Standard Model is best model (working title)

William Kennedy DiClemente

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of The University of Pennsylvania

in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2019 Last compiled: January 1, 2019

I. Joseph Kroll, Professor, Physics

Supervisor of Dissertation

Joshua Klein, Professor, Physics

Graduate Group Chairperson

<u>Dissertation Committee</u>

(Committee Prof. 1), Professor, Physics

(Committee Prof. 2), Associate Professor, Physics

(Committee Prof. 3), Professor, Physics

(Committee Prof. 4), Professor, Physics

I. Joseph Kroll, Professor, Physics

Standard Model is best model (working title)

Acknowledgements

I'd like to thanks the Ghosts of Penn Students Past for providing me with such an amazing thesis template.

ABSTRACT

Standard Model is best model (working title)

William Kennedy DiClemente

J. Kroll

This is the abstract text.

Contents

List of Tables

List of Figures

Preface

This is the preface. It's optional, but it's nice to give some context for the reader and stuff.

Will K. DiClemente

Philadelphia, February 2019

CHAPTER 1

Introduction

The Standard Model (SM)[1] has been remarkably successful...

[1] Here's a footnote.

Chapter 2

Theoretical Framework

(Some example introductory text for this chapter)...

2.1 Introduction to the Standard Model

Modern particle physics is generally interpreted in terms of the Standard Model (SM). This is a quantum field theory which encapsulates our understanding of the electromagnetic, weak, and strong interactions...

2.2 Electroweak Mixing and the Higgs Field

When the theory of the electroweak interaction was first developed [2, 3], the W and Z bosons were predicted to be massless (a typical mass term in the Lagrangian would violate the SU(2) symmetry). However, these were experimentally observed to have masses...

Chapter 3

LHC and the ATLAS Detector

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [4] is...

3.2 The ATLAS Detector

ATLAS is a general-purpose particle detector...

3.2.1 The Inner Detector

The Inner Detector serves the primary purpose of measuring the trajectories of charged particles...

3.2.1.1 Pixel Detector

The Pixel detector consists of four cylindrical barrel layers and three disk-shaped endcap layers...

3.2.1.2 Semiconductor Tracker

The Semiconductor Tracker uses the same basic technology as the Pixels, but the fundamental unit of silicon is a larger "strip"...

3.2.1.3 Transition Radiation Tracker

The Transition Radiation Tracker is the outermost component of the ID...

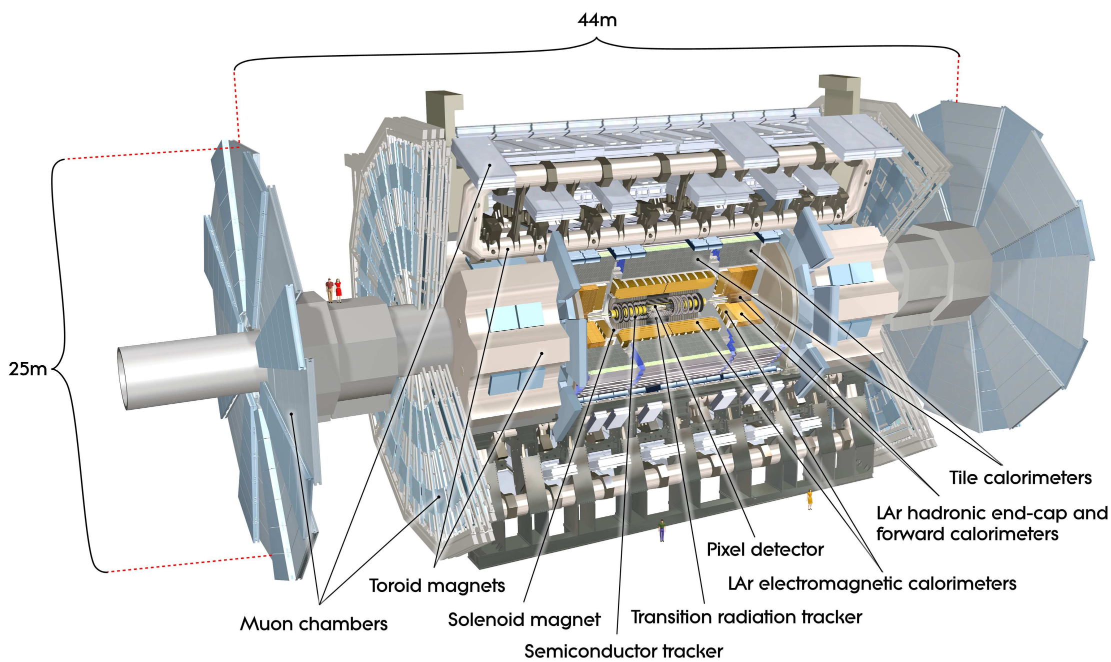

Figure 3.1: General cut-away view of the ATLAS detector [5].

3.2.2 The Calorimeters

ATLAS includes two types of calorimeter system for measuring electromagnetic and hadronic showers. These are the Liquid Argon (LAr) calorimeters and the Tile calorimeters. Together, these cover the region with $|\eta| < 4.9$...

3.2.2.1 Liquid Argon Calorimeters

The Liquid Argon system consists of...

3.2.2.2 Tile Calorimeters

The Tile calorimeter provides coverage for hadronic showers...

Chapter 4

Alignment of the ATLAS Inner Detector

In order for the subdetectors of the ID to operate at their designed precisions, it is essential that the locations of the sensors be known as precisely as possible. Differences between the expected and actual positions of a sensor can result in displaced particle hits and degrade track reconstruction quality. These misalignments can occur for any number of reasons, including but not limited to elemnts shifting during maintenance periods or cycles in ATLAS's magnetic field, or simply small movements during normal detector operations. Since it is not practical to physically realign hundreds of thousands of detector elements to μm precision by hand, an iterative track-based alignment algorithm is used to determine the physical positions and orientations of these elements [6]. The effects of misalignments and the steps taken to correct and monitor them are detailed in this chapter.

4.1 Effects of Misalignment

Hello world!

4.2 The Alignment Method

Hello world!

4.3 Momentum Bias Corrections

Hello world!

4.4 Alignment of the IBL

Hello world!

4.5 Alignment Monitoring

Hello world!

Chapter 5

Same-sign WW @ $\sqrt{s} = 13$ TeV

5.1 Theoretical motivation

Hello world!

5.2 Signal definition

Hello world!

5.3 Background estimations

Hello world!

5.4 Cross section measurement

Hello world!

Chapter 6

Prospects for same-sign WW at the High Luminosity LHC

On December 3, 2018, Run 2 of the LHC officialy ended, and the collider was shut down to begin the first of two scheduled extended maintenance periods [7]. During these two long shutdowns, the Phase-I and Phase-II upgrades of the LHC and ATLAS will occur in order to prepare for the High-Luminosity LHC (HL-LHC) which is scheduled to begin operation in 2026 [8].

The HL-LHC is planned to run at a center-of-mass energy of $\sqrt{s} = 14$ TeV with an instantaneous luminosity of $\mathcal{L} = 5 \times 10^{34}$ cm^{-2}s^{-1} with up to 200 collisions per beam-crossing. Over the course of operation, the HL-LHC is expected to collect a total integrated luminosity of $\mathcal{L} = 3000$ fb^{-1} by 2035 [9].

These run conditions are much harsher than what ATLAS has experienced so far, and as a result there are several planned upgrades to the detector. Most notably, the entire ID will be replaced with an all-silicon tracker which will extend the coverage from $|\eta| \leq 2.7$ up to $|\eta| \leq 4.0$. This will allow for reconstruction of charged particle tracks which can in turn be matched to clusters in the calorimeters for electron identification or forward jet tagging [10].

The upgraded detector combined with the higher beam energy and the considerable increase in integrated luminosity means that many analyses with low signal statistics in Run 2 have the potential to be greatly improved with the HL-LHC. While the ATLAS 13 TeV $W^{\pm}W^{\pm}jj$ cross section measurement certainly did not suffer greatly from low statistics TODO: –reword–, the accuracy of the measurement can still be improved at the HL-LHC. Of particular interest is the longitudinal polarization of the W bosons due to its sensitivity to electroweak symmetry breaking [11].

The analysis detailed in this chapter is based off of the 2018 public ATLAS $W^{\pm}W^{\pm}jj$ prospects study [12] which is itself an extension of the 2017 ATLAS study [13]. TODO: mention CMS's study + yellow report?

6.0.1 Analysis Overview

The experimental signature of interest here is identical to the 13 TeV analysis detailed in Chapter 5: two prompt leptons (electrons or muons) with the same charge, missing transverse energy, and two jets. Once again the two leading jets are required to have a large angular separation and a high combined invariant mass to preferentially select EWK VBS production over QCD $W^{\pm}W^{\pm}jj$ events.

Background processes that can mimic the signal are again similar to the 13 TeV analysis. The dominant source of prompt background from WZ+jets events where both bosons decay leptonically. If the lepton from the Z-decay with opposite charge from the W falls outside of the detector acceptance or is not identified, the remainder could appear to be a $W^{\pm}W^{\pm}jj$ signal event. To a lesser extent, ZZ+jets events can enter the signal region in much the same way provided two leptons are "lost". Other prompt sources include $t\bar{t} + V$ and and multiple parton interactions, however these processes do not contribute much. The upgrades to the ATLAS detector are expected to reduce the size of these prompt contributions due in large part to the increased detector acceptance from the forward tracking. Jets mis-reconstructed as leptons or leptons from hacronic decays (such as $t\bar{t}$ and W+jets production) comprise the non-prompt lepton background. Lastly, events with two prompt, opposite-charge electrons can contribute provided one of the electrons is mis-reconstructed as the wrong charge.

In this analysis, the EWK production of $W^{\pm}W^{\pm}jj$ is studied in the context of the planned HL-LHC run conditions and upgraded ATLAS detector. An optimized event selection (referred to as the *optimized selection*) is also explored in an effort to gain increased signal significance over the *default selection*. The cross section of the inclusive EWK production is measured for both the default and optimized selections, and the extraction of the longitudinal scattering significance is measured with the optimized selection.

6.1 Theoretical motivation

The theoretical motivation for studying the ssWW process is detailed in Section 5.1. The particular interest in polarization is the potential for the scattering amplitude of longitudinally polarized weak bosons to diverge linearly as the center of mass energy increases, ultimately violating unitarity around 1 TeV [14]. In the Standard Model, the Higgs boson cancels these divergences. However, as the Higgs is recently discovered it is still extremely to study the mechanism of electroweak symmetry breaking (EWSB), and the longitudinal scattering of W bosons is expected to be one of the most sensitive tests of EWSB [11].

6.1.1 Experimental sensitivity to longitudinal polarization

There are three possible polarization states for a massive vector boson: two transverse (+ or $-$) and one longitudinal (0). Therefore, in a system with two W bosons, the overall polarization can be purely longitudinal (00), purely transverse ($++$, $--$, and $+-$), or mixed ($+0$ and -0). The three combinations will be referred to as LL, TT, and LT respectively.

In order extract the longitudinal scattering component, it is necessary to find variables that distinguish the LL from the TT and LT. Several variables were studied, and those with the best discriminating power between the polarizations were the leading and subleading lepton p_{T} as well as the azimuthal separation ($|\Delta\phi_{jj}|$) of the two VBS jets. The LL events preferred lower p_{T} for both signal leptons (see Figure 6.1), which motivates keeping these two cuts as low as possible in the event selection in order to preserve as much longitudinal polarization as possible. In the case of $|\Delta\phi_{jj}|$, the LL events generally had a larger dijet separation (see Figure 6.2), and this variable is used in a binned likelihood fit to extract the longitudinal scattering significance.

6.2 Monte Carlo samples

As no real HL-LHC data will be available for many years, all processes in this prospects study must be simulated using Monte Carlo (MC) generators. Signal and background processes were generated at $\sqrt{s} = 14$ TeV, and the event yields scaled to the anticipated HL-LHC integrated luminosity of $\mathcal{L} = 3000$ fb^{-1}.

TODO: Consider putting all this in a table

The signal sample consists of both VBS and non-VBS electroweak (EWK) $W^{\pm}W^{\pm}jj$ production, and it is sumulated with the `Madgraph5_aMC@NLO` generator [15] using the `NNPDF3.0` PDF set [16] and interfaced with `PYTHIA v8` [17] for hadronization and parton showering. To study the longitudinal polarization more directly, two additional `Madgraph5_aMC@NLO` $W^{\pm}W^{\pm}jj$ samples are used: one containing only the longitudinal contribution (LL) and a second containing the transverse (TT) and mixed (LT) contributions.

There are many other processes that can produce the same final state as the $W^{\pm}W^{\pm}jj$ and must also be accounted for using MC simulations. WZ events are generated using `SHERPA v2.2.0` [18, 19,

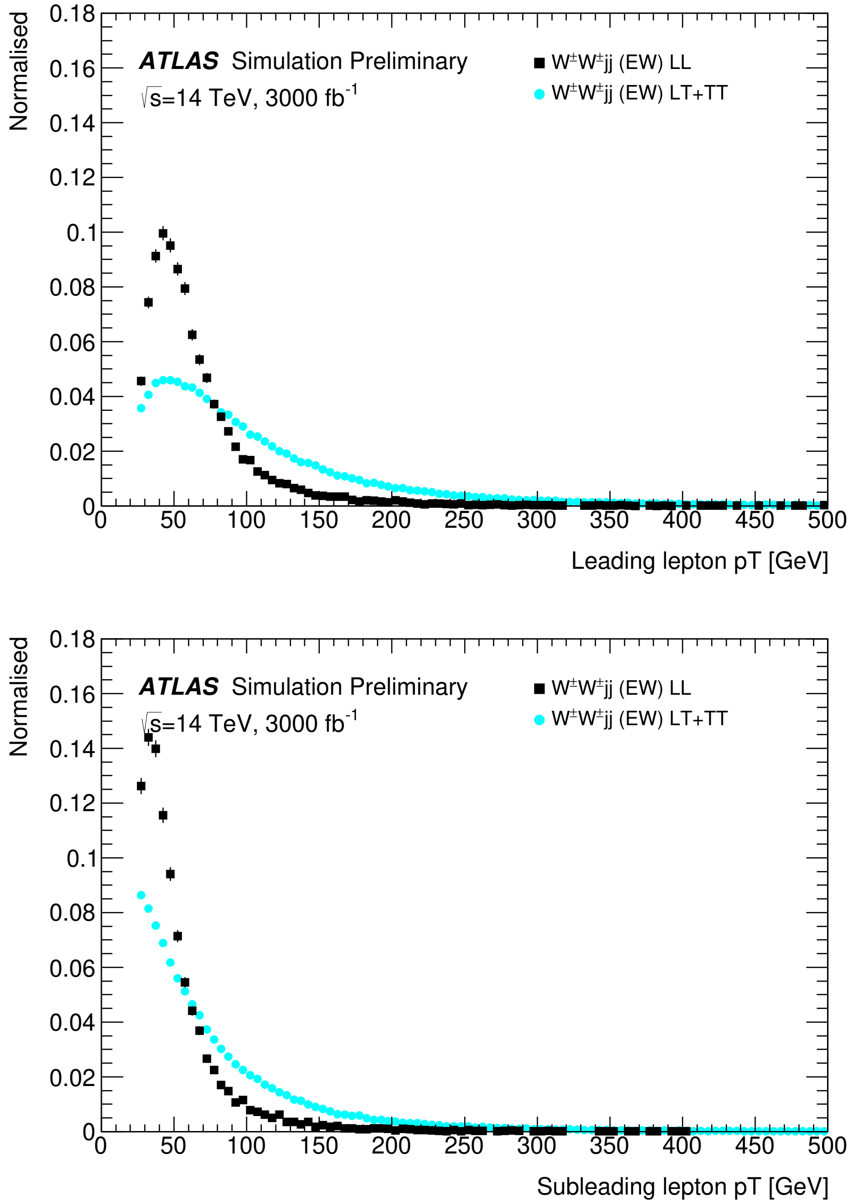

Figure 6.1: Comparison of the leading (top) and subleading (bottom) lepton p_{T} distributions for purely longitudinal (LL, black) and mixed polarization (LT+TT, cyan) $W^{\pm}W^{\pm}jj$ events. Plots from [1].

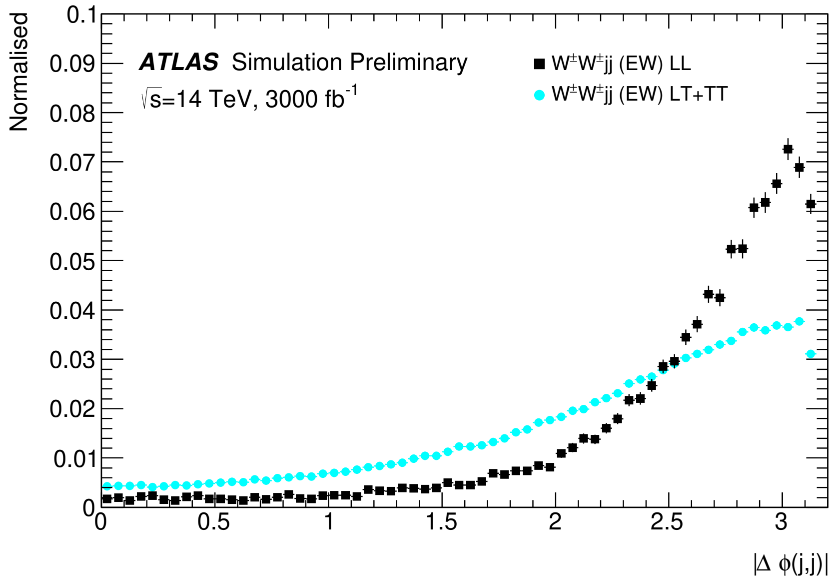

Figure 6.2: Comparison of the azimuthal dijet separation ($|\Delta\phi_{jj}|$) for purely longitudinal (LL, black) and mixed polarization (LT+TT, cyan) $W^{\pm}W^{\pm}jj$ events. Plot from [1].

20], which includes up to one parton at next-to-leading order (NLO) in the strong coupling constant α_s and up to three additional partons at leading order (LO). Both EWK and QCD production are included in these samples. ZZ events are generated using `SHERPA v2.2.2` with up to two additional partons in the final state. Triboson backgrounds $VVV, V = W, Z$ where the bosons can decay leptonically or hadronically are simulated with `SHERPA v2.2.2` with up to two additional partons in the final state. W+jets backgrounds are generated for electron, muon, and tau final states are generated at LO with `Madgraph5_aMC@NLO` and the `NNPDF3.0` set with showering from `PYTHIA v8` . Z+jets events are generated using `POWHEG-BOX` [21] and the `CT10` PDF set [22] interfaced with `PYTHIA v8` . Finally, $t\bar{t}$ and single-top events are generated using `POWHEG-BOX` with showering from `PYTHIA v6` .

Since the MC samples used in the analysis are generated at particle-level and have not been run through the typical full simulation of the ATLAS detector, smearing functions are instead used to estimate detector effects. These are derived from a `GEANT4` simulation of the upgraded ATLAS detector [23]. In addition, pileup events are fully simulated.

6.3 Background estimations

In this analysis, all background contributinos are estimated using MC simulations. Backgrounds such as electron charge misidentification and fake electrons from jets–which are traditionally estimated using data-driven techniques–are estimated using a set of parameterization functions applied to the MC. These functions calculate the probability that an electron is assigned the wrong charge or a jet is mis-reconstructed as an electron parameterized by the $p_{\rm T}$ and η of the electron or jet. The probabilites are derived from studies on expected electron performance with the upgraded ATLAS detector [24].

Processes involving two W and Z bosons are grouped together as *diboson* backgrounds, with the exception of $W^{\pm}W^{\pm}jj$ events produced via QCD interactions, which are kept separate. Similarly, all backgrounds with three vector bosons are combined and labeled as *triboson*. Any W+jets or top events that pass selection and do not contain a fake electron, as well as any Z+jets events without an electron identifed as having its charge misidentified are combined as *other non-prompt* backgrounds.

6.3.1 Truth-based isolation

Since the MC samples used in this analysis have not been run through a full detector simulation, they lack any kind of particle isolation variables (since they require, for example, information on the calorimeter response). Generally, this is not a large concern, as at truth-level, high $p_{\rm T}$ signal leptons tend to be well isolated to begin with. However, isolation is one of the most powerful tools for rejecting leptons from non-prompt sources such as top events, which are produced in association with additional nearby particles from b and c quark decays. In the absence of any sort of isolation requirement, contributions from top backgrounds (including single top, $t\bar{t}$ and $t\bar{t}+V$) were more than an order of magnitude higher than expected.

As a result, it was necessary to create an analogue to the isolation information that is available in fully-simulated samples. Track- and calorimeter-based isolation variables were constructed by summing the momentum and energy, respectively, of stable truth particles with $p_{\rm T} > 1$ GeV within a specified radius of each signal lepton. For the track-based isolation, only charged truth particles were used; both charged and neutral particles (excluding neutrinos) were included for the calorimeter-based isolation. Ultimately, a set of isolation cuts were chosen that are similar to those recommended by ATLAS for Run 2 analyses. The truth-based isolation requirements are listed in Table 6.1.

The truth-based isolation requirement reduced the top background by over 99% and reduced the contribution of top events to the total background from 83% to 2%. Additional details on the

	Electron Isolation	Muon Isolation
Track-based isolation cone size	$\Delta R < 0.2$	$\Delta R < 0.3$
Track-based isolation requirement	$\sum p_{\mathrm{T}}/p_{\mathrm{T}}^{e} < 0.06$	$\sum p_{\mathrm{T}}/p_{\mathrm{T}}^{\mu} < 0.04$
Calorimeter-based isolation cone size	$\Delta R < 0.2$	$\Delta R < 0.2$
Calorimeter-based isolation requirement	$\sum E_{\mathrm{T}}/p_{\mathrm{T}}^{e} < 0.06$	$\sum E_{\mathrm{T}}/p_{\mathrm{T}}^{\mu} < 0.15$

Table 6.1: Truth-based isolation requirements for electrons and muons.

truth-based isolation studies are presented in Appendix A.

6.4 Object and Event selection

6.4.1 Object selection

Electrons and muons are preselected to have $p_{\mathrm{T}} > 7$ and 6 GeV, respectively, and $|\eta| \leq 4.0$. The likelihood of a given lepton to pass the trigger or identification requirements is estimated by estimating an efficiency dependent on the p_{T} and η of the lepton. The leptons are also required to pass the isolation criteria detailed in Table 6.1. Jets that have been tagged as a fake electron by the functions described earlier in Section 6.3 are treated as electrons for the purpose of the object selection and are subject to the same criteria. In order to be considered a signal lepton, an additional requirement of $p_{\mathrm{T}} > 25$ GeV is applied on top of the preselection. The two highest p_{T} leptons passing this selection are chosen to be the leading and subleading signal leptons.

Jets are clustered using the anti-k_t algorithm [25] from final-state particles within a radius of $\Delta R = 0.4$ (excluding muons and neutrinos). Jets are required to have $p_{\mathrm{T}} > 30$ GeV and lie within $|\eta| < 4.5$, with an additional cut of $p_{\mathrm{T}} > 70$ GeV for jets above $|\eta| \geq 3.8$ in order to suppress jets from pileup interactions. Jets overlapping with a preselected electron within $\Delta R_{e,j} < 0.05$ are removed in order to prevent double counting. The two highest p_{T} jets are defined as the leading and subleading *tag jets*.

6.4.2 Event selection

The default event selection is summarized in Table 6.2 and described here. Exactly two signal leptons are required with the same electric charge and separated from each other by 0.3 in ΔR. In order to suppress contributions from Drell-Yan backgrounds, the two signal leptons must have an invariant mass m_{ll} greater than 20 GeV. Additionally, if both signal leptons are electrons, their mass must be at least 10 GeV from the Z-boson mass in order to reduce background from Z-boson decays[2]. The event is required to have at least 40 GeV of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) to account for the two neutrinos from the W decays. Events with additional preselected leptons are vetoed, which greatly reduces WZ and ZZ backgrounds. Both tag jets are required to not overlap with the signal leptons, and there is a veto on events with one or more b-jets. In order to preferentially select VBS production, the tag jets are also required to have a large separation between them and a large invariant mass. Finally, a cut on the lepton centrality, ζ, defined in Equation 6.1 enhances the EWK $W^{\pm}W^{\pm}jj$ signal.

$$\zeta = \min[\min(\eta_{\ell 1}, \eta_{\ell 2}) - \min(\eta_{j1}, \eta_{j2}), \max(\eta_{j1}, \eta_{j2}) - \max(\eta_{\ell 1}, \eta_{\ell 2})] \tag{6.1}$$

Selection requirement	Selection value
Lepton kinematics	$p_{\mathrm{T}} > 25$ GeV
	$\lvert\eta\rvert \leq 4.0$
Jet kinematics	$p_{\mathrm{T}} > 30$ GeV for $\lvert\eta\rvert \leq 4.5$
	$p_{\mathrm{T}} > 70$ GeV for $\lvert\eta\rvert > 3.8$
Dilepton charge	Exactly two signal leptons with same charge
Dilepton separation	$\Delta R_{l,l} \geq 0.3$
Dilepton mass	$m_{ll} > 20$ GeV
Z boson veto	$\lvert m_{ee} - m_Z\rvert > 10$ GeV (ee-channel only)
$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV
Jet selection	At least two jets with $\Delta R_{l,j} > 0.3$
b jet veto	$N_{\text{b-jet}} = 0$
Dijet separation	$\Delta\eta_{jj} > 2.5$
Trilepton veto	No additional preselected leptons
Dijet mass	$m_{jj} > 500$ GeV
Lepton-jet centrality	$\zeta > 0$

Table 6.2: Summary of the signal event selection.

6.5 Selection optimization

As mentioned earlier, the HL-LHC will feature forward tracking, an increase in center of mass energy, and a higher integrated luminosity. Therefore, this study is an excellent time to see if there are new optimizations to the signal event selection that can improve the signal to background ratio.

[2] The electron charge mis-ID rate is high enough that contributions from $Z \to ee$ backgrounds are non-negligible.

6.5.1 Random grid search algorithm

The chosen method for optimizing the event selection is a cut-based algorithm known as the Random Grid Search (RGS) [26]. Consider a simple case of two variables x and y chosen to differentiate the signal from the background. In order to be considered a signal event, a given event would be required to pass a *cut point* $c = \{x > x_c, y > y_c\}$. A simple method to choose the optimal cut point (i.e. the "best" values of the cuts x_c and y_c) would be to construct an $n \times m$ rectangular grid in x and y consisting of points $(x_0, y_0), (x_1, y_1), ..., (x_n, y_m)$, as in Figure 6.3. One can then choose a cut point $c_k = \{x > x_i, y > y_j\}$ that maximizes the signal significance as measured by a chosen metric. This would be considered a *regular* or *rectangular* grid search.

While effective in principle, this rectangular grid search comes with two major drawbacks:

1. The algorithm does not scale well as the number of variables to be optimized–the dimensionality of the grid–increases. In the case of a square grid with N bins per variable v, the number of cut points to be evaluated grows as N^v.

2. Signal and background samples are rarely evenly distributed over the entire grid, resulting in many cut points being sub-optimal and evaluating them would be a waste of computing resources.

To combat these limitations, the RGS algorithm constructs a grid of cut points directly from the signal sample itself. In the two-dimensional example, this means that the variables x_i and y_j making up the cut point $c_k = \{x > x_i, y > y_j\}$ take their values directly from a given signal event. This has the benefit of creating a *random grid* of cut points that is by construction biased towards regions of high signal concentration. This reduces the need for exponentially increasing numbers of cut points while ensuring that computing resources are not wasted in regions with few to no signal events. An example of the the two-dimensional random grid is shown in Figure 6.4.

Once the random grid of cut points is constructed, the optimal cut point can be chosen using whatever metric the analyzer chooses, such as signal to background ratio. For the purpose of the $W^\pm W^\pm jj$ upgrade study, the optimal cut point is the one that mazimizes the signal significance Z defined as in Equation 6.2 [27].

$$Z = \sqrt{2\left[(s+b)\ln\left(\frac{s+b}{b_0}\right) + b_0 - s - b\right] + \frac{(b-b_0)^2}{\sigma_b^2}} \tag{6.2}$$

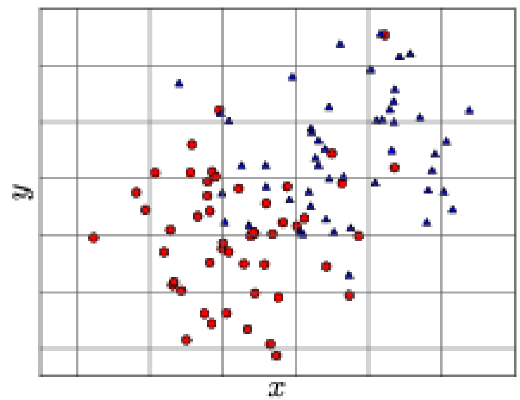

Figure 6.3: A visual representation of a rectangular grid search algorithm. The signal events are the blue triangles, and the red circles are the background events. TODO: replace with own figure

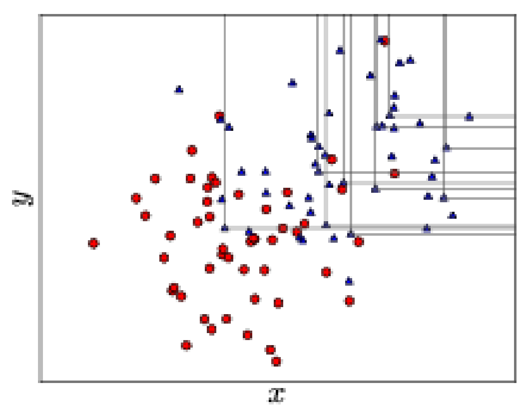

Figure 6.4: A visual representation of a random grid search algorithm. The signal events are the blue triangles, and the red circles are the background events. TODO: replace with own figure

where s and b are the number of signal and background events, respectively, σ_b is the total uncertainty on the background, and b_0 is defined as:

$$b_0 = \frac{1}{2}\left(b - \sigma_b^2 + \sqrt{(b - \sigma_b^2)^2 + 4(s+b)\sigma_b^2}\right) \tag{6.3}$$

In the case where the backround is known precisely (i.e. $\sigma_b = 0$), Equation 6.2 simplifies to

$$Z = \sqrt{2\Big(b\big[(1 + s/b)\ln(1 + s/b) - s/b\big]\Big)} \tag{6.4}$$

which further reduces to the familiar $Z = s/\sqrt{b}$ for the case when $s << b$.

6.5.2 Inputs to the optimization

In order to train the RGS, signal and background samples were prepared from events passing the event selection outlined in Table 6.2 up through the b-jet veto. The signal sample was chosen to be the longitudinally polarized $W^{\pm}W^{\pm}jj$ EWK events, and the transverse and mixed polarizations were treated as background along with $W^{\pm}W^{\pm}jj$ events from QCD interactions and the traditional backgrounds listed in Section 6.3. Splitting the inclusive $W^{\pm}W^{\pm}jj$ EWK events by polarization allows the optimization to favor the longitunally polarized events as much as possible, even though they both contribute to the EWK signal.

The following variables were chosen for optimization:

- Leading lepton p_{T}
- Dilepton invariant mass (m_{ll})
- Leading and subleading jet p_{T}
- Dijet invariant mass (m_{jj})
- Lepton-jet centrality (ζ)

Subleading lepton p_{T} was omitted as it is desirable to keep the cut value as low as possible due to its sensitivity to the longitudinal polarization (as discussed in Section 6.1.1). Additionally, the dijet separation $\Delta\eta_{jj}$ was included in the optimization originally, however it was dropped from the list due to the cut value being motivated by differences between EWK and QCD produced $W^{\pm}W^{\pm}jj$ events.

Two additional constraints were imposed when selecting the optimal cut point:

1. At least 1000 signal events must survive in order to prevent the optimization from being too aggressive and unnecssarily reducing signal statistics.

2. The dijet invariant mass may only vary within a 50 GeV range of the default value (from 450-550 GeV) due to the cut being physically motivated by the VBS event topology (TODO: reference where this is discussed in the 13TeV section).

Lastly, the decision was made to use calculate the signal significance without taking into account the uncertainty of the background using Equation 6.4. This was due to the fact that the statistical uncertainties of the fake electron and charge-misID backgrounds were quite large, and if Equation 6.2 were used instead, the optimization would cut unreasonably hard against these backgrounds. Since Monte Carlo statistics is not expected to be a limiting factor when this analysis is performed at the HL-LHC, it is more realistic to simply ignore these large statistical uncertainties for the purpose of the selection optimization.

6.5.3 Results of the optimization

Ultimately, the random grid was constructed from over 38,000 LL-polarized $W^{\pm}W^{\pm}jj$ events in the variables listed above. After applying the constraints, an optimal cut point was chosen which reduced the total background from 9900 to 2310 while reducing the signal from 3489 to 2958. This corresponds to an increase in signal significance from $Z = 33.26$ to $Z = 52.63$ as calculated by Equation 6.4. The updates to the event selection are listed in Table 6.3.

The large reduction in the background is primarily a result of the increase in the leading and subleading jet p_{T} from 30 GeV to 90 GeV and 45 GeV, respectively. As can be seen in Figure 6.7, this increase removes a significant portion of the backgrounds from jets faking electrons and charge mis-ID. Additionally, the loosening of the lepton-jet centrality cut ζ allows more signal events to survive the event selection (see Figure 6.9). Other changes to the event selection are minor and do not individually have a large impact on the signal or background yields.

The full event yields after optimization as well as the cross section measurement are detailed alongside those using the default selection in Section 6.6.

TODO: It's a bit awkward to reference the results of the default/optimized before they're properly presented. Maybe move the sections around? not sure...

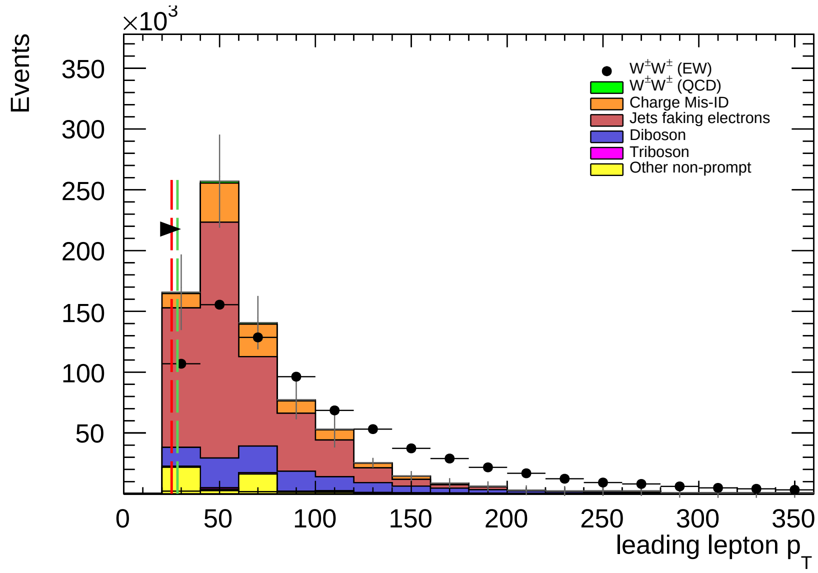

Figure 6.5: Leading lepton p_{T} distribution. The default and optimized cuts are represented by the red and green dashed lines, respectively. The $W^{\pm}W^{\pm}jj$ EWK signal (black points) is normalized to the same area as the sum of the backgrounds (colored histogram). TODO: Move to appendix or omit

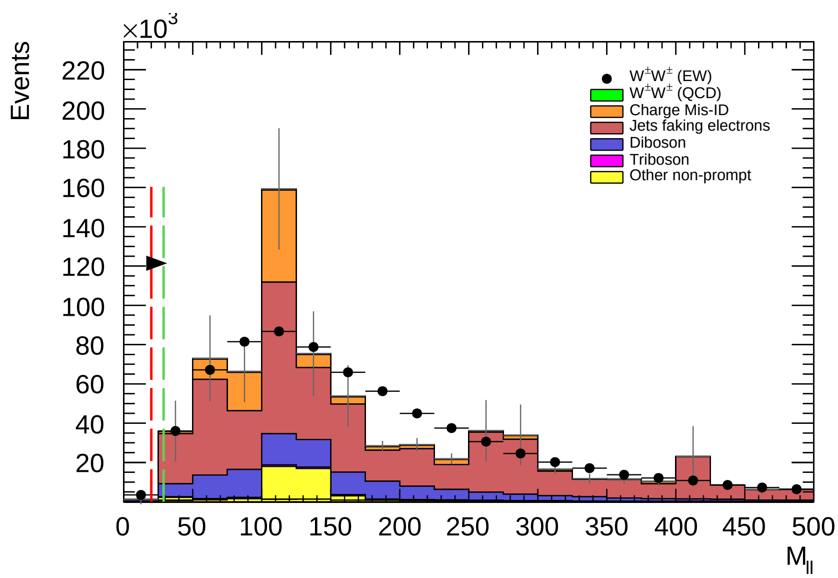

Figure 6.6: Dilepton invariant mass distribution. The default and optimized cuts are represented by the red and green dashed lines, respectively. The $W^{\pm}W^{\pm}jj$ EWK signal (black points) is normalized to the same area as the sum of the backgrounds (colored histogram). TODO: Move to appendix or omit

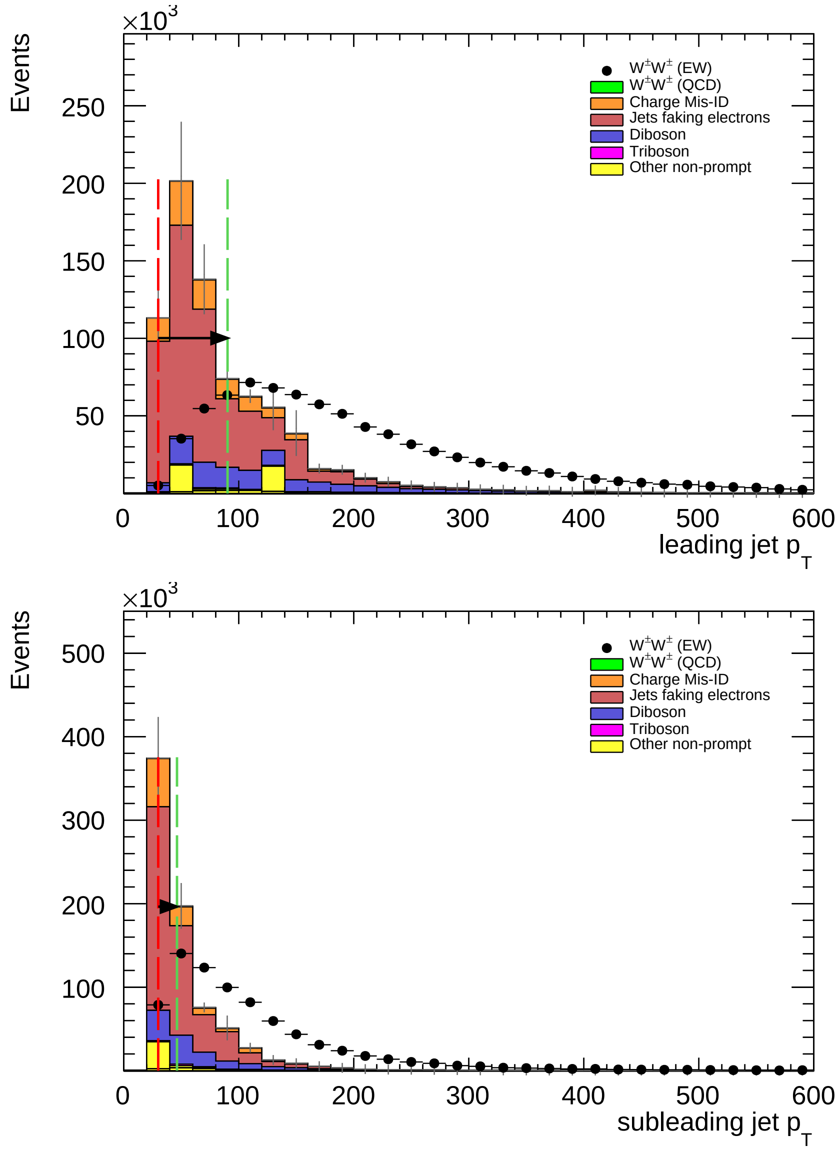

Figure 6.7: Leading (top) and subleading (bottom) jet p_{T} distributions. The default and optimized cuts are represented by the red and green dashed lines, respectively. The $W^{\pm}W^{\pm}jj$ EWK signal (black points) is normalized to the same area as the sum of the backgrounds (colored histogram).

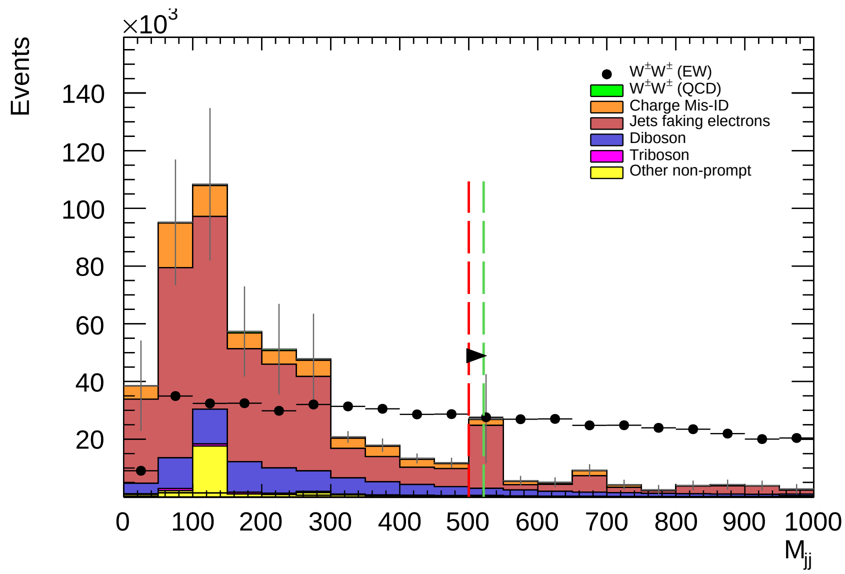

Figure 6.8: Dijet invariant mass distribution. The default and optimized cuts are represented by the red and green dashed lines, respectively. The $W^{\pm}W^{\pm}jj$ EWK signal (black points) is normalized to the same area as the sum of the backgrounds (colored histogram). TODO: Move to appendix or omit

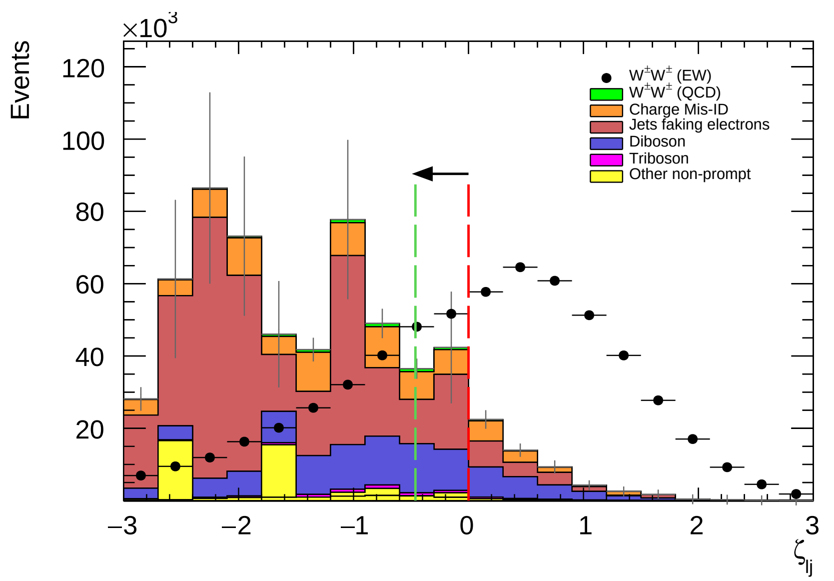

Figure 6.9: Lepton-jet centrality distribution. The default and optimized cuts are represented by the red and green dashed lines, respectively. The $W^{\pm}W^{\pm}jj$ EWK signal (black points) is normalized to the same area as the sum of the backgrounds (colored histogram).

Selection requirement	Selection value
Lepton kinematics	$p_\mathrm{T} > 28$ GeV (leading lepton only)
Jet kinematics	$p_\mathrm{T} > 90$ GeV (leading jet) $p_\mathrm{T} > 45$ GeV (subleading jet)
Dilepton mass	$m_{ll} > 28$ GeV
Dijet mass	$m_{jj} > 520$ GeV
Lepton-jet centrality	$\zeta > -0.5$

Table 6.3: Updates to the $W^\pm W^\pm jj$ event selection criteria after optimization. Cuts not listed remain unchanged from the default selection in Table 6.2.

6.6 Results

6.6.1 Event yields

After applying the full event selection, the analysis is broken down into four channels based off of the flavor of the signal leptons: $\mu\mu$, ee, μe, and $e\mu$. The full signal and background event yields are shown in Table 6.4 for each channel separately and combined using the default event selection. 3489 EWK $W^\pm W^\pm jj$ events are expected compared to 9900 background events. The dominant sources of background are jets faking electrons followed by charge misidentification and diboson processes. Triboson events, QCD $W^\pm W^\pm jj$, and other non-prompt sources make up approximately 5% of the total background combined.

	All channels	$\mu\mu$	ee	μe	$e\mu$
$W^\pm W^\pm jj$ (QCD)	206.4	91.1	22.8	38.4	54.1
Charge Misidentification	2300	0.0	2100	90	160
Jets faking electrons	5000	0.0	3400	1200	340
$WZ + ZZ$	2040	500	438	423	680
Tribosons	115	47	15.4	21.6	31.2
Other non-prompt	210	110	20	60	27
Total Background	9900	750	6000	1900	1290
Signal $W^\pm W^\pm jj$ (EWK)	3489	1435	432	679	944

Table 6.4: Signal and background event yields using the default event selection for an integrated luminosity of $\mathcal{L} = 3000\ \mathrm{fb}^{-1}$. Events containing a fake or charge-flipped electron are removed from their respective sources and combined into a single entry each.

The event yields for the optimized selection detailed in Section 6.5.3 are listed in Table 6.5. After optimization, 2958 signal events and just 2310 background events are expected. Diboson events now are the primary source of background, as the optimization greatly reduces the fake and charge misidentification backgrounds. As discussed earlier, the increase in the leading and subleading jet p_T cuts as well as the loosening of the centrality cut are most responsible for the changes in the signal and background yields; distributions of these quantities using the default and the optimized event selections can be found in Figures 6.10, 6.11, and 6.12, respectively.

	All channels	$\mu\mu$	ee	μe	$e\mu$
$W^{\pm}W^{\pm}jj$ (QCD)	168.7	74.6	19.7	32.2	42.2
Charge Misidentification	200	0.0	11	30	160
Jets faking electrons	460	0.0	130	260	70
$WZ + ZZ$	1286	322	289	271	404
Tribosons	76	30.1	9.6	15.1	21.6
Other non-prompt	120	29	16.6	50	19
Total Background	2310	455	480	660	710
Signal $W^{\pm}W^{\pm}jj$ (EWK)	2958	1228	380	589	761

Table 6.5: Signal and background event yields using the optimized event selection for an integrated luminosity of $\mathcal{L} = 3000$ fb^{-1}. Events containing a fake or charge-flipped electron are removed from their respective sources and combined into a single entry each.

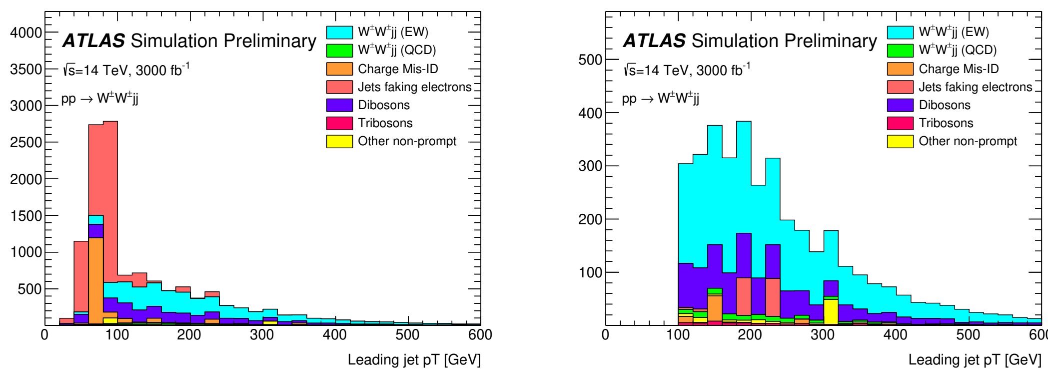

Figure 6.10: p_{T} distributions for the leading jet using the default (left) and optimized (right) event selections for all channels combined.

It is important to note, however, that the MC sample used to estimate Z+jets events suffers from poor statistics which results in large per-event weights once scaled to $\mathcal{L} = 3000$ fb^{-1}. This sample contributes heavily to the fake and charge misidentification backgrounds, and a handful of these events being cut out by the optimization contributes has a large effect on the dramatic reduction of these backgrounds. As a result, these particular optimized results are likely overly optimistic. However, given proper MC statistics, it is still expected that the optimization will outperform the default selection.

6.6.2 Uncertainties

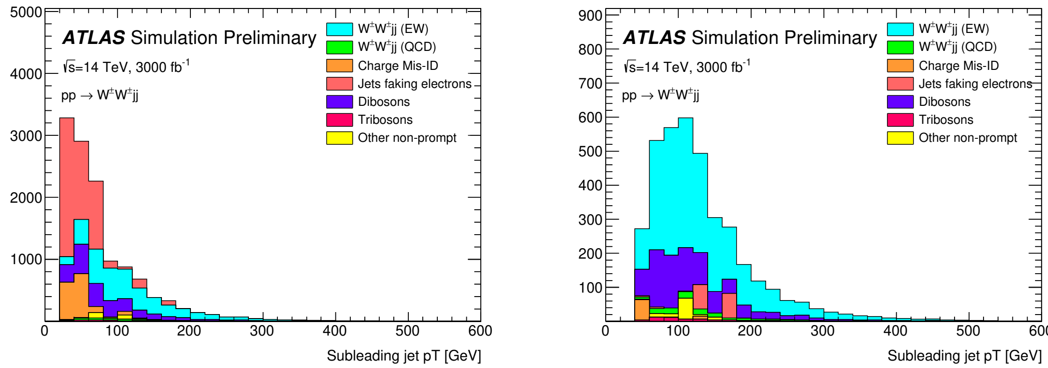

Figure 6.11: p_{T} distributions for the subleading jet using the default (left) and optimized (right) event selections for all channels combined.

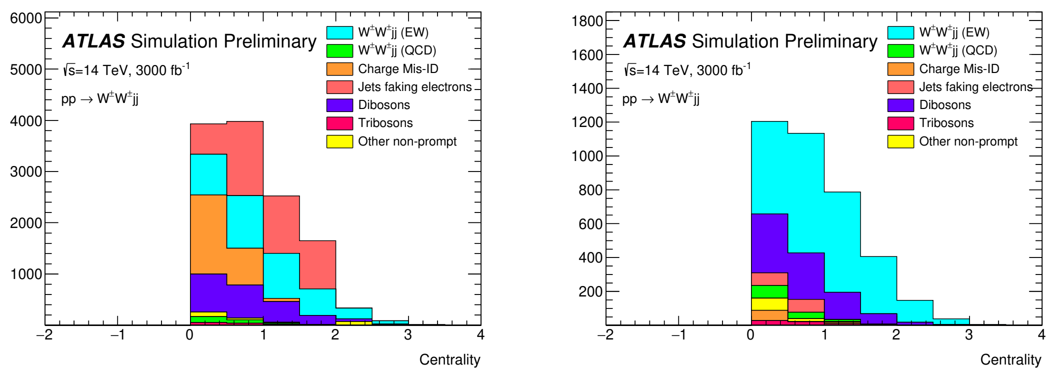

Figure 6.12: p_{T} distributions for lepton-jet centrality ζ using the default (left) and optimized (right) event selections for all channels combined.

TODO: Ask for details on how some of these uncertainties were calculated – specifically the fakes and charge mis-ID The uncertainties considered for the analysis are summarized in Table 6.6. Values for experimental systematics on the trigger efficiency, lepton and jet reconstruction, and flavor tagging are taken directly from the 13 TeV analysis TODO: (cite or reference?). The rate uncertainties for the background processes are halved from the 13 TeV values.

6.6.3 Cross section measurement

The cross section is calculated using the same method as in the 13 TeV analysis, detailed in Chapter 5. TODO: update from chapter reference to subsection reference (once it's written)... Once again, each of the four lepton flavor channels is further split by charge (i.e. $\mu\mu \rightarrow \mu^+\mu^+ + \mu^-\mu^-$),

Source	Uncertainty (%)
$W^{\pm}W^{\pm}jj$ (EWK)	3
Luminosity	1
Trigger efficiency	0.5
Lepton reconstruction and identification	1.8
Jets	2.3
Flavor tagging	1.8
Jets faking electrons	20
Charge misidentification	25
$W^{\pm}W^{\pm}jj$ (QCD)	20
Top	15
Diboson	10
Triboson	15

Table 6.6: Summary of estimated experimental and rate uncertainties.

as this increases the sensitivity of the analysis. Each channel's m_{jj} distribution is combined in a profile likelihood fit to extract the EWK $W^{\pm}W^{\pm}jj$ production cross section. The expected cross section calculated using the default event selection is:

$$\sigma^{\text{expected}}_{W^{\pm}W^{\pm}jj} = 16.89 \pm 0.36\ (\text{stat}) \pm 0.53\ (\text{theory}) \pm 0.84\ (\text{syst})\ \text{fb} \tag{6.5}$$

The expected cross section calculated using the optimized event selection is:

$$\sigma^{\text{expected}}_{W^{\pm}W^{\pm}jj} = 16.94 \pm 0.36\ (\text{stat}) \pm 0.53\ (\text{theory}) \pm 0.78\ (\text{syst})\ \text{fb} \tag{6.6}$$

The optimized selection should not change the measured value of the cross section, and indeed both are consistent with within uncertainties. The sytematic uncertainty is reduced by approximately 7% with the optimized selection. Projections of the total uncertainty on the cross section as a function of integrated luminosity made by TODO: how was this made? is shown in Figure 6.13.

6.6.4 Longitudinal scattering significance

TODO: get some details on how this was all done... The longitudinal scattering significance is extracted from the $|\Delta\phi_{jj}|$ distribution using a simultaneous binned likelihood fit. In order to increase sensitivity, the $|\Delta\phi_{jj}|$ distribution was split into two bins in m_{jj}, and an additional cut on the pseudorapidity of the subleading lepton was applied ($|\eta| < 2.5$) to reduce background from fake and charge misidentification. The $|\Delta\phi_{jj}|$ distributions used in the fit are shown in Figure 6.14. Due to limited statistics, the four lepton flavor channels were not split by charge. The expected significance of the $W^{\pm}_L W^{\pm}_L jj$ process is 1.8σ with a precision of 47% on the measurement. Projections of the expected significance as a function of integrated luminosity is shown in Figure 6.15.

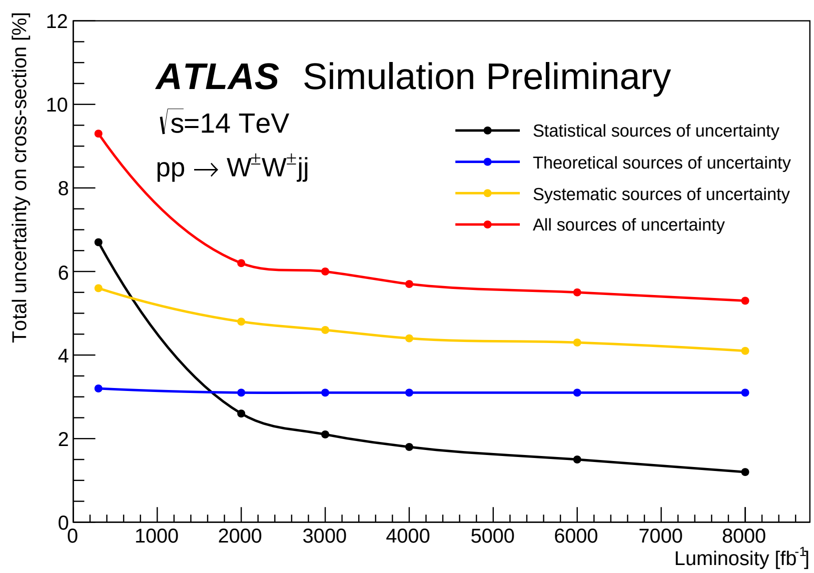

Figure 6.13: Projections of the statistical (black), theoretical (blue), systematic (yellow), and total (red) uncertainties on the measured cross section as a function of integrated luminosity using the optimized event selection.

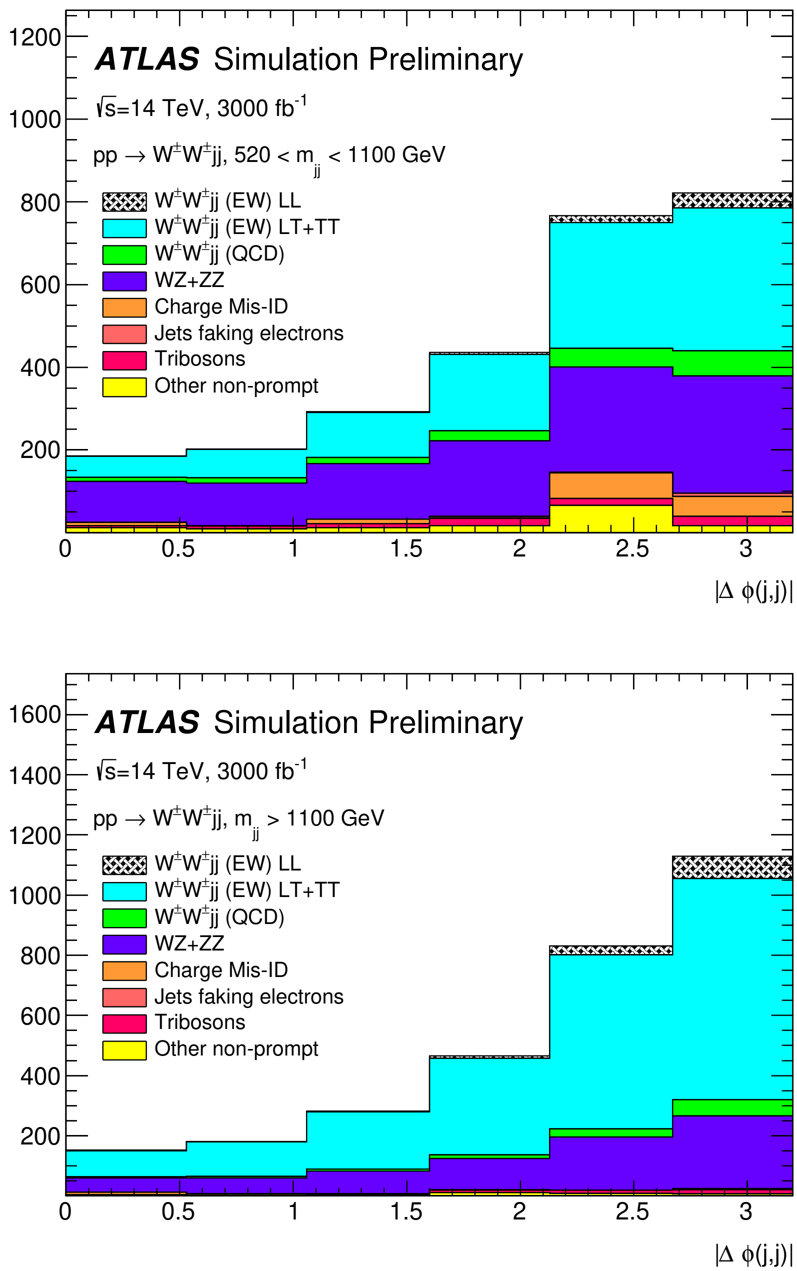

Figure 6.14: Dijet azimuthal separation ($|\Delta\phi_{jj}|$) for the low m_{jj} region ($520 < m_{jj} < 1100$ GeV, top) and the high m_{jj} region ($m_{jj} > 1100$ GeV, bottom). The purely longitudinal (LL, gray) is plotted separately from the mixed and transverse (LT+TT, cyan) polarizations.

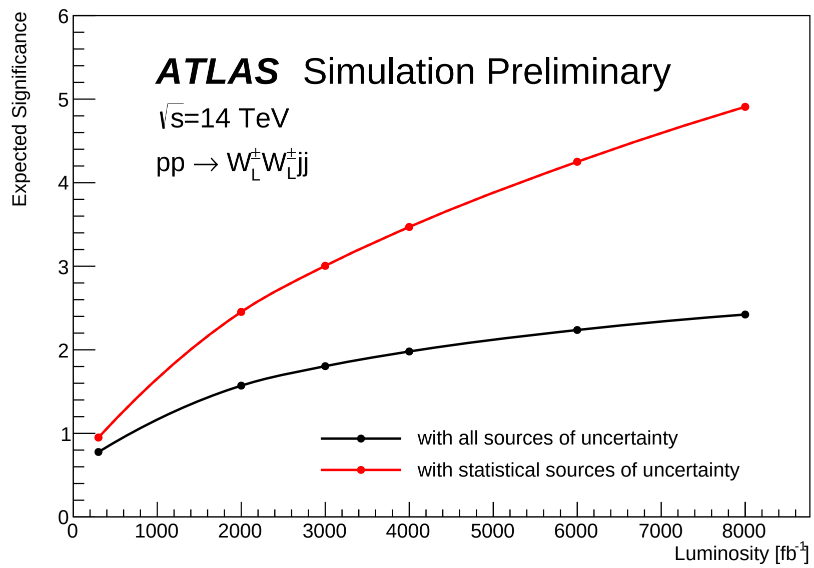

Figure 6.15: Projections of the expected longitudinal scattering significance as a function of integrated luminosity when considering all sources of uncertainties (black) or only statistical uncertainties (red).

Chapter 7

Conclusion

Here's where you wrap it up.

Looking Ahead

Here's an example of how to have an "informal subsection".

Appendix A

Additional material on truth isolation

yields by type	all channels	$\mu\mu$	ee	μe	$e\mu$
signal	4011	1583.2	531.7	793.1	1103.1
ww qcd	252.6	105.8	30.4	48	68.4
charge flip	2528.4	0.0	2075.4	255.1	197.8
fakes	7135.4	0.0	4675.1	1904.3	555.9
diboson	2370.4	581.2	491.8	517.9	779.6
triboson	125.5	49.1	17.8	24.6	34.1
top	90150.5	26618	15301.6	25277.9	22953.1
z+jets	241.2	0.0	0.0	0.0	241.2
w+jets	31.4	3.9	7.6	13.2	6.7
total bkg	102803.9	27354	22592	28027.8	24830.1
signal	4011	1583.2	531.7	793.1	1103.1

Table A.1: Event yields prior to applying any form of truth-based isolation criteria.

yields by type	all channels	$\mu\mu$	ee	μe	$e\mu$
signal	3470.5	1427.3	428.8	675.8	938.7
ww qcd	205.8	90.8	22.7	38.3	54
charge flip	2398.3	0.0	2104.6	95.8	197.9
fakes	4309.7	0.0	3390.6	750.8	168.3
diboson	1552.4	311.3	355.6	346.8	538.7
triboson	115	46.8	15.4	21.6	31.2
top	156.9	42.3	14.8	76.6	23.3
z+jets	0.0	0.0	0.0	0.0	0.0
w+jets	0.3	0.0	0.0	0.3	0.0
total bkg	8738.1	491.3	5903.7	1329.8	1013.4
signal	3470.5	1427.3	428.8	675.8	938.7

Table A.2: Event yields after applying a test version of the truth-based isolation.

TODO: Add tables for tight vs loose working point, information on the necessity of TRUTH1++

Bibliography

[1] K. J. Potamianos, W. K. Di Clemente, M.-A. Pleier, C. A. Lee, J. I. Kroll, S. Yacoob, and M. Leigh, *Prospects for the measurement of the $W^{\pm}W^{\pm}$ scattering cross section and extraction of the longitudinal scattering component in pp collisions at the High-Luminosity LHC with the ATLAS experiment.*, Tech. Rep. ATL-COM-PHYS-2018-1479, CERN, Geneva, Oct, 2018. https://cds.cern.ch/record/2644264. (document), 6.1, 6.2

[2] S. L. Glashow, *The Renormalizability of Vector Meson Interactions*, Nucl. Phys. **10** (1959) 107–117. 2.2

[3] A. Salam and J. C. Ward, *Weak and Electromagnetic Interactions*, Nuovo Cimento **11** (1959) 568–577. 2.2

[4] L. R. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001. https://cds.cern.ch/record/1129806. This report is an abridged version of the LHC Design Report (CERN-2004-003). 3.1

[5] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. 3.1

[6] ATLAS Collaboration Collaboration, *Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions at $\sqrt{s} = 7\ TeV$*, Tech. Rep. ATLAS-CONF-2011-012, CERN, Geneva, Mar, 2011. https://cds.cern.ch/record/1334582. 4

[7] R. Steerenberg, *LHC Report: Another run is over and LS2 has just begun...*, https://home.cern/news/news/accelerators/lhc-report-another-run-over-and-ls2-has-just-begun, 2018. Accessed: 2018-12-14. 6

[8] *Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment*, Tech. Rep. CERN-LHCC-2011-012. LHCC-I-020, CERN, Geneva, Nov, 2011. http://cds.cern.ch/record/1402470. 6

[9] G. Apollinari, I. Bjar Alonso, O. Brning, M. Lamont, and L. Rossi, *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2015. https://cds.cern.ch/record/2116337. 6

[10] ATLAS Collaboration Collaboration, ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, Cern-lhcc-2015-020, Geneva, Sep, 2015. http://cds.cern.ch/record/2055248. 6

[11] D. Espriu and B. Yencho, *Longitudinal WW scattering in light of the "Higgs boson" discovery*, Phys. Rev. D **87** (2013) 055017, arXiv:1212.4158 [hep-ph]. 6, 6.1

[12] ATLAS Collaboration Collaboration, *Prospects for the measurement of the $W^{\pm}W^{\pm}$ scattering cross section and extraction of the longitudinal scattering component in pp collisions at the High-Luminosity LHC with the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2018-052, CERN, Geneva, Dec, 2018. http://cds.cern.ch/record/2652447. 6

[13] ATLAS Collaboration Collaboration, *Studies on the impact of an extended Inner Detector tracker and a forward muon tagger on $W^{\pm}W^{\pm}$ scattering in pp collisions at the High-Luminosity LHC with the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2017-023, CERN, Geneva, Dec, 2017. https://cds.cern.ch/record/2298958. 6

[14] B. W. Lee, C. Quigg, and H. B. Thacker, *The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass*, Phys. Rev. Lett. **38** (1977) 883–885. 6.1

[15] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph]. 6.2

[16] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph]. 6.2

[17] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv:1410.3012 [hep-ph]. 6.2

[18] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph]. 6.2

[19] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorization*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph]. 6.2

[20] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv:0903.1219 [hep-ph]. 6.2

[21] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph]. 6.2

[22] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph]. 6.2

[23] S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303. 6.2

[24] ATLAS Collaboration Collaboration, *Expected performance for an upgraded ATLAS detector at High-Luminosity LHC*, Tech. Rep. ATL-PHYS-PUB-2016-026, CERN, Geneva, Oct, 2016. http://cds.cern.ch/record/2223839. 6.3

[25] M. Cacciari, G. P. Salam, G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0**4 (2008) 063, arXiv:0802.1189 [hep-ph]. 6.4.1

[26] P. C. Bhat, H. B. Prosper, S. Sekmen, and C. Stewart, *Optimizing Event Selection with the Random Grid Search*, Comput. Phys. Commun. **2**28 (2018) 245–257, arXiv:1706.09907 [hep-ph]. 6.5.1

[27] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C**71 (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J.C73,2501(2013)]. 6.5.1