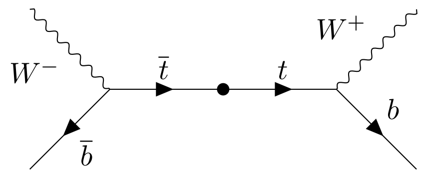
Top quark pair production and W boson charge asymmetry

EPP report, June 2021

Luca Naterop

1 Introduction

Top quarks, initially observed at the Fermilab Tevatron collider [1], can be produced at $\sqrt{s} = 7$ TeV at the LHC [2]. Almost always, they decay via the electroweak interactions to W bosons and b quarks. Such an event can be pictured as



(1)

where the W bosons can decay to $q\bar{q}$ or to a lepton and a neutrino. In this report, we analyze 50 pb^{-1} data recorded from the Compact Muon Solenoid (CMS) experiment in order to measure the $t\bar{t}$ cross section in the semileptonic (muon) channel, which is done in section 2. In section 3, we measure the W boson charge asymmetry as a function of the isolated muon rapidity.

The full python3 code that was used to produce the results in this report is located in the github repository at `https://github.com/lucnat/epp-project`.

2 Top production cross section

2.1 Event selection and signal reconstruction

In order to estimate a cross section from the data, the first step is to discriminate events which stem from the process of interest

$$t\,\bar{t} \;\rightarrow\; \mu\,\nu\,b\,\bar{b}\,q\,\bar{q} \tag{2}$$

from background events, which consists of events with a similar signature. To achieve this, we use tagged Monte-Carlo (MC) data and plot these as a function of a number of discriminator variables, on which cuts are applied. In the present case, an event is treated as a signal event if there is exactly one isolated muon with transverse momentum $p_T > 25$ GeV, where an isolated muon is one for which the relative isolation $I_{\text{muon}}/p_T < 0.03$. Additionally, in order to eliminate Drell-Yan background, we ask for at least four hadronic jets, of which at least one must have $p_T > 25$ GeV and at least one must be B-tagged. This still leaves a considerable amount of QCD background, most of which is eliminated by also requiring a minimum missing transverse energy of $\not{E} > 20$ GeV. Finally, events are required to be triggered using the variable triggerIsoMu24. The resulting signal and background after the selection is shown in figure 1. The number of events is estimated as $N_{\text{sig}}/a\epsilon$, where N_{sig} is the number of signal events passing the selection, a is the acceptance and ϵ the trigger efficiency. Acceptance is estimated by applying our to MC data and taking the ratio of signal events before and after applying the cuts. Trigger efficiency is the ratio of of triggered and untriggered signal events while imposing the cuts.

2.2 Uncertainties

Bins in histograms are assumed to be independently Poisson distributed. Therefore bins of height N are assumed to have error $\sqrt{N}$, and the error of an integrated histogram is obtained through Gaussian error propagation from the bins to the integral. For acceptance, the error is given through propagation from error of the counted events passing the selection and the error of the total $t\bar{t}$ signal which is $\sqrt{7928.61}$. The error on the trigger efficiency also stems from Poisson-distributed bins.
To obtain another estimation of the error for acceptance and trigger efficiency, one can also assume a binomial process. For the MC data this is applicable since there is a fixed number of trials, which can be assumed to pass with a certain probability. The number of successes in fixed amount of trials then follows a binomial distribution. For the trigger efficiency with m out of n events passing, one then obtains the binomial error estimate

$$\sigma_\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{n}} \tag{3}$$

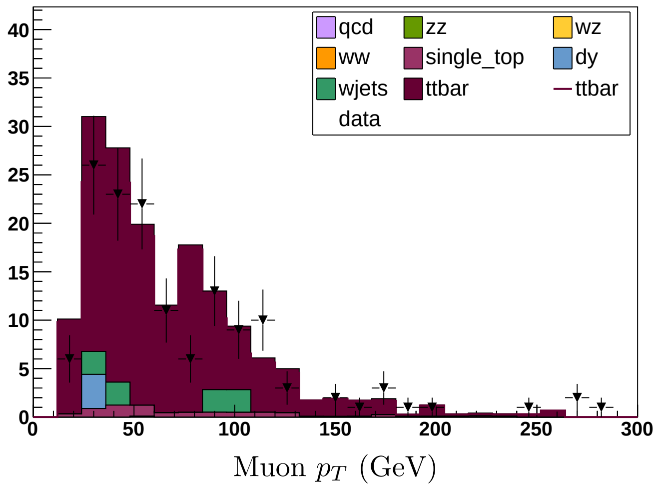


Figure 1: The distribution of muon p_T events passing the selection criteria

where $\epsilon = m/n$ is the expectation value.
In this report, the error was compared with the Poisson error and as a conservative estimate the larger error was used for further analysis (which was always the Poisson error).
To conclude, the statistical error of the resulting cross section is obtained from the error on events, acceptance and trigger efficiency. As the only source of systematic error, a 10% relative error on luminosity is assumed. Since the cross section is multiplicative in the luminosity, it will have the same relative systematic error.

2.3 Results

The measured cross section is $\sigma = 173 \pm 21(\text{stat}) \pm 17(\text{sys})\,\text{pb}$ which agrees well with the theory prediction [3] of $\sigma_{\text{pred}} = 174 \pm 12$ pb. The number of observations and background events are $N_{\text{obs}} = 141 \pm 12$ and $N_{\text{bkg}} = 50 \pm 5$ respectively.
The acceptance is $a = 15.9 \pm 0.8$‰ (per mille), which results from 125 ± 6 out of 7929 ± 89 accepted MC signal events. The trigger efficiency is $\epsilon = 88.2 \pm 2.7\%$ based on 126 ± 6 out of 142 ± 6 events passing the trigger criteria.

3 W boson charge asymmetry

3.1 Introduction

In this part, the W boson charge asymmetry in the muonic channel of W production is measured. The charge asymmetry $\mathcal{A}$ can be defined as

$$\mathcal{A} = \frac{N_+ - N_-}{N_+ + N_-} \tag{4}$$

where $N_{+,-}$ are the number of events with a $\mu^{+,-}$ respectively. Since we expect a dependence from muon pseudorapidity η, we measure $\mathcal{A}(\eta)$ in different intervals of η.

3.2 Event selection and signal reconstruction

The $W \to \mu\nu$ event candidates are selected by requiring at least one isolated muon with $p_T > 25$ GeV and relative isolation (as defined in the last section) < 0.1. In addition, the muon pseudorapidity must fulfill $\eta < 2.0$. Furthermore, we ask for undetected neutrinos by requiring $\not{E}_T > 20$ GeV. This helps to suppress a lot of QCD and Drell-Yan background, since a large $\not{E}_T$ means that the neutrino comes from a primary vertex, and not from somewhere down the line. An overview of the events passing the selection is shown in figure 2 where we show for μ^+ and μ^- the pseudorapidity η as well as the reconstructed W mass. The latter is reconstructed as

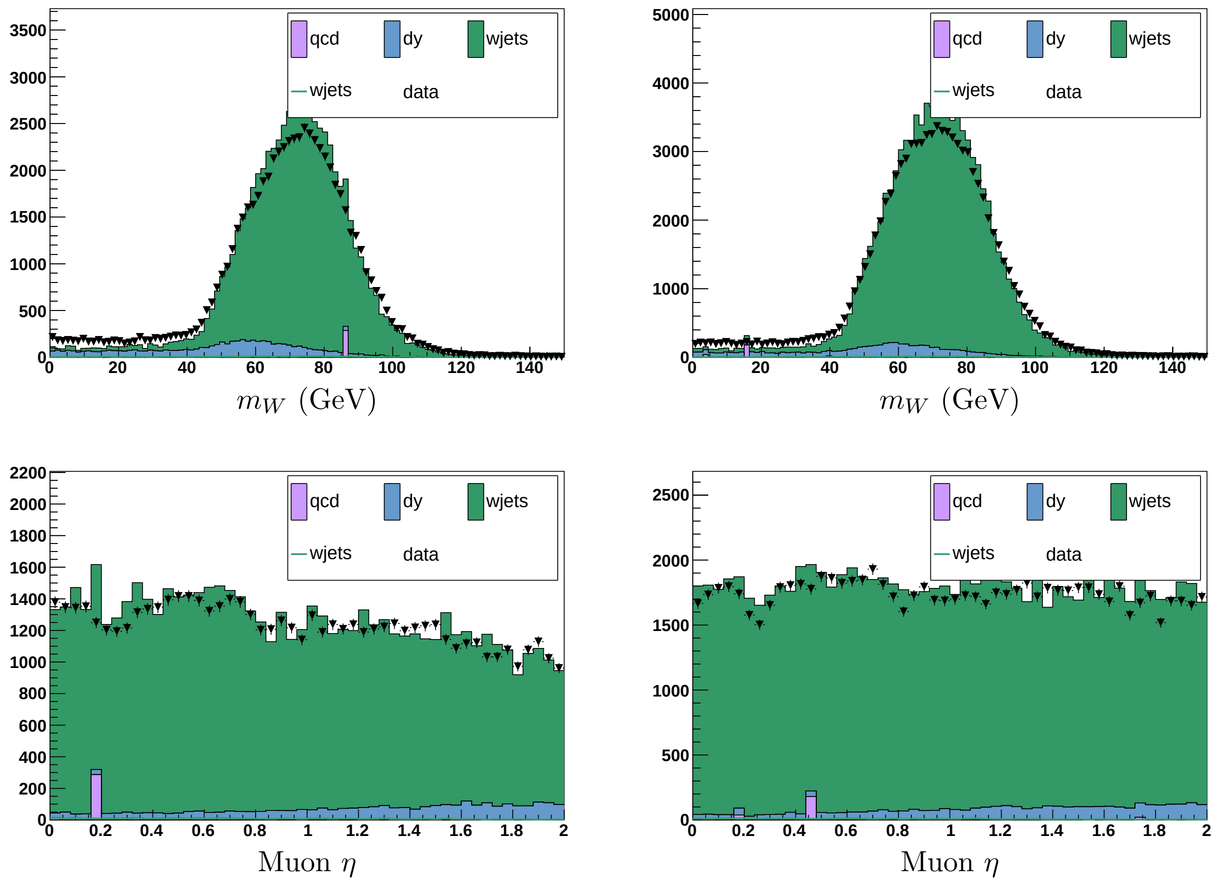


Figure 2: The distribution events passing the selection criteria for μ^- (left) and μ^+ (right). Top: The distribution of the reconstructed W mass m_W. Bottom: The distribution of muon pseudorapidity η.

$$m_W^2 = (p_{\mu,T} + \not{p}_T)^2 \tag{5}$$

where $p_{\mu,T}$ is the transverse muon four-momentum and $\not{p}_T$ the transverse missing four-momentum. In each bin, we account for acceptance and trigger efficiency separately. The estimation of the uncertainties for number of events, acceptances and trigger efficiencies follows the same procedure outlined in the last section.

3.3 Results

The measured $\mathcal{A}(\eta)$ is summarized in figure 3. Since many raw numbers are used for the result, we don't list them here. The error of $\mathcal{A}(\eta)$ is obtained through Gaussian propagation by assuming independence of N_- and N_+. It can be observed that the asymmetry is approximately 15 ± 3 % at

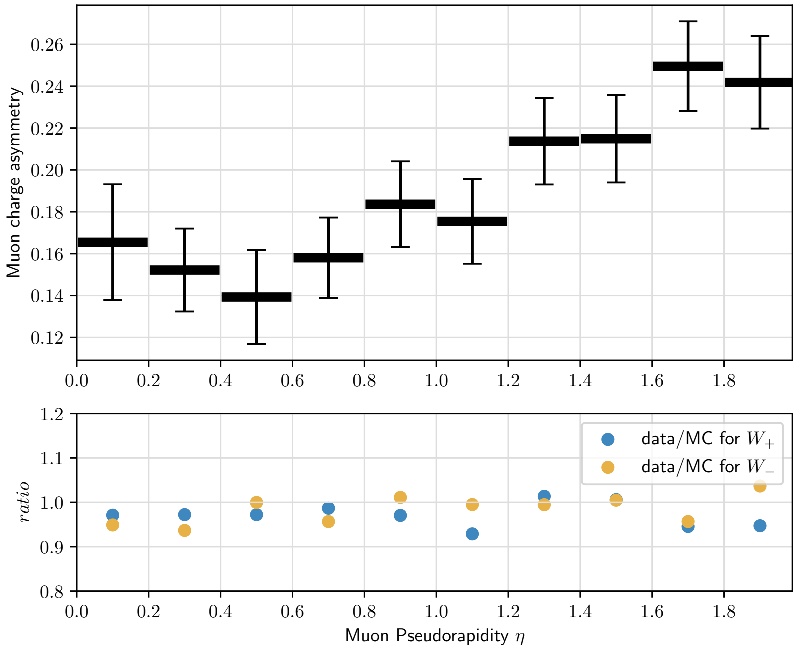


Figure 3: Top: The W boson charge asymmetry $\mathcal{A}(\eta)$ measured in different intervals of η. Bottom: The ratio of data and MC events for W^+ and W^- separately.

small values of η and increases towards approximately 24 ± 2 % for $1.8 < \eta < 2.0$. Due to the large quantity of raw values used to compute the results, we do not list them here, instead we refer to the code used for the analysis which can be found in the project's github repository (see section 1).

3.4 Discussion and Outlook

As it can be observed in figure 3, the agreement between the data and MC events can be considered sufficient for the asymmetry error targeted in this report. To make the comparison more precise, the ξ^2/DoF was computed in each bin, where a maximum value of $\xi^2/DoF = 3.8$ was observed. Nevertheless, there is some discrepancy, for example in the region $1.0 < \eta < 1.2$ for W_+. As for the reason of the discrepancy, it can be speculated that it results most likely from the finite order approximation of the monte carlo data. It should then decrease if one uses a more accurate event generator for the MC data (e.g. one that goes to higher order in perturbation theory).
To deal with this, it would be advantageous to perform a maximum-likelihood fit to separate signal from background with some model. However, the fit should not be performed in the η distribution since both signal and background seem to be more or less linear in that variable. In the η distribution it would therefore be close to impossible to separate signal from background. Instead, one could for example use the reconstructed W mass from the top row of fig 2, where it might be possible to use a Gaussian model for the signal, and perhaps an error function plus Gaussian for the background. One could then compare the two measurements.

References

[1] T.C. Collaboration, *Observation of top quark production in pbar-p collisions*, `arXiv:hep-ex/9503002`.

[2] L. Evans and P. Bryant, eds., *LHC Machine*, *JINST* **3** (2008) S08001.

[3] M. Czakon, P. Fiedler and A. Mitov, *The total top quark pair production cross-section at hadron colliders through* $o(\alpha_s^4)$, `arXiv:1303.6254`.