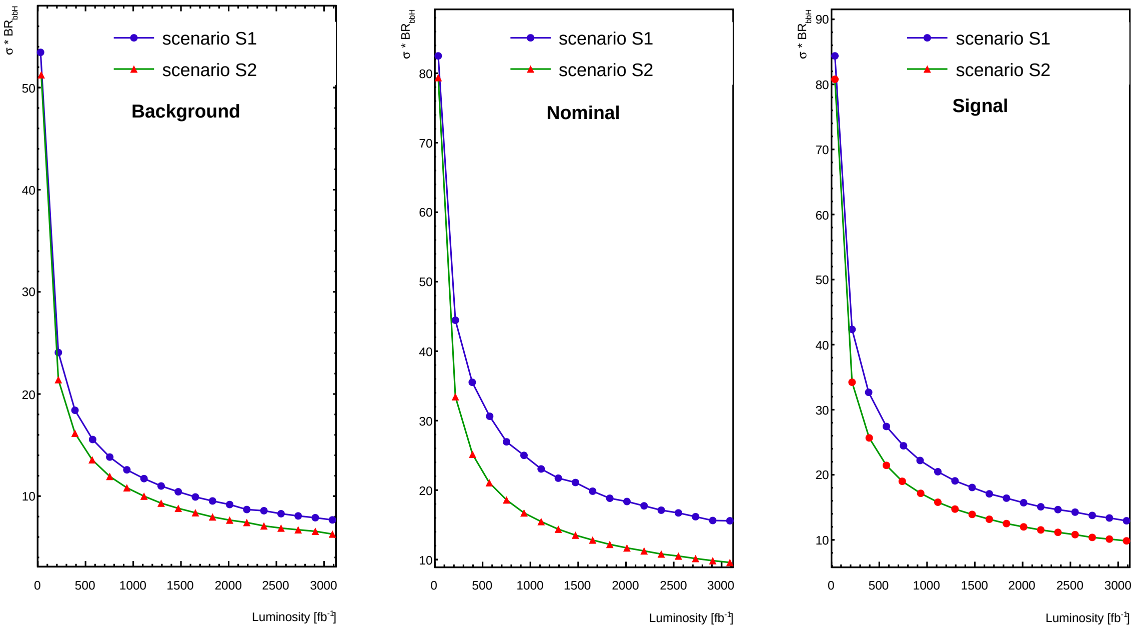

Figure 13: Comparison between scenario S1 and S2 for the uncertainties.

In Fig. 13 we can see that for all the three different cases to scale the 2016 analysis the scenario S2 is the one with the lowest limit, which was expected, since this is the optimistic case.

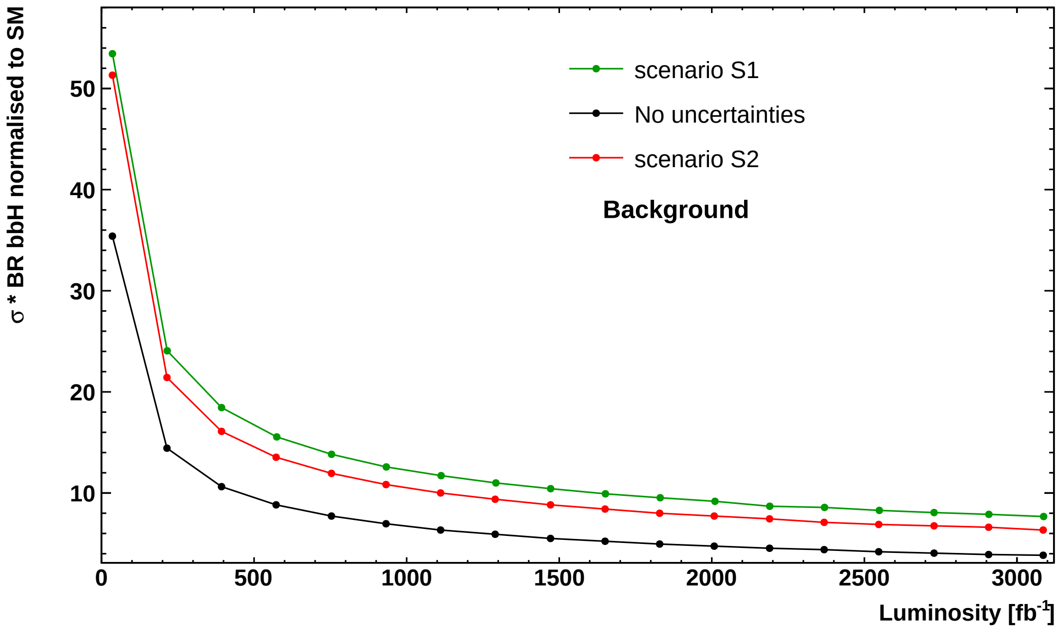

Figure 14: Comparison between scenario S1 and S2 for the uncertainties and the case when we assume no uncertainties.

In Fig. 14 one can see better what is the effect of choosing different scalings for the uncertainties. The black line is the limit as a function of the luminosity when we set uncertainties to zero. As expected, this case has lowest limits than the optimistic scenario (S2).

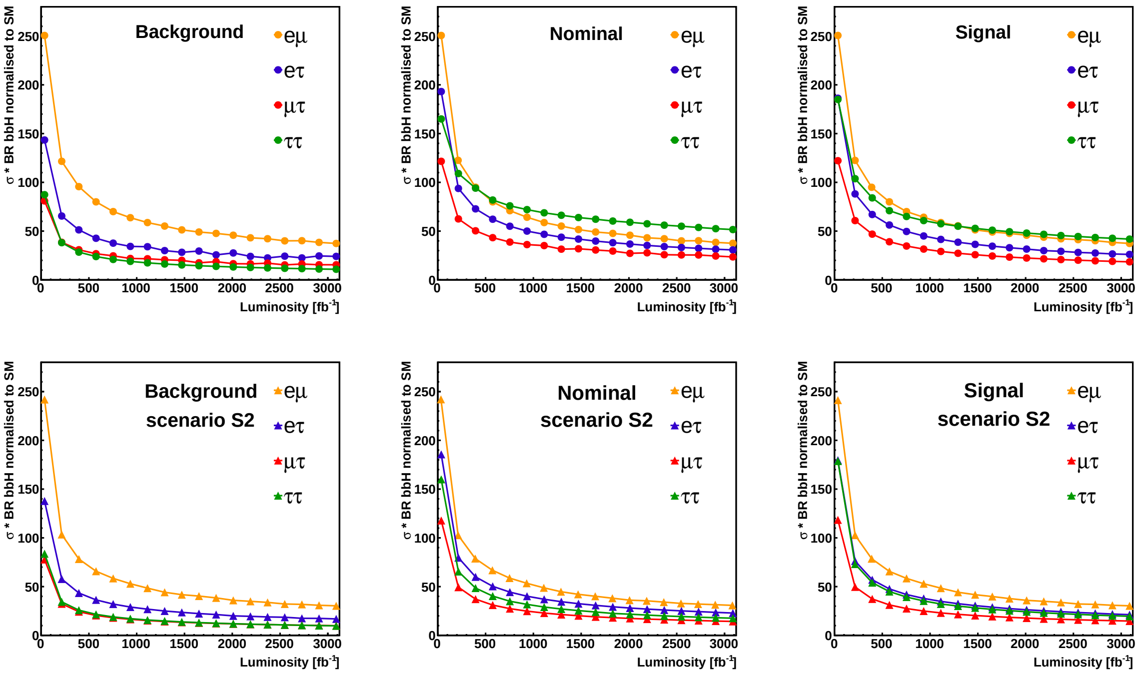

Figure 15: Contribution of different channels to the limits in the six cases of scaling.

In Fig. 15 it is possible to see which final state of the Higgs is most contributing to the limit in each of the six possible combinations for scaling. For the best case, the final state $\tau\tau$ is contributing the most.

6 Conclusions

- A projection study was done to determine the possibilty of the measurement of the production of the Higgs (125GeV) boson in association with a pair of b-quarks in the $\tau\bar{\tau}$ final state in the region of HL of the LHC.
- The best limit we can have is 6.5 times the SM prediction, when we scale using the background and the scenario S2 for the uncertainties.
- There are a lot of different improvements that can be applied to this analysis in order to have a better result, for instance we can choose a better discriminant variable and study this channel of interest in different decay final states.

7 Acknowledgements

First I would like to greatly thank my supervisor, Adinda de Wit, for being so supportive and helpuful throughout the project and ensuring that I would have a nice time doing the project. Mission accomplished! I would also like to thank Olaf Behnke and the other members of the Summer Student organization team for organizing such an amazing project and giving me the opportunity to be part of it. Finally I would like to thank all my Summer Student's colleagues for all the unforgettable moments.

References

[1] CMS collaboration. Technical proposal for the upgrade of the CMS detector through 2020. No. CMS-UG-TP-1. 2011.

[2] CMS Collaboration, Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, CERN-EP-2018-026, arXiv:1803.06553 [hep-ex], 2018