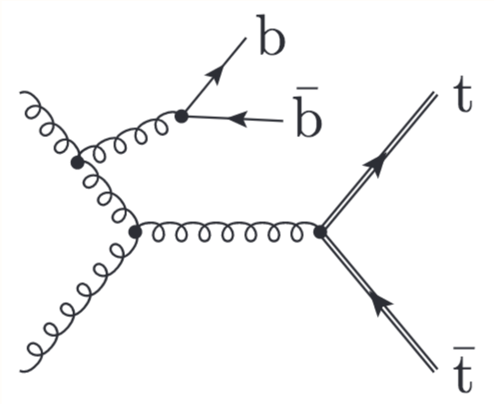
ttbb measurements in the all-hadronic channel

Mikuni, V. M., Canelli, F. ,Wertz, S.

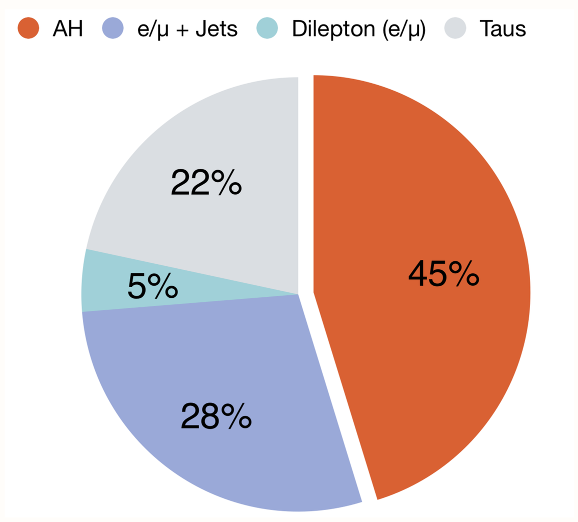
Introduction

The precise measurement of the associated production of **top quarks + b-jets** is an interesting and challenging task. The theoretical description of this process requires an accurate understanding of **Q**uantum **C**hromo**D**ynamics (**QCD**) [1],[2],[3] together with input from real data collected by experiments. Understanding **ttbb** is crucial for gaining insights into the **ttH(bb)** process.



The all-hadronic channel

In the all-hadronic channel, which **both W's** originating from the top quarks, decay to **jets**. This channel has the advantage of being **reconstructible** (all particles can be detected) and having the largest branching fraction. However, this channel is the one with the largest background due to the large number of **Multijet** events at the LHC.



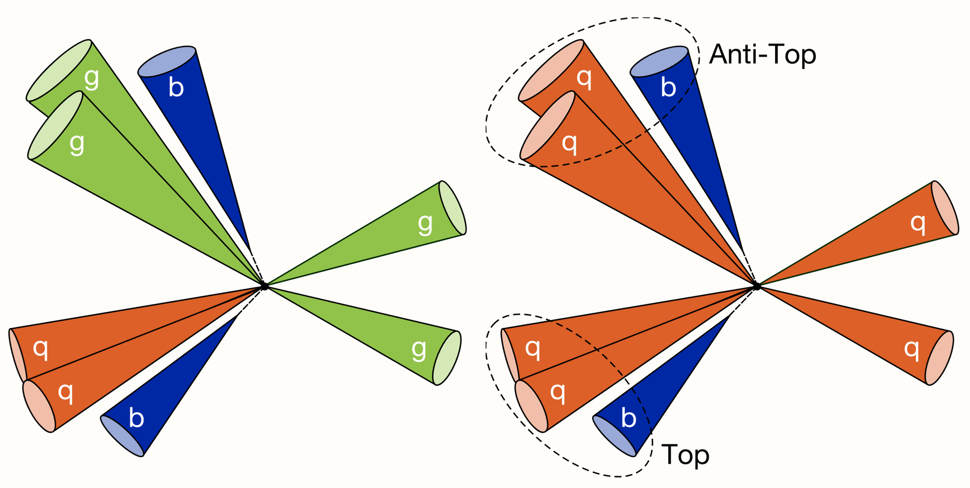
Event Selection

For the measurement we require every event to have:

- At least 6 jets with $p_T >$ **40 GeV**
- At least 2 additional jets with $p_T >$ **30 GeV**
- All jets with $|\eta| <$ **2.4**
- **HT** $>$ **500 GeV**
- At least **2 b-jets**
- **No leptons**

Multijet rejection BDT

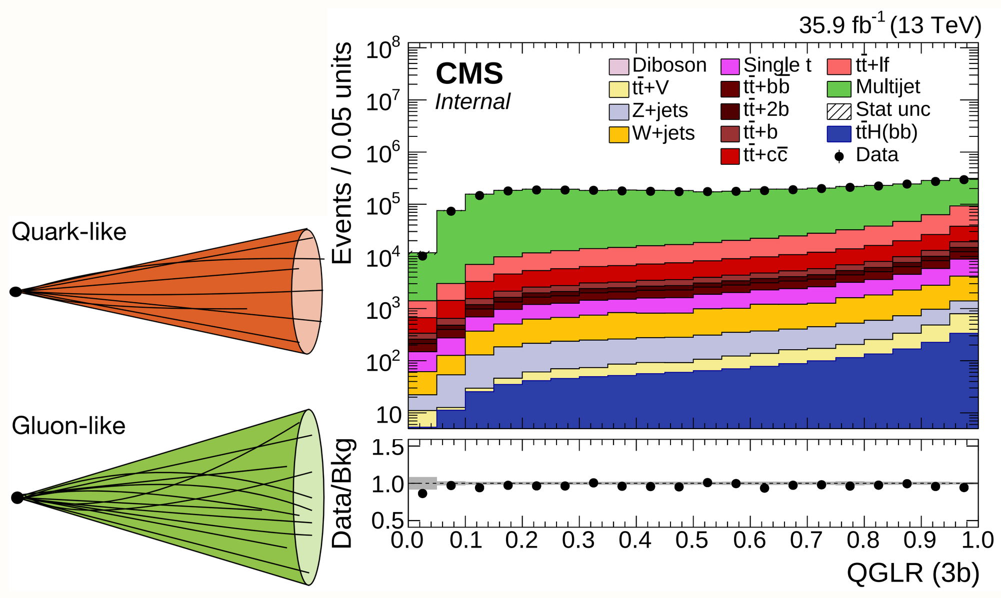
The main background for this analysis is the **Multijet** background produced at the **LHC**.



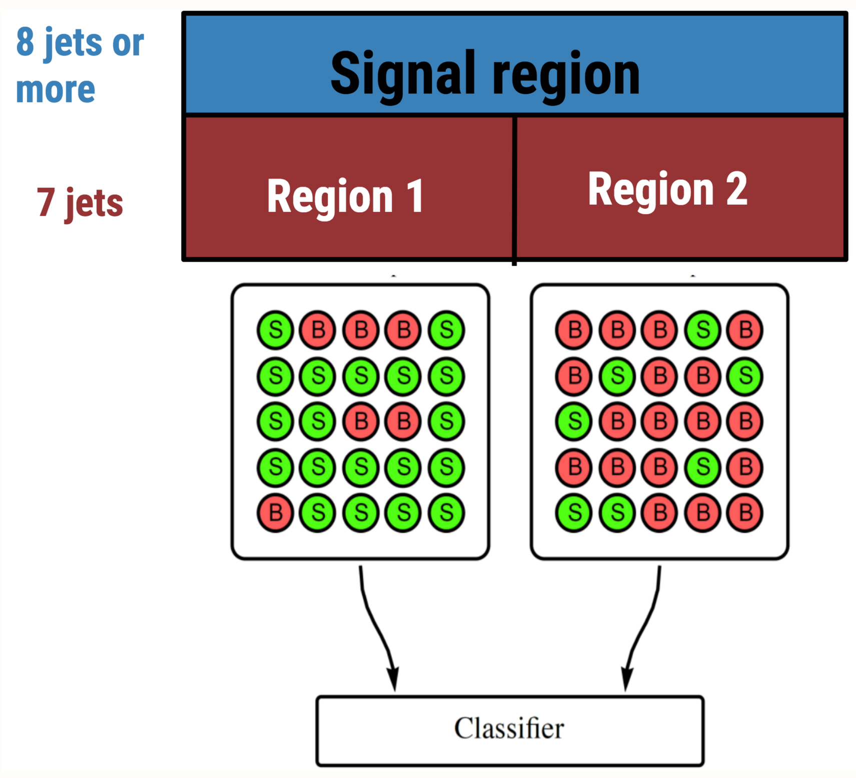
To suppress this background, a **B**oosted **D**ecision **T**ree (**BDT**) is used. Since the simulation for the Multijet background is not perfect, a **data driven** method to train the classifier was used. The method, developed by [4], learns to separate **2 mixed regions** with different signal purities, as proxies to the actual signal and background. The resulting classifier from this training converges to the optimal classifier that separates signal from background.

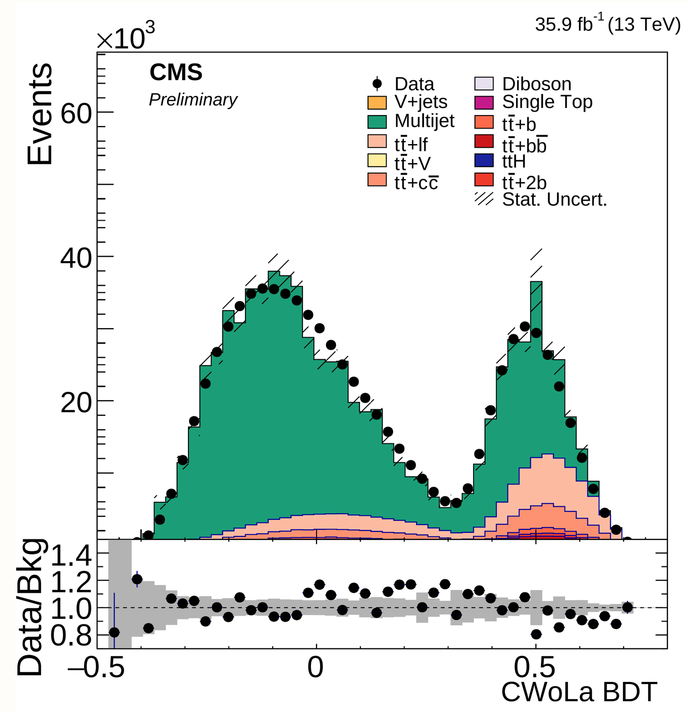
To create the regions, a cut on the **Q**uark-**G**luon **L**ikelihood **R**atio (**QGLR**) is applied. The likelihoods are built by comparing the **2 hypotheses**:

- **4** of the jets in the event originate from quarks ($t\bar{t}$ **like**)
- **0** jets in the event are originated from quarks (**Multijet like**)



To separate the events used for training from the measurement, the region with **exactly 7 jets** was used for the training.

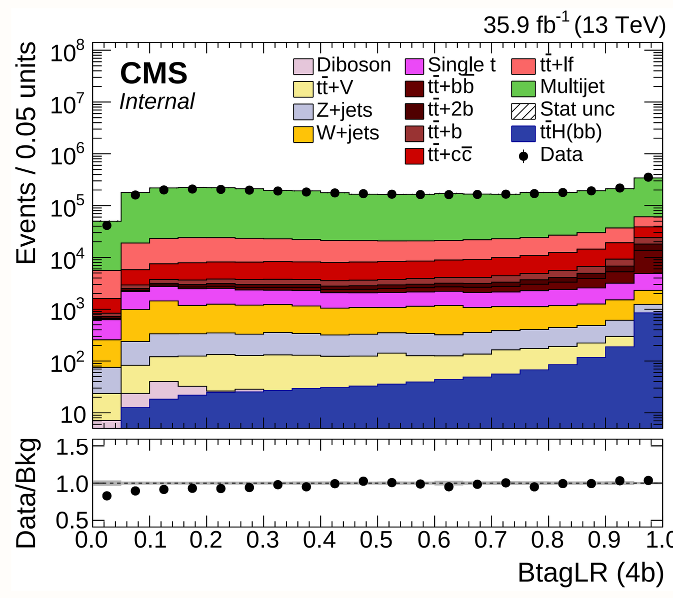




Main observable

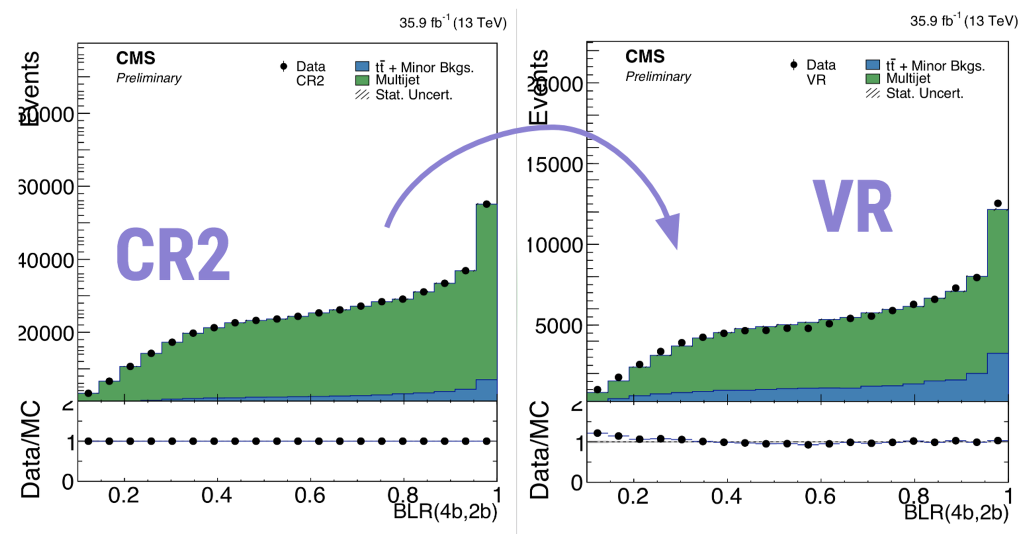
The final measurement is done by fitting the **B**-tagging **L**ikelihood **R**atio (**BLR**). In order to separate the $t\bar{t}b\bar{b}$ component from the other $t\bar{t}+$ **light jets** components a the BLR was built using the **2 competing hypothesis**:

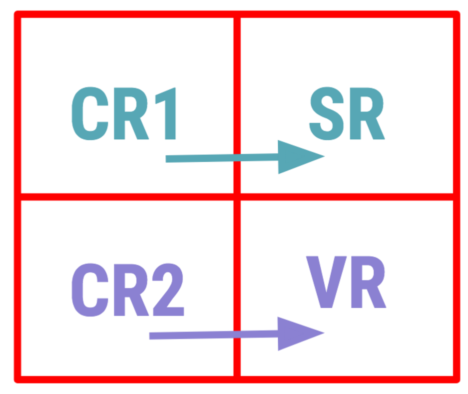
- **4** of the jets are **b-jets** ($t\bar{t}b\bar{b}$ **like**)
- Only **2** of the jets are **b-jets** ($t\bar{t}+$ **light jets like**)



Data driven background estimation

To estimate the **Multijet** background, a data driven approach is used. To assess this contribution in the **S**ignal **R**egion (**SR**) we define a **C**ontrol **R**egion (**CR**). This region has the property of being **independent** and **uncorrelated** with the SR. In this way, we can estimate the Multijet distribution as the difference between the data and the other backgrounds in the CR. To verify if the method works, we use a separate **V**alidation **R**egion (**VR**) with its own **C**ontrol **R**egion (**CR2**).



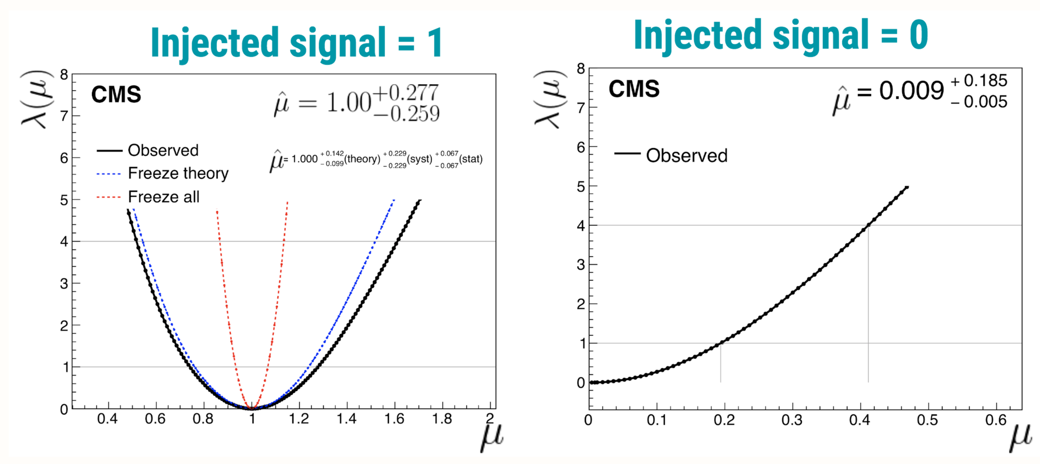


Preliminary Results

Preliminary, blind results for the **ttbb** cross section measurement are obtained by performing a **binned maximum likelihood fit** to the BLR distribution. To calculate the expected significance, a toy study is performed. The quantity measured is the **signal strength** defined as:

$$\mu = \frac{\sigma_{\text{meas}}}{\sigma_{\text{pred}}}$$

where σ_{meas} and σ_{pred} are the measured and the predicted cross sections respectively. By using pseudo-data with different values of μ to the toy study, the distribution of the Likelihood $\lambda(\mu)$ shows an expected sensitivity to be of **30**% for the case of $\mu = 1$, the case of simulation in perfect agreement with the measurement.



Conclusions and future plans

The preliminary results show that the setup developed is working, while also giving an estimate of the expected sensitivity of the analysis. The next step is to finish the measurement using data events. For the future, using more data collected by the **CMS Experiment**, the cross section measurement of this process could also be performed differentially.

References

[1] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, Phys. Rev. Lett. **103**, 012002 (2009)

[2] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, R. Pittau and M. Worek, JHEP **0909**, 109 (2009)

[3] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, JHEP **1003**, 021 (2010)

[4] E. M. Metodiev, B. Nachman and J. Thaler, JHEP **1710**, 174 (2017)

Physics Open Day, University of Zurich, 2018