KSETA Work Report 2022

Rufa K M Rafeek, PhD student, ETP

Title:

- Simultaneous measurement of tt+X (X=bb/cc) processes in the semileptonic channel at the CMS experiment,

- Acceptance Study for early Run-3 Analysis

Supervisor: Ulrich Husemann

Period of KSETA membership: 4 years

Funding: Type, period: DAAD Scholarship, 4 years

Start and (presumable) time of doctorate: October 2021 – September 2024

Main part: Presentation of the PhD project during the above period of membership and research results obtained so far

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors at the LHC.

In the initial part of this report, a brief description of involvement in the scientific exploitation of the LHC collider data taken with the CMS detector designed with the full Run-2 data (2015-2018), is given. The top quark anti-quark pairs ($\bar{t}t$) are produced in association with other particles (X) where X can be the Higgs boson, Z/W boson or QCD-initiated heavy flavour jets (bb/cc). In this analysis, the semileptonic decay mode is used, which contains a single isolated lepton (muon or electron).

The measurement of tt+X processes is important as it is a direct probe of the coupling of standard model particles to the top quark, gives better understanding of the background of many SM as well as Beyond SM processes and can also be a bridge to unravel new physics. The associated production of top quarks with a Higgs boson is important to probe the coupling of Higgs to the top, an up type quark and the heaviest known elementary particle, thus also a step to probe its coupling to fermions. The ttZ process can occur via Z boson radiation from the final state quarks giving access to the coupling of the Z boson to the top quark via final state radiation. The ttH and ttZ processes are considered where H and Z decays to a pair of b-quarks, thus having the same signature, ttbb with at least 6 jets with 4 b-quark jets. The tt pairs can also be produced with additional jets, where the jets can be from quarks or gluons. Here, we focus on events with at least two additional b-jets (ttbb) and two additional c-jets (ttcc). These high jet multiplicity final states create ambiguities in the reconstruction and identification of processes thus making them hard to differentiate from each other.

Signal and background separation is very crucial for any analysis. The signal and background regions are defined by classifying events by their jet and b-jet multiplicities. The major background is tt production associated with the light flavour (LF) jets. The signal, tt+X, seemed to be concentrated more in the region with at least 6 jets with 4 b-jets since the signals demand a large number of b-jets and the similar kinematic features of the b-jet and c-jet leads to the mistagging of a c-jet as a b-jet.

The ttH and ttZ events are separated by different boson decay as they have different initial states, H → bb/cc and Z → bb/cc, thus those events are concentrated in region with at least 6 jets with at least 4 b-jets. The same signal region is applicable to ttbb as it shares the same signature with ttH and ttZ with the bosons decaying to bb.

For ttcc process, the region is selected with a looser cut on b-jets as the process at gen-level is expected to have 2 b-jets and 2 c-jets when compared to the other signals. The c-tagging is more challenging as the properties of the c-jet are in between that of the b and the light jet. Thus, two-dimensional working points are used to separate the c-jets from the b-jets and light flavour (LF) jets. The c-tagging algorithm used has an output structure which predicts the probabilities for each jet to contain one or more c hadrons, one or two b hadrons or no b or c hadrons. This 2D distribution is separated into 3 Working Points (WP) called loose, tight and medium, based on the c-tagging efficiency. The medium is the trade-off WP as it has a better c-tag efficiency and lower b hadron and light mis-tag efficiency when compared to the other WPs. Thus, the ttcc contributions are enriched in region with at least 6 jets and 2 b-jets. An additional requirement on c-jets, that is, the region with at least 6 jets, 2 b-jets and at least 2 c-tight tagged jets result in more contribution from ttcc with less background contamination from ttbb and ttLF processes. The performance of these regions with c-tags to the analysis has to be investigated more and will be done in the later part of the analysis.

Acceptance Study for early Run-3 Analysis

The LHC has started its third data taking period (Run-3) in July 2022. The measurement of the cross sections of the Drell-Yan (DY) process, $\bar{q}q \rightarrow Z/W$ to a pair of leptons, with the Run 3 commissioning dataset is the first showcase of CMS Run 3 analyses. The theoretical prediction of cross sections and their uncertainties are essential for the modeling and interpretation of this dataset.

The goal is to measure the inclusive and fiducial cross sections of the Drell-Yan (DY) process and the theory predictions are used for calculating the cross sections of Z and W bosons at the centre of mass energy of 13.6 TeV to get the acceptance and compare different models for uncertainty.

The Drell-Yan Turbo (DYTurbo) tool is used to calculate and compare the Z and W cross sections at different orders of QCD perturbation theory by transverse momentum (qT) resummation. DYTurbo calculates the cross section at the NNLO+NNLL (next-to-next-leading-order and next-to-next-leading logarithmic) accuracy with leading-order (LO) QED. The cross section is calculated as a function of qT , rapidity, y and the invariant mass, m of the lepton pair. For the inclusive cross section calculations, a Fortran based tool called, Fully Exclusive W and Z Production (FEWZ) is used for the theory prediction (with scale and PDF uncertainties). It calculates the DY process cross section at next-to-next-leading-order (NNLO) in QCD perturbation theroy, no resummation, but with NLO (next-to-leading-order) QED corrections. The theoretical uncertainties associated with muon and electron channel acceptance are also studied. The comparison between these tools is performed to identify the tool which provides the best prediction of the SM cross section.

For the inclusive cross section calculation, the lepton pairs are selected with invariant mass (mll) between 60 GeV and 120 GeV. For acceptance calculations, the scale and PDF uncertainties are calculated using MadGraph by counting the number of generated events, the resummation uncertainty is taken from the kinematic acceptance with MADGRAPH5a MC@NLO and is compared to the DYTurbo acceptance predictions with NLL (next-to-leading-logarithmic) accuracy matched with NLO accuracy. The fiducial region for Z boson acceptance is defined with leptons with pT > 25 GeV and a pseudorapidity of $|\eta| < 2.4$ for the muon channel and $|\eta| < 2.5$ for the electron channel and dilepton invariant mass, 60GeV < mll < 120 GeV. The accep-

tance comparison within DYTurbo between NLO+NLL and NNLO+NNLL tells us whether we should apply a correction for the higher order corrections on the acceptance.
The MG sample is generated at NLO+LL. The comparison between MG and DYTurbo tells us whether resummation makes a significant difference and whether there should be a correction. The acceptance obtained from MadGraph has no visible difference with that calculated using DYTurbo at NLO+NLL accuracy. The total cross section is calculated using DYTurbo and FEWZ at NNLO+NNLL accuracy and compared. The value calculated using DYTurbo is found to align with that from FEWZ result and has no significant change in the cross section observed with respect to the electroweak parameters.
Thus, since the differences turned out negligible, we can take the acceptance from MG, and use the cross section prediction from DYTurbo NNLO+NNLL for our analysis.

Participation in the following courses/events of the school:

- KSETA Topical courses
- Plenary workshop conducted by KSETA at Durbach, March 2022

Participation in conferences, conventions, etc. in Germany and abroad

- DPG Spring Conference 2022, Heidelberg
  Presented a talk on the updates of my ongoing analysis.
- Participated in the CMS induction course at CERN, June 2022
- Attended the FSP meeting (CMS), Aachen, September 2022