WW Same-Sign Double Parton Scattering
CMS Experiment

Eleanor Fascione
Supervisors: M. Dünser and G. Petrucciani

Abstract

Data collected by the CMS detector of 13 TeV centre of mass proton-proton collisions, provided during run 2 of the LHC, constitutes an integrated luminosity of $12.9\,\text{fb}^{-1}$. This provides a tool to probe a previously unmeasured process, double parton scattering (DPS) in same-sign WW. Leptonic decay to a muon and associated neutrino was selected for each W boson. In this manner complicated backgrounds were all but negated. A BDT was trained against the most kinematically similar background, and appropriate statistical measures were used to estimate the contribution of fake leptons from W+jets and, to a lesser extent, ttbar. An expected uncertainty on the signal of 52% was extracted, or 2σ significance. An in-progress analysis will determine a limit on the production cross section.

August 19th 2016

1 Introduction

As one of the foremost tools of exploring the Standard Model, the Large Hadron Collider (LHC) uses proton-proton collisions to gain insight into the various processes and particles that constitute the universe. Typical analyses concentrate on single parton interactions within the proton-proton collisions; however, this report focusses on the measurement of a process in which two partons interact simultaneously, producing two distinct hard scatters. The high energy of the LHC events allows for precise measurement of these multiple hard scatterings.

In a naive model of DPS the cross section of any DPS process A+B is given by

$$\sigma_{A+B}^{DPS} = \frac{m}{2}\frac{\sigma_A \times \sigma_B}{\sigma_{eff}} \tag{1}$$

where σ_{eff} is $\approx$ 15 mb as measured by CMS and ATLAS. From this formula σ_{WW}^{DPS} is calculated as 126.2 fb, whereas the cross section obtained from pythia 8 is 147.6 fb. Both these cross sections have W-¿l nu. With only this naive approximation, close agreement to theory can be seen.

Same-sign WW production via DPS is an excellent candidate for exploring DPS due to the relative ease of separating signal from background. By narrowing the analysis to muon and associated neutrino decay only, the complicated and unsightly business of jets associated with W to electron or quark decay can be avoided.

The analysis utilizes data from 13 TeV centre of mass collisions from LHC Run 2, with integrated luminosity $12.9\,\text{fb}^{-1}$. The two main backgrounds are WZ, which looks similar to WW DPS when one of the leptons from Z decay is lost, and fakes from W+jets where one jet fakes a lepton of the same charge. To a lesser extent fakes are also observed from ttbar. A multivariate analysis was performed in order to distinguish WZ from DPS WW, which have numerous similarities and can be difficult to separate. Attention was paid to the significant background of fake leptons by performing an estimation from data.

2 The Experiment

2.1 LHC and CMS Detector

Operating at 13 TeV centre of mass energy, the LHC is a 27 km circular accelerator that produces proton bunch crossings at a frequency of 40 MHz. The byproducts of the high energy collisions during these crossings are analyzed in an attempt to gain insight into the nature of the collisions and the subatomic particles involved.

CMS is a general purpose detector at the LHC. As the name would suggest, a key feature of the apparatus is the superconducting niobium-titanium alloy solenoid, which provides a maximum magnetic field of 3.8 T.

Surrounding the beam pipe while remaining within the bulk of the magnet are, in order of radial distance from the beam pipe, a layered silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter (HCAL). Each of these components is composed of a barrel and two endcaps, allowing the detector to be hermetically sealed.

Finally, high precision measurements of muons is made by a muon chamber system. Three types of gas-filled detectors are encased in the return yoke of the magnet. The combination of tracking and muon chambers results in a momentum resolution of roughly 1% for muons considered in this analysis. For more information see [1] and [2].

Information about the transverse momentum, the projection of particle momentum onto the transverse plane, $\vec{p}_T$, is obtained from the tracking detectors. Conservation of momentum gives $\vec{p}_T^{miss}$, the negative of the vector sum of all particle momenta from a given event. The HCAL provides valuable information about the missing energy in the transverse plane, E_T^{miss}, which is the magnitude of the aforementioned $\vec{p}_T^{miss}$.

Several important variables arise from the geometry of the detector. ϕ, the azimuthal angle, is measured in the xy plane, and the polar angle θ is measured from the positive z direction. Of more use is the pseudo-rapidity, $\eta = -\log[\tan(\frac{\theta}{2}]$. Both η and ϕ will be of use in the following analysis.

3 Event Selection and WZ MVA

There are several basic characteristics that a same sign WW DPS event should exhibit, and these provide the basis for quality cuts on the data. Naturally only events with at least two muons are selected, requiring the leading muon to have transverse momentum greater than 25 GeV, and the sub-leading muon to surpass 20 GeV. If a third lepton is present then it must have a transverse momentum below 5 GeV for muons, 7 GeV for electrons, and 20 GeV for hadronically decaying taus. Leptons are selected using a lepton MVA set at 0.75. The missing transverse energy had to exceed 15 GeV.

At most one jet is allowed, with $p_T > 30$, and $|\eta| < 2.4$, and b-jets are vetoed with $p_T > 25$, $|\eta| < 2.4$, and a loose working point of the CSV discriminator.

A TMAV [3] BDT was trained to distinguish WZ from the DPS signal after a short study on observables in which the two processes differ in shape. Several input cuts were implemented in order to restrict backgrounds from top quark production.

Flavour selections were made for the purposes of the BDT. For the background only same sign lepton pairs were allowed; events with two muons, two electrons, or an electron and a muon were all accepted. This was loosened for the signal to increase the sample size by also allowing opposite sign pairs.

3.1 Observables Sensitive to DPS

The observables most sensitive to DPS WW production were chosen based on their background rejection vs. signal efficiency.

- MT_2 of both leptons and E_T^{miss}
- Transverse momentum of leading muon, $p_T^{\mu 1}$
- Transverse momentum of subleading muon, $p_T^{\mu 2}$
- Missing transverse energy, E_T^{miss}
- $m_T(\mu_1, \mu_2)$
- $m_T(\mu_1, E_T^{miss})$
- $\eta_1 * \eta_2$
- $|\eta_1 + \eta_2|$
- Transverse angular separation between the leading muon and E_T^{miss}, $|\Delta\phi(\mu_1, E_T^{miss})|$
- Transverse angular separation between the leading muon and the subleading muon, $|\Delta\phi(\mu_1, \mu_2)|$
- Transverse angular separation between the vector sum of the muons and the subleading muon, $|\Delta\phi(ll, \mu_2)|$

Where $m_T(a, b) = \sqrt{2 \cdot p_a \cdot p_b \cdot (1 - \cos(\Delta\phi(a, b)))}$ is the transverse mass of objects a and b.

3.2 BDT MVA

MVA analysis using BDT increases sensitivity to DPS by training against WZ, one of the more difficult backgrounds to discriminate from DPS. By training selectively on the WZ background only, rather than on a variety of backgrounds, this ensures maximal rejection against the kinematically most similar background WZ.

The power of the BDT can be summarized by noting that at 55% background rejection 90% signal efficiency is retained.

Figure 1 shows the discrimination power of the trained BDT variable. The WZ background is chosen to be flat and the signal is heavily concentrated towards a value of 1.

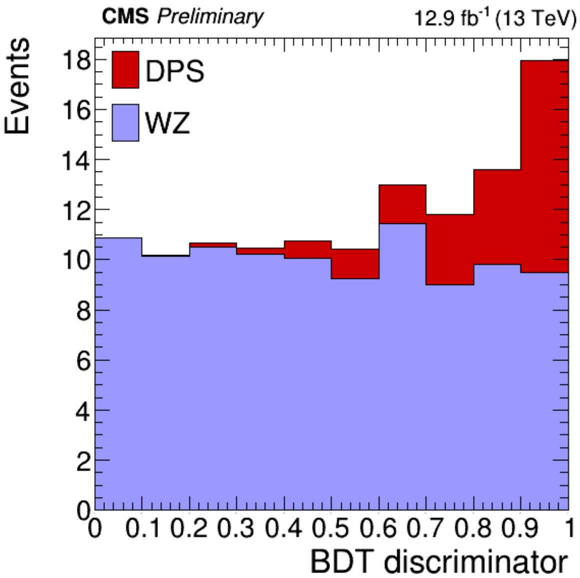

Figure 1: The BDT variable for MC WW and WZ data.

Simultaneously TMVA was used to evaluate a Fischer discriminant and rectangular cuts and the BDT was found to be superior to both. Figure 2 demonstrates each training variable and the difference between signal and background, with signal in blue and background in red.

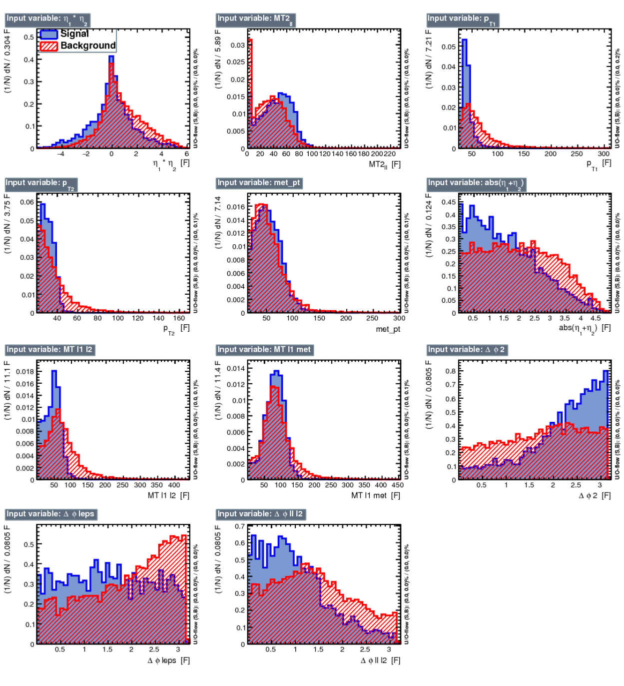

Figure 2: The BDT variable for MC WW and WZ data.

4 Signal Efficiency Estimation

When comparing MC simulation to data, it is necessary to ensure that all processes are properly scaled. It is not sufficient to simply consider the total integrated luminosity and the cross section of a particular process, as there are effects relating to the trigger efficiency and the efficiency of the detector to reconstruct, in this case, a prompt lepton.

4.1 Trigger and Reconstruction Efficiencies

In order to calculate the ratio of the data to MC efficiency, one can look at DY data with two leptons of opposite charge and the standard quality cuts. In this region DY is the only expected background. A fit of the MC is performed to the data in two scenarios: one where both leptons are tight leptons, and another where at least one lepton is tight. In the first case, the ratio of data to MC (i.e. the scale factor) is the ratio of the efficiency of the event, and is given by

$$\frac{\varepsilon_{trigger}\varepsilon_{l_1}\varepsilon_{l_2}}{1\cdot\varepsilon_{M,l_1}\varepsilon_{M,l_2}} = \text{Scale Factor}$$

where the scale factor is that on DY determined by the fitting algorithm, and $\varepsilon(l_i)$ is the lepMVA efficiency of the i-th lepton in data. Similarly, ε_{M,l_i} is the efficiency of the i-th lepton in MC. These are assumed to be equal for both leptons; that is, $\varepsilon_{l_1} = \varepsilon_{l_2}$, and the same for MC. $\varepsilon_{trigger}$ is the efficiency of the trigger, which is not present in MC. This leaves

$$\varepsilon_{trigger}\frac{\varepsilon^2}{\varepsilon_M^2} = \text{Scale Factor}$$

In the second case, the efficiency of the leptons is assumed to be equal again, but only one lepton is necessarily tight. Therefore

$$\frac{\varepsilon_{trigger}[\varepsilon_{l_1}+\varepsilon_{l_2}-\varepsilon_{l_1}\varepsilon_{l2}]}{1\cdot[\varepsilon_{M,l_1}+\varepsilon_{M,l_2}-\varepsilon_{M,l_1}\varepsilon_{M,l2}]} =$$

$$\varepsilon_{trigger}\frac{2\varepsilon-\varepsilon^2}{2\varepsilon_M-\varepsilon_M^2} = \text{Scale Factor}$$

Using the fit to the data and the two equations, the ratio of $\varepsilon/\varepsilon_M$ can be determined. This ratio was determined to be 0.987, and is applied to all MC. The trigger efficiency for data can also be determined from the scale factor.

$$\varepsilon_{trigger} = \frac{\text{Scale factor from Z} \to \text{2 tight leptons}}{0.987^2}$$

4.2 WZ Scale Factor

Knowing these factors now allows for study of a selection on 3 leptons, from which the scale factor for WZ can be determined. With three leptons, the 0.987 ratio must be cubed. Three different cases were examined, triple muon events both with and without the trigger efficiency, as well as muon-muon-electron events with the trigger efficiency as the trigger efficiency for events with three muons is bound to be between that of events with two muons and 100%. From this the median value was taken as the WZ scale factor. This results in 0.954.

The systematic uncertainty was taken as the ratio of the maximum difference between the median and the extremes and the median itself. This gives a final scale factor and uncertainty on WZ of $0.954 \pm 16\%$.

4.3 ZZ Scale Factor

In fitting WZ, one must first ensure that ZZ is scaled properly, as that is still a significant background even in the 3 lepton region. In a 4 lepton control region, ZZ is the only expected background. Several cuts were put in place to select pure ZZ events, namely restricting mZ1 and mZ2 to be within 60 and 120 GeV, and ensuring that the mass of the four leptons from the event be greater than 182.4. The resulting fit gives a scale factor of $1.24 \pm 19\%$, which is expected as the cross section for ZZ is now known to be 20% higher than that used in the production of the MC simulation.

5 Fakes

5.1 Prompt-Fake

In any experiment there is a disconnect between the true physics of an event and what is reconstructed in the detector. For any given event with leptons, each lepton is either 'prompt', meaning that the particle was indeed a lepton, or 'fake', meaning that the particle was not a lepton from bosonic decay. In simulation it is possible to distinguish the prompt from the fake leptons, but in experimental data this is impossible, hence the disconnect between physics and experiment.

To account for this problem, it is not necessary to have an understanding of the fake leptons on a per-lepton basis, but instead to determine an overall 'fake rate', that is, the probability that a fake lepton, given that it passes a series of quality cuts, will pass an MVA cut at a specific level (i.e. it is a tight lepton). The MVA cut is a measure of how 'good' a given lepton is. Most true prompt leptons will have a very high value close to 1, whereas fake leptons will have low or negative values. This provides discriminating power on the lepton events, and thus a basis for determining the fake rate.

Determining the true fake rate for a particular background would require isolating a pure sample of the background; however, this is not necessarily possible in real data. Therefore, it is sufficient to choose a region that is known to be rich in fake leptons, and low in all other backgrounds. For any remaining background in the control region, it is possible to perform a background subtraction using Monte Carlo data.

The measurement region used required one lepton and a single jet away from the lepton with transverse momentum > 30 GeV. $E_T^{miss} < 15$ is also required, and $|\eta| < 2.4$ for both objects. Here there will be some background from W and Z bosonic decay, which will produce true leptons. In order to properly scale the backgrounds to the data, a fit is performed to the transverse mass of the lepton and E_T^{miss} (without the $E_T^{miss} < 15$ requirement, scaling both QCD and the combined W and Z Monte Carlo backgrounds to obtain the best fit to the real data. The scale factor on the bosonic background is then used to subtract the W and Z from the data.

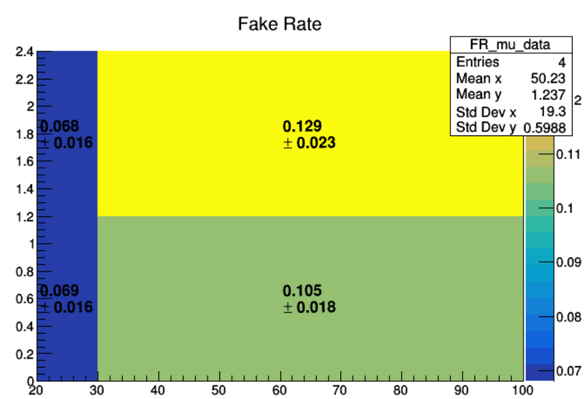

Figure 3: Fake rate map from QCD fake-rich region

The fake rate is then the number of leptons in the region that are tight over all of the leptons. In this manner the fake rate is extracted in four different regions, based on p_T and η of the lepton in question. The η bins separate the barrel and the end caps of the detector, and the p_T bins are based on a plot of the fake rate in MC QCD, which is low between 20-30 GeV and relatively constant between 30-100 GeV. The fake rate map is shown in Figure 3.

Once the fake rate is obtained, it it applied to tight, not-tight events. The leptons are sorted into one of four bins, and are given a weight based on this bin. The weight is

$$w = \frac{f}{1-f}$$

where f is the fake rate in the given bin.

Comparison can be made in simulation, by applying the fake rate derived from QCD simulated events to simulated W+jets events. This prediction is then compared to the W+jets events that are known to have a fake lepton. Good agreement is an indicator that the method works. Fakes from W+jets in truth matched simulation yields 14.93 ± 5.73 events, whereas W+jets with fake rate, estimated from tight-not-tight gives 16.02 ± 1.37, showing strong agreement, and yielding a systematic uncertainty of ≈8%.

To further gauge the systematic uncertainty on the fakes, the scale factor determined on the W and Z backgrounds is varied by 10%, and the p_T of the jet is also adjusted. These adjustments are used to determine an uncertainty on the fake rate map. The upper and lower values are then used to determine limits on the number of fakes, resulting in 10% uncertainty. The final systematic uncertainty on the number of fakes is the 13%.

To ensure that the method used is correct, comparison between data and the MC backgrounds is made by plotting several kinematic variables. As no scale factors had yet been implemented, only the overall shape was important. From both checks it was seen that not only is the method valid, but the relative uncertainty of 13% is sufficient.

6 Results

Once all scale factors and the fake rate have been estimated, a plot of the BDT discriminator can be produced with all significant backgrounds in the signal region, and compared with data up to a value of 0.7 (to remain blinded in the high-sensitivity signal region).

The agreement between simulation and data is very good, which is promising for the future of the analysis when examining the currently blinded region. Table 1 gives the number of events for each simulated background as well as signal that pass the BDT > 0.7 cut, along with the standard quality cuts. Note that ZZ, WWW, and SPS are included in Rare MC.

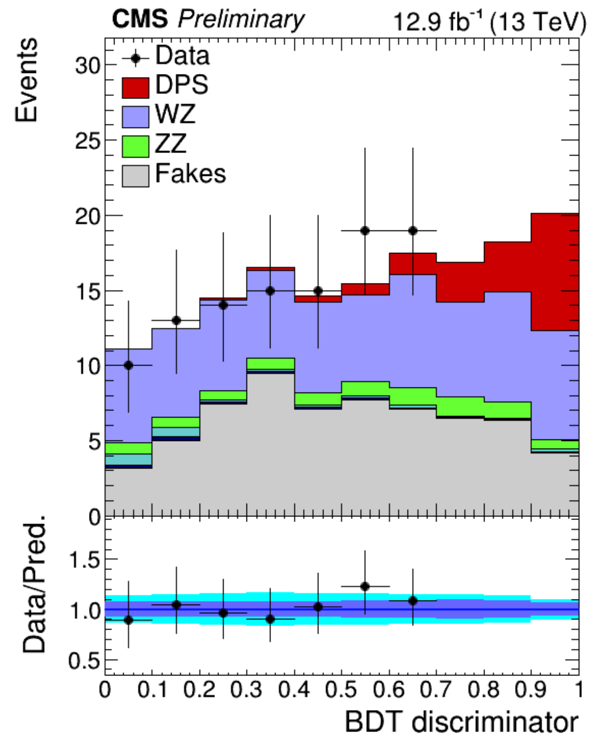

Figure 4: Final BDT for simulation and data (<0.7)

Table 1: Simulation for BDT > 0.7

DPS	13.8 ±	0.6 (stat)			
WZ	20.8 ±	0.7 (stat) ±	3.3 (syst)
Fakes	18.4 ±	2.2 (stat) ±	2.4 (syst)
Rare MC	3.6 ±	0.2 (stat) ±	0.5 (syst)

Table 2: Simulation and data for BDT < 0.7

DPS	2.8 ±	0.3 (stat)	
WZ	43.3 ±	1.0 (stat) ±	6.9 (syst)
Fakes	44.3 ±	2.6 (stat) ±	5.8 (syst)
Rare MC	8.9 ±	0.3 (stat) ±	0.9 (syst)
All MC	102 ±	3.1 (stat) ±	13.3 (syst)
Data	105		

From a fit of the signal and backgrounds (the latter floating within their uncertainty) to pseudo-data, the expected uncertainty on DPS from same-sign WW is 52%, corresponding to a σ of 2. This is a vast improvement on work done in the past by CMS [4], which never reached appropriate sensitivity and was only able to produce a weak limit of σ_{eff}, which is more readily constrained by other processes.

Future plans for the analysis include extending the event selection to muon-electron events, which would increase the statistics but requires a more in-depth study of fake leptons.

7 Conclusion

Same-sign WW DPS in proton-proton collisions was studied using 12.9 fb^{-1} of 13 TeV data. A BDT classifier was trained to discriminate against the kinematically similar WZ background with a missed lepton from Z decay, and the fake rate method was applied to data to estimate the fake contribution from W+jets and ttbar. The reconstruction, ID, and trigger efficiencies were approximated to produce a data:MC scale factor, and other backgrounds were appropriately scaled. An uncertainty of 52% is expected on signal, or 2σ significance.

8 Acknowledgements

Thank you to Marc and Giovanni for being brilliant, knowledgable, and engaging supervisors, from whom I have learned a great deal and whose guidance has led to an analysis with exciting prospects. It has been an absolute pleasure working with you; the past two months have been invaluable.

I'd like to extend my sincere gratitude towards CERN for hosting me and providing an opportunity to explore the frontier of high energy physics as a part of their Summer Student Program, and to the Institute of Particle Physics (IPP) for selecting and supporting me throughout this experience.

References

[1] CMS Collaboration. The CMS experiment at the CERN LHC. JINST, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[2] CMS Collaboration. Physics technical design report, vol. 1: Detector performance and software. CERN/LHCC, 2006-001, 2006

[3] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, “TMVA: Toolkit for Multivariate Data Analysis,” PoS A CAT 040 (2007) [physics/0703039].

[4] CMS collaboration, Chatrchyan, S., Khachatryan, V. et al. J. High Energ. Phys. (2014) 2014: 32. doi:10.1007/JHEP03(2014)032