

# Run IX Di-Jet Cross Section and $A_{LL}$ Analysis Note

Brian S. Page

August 12, 2016

# Contents

<b>Contents</b>	<b>1</b>
<b>List of Figures</b>	<b>4</b>
<b>List of Tables</b>	<b>7</b>
<b>1 Data Sample and Calibration</b>	<b>8</b>
1.1 Data Quality Assurance . . . . .	8
1.2 Trigger . . . . .	10
1.3 Integrated Luminosity . . . . .	13
1.4 Average JP1 Pre-Scale . . . . .	14
1.5 Detector Calibrations . . . . .	14
1.5.1 TPC Calibration . . . . .	14
1.5.2 BEMC Calibration . . . . .	16
1.5.3 EEMC Calibration . . . . .	18
<b>2 Simulation Sample</b>	<b>21</b>
2.1 Three Levels of Simulation . . . . .	22
2.2 Z-Vertex Correction . . . . .	23
2.3 Data - Simulation Comparison . . . . .	23
<b>3 Jet and Di-jet Reconstruction</b>	<b>31</b>
3.1 TPC and Calorimeter Requirements . . . . .	31
3.2 Di-jet Selection and Cuts . . . . .	33
3.3 Di-jet Mass Definition . . . . .	35
3.4 Change in Di-jet Selection . . . . .	36
<b>4 Di-jet Cross Section</b>	<b>41</b>
4.1 Trigger Selection . . . . .	41

4.2	Unfolding . . . . .	43
4.2.1	Response Matrix . . . . .	44
4.2.2	Fakes and Efficiency Corrections . . . . .	45
4.2.3	RooUnfold Package . . . . .	46
4.3	Systematic Errors . . . . .	50
4.3.1	Detector Response Systematics . . . . .	50
4.3.2	Unfolding Systematics . . . . .	53
4.3.3	Time Variation Systematic . . . . .	54
4.3.4	Total Systematic . . . . .	63
4.4	Underlying Event and Hadronization Corrections . . . . .	64
4.4.1	UEH Correction Systematic . . . . .	68
4.5	Trigger Sample Comparison . . . . .	71
4.6	Results . . . . .	71
<b>5</b>	<b>Di-jet Helicity Asymmetries</b>	<b>76</b>
5.1	Beam Polarizations . . . . .	77
5.1.1	Residual Transverse Component . . . . .	77
5.1.2	Polarization Decay . . . . .	77
5.2	Spin Patterns . . . . .	79
5.3	Relative Luminosity . . . . .	80
5.4	Trigger Selection and Binning . . . . .	81
5.5	False Asymmetries . . . . .	83
5.6	Invariant Mass Shift . . . . .	85
5.7	Trigger and Reconstruction Bias Correction . . . . .	88
5.8	Systematics . . . . .	91
5.8.1	Mass Shift Systematic . . . . .	93
5.8.2	Jet Energy Scale Systematic . . . . .	93
5.8.3	Pythia to NLO Systematic . . . . .	94
5.8.4	Trigger and Reconstruction Bias Systematic . . . . .	95
5.8.5	Residual Transverse Component Systematic . . . . .	95
5.8.6	Total Systematic Error . . . . .	95
5.9	Results . . . . .	96
5.9.1	Comparison to Theory . . . . .	99
<b>6</b>	<b>Theory Code</b>	<b>102</b>
6.1	Calculation . . . . .	102

6.2	Modifications . . . . .	102
6.3	Results . . . . .	104
6.4	Di-jet mass definition . . . . .	107
6.4.1	Jet merging scheme . . . . .	107
6.4.2	Di-jet mass . . . . .	109
6.5	Di-jet definition cut ordering . . . . .	110
6.6	Using the code . . . . .	111
6.6.1	Installation . . . . .	111
6.6.2	Old version . . . . .	112
<b>A</b>	<b>List of Runs and Fills</b>	<b>113</b>
A.1	Cross Section Runs . . . . .	113
A.2	$A_{LL}$ Runs . . . . .	116
A.3	Cross Section Fills . . . . .	116
A.4	$A_{LL}$ Fills . . . . .	117
<b>B</b>	<b><math>A_{LL}</math> and Mass Formula Derivation</b>	<b>118</b>
B.1	Asymmetry Formula Derivation . . . . .	118
B.2	Asymmetry Error Formula Derivation . . . . .	120
B.3	Invariant Mass Formula . . . . .	122
<b>C</b>	<b>Code Location</b>	<b>124</b>
C.1	Di-jet Trees . . . . .	124
C.2	Cross Section Code . . . . .	125
C.3	Asymmetry Code . . . . .	126
	<b>Bibliography</b>	<b>128</b>

# List of Figures

1.1	QA Plots . . . . .	9
1.2	L2Jet Output Example . . . . .	12
1.3	TPC Sector Misalignment . . . . .	15
1.4	Track $\Phi$ Vs Transverse Momentum . . . . .	16
1.5	Initial 2009 Gain Data-Simulation Comparison . . . . .	19
1.6	Modified 2009 Gain Data-Simulation Comparison . . . . .	20
2.1	Vertex Re-weighting . . . . .	24
2.2	Vertex Re-weighting Result . . . . .	25
2.3	Vertex Matching . . . . .	26
2.4	Jet $p_T$ : Data - Simulation Comparison . . . . .	26
2.5	Di-jet $\Delta\eta$ : Data - Simulation Comparison . . . . .	27
2.6	Di-jet $\cos(\Delta\phi)$ : Data - Simulation Comparison . . . . .	28
2.7	Jet $\eta$ : Data - Simulation Comparison . . . . .	29
2.8	Jet $\phi$ : Data - Simulation Comparison . . . . .	29
2.9	Di-jet Mass: Data - Simulation Comparison . . . . .	30
3.1	NHits Vs Eta . . . . .	32
3.2	Opening Angle Cut . . . . .	35
3.3	Neutral Fraction Cut . . . . .	36
3.4	Pt Balance Cut . . . . .	37
3.5	Eta Cut Comparison: Di-jet Mass . . . . .	38
3.6	Eta Cut Comparison: Delta Phi . . . . .	39
3.7	Eta Cut Comparison: Asymmetry . . . . .	40
4.1	Response Matrices . . . . .	44
4.2	Dijet Matching Efficiency . . . . .	46
4.3	Bin-by-Bin to SVD Comparison . . . . .	48

4.4	Bayesian to SVD Comparison . . . . .	49
4.5	Track Inefficiency Systematic . . . . .	51
4.6	Track $p_T$ Systematic . . . . .	52
4.7	Tower Energy Scale Systematic . . . . .	53
4.8	SVD Cutoff Value Systematic . . . . .	54
4.9	SVD Bin Systematic . . . . .	55
4.10	Mass Spectrum Time Variation . . . . .	56
4.11	Average Di-jet Mass Vs Run . . . . .	57
4.12	Average NHitsFit/Pos Vs Run . . . . .	58
4.13	Time Dependent Mass Shift . . . . .	59
4.14	Vernier Scan Cross Sections . . . . .	60
4.15	Di-Jet Over BBCMB-Cat2 Ratio . . . . .	61
4.16	Di-Jet Vs BBCMB-Cat2 . . . . .	62
4.17	Di-jet Luminosity Systematic Estimation . . . . .	63
4.18	Combined Systematic Errors . . . . .	65
4.19	Particle/Parton Yield Ratios . . . . .	66
4.20	UEH Correction Scheme Comparison . . . . .	67
4.21	UEH Correction Tune Dependence . . . . .	68
4.22	UEH Correction Scale Uncertainty . . . . .	69
4.23	UEH Correction PDF Error Sets . . . . .	70
4.24	Raw Di-jet Yields . . . . .	72
4.25	L2JetHigh Vs JP1 Cross Section Comparison . . . . .	73
4.26	Di-jet Cross Section . . . . .	75
5.1	Fill-by-Fill Polarization . . . . .	78
5.2	Spin-4 Values Vs Bunch Crossing . . . . .	80
5.3	Bunch Crossing Spectrum . . . . .	82
5.4	Full Sample False Asymmetries . . . . .	85
5.5	East-West Sample False Asymmetries . . . . .	86
5.6	East-East+West-West Sample False Asymmetries . . . . .	87
5.7	False Asymmetry ( $A_L^Y$ ) for Each Run . . . . .	88
5.8	Mass Integrated False Asymmetries . . . . .	89
5.9	Full Sample Mass Shifts . . . . .	90
5.10	$\Delta A_{LL}$ Values for Trig and Reco Bias . . . . .	92
5.11	PYTHIA/NLO Cross Section Ratio . . . . .	94
5.12	$A_{LL}$ Result . . . . .	99

5.13	Mass Integrated Asymmetry values . . . . .	100
6.1	The gluon component of the polarized PDF sets fit with RHIC data (NNPDF-pol1.1, DSSV2014) and without (NNPDFpol1.0, DSSV2008). . . . .	104
6.2	Comparison of the two pQCD calculations of the unpolarized di-jet cross section at $\sqrt{s} = 200$ GeV for CT10 PDF set. . . . .	105
6.3	Theory prediction for unpolarized di-jet cross section for CT10 PDF set compared to the STAR data corrected to the parton level at $\sqrt{s} = 200$ GeV.	105
6.4	Di-jet $A_{LL}$ theory predictions against the STAR data at $\sqrt{s} = 200$ GeV. This plot needs an update to show DSSV2014/MRST2008 (small difference).	106
6.5	Correspondence between the used PDF sets . . . . .	106
6.6	Polarized, unpolarized di-jet cross sections and di-jet $A_{LL}$ for NNPDF-pol1.1/NNPDF2.3 for various renormalization and factorization scale factors at $\sqrt{s} = 200$ GeV. . . . .	106
6.7	Di-jet $A_{LL}$ theory predictions for NNPDFpol1.1 with PDF and scale uncertainties against STAR data at $\sqrt{s} = 200$ GeV. . . . .	107
6.8	NLO pQCD theory prediction for unpolarized di-jet cross section spectrum for different definitions definitions of mass (x axis). The renormalization and factorization scales are taken to be $\mu_R = \mu_F = \sqrt{(p_1^\nu + p_2^\nu)^2}$ for all three curves. . . . .	110
6.9	Comparison of NLO pQCD theory predictions for different di-jet definition cuts. . . . .	111

# List of Tables

1.1	Trigger IDs . . . . .	10
1.2	Jet Patch Positions . . . . .	11
1.3	Jet Patch Thresholds . . . . .	11
4.1	Cross Section Binning . . . . .	42
4.2	Systematic Error Values . . . . .	64
4.3	UEH Correction and Systematic Error Values . . . . .	70
4.4	Theoretical Cross Section and Systematic Error Values . . . . .	72
4.5	Di-jet Cross Section Values . . . . .	74
5.1	Residual Transverse Components . . . . .	79
5.2	Run IX Spin Patterns . . . . .	79
5.3	Spin-4 Values . . . . .	80
5.4	$A_{LL}$ Binning . . . . .	82
5.5	Helicity State Events by Trigger . . . . .	83
5.6	Helicity State Events by Topology . . . . .	84
5.7	Mass Shifts by Trigger . . . . .	87
5.8	Mass Shifts . . . . .	89
5.9	Trigger and Reconstruction Bias Correction . . . . .	91
5.10	Total $A_{LL}$ Systematics . . . . .	96
5.11	Eaet-East+West-West $A_{LL}$ Systematics . . . . .	96
5.12	East-West $A_{LL}$ Systematics . . . . .	97
5.13	Combined $A_{LL}$ Results . . . . .	98
5.14	East-East+West-West $A_{LL}$ Results . . . . .	98
5.15	East-West $A_{LL}$ Results . . . . .	98
5.16	Vals . . . . .	101



# Chapter 1

## Data Sample and Calibration

The data used in these analyses were collected during Run IX which took place during the 2009 calendar year. The data were taken during  $pp$  running with a center of mass energy of  $\sqrt{s} = 200$  GeV.

### 1.1 Data Quality Assurance

The run selection and QA were completed in several steps, the first of which was an initial human QA. A list was created of all runs longer than 1 minute which were not test or setup runs. This list was divided among several members of the STAR collaboration who looked at the Shift Log and several sets of diagnostic plots for each run, noted any problems with the detector, and discarded runs which contained major problems. From the remaining runs, only those with the trigger setups `production2009_200Gev_Hi`, `production2009_200Gev_Lo`, and `production2009_200Gev_Single` were retained. The trigger setup denotes the specific triggers included in the run: `production2009_200Gev_Hi` contains only the L2JetHigh triggered events, `production2009_200Gev_Lo` contains only the JP1 triggered events, and `production2009_200Gev_Single` contains both L2JetHigh and JP1 triggered events. The L2JetHigh and JP1 triggers will be discussed further in section 1.2. Runs which were shorter than 3 minutes or which did not include the TPC, BEMC, and EEMC were discarded. The resulting run list contained 1269 runs.

The next step in the QA process was to look at various quantities of interest to the analyses on a run-by-run basis and identify outliers. The quantities looked at are things which could be affected by the state of the detector, such as the number of tracks reconstructed

or the energy deposited in the calorimeter in a given event. For each run, the per event averages of the quantities looked at are plotted and the mean over all runs is found. Individual runs for which the per event mean of some quantity deviates too much from the global mean are investigated further, and runs which contain means farther than  $5\sigma$  from the global mean are discarded. Figure 1.1 gives an example of the type of plots used in this part of the QA. There were 1112 runs remaining after the run-by-run QA was performed.

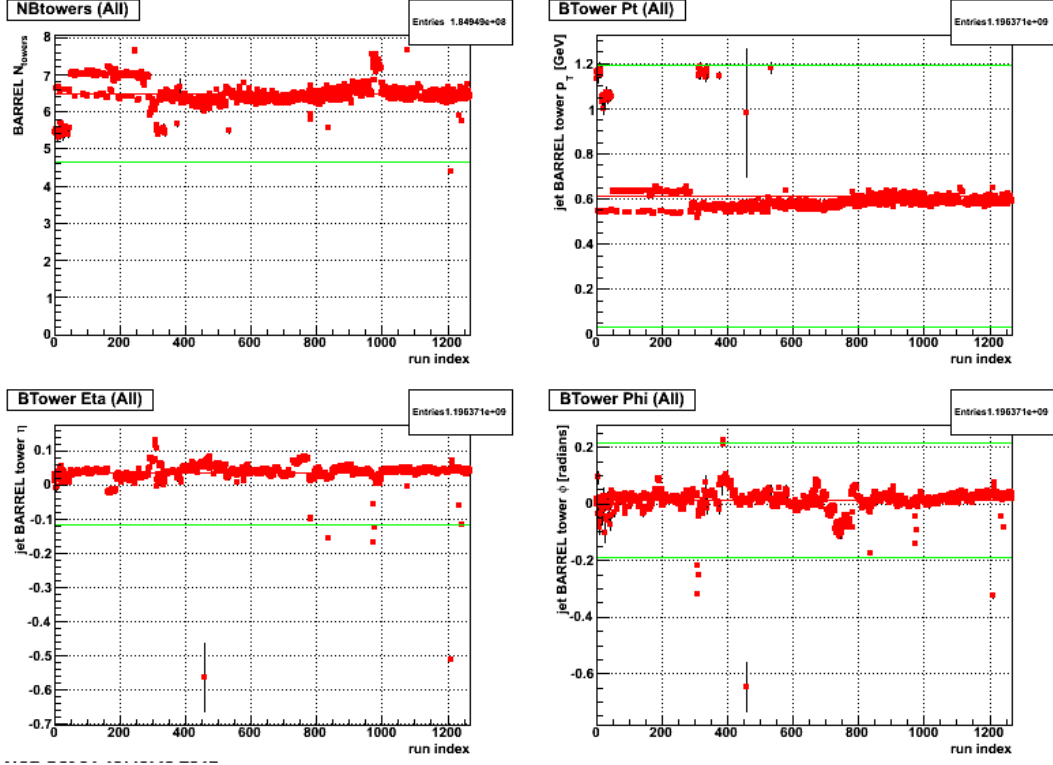


Figure 1.1: QA plot example. Each red point shows the per event mean value of the quantity specified for one run. The red line shows the mean of the points for all runs and the green lines show the  $\pm 5\sigma$  lines. Points lying outside the green lines are discarded.

The calculation of  $A_{LL}$  requires three pieces of information for each run: the polarization values for each beam, the relative luminosity values, and the jet yields for each helicity combination, which requires valid spin bit values. Relative luminosity is described in section 5.3 and the spin bits are described in section 5.2. The last step in the QA and run selection process is to remove runs in which any of the above information is missing. Removing these leaves 1029 runs. It was also found later that the production2009\_200Gev\_Lo trigger setup caused an issue with the analyses, so all fifty of these

runs were removed, leaving 979 runs.

For the cross section analysis, it was decided that only runs which had a corresponding embedded simulation run (see section 2) should be used. Embedded simulation samples were produced for 867 runs; of these, 773 matched a data run which had passed all other QA. These 773 runs were used in the cross section analysis. The complete list, along with a list of fills they are from, can be found in appendix A.

Unlike the cross section analysis, the  $A_{LL}$  measurement is statistically limited so it was decided to use all runs from day 120 onward, regardless of whether or not it had an associated embedded simulation run. There were 888 runs used in the asymmetry analysis. The additional runs and fills used in the asymmetry analysis can also be found in appendix A.

## 1.2 Trigger

The two L0 triggers used in the di-jet analyses are JP1 and L2JetHigh. The BBCMB-Cat2 trigger is also used in the determination of the integrated luminosity for the data sample. See table 1.1.

Trigger Name	Trigger ID
JP1	240410
	240411
L2JetHigh	240650
	240651
	240652
BBCMB-Cat2	240114
	240124
	240224

Table 1.1: Trigger IDs used in the analysis

The L2JetHigh and JP1 triggers are based on jet patches. In Run IX, there were 30 jet patches each spanning an area of  $1 \times 1$  in  $\eta - \phi$  space (See table 1.2). Each jet patch has 3 associated ADC thresholds (from lowest to highest): TH-0, TH-1, and TH-2 (See table

1.3). The JP1 trigger fires if one of the jet patch ADCs is above the TH-1 threshold. The L2JetHigh trigger fires if one jet patch ADC is above the TH-2 threshold or if two jet patches above the TH-0 threshold are adjacent in azimuth. Note: this so called adjacent jet patch logic is not implemented for the jet patches which span the service gap between the BEMC and EEMC.

Events which pass the L2JetHigh trigger are passed to the jet algorithm at Level-2. The Level-2 trigger is a set of software based algorithms which can make more sophisticated trigger decisions than what is possible at L0. Because the L2 jet algorithm is implemented in software, it is free to scan all possible  $1 \times 1$  regions in the calorimeters, not just the patches defined in hardware. The algorithm finds the patch with the largest  $E_T$  and then finds the patch with the next highest  $E_T$  which is at least 30 degrees away in azimuth from the first patch.

	$\eta$ - Position
BEMC-East	$-1 < \eta < 0$
BEMC-Overlap	$-0.6 < \eta < 0.4$
BEMC-West	$0 < \eta < 1$
BEMC-EEMC-Overlap	$0.4 < \eta < 1.4$
EEMC	$1 < \eta < 2$

Table 1.2: The  $\eta$  ranges spanned by the each of the 5 categories of jet patches.

	TH-0	TH-1	TH-2
BEMC-EEMC-Overlap	19	26	34
BEMC	20	28	36
EEMC	18	25	32

Table 1.3: Typical ADC thresholds for jet patches. Approximate transverse energy value is  $(\text{ADC}-1) \times 0.236$ .

The L2Jet algorithm defines three trigger categories: monojet, dijet, and random. The monojet trigger is satisfied if the high patch transverse energy is above the monojet

threshold, typically set around 6.5 GeV. The dijet trigger is satisfied if both patch transverse energies are above separate dijet thresholds which are somewhat lower than the monojet threshold. There are three sets of dijet thresholds and which set is used is based on the sum of the energy weighted  $\eta$  positions of the two patches. This functionality was included because patches at forward pseudorapidities tend to have lower transverse energies so lower thresholds are needed in the forward region to keep the dijet accept rate somewhat constant across the detector. The random trigger, as its name suggests, accepts random events at a set rate regardless of whether or not the event would have passed either the monojet or dijet conditions. The random trigger accept rate can be set to 100% to force L2Jet to accept all events it sees. A given event can satisfy multiple L2Jet trigger categories simultaneously but events which don't satisfy any of the above categories are dropped and not used in these analyses. An example of the kinematic distributions of events passing the dijet trigger condition can be seen in figure 1.2.

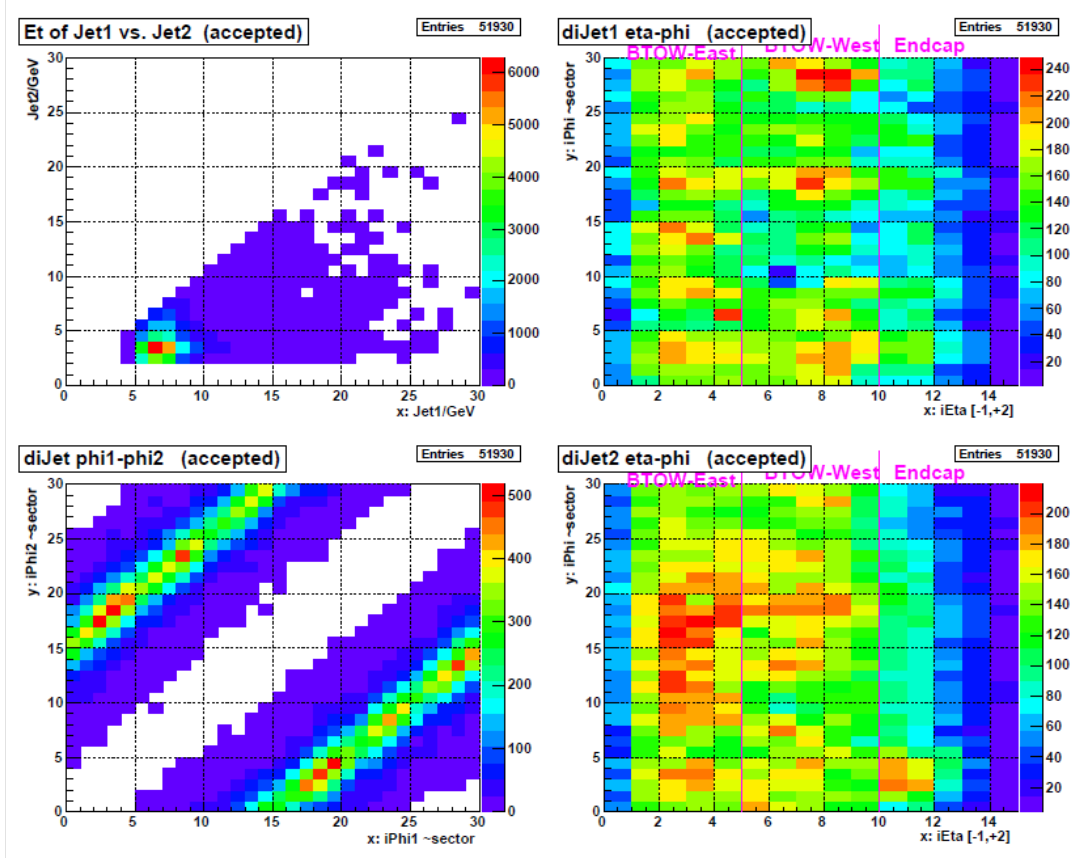


Figure 1.2: Example histograms from the L2Jet algorithm for events passing the dijet condition. Clockwise from upper left: the  $p_T$  vs  $p_T$  of the two jet patches, the energy weighted  $\phi$  vs  $\eta$  position of the high jet patch, the energy weighted  $\phi$  vs  $\phi$  of the two jet patches, and the energy weighted  $\phi$  vs  $\eta$  position of the low jet patch.

### 1.3 Integrated Luminosity

A critical component of the cross section analysis is the determination of the integrated luminosity of the data sample, which is (roughly speaking) a measure of how many collisions could have occurred in the data being analyzed.

The cross section ( $\sigma$ ) for a given final state is proportional to the probability that a single collision will produce said final state. Thus, if a final state is observed  $N$  times, that number needs to be normalized by the integrated luminosity in order to determine the cross section. The relationship can be written as [1]:

$$\sigma = \frac{N}{\int \mathcal{L}(t) dt} \quad (1.1)$$

The form of equation 1.1, with the integrated luminosity as a known quantity, is essentially what is used to find the di-jet cross section. If there is some other process for which the cross section has previously been determined, the above equation can be inverted to find the integrated luminosity. At STAR, the absolute BBC trigger cross section was measured using the Vernier Scan technique [2] and was found to have a value of  $26.1 \pm 0.2(\text{stat}) \pm 1.8(\text{sys})$  mb [3]. As will be discussed in section 4.3.3, a new determination of the BBCMB-Cat2 trigger cross section was performed using Vernier scans taken during the Run 9 200 GeV period. The new BBCMB-Cat2 trigger cross section was found to be  $24 \pm 5.4\%$ . Using this cross section, the integrated luminosity of the data sample can be found by counting the number of BBC coincidence triggers and scaling by the absolute BBC cross section.

There are two factors which need to be taken into account when using the BBC to determine the integrated luminosity: dead time and prescale. The L2JetHigh and JP1 L0 triggers have an associated dead time which is related to the detectors which these triggers read out. The BBC coincidence trigger which is used in the integrated luminosity determination must have the same dead time conditions as the jet triggers, or it won't sample the same events seen by the jet triggers. The BBCMB-Cat2 trigger was set up so as to have the same dead time conditions as the L0 jet triggers and is thus used to determine the integrated luminosity. The second issue which needs to be addressed is the prescale associated with the BBCMB-Cat2 trigger. The number of BBCMB-Cat2 triggered events recorded in the data files is actually scaled by a factor of  $1.0 \times 10^5$  from the scalar system. The number of BBCMB-Cat2 triggers must be multiplied by this prescale factor to get the actual number of BBC coincidence events needed in order to

find the integrated luminosity.

The integrated luminosity of the data set used in the cross section measurement was determined to be  $18.6 \text{ pb}^{-1}$ . The integrated luminosity of the data used in the  $A_{LL}$  analysis was determined to be  $21.2 \text{ pb}^{-1}$ .

## 1.4 Average JP1 Pre-Scale

The JP1 trigger was satisfied at a lower energy threshold and thus fired at a higher rate than the L2JetHigh trigger. In order to keep the JP1 trigger from filling up the available trigger bandwidth, it was prescaled throughout the run. The prescale factor changed from run to run based on a measure of the instantaneous luminosity seen at STAR at the beginning of that run. This prescale factor must be taken into account in the calculation of the JP1 cross section, to convert from the number of events seen to a ‘true’ number of events produced from the  $pp$  collisions. (Note: the L2JetHigh trigger was always take-all).

Because the JP1 prescale factor changed from run to run, it is the event-weighted average prescale which must be used in the determination of the cross section. The average prescale was found by summing the prescale factors of all events firing the JP1 trigger and dividing by the total number of JP1 events. All JP1 events entering the analysis were used in the determination of the average prescale whether or not that event yielded a di-jet.

The properly weighted average JP1 prescale for the data sample was found to be 5.16

## 1.5 Detector Calibrations

The analyses presented here rely primarily on data from three sub-components of the STAR detector, the TPC, BEMC, and EEMC. These sub-systems must all be calibrated in order to yield useful data.

### 1.5.1 TPC Calibration

There are a number of calibrations the TPC must undergo in order for it to realize its full capability. The most important calibrations for these analyses are those which affect the  $p_T$ s of reconstructed tracks. Two calibrations which can affect the track  $p_T$  are the

corrections for Space Charge and Grid Leak (SC&GL) [4]. The SC&GL calibrations correct for the electromagnetic distortion caused by drifting charges from the ionized gas. As part of the SC&GL calibration, the orientations of the inner and outer TPC sectors with respect to each other are allowed to vary. Misalignment of the sectors can change the curvature, and therefore the  $p_T$ , of tracks. See 1.3 for a cartoon showing how this could be possible.

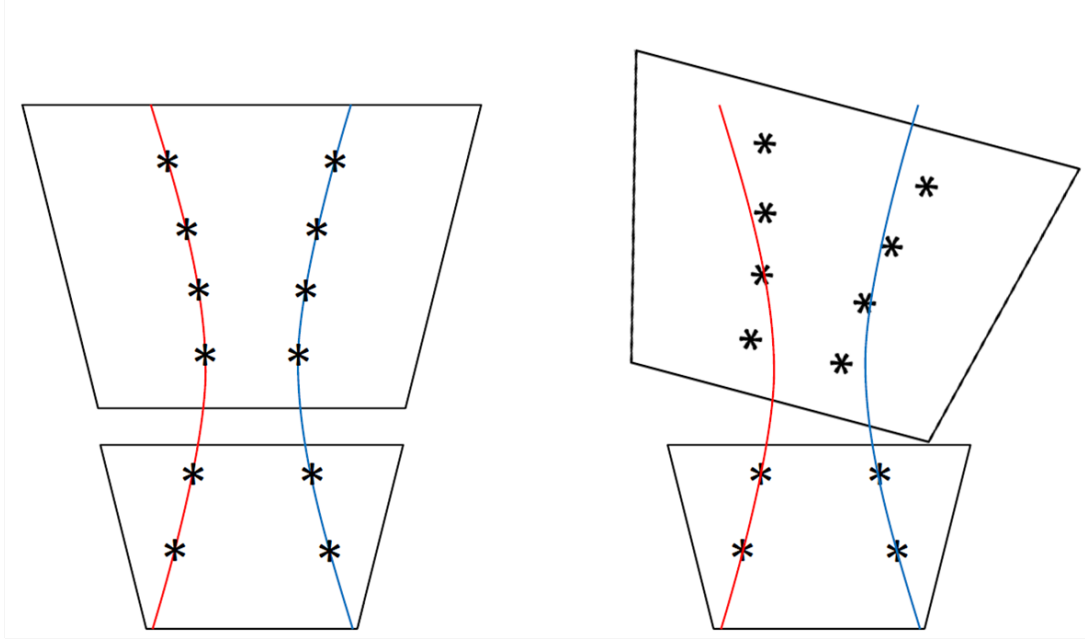


Figure 1.3: Cartoon showing how TPC sector misalignment could affect track  $p_T$ . The left side of the figure shows properly aligned TPC inner and outer sectors with positively and negatively curved particle trajectories (red and blue curves). The black stars represent padrow hits caused by the particles. The right side of the figure shows an improperly aligned TPC sector. The particle trajectories and padrow hits are in the same position, but the misaligned sectors make the track reconstruction think the padrow hits have more / less curvature.

The SC&GL calibration is performed for each of the 24 TPC sectors for both the Reverse Full Field (RFF) and Full Field (FF) magnetic field configurations. The original production (in library P10ic) exhibited an excess of very high  $p_T$  tracks in an isolated region of the TPC in the FF runs but not in the RFF runs (see figure 1.4). The region in question is sector 20, which has had calibration problems in the past.

It was decided that the best way to remedy this issue was to redo the TPC calibrations



and reproduce both the RFF and FF data. For the new calibrations (P11id) the TPC geometry was fixed to the values from the year 2000. When the SC&GL calibration was performed, the TPC alignments were no longer free parameters. The new calibration and production removed the excess of high  $p_T$  tracks in sector 20 in the FF data 1.4.

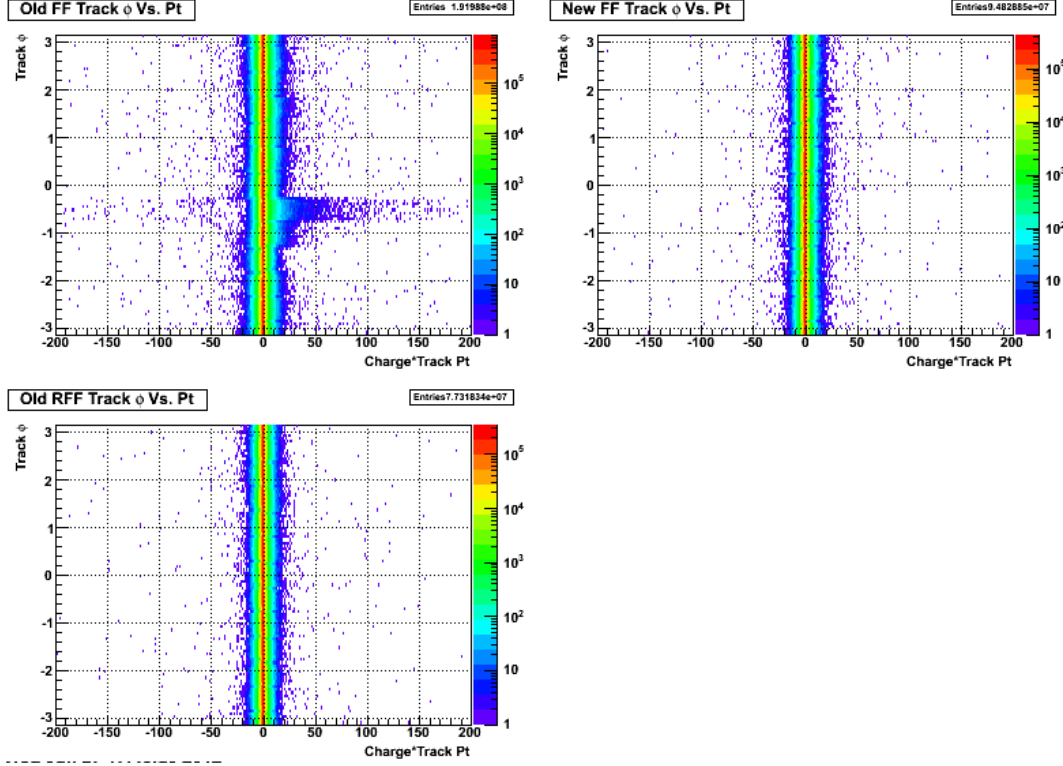


Figure 1.4: This figure shows the track phi vs the charge signed transverse momentum. The upper left panel shows the FF tracks from the P10ic production and a large excess of high  $p_T$  positively charged tracks can be seen at  $\phi = -0.5$ . For comparison, the bottom left panel shows the RFF tracks from the same production and the high  $p_T$  excess is clearly not present. The upper right panel shows the FF tracks from the P11id production which behave much more like the P10ic RFF tracks.

### 1.5.2 BEMC Calibration

The BEMC is a lead-scintillator sampling calorimeter and is used to measure the transverse energy of electromagnetically showering particles (primarily neutral pions). As particles deposit energy in the scintillator, light is produced, detected by a photo multiplier tube, and converted into a digital signal. The calibration is used to determine a gain factor which relates the size of the digital signal to the amount of transverse energy which was deposited in the calorimeter. The calibration was set such that a transverse

energy deposit of 60 GeV would appear in channel 4096, that is it would saturate a 12-bit ADC.

A two step process is used to calibrate the BEMC (details on the calibration can be found in [5] [6]). First, a relative calibration for each tower was found using the response from minimum ionizing particles (MIPs). MIPs are hadrons (mostly charged pions) which do not shower in the calorimeter. The amount of energy a MIP deposits in the calorimeter is roughly independent of particle species or momentum and was found from test beam data and simulation to have the functional form [7]:

$$\text{MIP} = (264 \pm 4_{\text{stat}} \pm 13_{\text{sys}} \text{MeV}) \times (1 + 0.056\eta^2) / \sin \theta \quad (1.2)$$

where  $\eta$  and  $\theta$  are the pseudorapidity and polar angle, respectively, of a given tower. MIPs are copiously produced in  $pp$  collisions, so nearly all towers see enough MIP hits to form a useful spectrum.

The second step in the process is the determination of the absolute calibration, which is done using electrons. Electron tracks are selected using  $dE/dx$  measurements from the TPC and only those tracks which enter and exit the same tower are used in the calibration. The number of electrons identified for the calibration is much lower than the number of MIPs, so an absolute calibration can not be done for each tower using electrons. To aggregate sufficient statistics, the electrons which strike towers read out by the same electronics crate at a given pseudorapidity (there are 8 towers at a given  $\eta$  which are read out by the same crate and there are 40  $\eta$  rings) are grouped together and the ratio of the energy deposited in a calorimeter tower by an electron to the momentum of the electron as measured by the TPC is formed. The distribution of this ratio for a given crate-pseudorapidity group is then fit with a Gaussian and the offset of the mean from unity is used as the correction to the absolute scale for that particular crate-pseudorapidity grouping. Using the absolute corrections from the electrons and the relative corrections from the MIPs, the gain factors for each tower were found.

The overall uncertainty on the tower energy scale of 3.7% was reached by combining the error quoted in the calibration note [6] due to observed differences in charge sign and TPC field direction with several other sources of uncertainty such as tower edge effects and the linearity of the calorimeter response with respect to energy (private communication with Will Jacobs). The uncertainty on the tower energy scale was also determined as part of the 2009 STAR  $W$  cross section analysis [8]. In that analysis, the BEMC calibration was

determined by comparing the measured lepton  $E_T$  spectrum in the  $W$  signal region with a series of ‘template’ distributions which were generated from GEANT particle level lepton energies multiplied by different energy scale factors. The energy scale factor corresponding to the template which best described the data, as determined by the maximum likelihood method, was used as a correction factor in the  $W$  cross section analysis. The quoted uncertainty on the BEMC calibration from this method was 3.6%.

### 1.5.3 EEMC Calibration

Like the BEMC, the EEMC must be calibrated in order to relate an observed ADC signal to the actual transverse energy deposited in the calorimeter. Because of the falling TPC tracking efficiency in the EEMC region, calibration methods which rely on the TPC, such as the electron  $E/p$  method used in the BEMC, are of limited use. In addition, an absolute calibration based on the reconstruction of the neutral pion invariant mass was not practical before 2009 due to problems with the simulation of the EEMC. Given these issues, the MIP method was chosen for the calibration of the EEMC.

The lack of particle identification from the TPC over a significant amount of the EEMC means that charged hadron tracks could not be tagged as possible MIP candidates, so a calorimeter based MIP identification / calibration scheme was developed. A MIP is identified by requiring an isolated amount of energy, consistent with the expected MIP response, in all layers of a given calorimeter tower. Isolation is enforced by requiring energy deposits in two neighboring SMD strips in each plane while demanding several empty strips on either side. The coincidence between the fired SMD strips in the two layers can also be used to place fiducial cuts on the MIP candidate. With MIP candidates selected, the calibration of a given layer (preshower layers 1 and 2, SMD layers U and V, the postshower layer, and the entire tower) in a given EEMC tower proceeds by requiring an energy deposit consistent with a MIP be present in all other layers. This is done for each layer in the tower and then the entire process is iterated until the gains for each layer converge.

The gains initially used in the 2009 analyses were based on a MIP calibration done using  $pp$  data taken in 2006 (these 2006 gains were adjusted to take into account changes in the PMT high voltages which were made at the end of the 2007 run). Preliminary investigations of neutral pions in the EEMC showed that the reconstructed pion mass from the 2009 data was roughly 8% lower than the pion mass reconstructed using the

2006 data, which matched the PDG value of 135 MeV. In addition, the simulation sample created for the 2009 inclusive jet analysis, which used the initial 2009 gains, showed a large excess of events as compared to the data in the EEMC region (see figure 1.5). Both of these observations were qualitatively consistent with a drop of the physical gain of the calorimeter of roughly 7% to 10% as compared to 2006.

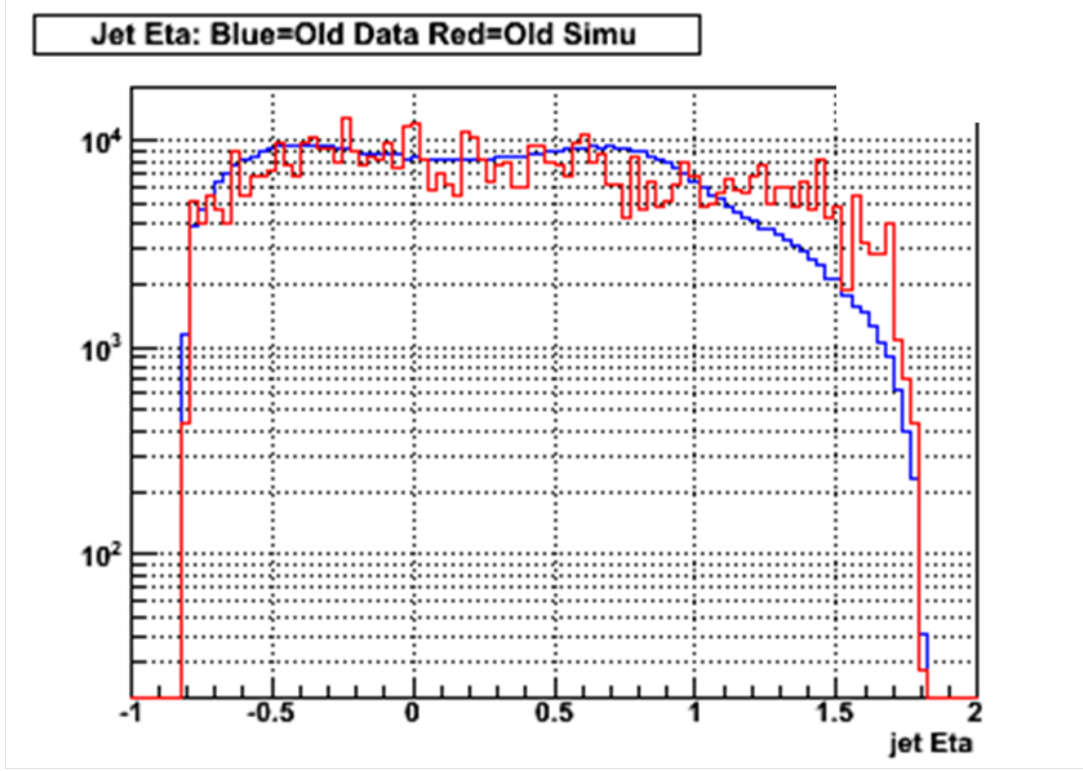


Figure 1.5: Data (blue) and simulation (red) comparison for the jet  $\eta$  spectrum using the initial gain values from 2009. Note that this comparison includes all jets found, not just those which are identified as a di-jet.

Based on the above observations, the EEMC gains initially used for the 2009 analyses were lowered by 7.7% for each tower. The inclusive jet simulation was rerun using the lower gains and the agreement with data became much better (see figure 1.6). This simulation, which used the lowered EEMC gains, was used in these analyses.

The EEMC gain drift between 2006 and 2009 does not greatly effect the mid-rapidity results presented here as the jet thrust axes were restricted to the BEMC region, however, it will be of primary importance for the future extension of the di-jet cross section and  $A_{LL}$  into the EEMC. The data / simulation agreement shown in figure 1.6 is for a limited

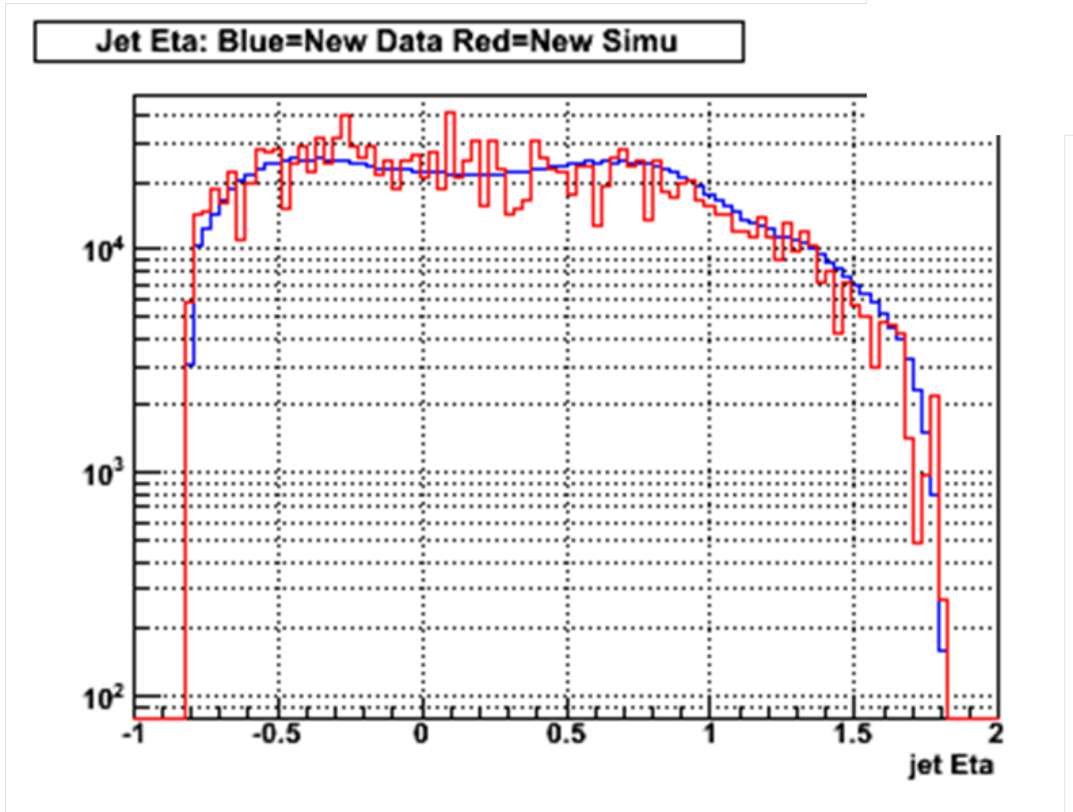


Figure 1.6: Data (blue) and simulation (red) comparison for the jet  $\eta$  spectrum using the 2009 EEMC gains which were lowered by 7.7%. Note that this comparison includes all jets found, not just those which are identified as a di-jet.

statistics sample and does not take into account issues such as trigger grouping. In addition, there may be pseudorapidity and azimuthal dependencies to the tower gain changes which were not represented in the overall gain shift of 7.7%. More work will need to be done to quantify and mitigate the effect of this gain shift for future EEMC analyses.

# Chapter 2

## Simulation Sample

Accurate simulation is vital to both the cross section and  $A_{LL}$  analyses to correct for various detector effects so that comparisons to theory can be made. The simulation sample used for these analyses was originally produced for the 2009 inclusive  $A_{LL}$  analysis [9].

The simulation sample consists of roughly 21 million  $pp$  collisions at  $\sqrt{s} = 200$  GeV generated using the Pythia [10] event generator in ten separate partonic  $p_T$  bins. Version 6.426 of Pythia with the Perugia 0 tune [11] was used. The response of the STAR detector to the Pythia events is modeled using the GSTAR package which is based on GEANT 3 [12]. The Pythia events are also embedded into real zero bias events to simulate the effect of pileup. These zero bias events were acquired by triggering the STAR detector on random bunch crossings, and were taken from the same runs that were used in the analysis in order to accurately simulate the detector conditions throughout the run. The simulation is broken into individual runs, and the embedding files used in that simulation run are taken from the corresponding data run so that a given simulation run will match the detector conditions seen during that data run.

A significant amount of computing time is needed to fully simulate the STAR detector response to a Pythia event. In order to reduce the amount of time needed to run the simulation, a trigger filter was used. The trigger filter rejects events which would not have fired the JP1, AJP, or BHT3 (BHT3 is a Barrel high tower trigger) triggers before the detector response is simulated. The trigger filter rejected roughly 91.5% of all Pythia events; however, the full Pythia record for the rejected events is saved so that corrections to the unbiased sample can be made. The embedded simulation would have taken a

prohibitively long time to generate without the use of the trigger filter.

For more detail about the simulation sample, including tables listing simulation statistics, see Section 8 of [9].

## 2.1 Three Levels of Simulation

The simulation consists of three distinct levels of information corresponding to the partonic hard scattering, the hadronization of the scattered partons, and the response of the detector to the final state particles. These divisions will be referred to as the parton level, particle level, and detector level, respectively.

The parton level of the simulation contains information about the partons involved in the  $2 \rightarrow 2$  hard scattering event generated by Pythia. Various kinematic properties of the hard scattering, such as the  $Q^2$ , center of mass scattering angle, and momentum fractions of the incoming partons are stored. When the jet finder is run on the parton level of the simulation, only the partons involved in the hard scattering and partons which arise from initial or final state radiation are included. Partons arising from the underlying event or beam remnant are not included in the parton-level jet finding.

The partons generated by Pythia propagate and hadronize to form stable, color-neutral particles. The particle level of the simulation records the kinematic information, particle id, and index of the parent parton for all of these stable particles. When the jet finder is run on the particle level, all stable particles are used, including those which arise from the underlying event and beam remnant.

The last level of the the simulation records the detector response to the stable particles from the previous level. As the particles traverse the GSTAR model of the detector, they interact with the various volumes in ways consistent with how that particular particle would behave in a specific material. This interaction includes ionizing the gas in the TPC and depositing energy in the scintillator of the calorimeters. This, along with a simulation of the detector readout electronics, allows the simulation to respond to particles as the real detector would. When the jet finder is run on the detector level simulation, it constructs jets from the simulated response of the TPC and calorimeter towers, and their readout electronics.

## 2.2 Z-Vertex Correction

The simulated events are thrown with some distribution of z-vertex position which may not exactly match the z-vertex distribution of the data. In order to achieve the best matching between the data and simulation, the z-vertex distribution of the simulation should be altered to match that of the data. This is done by re-weighting the simulation. The re-weighting is performed separately for the two trigger categories used in the cross section analysis, and is only performed on the simulated events which pass the trigger filter and undergo the full detector reconstruction.

The re-weighting is done by taking the ratio of the z-vertex distribution observed in the data to that of the simulation. This ratio is then fit with a 4<sup>th</sup> order polynomial, as shown in figure 2.1. When the simulation is used in the analysis, the function which fits the ratio is evaluated at the z-vertex of the simulated event and the value returned is multiplied by the weight associated with the partonic  $p_T$  bin in which the simulated event was generated. Figure 2.2 shows the z-vertex distributions from data and simulation after the z-vertex re-weighting has been applied.

It should also be noted that there is a matching condition placed on the z-vertices from the particle / parton level and the detector level. The vertex associated with the particle / parton level is the one which is set when Pythia is run, whereas at the detector level the vertex is found by emulating the vertex finder which is run on the actual data. This can cause a small difference between the vertex position at the particle / parton and detector levels. The fact that the simulation is embedded into physical zero-bias events can also change the detector level vertex position as more tracks are added to the event. Once in a while a zero-bias event will contain an actual hard scattering event, and the vertex of that event will be used as the detector level vertex, causing a large mismatch between the particle / parton vertex and the detector level vertex. To avoid these cases, only simulated events where the particle / parton level z-vertex and detector level z-vertex are within two centimeters are used, see figure 2.3.

## 2.3 Data - Simulation Comparison

In order to be useful for correcting detector effects, the simulated detector response must closely match the actual data for the quantities and kinematic regions of interest. The main kinematic quantity used in these analyses is the di-jet invariant mass. The di-jet



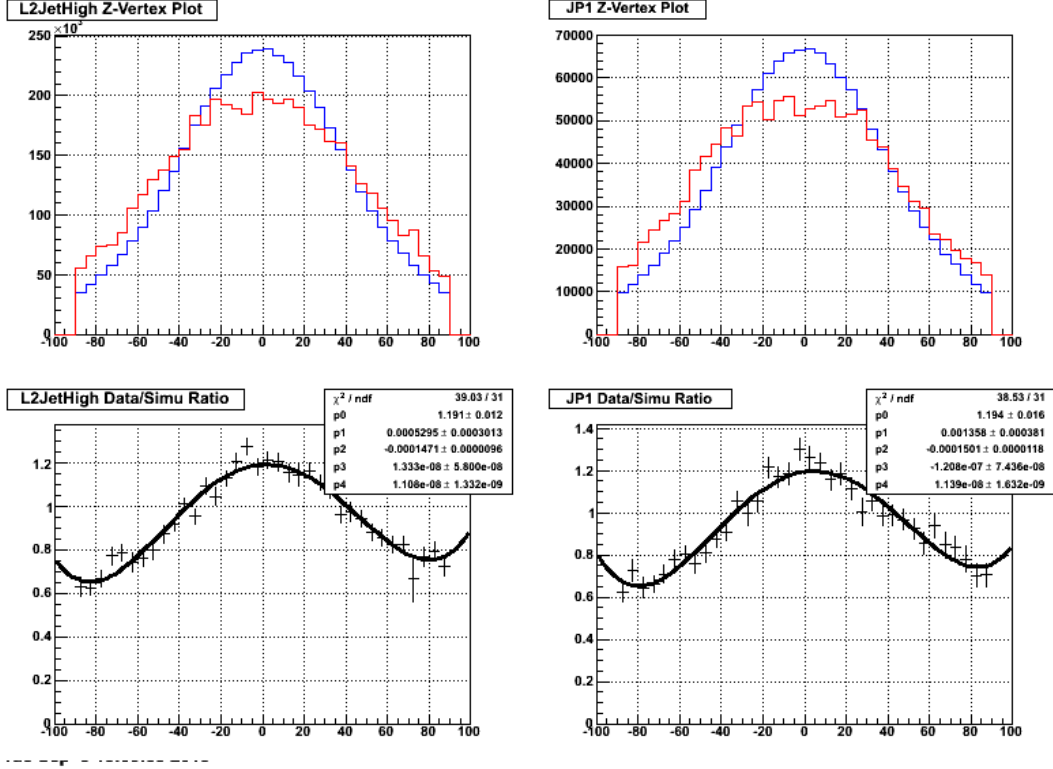


Figure 2.1: z-vertex distributions from the data (blue curves) and the simulation (red curves) for the L2JetHigh (left) and JP1 (right) trigger categories. Also shown are the data / simulation ratio fit with a 4<sup>th</sup> order polynomial.

mass depends mainly on the  $p_T$  of the two jets, the difference in pseudorapidity, and the difference in azimuthal angle between the two jets and can be written:

$$M = \sqrt{m_3^2 + m_4^2 + 2\sqrt{m_3^2 + p_{T3}^2}\sqrt{m_4^2 + p_{T4}^2}\cosh(\Delta\eta) - 2p_{T3}p_{T4}\cos(\Delta\phi)} \quad (2.1)$$

Figures 2.4, 2.5, and 2.6 show the comparisons between data and simulation for these quantities. Figures 2.7 and 2.8 show the comparisons between data and simulation for the jet pseudorapidity and jet azimuthal angle distributions. The good agreement found between data and simulation for jet  $\eta$  and  $\phi$  show that the detector conditions are well matched in the simulation, as the  $\phi$  spectrum in particular is sensitive to hardware failures in the TPC. Finally, figure 2.9 shows the comparison between the di-jet invariant mass as measured in data and in the embedded simulation, and very good agreement is seen. This gives confidence that the response of STAR is well understood, including the effects of hardware problems.

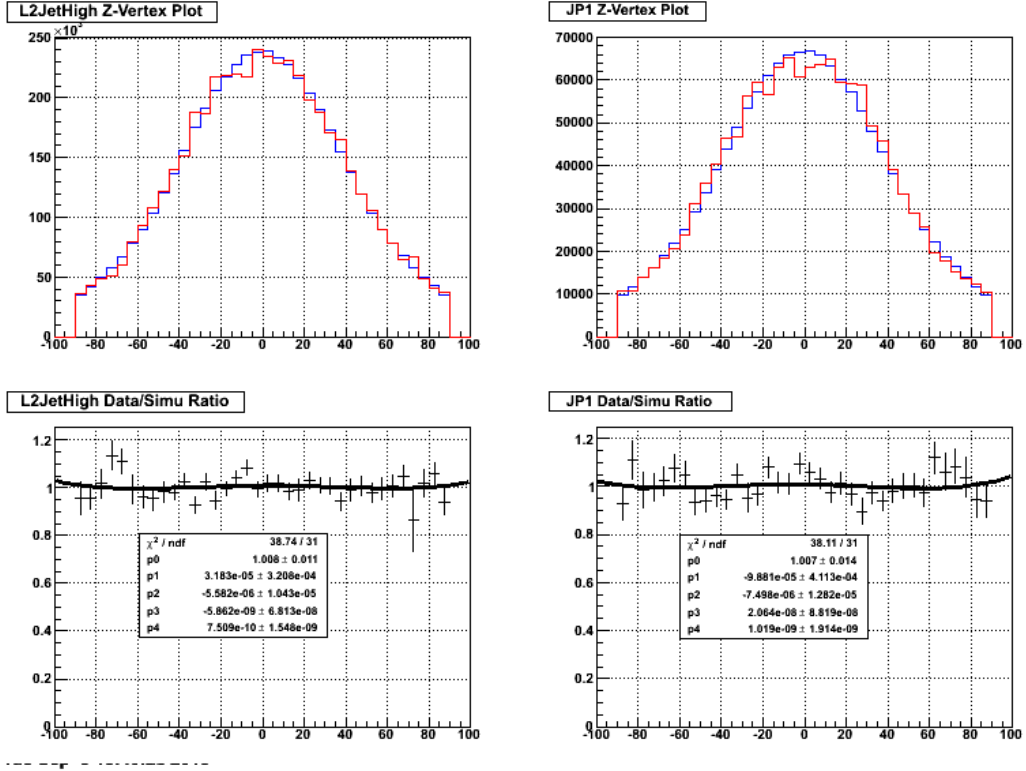


Figure 2.2: z-vertex distributions from the data (blue curves) and the simulation (red curves) after the vertex re-weighting has been applied.

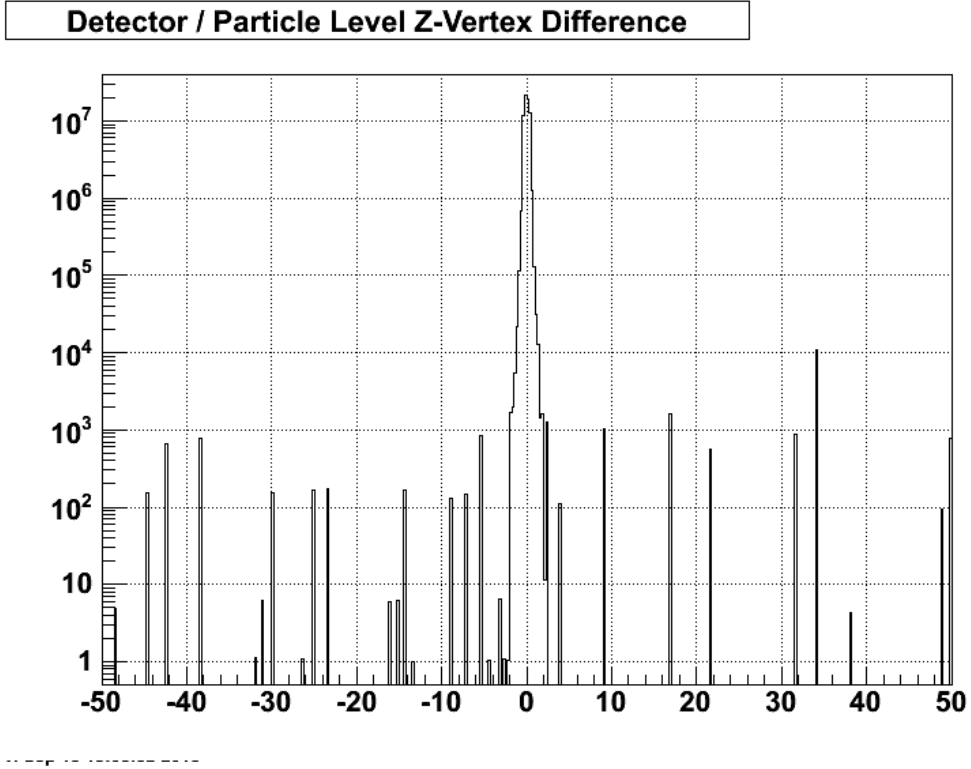


Figure 2.3: Difference in  $z$  between the particle level vertex and the detector level vertex.

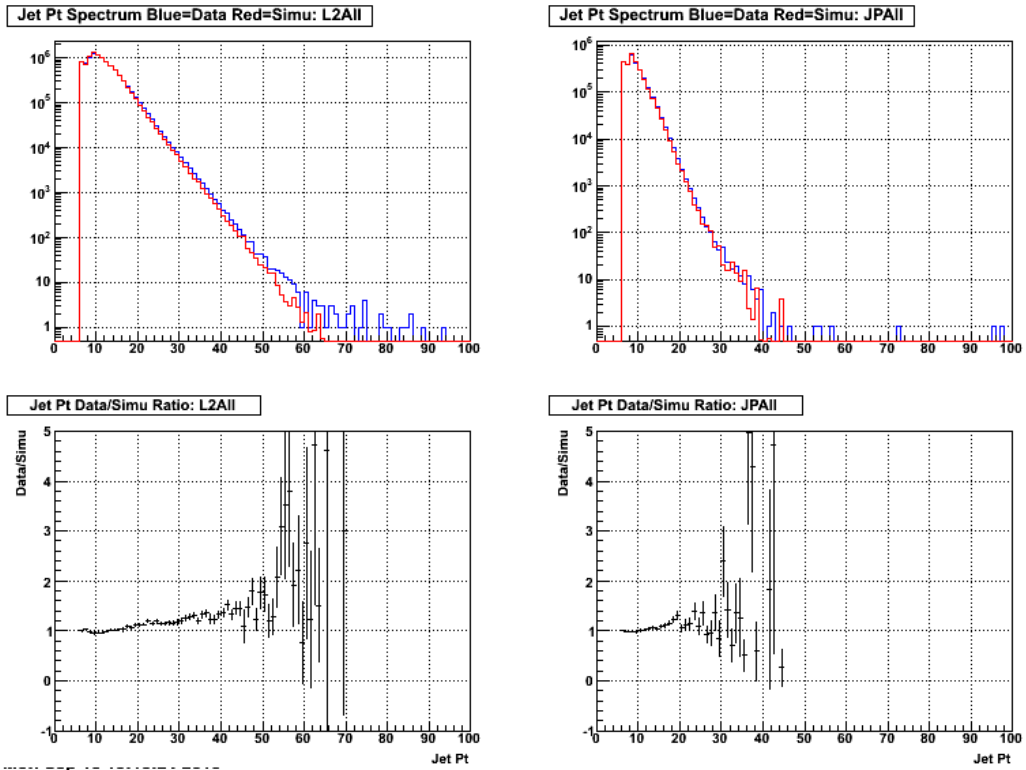


Figure 2.4: Data / simulation comparisons for the jet transverse momentum.

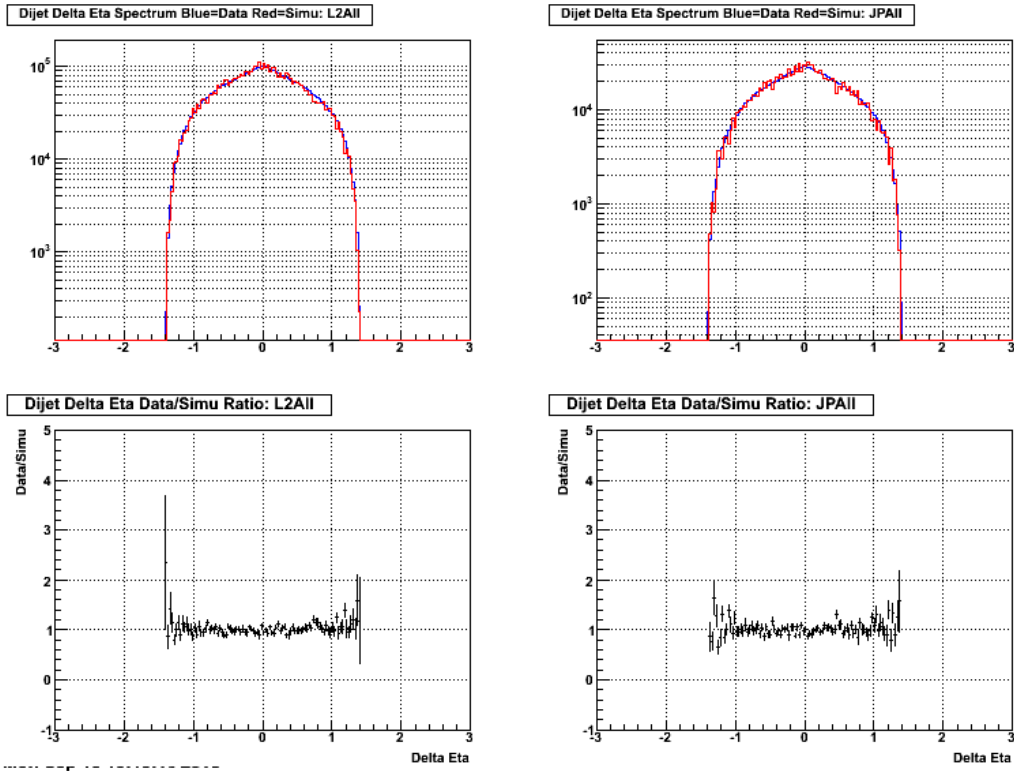


Figure 2.5: Data / simulation comparisons for the di-jet pseudorapidity difference.

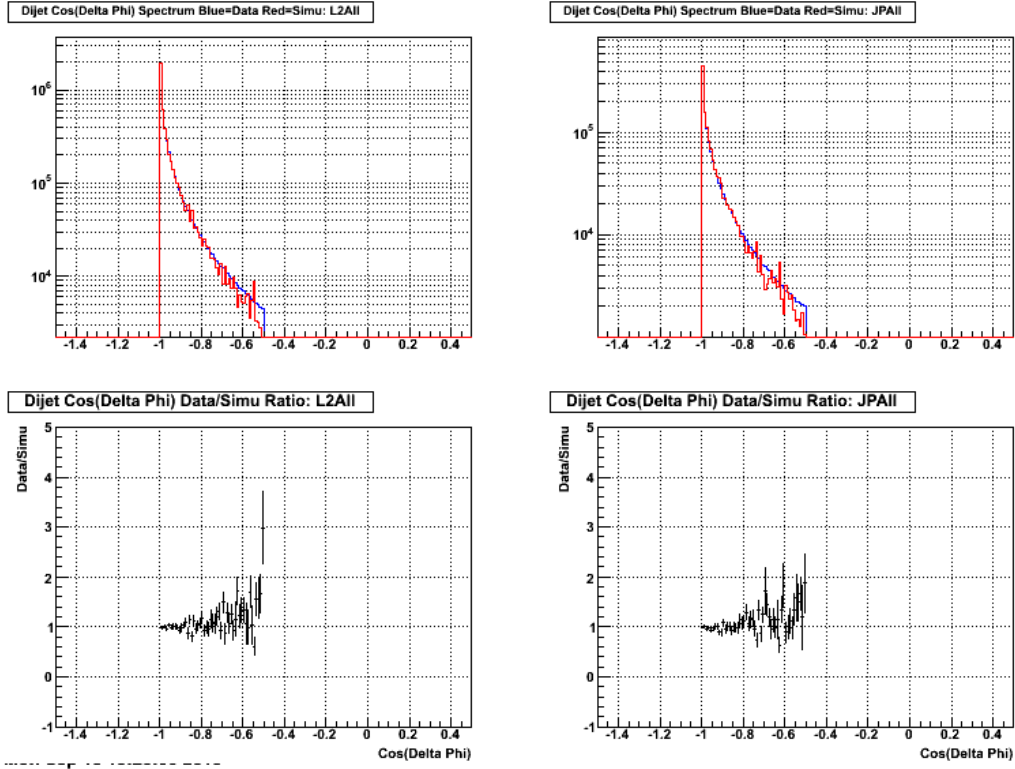


Figure 2.6: Data / simulation comparisons for the cosine of the di-jet azimuthal angle difference.

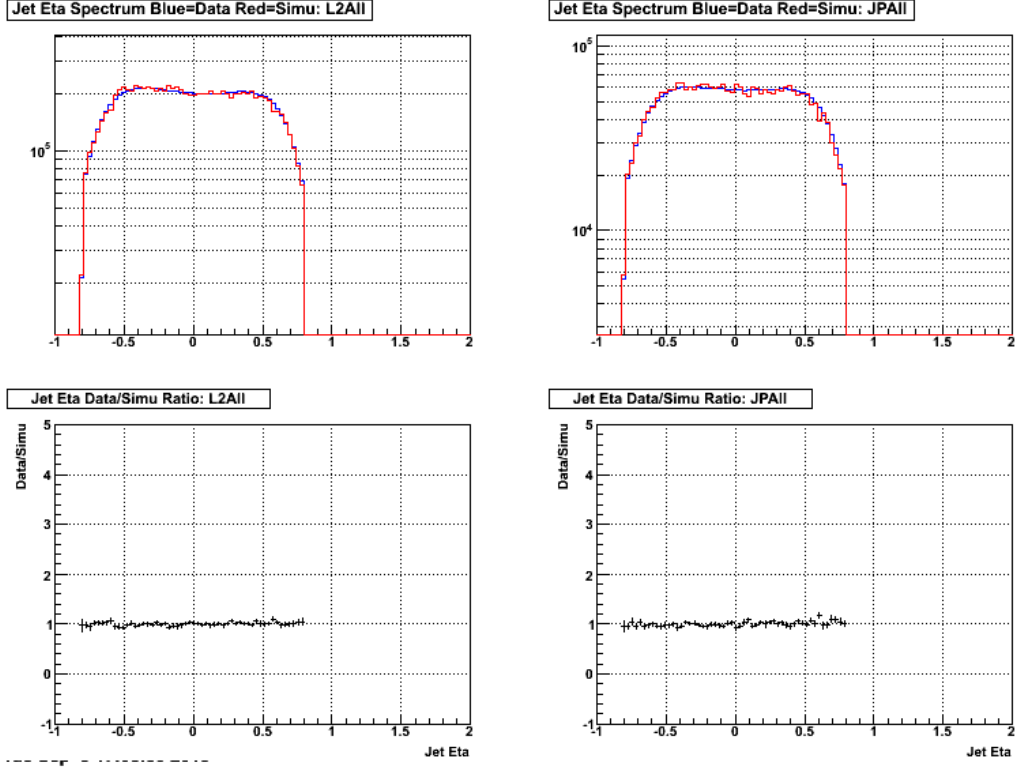


Figure 2.7: Data / simulation comparisons for the jet pseudorapidity.

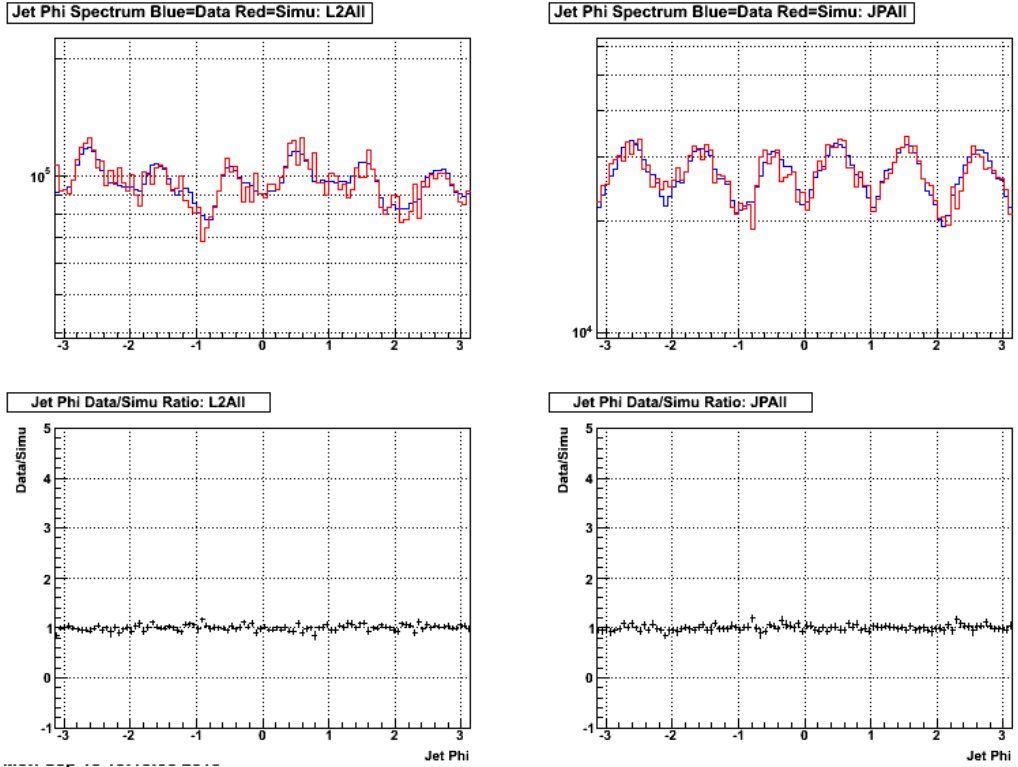


Figure 2.8: Data / simulation comparisons for the jet azimuthal angle.

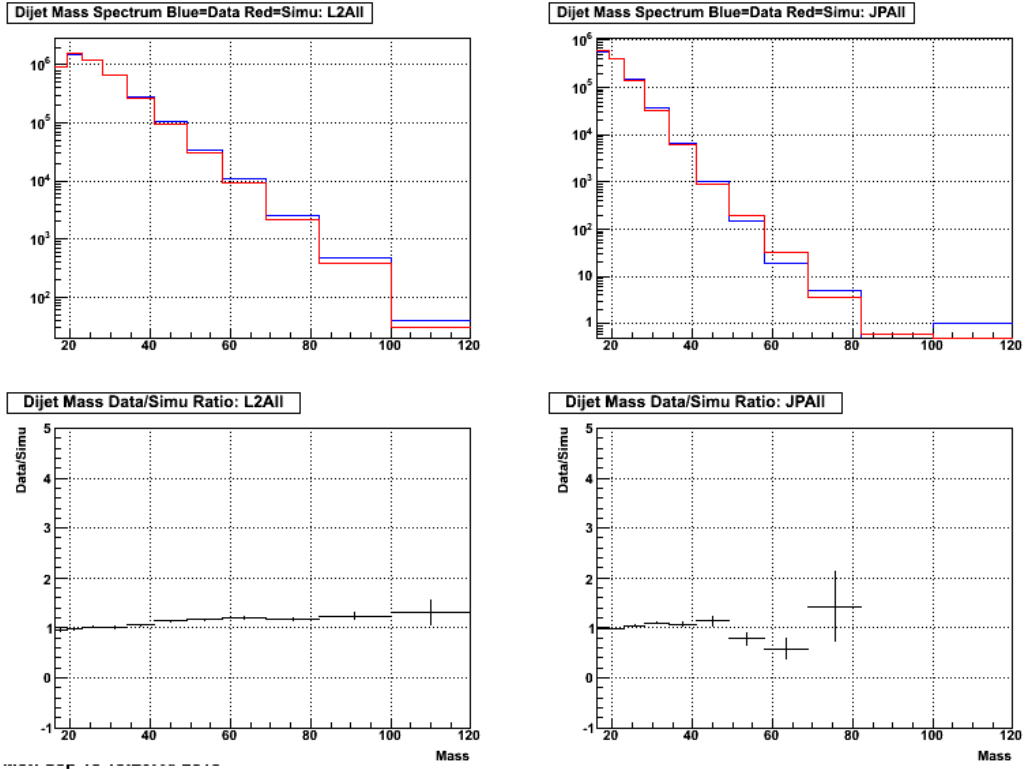


Figure 2.9: Data / simulation comparisons for the di-jet invariant mass.

# Chapter 3

## Jet and Di-jet Reconstruction

Jet-finding for the di-jet cross section and  $A_{LL}$  analyses utilize the anti- $k_T$  algorithm [13] (as implemented in the FastJet package [14]) with a radius of 0.6 and the E-scheme for particle recombination. This is the same jet definition as that used in the inclusive jet  $A_{LL}$  analysis [9]. Good overviews of different jet algorithms and jet finding in general can be found in [15] and [16]. The conditions on TPC tracks and calorimeter towers are nearly identical to the inclusive analysis with the changes meant to facilitate the future extension of the cross section and  $A_{LL}$  measurements into the EEMC.

### 3.1 TPC and Calorimeter Requirements

When acting on data (or full detector simulation), the jet-finder takes as input TPC tracks and calorimeter towers from the BEMC and EEMC. There are several requirements which tracks and towers must satisfy in order to be included in jet-finding.

TPC tracks must satisfy the following conditions in order to be used in jet-finding:

- Flag > 0
- Track  $p_T > 0.2$  GeV
- $|\eta| < 2.5$
- Number of padrow hits > 5
- Number of padrow hits divided by number of possible hits > 0.51
- Transverse momentum dependent DCA cut



Most conditions are self explanatory. The  $p_T$  dependent DCA cut rejects tracks which have a 3-D DCA greater than or equal to 2 cm if the track has a  $p_T$  less than 0.5 GeV/c and tracks with a DCA greater than or equal to 1 cm if the track has a  $p_T$  greater than 1.5 GeV/c. The cut is linearly interpolated between  $p_T$ s of 0.5 and 1.5 GeV/c.

The condition on the number of padrows differs from previous jet analyses at STAR which have required that tracks have more than 12 hit padrows in order to be included in jet-finding. Tracks which point at the EEMC deposit charge over fewer padrows due to the geometry of the TPC. Reducing the number of required padrows allows more tracks to be reconstructed and included in jet-finding. Note: the 5 point tracking does not extend over the full TPC, it is only implemented for tracks with  $\eta > 0.6$  see figure 3.1.

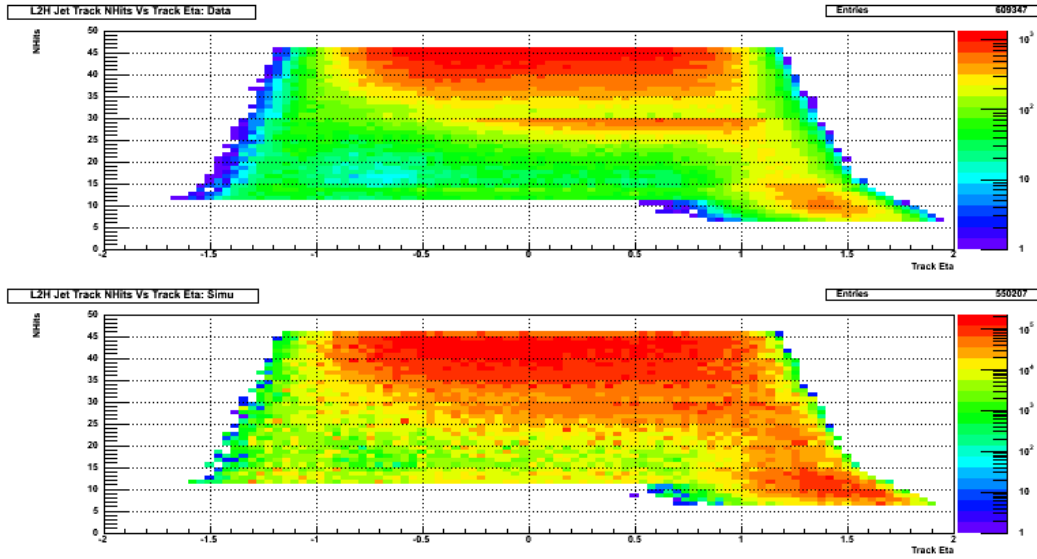


Figure 3.1: This figure shows the number of padrows hit by tracks at a given track pseudorapidity. The pseudorapidity region where 5 point tracking is implemented is visible for NHits  $\geq 12$  and  $\eta$  greater than 0.5. The top panel is data and the bottom panel is simulation.

Like with TPC tracks, there are several conditions the calorimeter towers must satisfy to be included in the jet-finding:

- Status = 1
- Tower transverse energy  $> 0.2$  GeV
- Tower ADC  $> 4$

- Tower  $\text{ADC} > 3 \times \text{RMS}$
- Hadronic subtraction scheme

The hadronic subtraction scheme is designed to prevent double counting the contributions from tracks and towers. The hadronic subtraction removes 100% of the  $p_T$  of a track from the tower that the track points to. If the track  $p_T$  is greater than the transverse energy of the tower, the tower  $E_T$  is set to zero.

## 3.2 Di-jet Selection and Cuts

Di-jets are two jets which arise from a single partonic hard scattering event. It is important to fully specify how di-jets are selected from the many jets which may be reconstructed from any given bunch crossing, as applying the cuts in different orders may lead to different individual jets being identified as the di-jet. The steps used to determine which jets in an event will be identified as a di-jet candidate are (in order):

1. Select highest ranked vertex in event (vertex must have rank  $> 0$ )
2. Require vertex to have  $|z| < 90\text{cm}$
3. Select all jets satisfying  $-0.8 \leq \eta \leq 0.8$  and  $-0.7 \leq \text{detector } \eta \leq 0.7$
4. Select two highest  $p_T$  jets
5. Require that at least one jet satisfies one of the trigger categories (see sec 4.1)

NOTE: The steps above apply when using the data and detector level simulation. When using the unbiased full Pythia sample, step 3 is omitted and the two highest  $p_T$  jets at any pseudorapidity are selected.

The two jets selected using the steps above define a di-jet candidate. There is only one di-jet candidate defined per event, even if the event contains enough individual jets to construct more two-jet pairs. The candidate di-jets constructed using the procedure above must pass a number of further conditions to be used in the analyses:

- Opening angle cut:  $\cos(\Delta\phi) \leq -0.5$
- Neutral fraction cut: One jet must have neutral fraction  $< 1.0$

- Asymmetric  $p_T$  cut: High  $p_T$  jet  $\geq 8.0$  GeV/c and low  $p_T$  jet  $\geq 6.0$  GeV/c
- $p_T$  balance cut

NOTE: Because there is no detector information, the neutral fraction cut and the  $p_T$  balance cut are not used when selecting di-jets at the particle or parton levels from the full unbiased Pythia sample.

The two jets arising from a partonic hard-scattering event should be roughly back-to-back in azimuth ( $\phi$ ). Jets which are close to each other in azimuth likely do not represent the  $2 \rightarrow 2$  scattering. To remove these events from the analysis, an opening angle cut was placed on the two jets included in the di-jet such that the azimuthal angle between them must be 120 degrees or greater. See figure 3.2.

Previous inclusive jet analyses have placed neutral energy conditions to remove jets comprised primarily of background energy. The cut was usually placed such that jets with greater than 95% of their  $p_T$  coming from the calorimeter towers are rejected. This kind of cut is not useful when studying jets in the EEMC because the falling TPC efficiency means that jets in this region will have few tracks and therefore large neutral fractions. It is unlikely that a background jet will be coincident with a physics jet so instead of putting a neutral energy cut on the individual jets, the requirement can be loosened so that only di-jet candidates where both jets have neutral fractions of 100% are rejected. See figure 3.3.

Theoretical calculations of di-jet quantities require an asymmetric cut on the transverse momenta of the two jets. Generally, the lowest  $p_T$  jets should differ by 2 GeV/c. To be able to compare to theory, an asymmetric  $p_T$  cut is placed on the data and simulation. The high  $p_T$  jet must have a  $p_T$  greater than or equal to 8 GeV/c and the low  $p_T$  jet must have a  $p_T$  greater than or equal to 6 GeV/c.

Plotting the  $p_T$  of one jet vs the  $p_T$  of the other jet in di-jet events reveals 'tails' of events with highly unbalanced  $p_T$ s (see figure 3.4). The jets in these tails are dominated by single tracks with very high  $p_T$ s. Unlike the high  $p_T$  tracks caused by the misalignment of the TPC discussed in the calibration section 1.5, these tracks are uniformly distributed throughout the TPC and are likely a result of finite resolution in the measurement of track curvature. These high  $p_T$  tracks are not a result of physics processes, so the jets which contain them should be removed from the analysis. This is done with a two step

cut. First, jets which contain a track with  $p_T \geq 30$  GeV are flagged and the event is dropped unless the ratio of the  $p_T$ s of two jets in the di-jet is between  $2/3$  and  $3/2$ . Figure 3.4 shows the effect of this  $p_T$  balance cut.

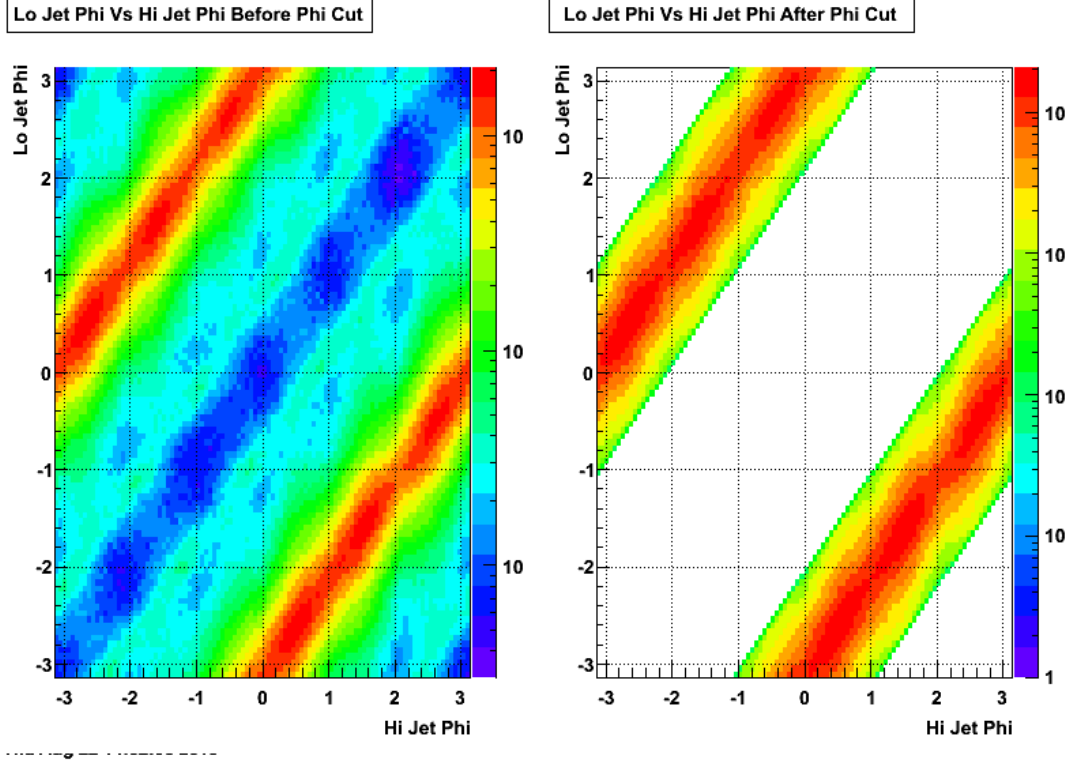


Figure 3.2: The left panel shows the jet- $\phi$  vs jet- $\phi$  correlation for di-jet candidates before the azimuthal opening angle cut was applied. The right panel shows the same correlation after the cut is applied.

### 3.3 Di-jet Mass Definition

The (squared) di-jet invariant mass is found by taking the square of the sum of the 4-momenta of the two jets which make up the di-jet:

$$M^2 = (\mathcal{P}_3 + \mathcal{P}_4)^2 \quad (3.1)$$

This can be expressed in terms of familiar jet quantities (see appendix PUT REF HERE) as:

$$M = \sqrt{m_3^2 + m_4^2 + 2\sqrt{m_3^2 + p_{T3}^2}\sqrt{m_4^2 + p_{T4}^2}\cosh(\Delta y) - 2p_{T3}p_{T4}\cos(\Delta\phi)} \quad (3.2)$$

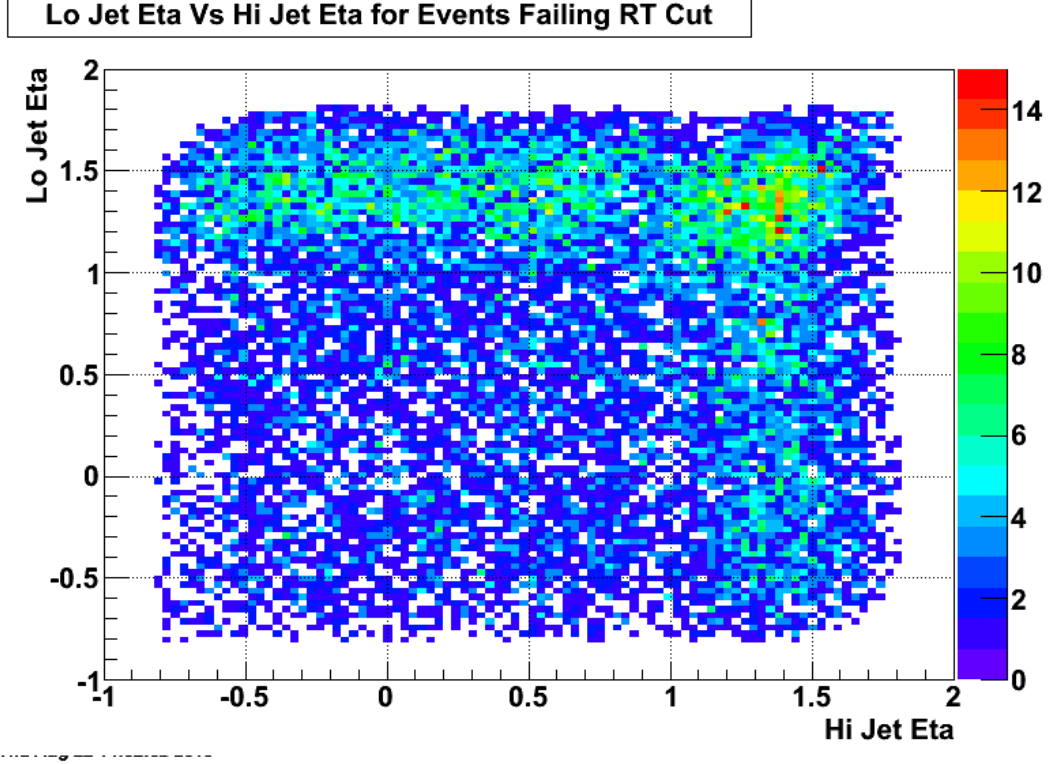


Figure 3.3: This figure shows the pseudorapidity of the low  $p_T$  jet Vs. that of the high  $p_T$  jet for di-jet candidates which fail the neutral fraction cut.

Equation 3.2 will be referred to as the full mass equation and, unless otherwise specified, is the expression used to calculate the di-jet mass from data and simulation. The theoretical calculations for the cross section and asymmetry on the other hand are performed assuming massless partons and neglecting the individual mass terms (i.e. only the  $2\mathcal{P}_3 \cdot \mathcal{P}_4$  term is considered). Neglecting the mass terms in equation 3.2 leads to the following simplified expression:

$$M = \sqrt{2p_{T3}p_{T4} [\cosh(\Delta y) - \cos(\Delta\phi)]} \quad (3.3)$$

Equation 3.3 will be referred to as the no mass equation and will be used in corrections back to the parton level.

### 3.4 Change in Di-jet Selection

The selection criteria used to identify di-jet candidates listed above were not the original criteria used during most of the course of the analysis. The third step used to be: Select all

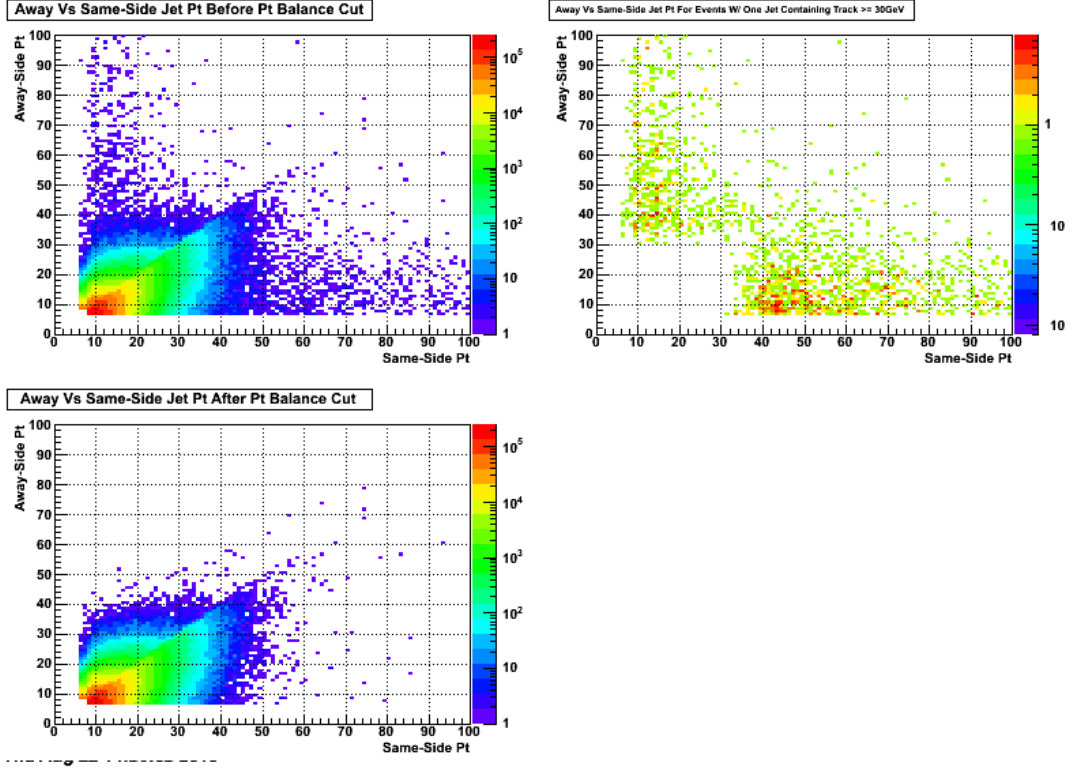


Figure 3.4: This figure shows the effect of the  $p_T$  balance cut. The top left panel shows the away vs same-side jet  $p_T$  for di-jet candidates before the balance cut is applied. The top right panel shows the  $p_T$ s for di-jet candidates in which at least one jet has a track with  $p_T \geq 30$  GeV. The bottom left panel shows the  $p_T$ s for events which pass the balance cut.

jets satisfying  $-0.8 \leq \eta \leq 1.8$  and  $-0.7 \leq \text{detector } \eta \leq 1.7$ . The condition restricting the jets to mid-rapidity was put in after the di-jet candidate had been identified. This meant that a 3-jet event in which the highest  $p_T$  jet was in the endcap and two lower  $p_T$  jets were in the barrel would be rejected as the jet in the endcap and one jet from the barrel would be chosen as the di-jet candidate. This event would then fail the further condition that both jets be at mid-rapidity. However, with the selection criteria listed in the above section, this event would be accepted (assuming the jets in the barrel are back-to-back), because only jets within the barrel are considered.

The original criteria, looking over an extended pseudorapidity range for the highest  $p_T$  jets, were chosen with the future endcap di-jet analysis in mind. With the original selection scheme, you are guaranteed to have unique barrel-barrel and barrel-endcap events. However, after discussion with the GPC and the DSSV group, it was decided that for the mid-rapidity analysis, only jets within the final acceptance should be considered

as potential di-jet candidates.

As the change to the new selection scheme came during GPC review, several checks were made to ensure that the impact on final results would be small. Overall, there are roughly 1% more di-jet events with the new criteria and changes to the di-jet mass are no greater than 2% at low mass. The di-jet  $\Delta\phi$  distribution becomes a little bit wider. The changes to the cross section and asymmetry are quite small as well.

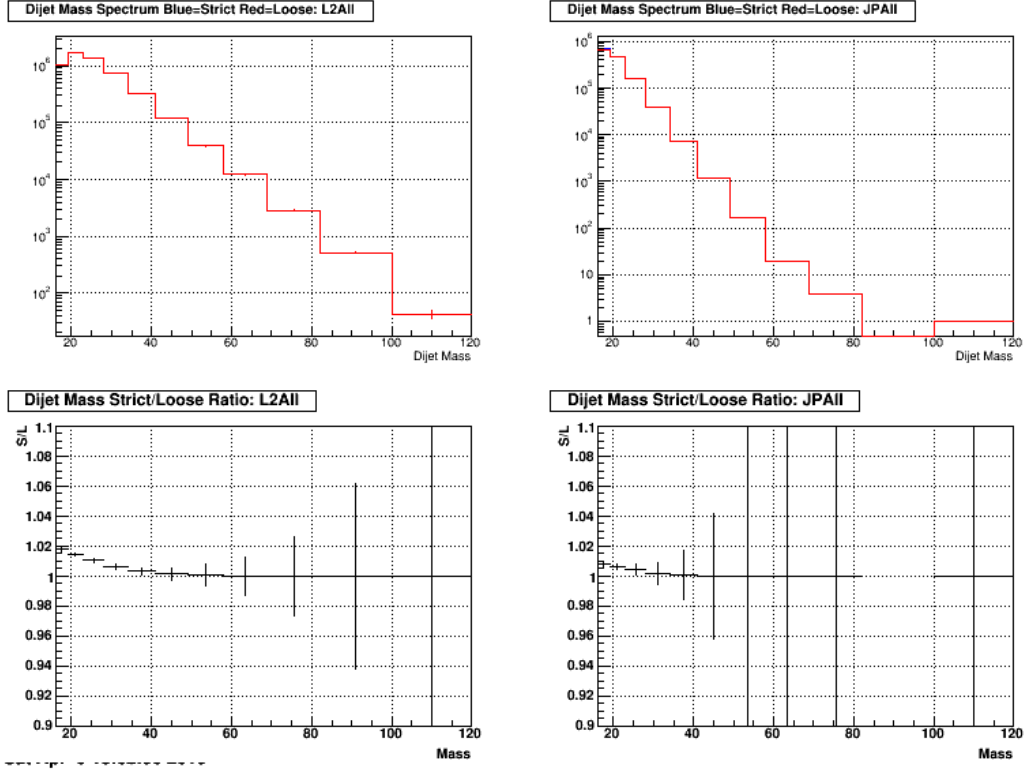


Figure 3.5: Comparison of di-jet mass using the original selection criteria (red) and new selection criteria (blue). Left column is L2JetHigh trigger and right column is JP1 trigger.

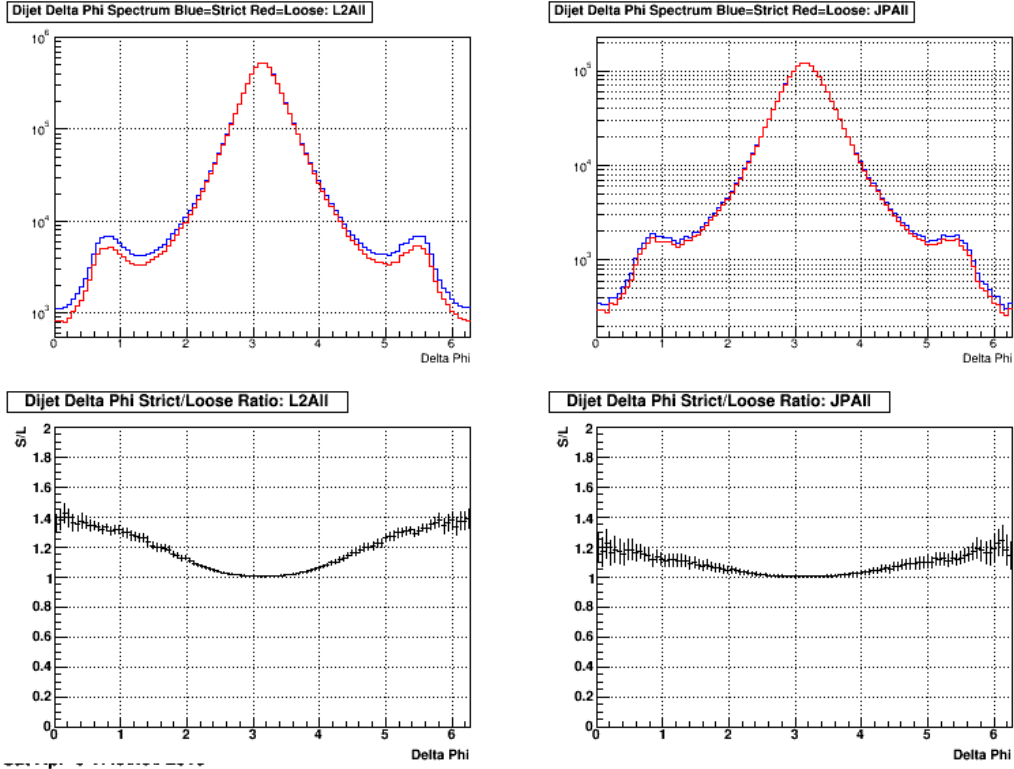


Figure 3.6: Comparison of di-jet  $\Delta\phi$  using the original selection criteria (red) and new selection criteria (blue). Left column is L2JetHigh trigger and right column is JP1 trigger.



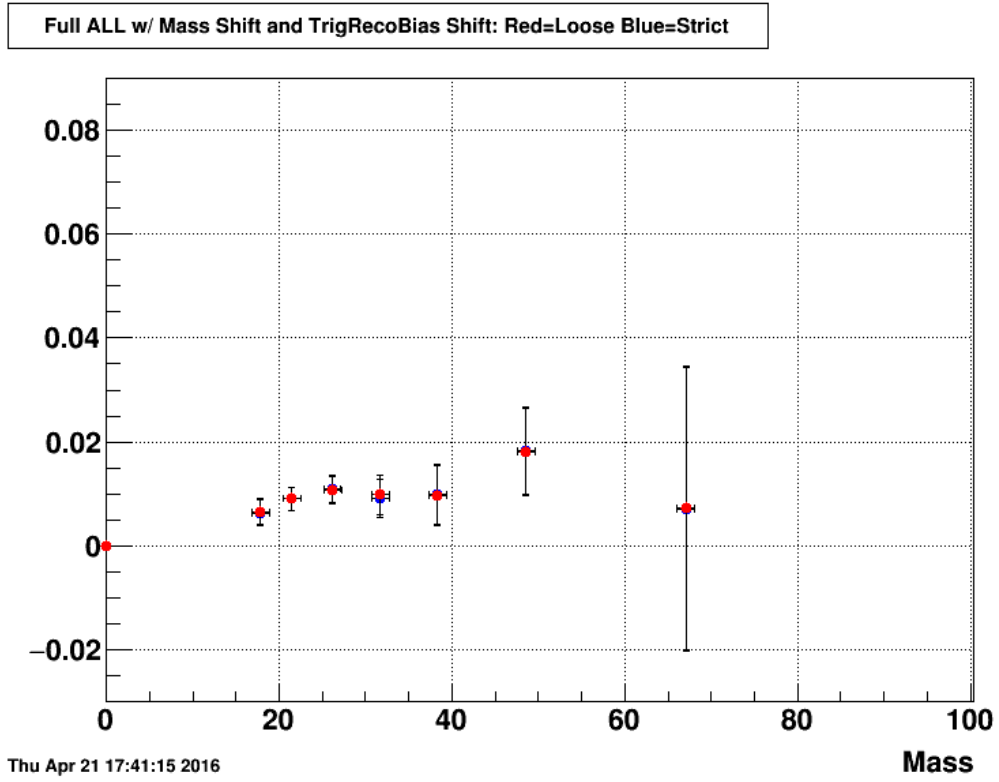


Figure 3.7: Comparison of di-jet  $A_{LL}$  (combined topology) using the original selection criteria (red) and new selection criteria (blue). All corrections have been applied.

# Chapter 4

## Di-jet Cross Section

The di-jet cross section is presented as a function of invariant mass for the L2JetHigh trigger configuration (described in section 4.1). The comparison between the cross section and theory serves to show that detector and jet reconstruction effects are well understood.

The cross section in a particular invariant mass bin is given by the following formula:

$$\frac{d^3\sigma}{dM d\eta_1 d\eta_2} = \frac{1}{\Delta M \Delta \eta_1 \Delta \eta_2} \frac{\langle \mathcal{P} \rangle}{\mathcal{L}} J \quad (4.1)$$

The  $\Delta M$  and  $\Delta \eta$  terms give the phase space volume with  $\Delta M$  being the width of a given invariant mass bin and  $\Delta \eta$  being the allowed jet pseudorapidity range which is 1.6. The  $\langle \mathcal{P} \rangle$  term is the average prescale for the JP1 trigger as discussed in section 1.4. This term is set to unity for the L2 trigger sample as the L0 L2JetHigh trigger was take-all throughout the run. The  $\mathcal{L}$  term is the integrated luminosity of the data sample as discussed in section 1.3. Finally,  $J$  is the corrected yield in a given mass bin.  $J$  is the raw yield corrected for efficiency, acceptance effects, and bin smearing due to finite detector resolution. In this analysis, all the corrections needed to take the raw yield to the corrected yield  $J$  are handled in the unfolding procedure which is described in section 4.2.

### 4.1 Trigger Selection

Section 1.2 described how events were selected for this analysis based on energy deposited in the calorimeters. This section will describe how the di-jet events are sorted into different trigger categories.

Bin	Mass Range (GeV)	$\frac{1}{\Delta M \Delta \eta}$ ( $\Delta \eta = 1.6$ )
1**	13-16	0.208
2*	16-19	0.208
3	19-23	0.156
4	23-28	0.125
5	28-34	0.104
6	34-41	0.089
7	41-49	0.078
8	49-58	0.069
9	58-69	0.057
10	69-82	0.048
11	82-100	0.035
12*	100-120	0.031

Table 4.1: This table shows the mass ranges and phase space factors for the cross section result. The starred bins are used when unfolding but are not shown in the final result. The double starred bin is only used when calculating the systematic on the unfolding due to bin choice.

A di-jet event is categorized based on the trigger conditions its individual jets satisfy. An individual jet can fall into one of three categories: L2JetHigh, JP1Lo, and JP1Hi. If a jet does not fall into one of these categories, it is considered untriggered. The conditions a jet must satisfy to be placed in one of the categories are listed below:

**L2JetHigh:** The jet must be geometrically matched to a jet patch which fired the L0 L2JetHigh trigger. The event must satisfy the mono-jet, di-jet, or random Level-2 jet trigger. The jet must also have  $p_T \geq 8.4$  GeV.

**JP1Lo:** The jet must be geometrically matched to a jet patch which fired the L0 JP1 trigger. The jet must also have  $p_T < 8.4$  GeV.

**JP1Hi:** The jet must be geometrically matched to a jet patch which fired the L0 JP1 trigger. The jet must also have  $p_T \geq 8.4$  GeV. Finally the L0 L2JetHigh trigger cannot have fired.

The geometrical matching condition is that the thrust axis of a jet must be within 0.6 in

both  $\eta$  and  $\phi$  of the center of the jet patch (or the  $E_T$  weighted centroid of two adjacent jet patches for the AJP component of L2JetHigh). The  $p_T$  conditions on the jets are designed to make sure no jet can be placed into more than one category.

There are eight unique, valid combinations of trigger categories: L2JetHigh-L2JetHigh, L2JetHigh-JP1Lo, L2JetHigh-Untriggered, JP1Lo-JP1Lo, JP1Lo-JP1Hi, JP1Lo-Untriggered, JP1Hi-JP1Hi, and JP1Hi-Untriggered. Because the cross section calculation requires the average prescale factor for the event sample, it was decided to split the sample into events which were take-all and those which were prescaled. The L0 L2JetHigh trigger was take-all throughout the run so all combinations which contain a jet in the L2JetHigh category are grouped together and are referred to as the L2JetHigh sample. The remaining combinations are also grouped together and referred to as the JP1 sample. The final cross section is computed from the L2JetHigh sample only and the JP1 sample is used as a consistency check. (The JP1 sample is used in the asymmetry result, section 5).

## 4.2 Unfolding

In order to make comparisons to theory, or to other experimental data, the di-jet cross section should be plotted as a function of the physical invariant mass. What is measured however, is an invariant mass which has been distorted by various detector effects such as finite acceptance and energy resolution. The true cross section is obtained via a process known as ‘unfolding’ in which detailed knowledge of the detector is used to correct the raw distribution back to the true distribution. This process can roughly be thought of in terms of a set of linear equations:

$$\mathbf{Ax} = \mathbf{b} \tag{4.2}$$

If the invariant mass is divided into  $\mathcal{N}$  bins, then  $\mathbf{x}$  and  $\mathbf{b}$  are vectors with dimension  $\mathcal{N}$  where the  $i^{\text{th}}$  component of the vector contains the particle and detector level cross sections respectively in the  $i^{\text{th}}$  mass bin. The  $\mathbf{A}$  term is a  $\mathcal{N} \times \mathcal{N}$  response matrix relating the particle (or parton) level to the detector level. The  $\mathbf{b}$  term is what is measured and  $\mathbf{A}$  is constructed from simulation so 4.2 must be solved for  $\mathbf{x}$ . These types of inversion problems can be ill-determined and unstable with respect to small variations in  $\mathbf{A}$  or  $\mathbf{b}$ . There are a number of techniques which can be used to unfold measured distributions in a stable way and several will be discussed in section 4.2.3.

### 4.2.1 Response Matrix

The response matrix encodes the response of the detector to a given physical input. In practice, the response matrix is a 2-dimensional histogram filled from simulation, see figure 4.1. The x-axis is the detector level di-jet invariant mass and the y-axis shows the corresponding particle or parton level di-jet mass, both axes use the binning shown in table 4.1.

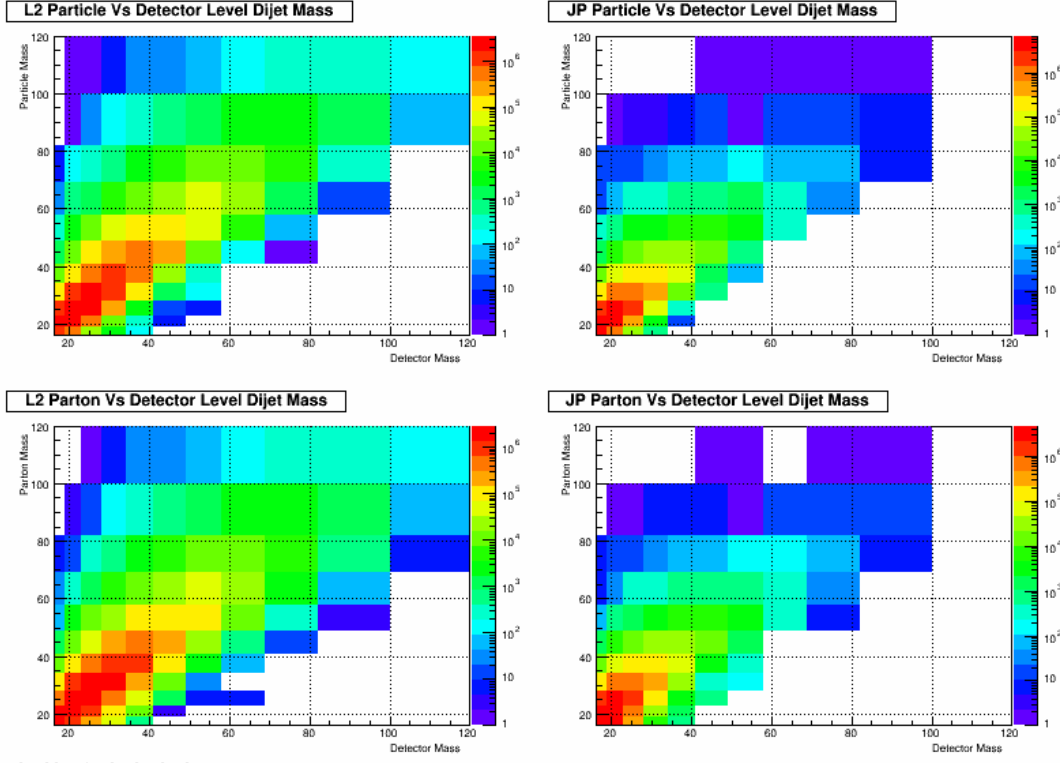


Figure 4.1: This figure shows the response matrices used in the analysis. The top panels shows the L2JetHigh (Left) and JP1 (Right) particle level vs detector level response matrices and the bottom panels show the parton level vs detector level matrices.

Because the response matrix relates the detector level to the particle or parton levels, only the simulation which passes the trigger filter can be used to populate this histogram as only those events undergo the full detector simulation (see section 2 for description of the simulation). The first step in populating the response matrix is to find a di-jet at the detector level. The same selection criteria used for the data are used to find di-jet candidates in the simulation and the same cuts are placed on the simulated candidates. Once the detector level di-jet has been found, the next step is to find the matching particle or parton level di-jet. A detector level di-jet is said to have a matching particle or parton

level di-jet if both detector level jets have matching particle or parton level jets.

For each detector level jet, the matching jet is found by looping over all particle or parton level jets and selecting the one which is closest in  $\eta - \phi$  space as given by the quantity  $\Delta R$ :

$$\Delta R = \sqrt{(\eta_{\text{Det}} - \eta_{\text{Par}})^2 + (\phi_{\text{Det}} - \phi_{\text{Par}})^2} \quad (4.3)$$

A particle or parton level jet is only considered matched to a detector level jet if  $\Delta R < 0.5$ . No kinematic cuts beyond the matching criteria are placed on the particle or parton level jets.

Once the particle or parton level di-jet which matches the detector level di-jet is found, the invariant masses of both di-jets are calculated. For each event containing matching di-jets, the response matrix bin which corresponds to the intersection between the detector level and particle or parton level mass is filled.

## 4.2.2 Fakes and Efficiency Corrections

Before the response matrix can be used in the unfolding to correct the data invariant mass spectrum back to the particle or parton level, a correction must be made to the data to remove fake events. Fake events are those which have a detector level di-jet with no corresponding particle or parton level di-jet. The correction used is just the ratio of detector level di-jets which have a matching particle / parton level jet to the total number of detector level di-jets in the simulation and is calculated for each mass bin and trigger type and for the particle and parton levels. Figure 4.2 shows the correction factors for the particle level and parton levels. The particle level correction factors are so close to unity that they are set identically to one in the analysis.

After the fake event correction is applied and the response matrix is used to unfold the data invariant mass spectrum back to the particle or parton level, an efficiency correction needs to be applied to account for particle or parton level di-jets which are missed. To find this correction factor, the full Pythia sample is needed. This sample contains all 21 million embedded events regardless of whether or not they passed the trigger filter. Note that the full Pythia sample only has particle and parton level information. Because the full Pythia sample contains all jets which arise from the generated hard scattering events,

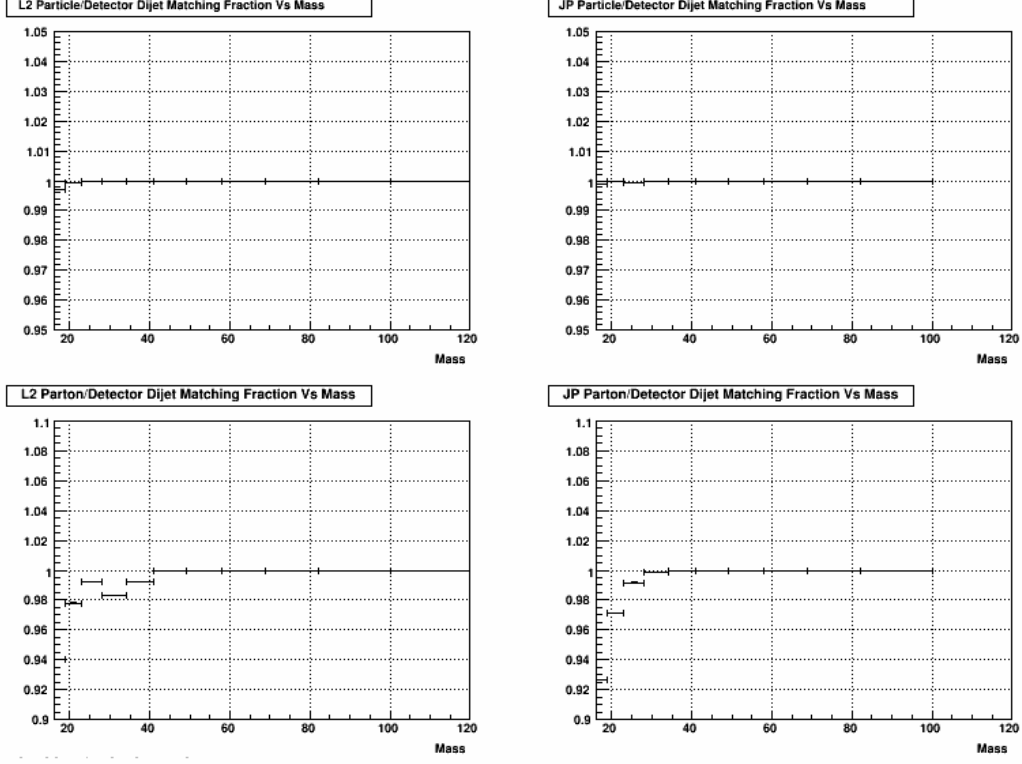


Figure 4.2: This figure shows the matching fraction, that is the ratio of the number of detector level di-jets which have matching particle or parton level di-jets to the total number of detector level di-jets. The upper row shows the matching to particle level and the bottom row shows the matching to parton level. The left column is for L2JetHigh di-jets and the right column is for JP1 di-jets

the difference between the full Pythia yield and the matched particle / parton yield gives a measure of the total efficiency. The efficiency correction should remove all remaining dependence on the specifics of the analysis such as trigger and acceptance effects.

### 4.2.3 RooUnfold Package

The unfolding and corrections described above are all handled by the RooUnfold package [17]. RooUnfold is a software package based on the ROOT [18] analysis framework which implements several unfolding methods including the bin-by-bin method, an iterative method based on Bayes' theorem [19], and one based on the Singular Value Decomposition method [20]. RooUnfold is run from a ROOT script and takes as input histograms containing the raw data, the response matrix, the total number of matched and unmatched di-jets from the simulation sample which passes the trigger filter, and the total number of di-jets from the full Pythia sample.

The Singular Value Decomposition (SVD) method was used to unfold the raw data for the cross section analysis. Further details on the SVD method can be found in the above reference and references therein. The SVD method allows for the expression of the response matrix in terms of a set of singular values and singular vectors. Once the response matrix is expressed in this way, it is relatively easy to identify and overcome problems which arise due to degeneracies in the problem and the set of linear equations in 4.2 can be solved in a stable way for  $\mathbf{x}$ .

The SVD routine requires a cutoff parameter to drop degenerate terms in the problem which may cause instability in the unfolded result. The cutoff parameter can run between zero and the number of bins in the problem. The smaller the cutoff parameter, the more closely the unfolded result will follow the original and the larger the cutoff parameter, the more the unfolded result will be dominated by unphysical statistical fluctuations. The cutoff value is the only parameter needed by the SVD method, but there are several other parameters required to specify the behavior of RooUnfold. These parameters set the number of toy simulations used in the propagation of errors and the error methods used. The parameters used in this analysis are:

- Cutoff Parameter = 6
- Number of Toys = 1000
- IncludeSystematics = 1
- ErrorMethod = 2

The SVD method is used to unfold the raw data, but the physical cross section should not depend on the unfolding method used. To confirm that the choice of unfolding method does not significantly affect the cross section, two other methods were tested: the bin-by-bin correction and the iterative Bayesian method. The bin-by-bin method is the simplest correction used. The yield in each di-jet mass bin is multiplied by a factor unique to that bin. This factor is simply the ratio of the number of particle or parton level di-jets found in the full Pythia sample to the number of particle or parton di-jets which were matched to a detector level di-jet in a given detector level mass bin. This method does not account for possible bin migration effects. A comparison of the cross section extracted using the



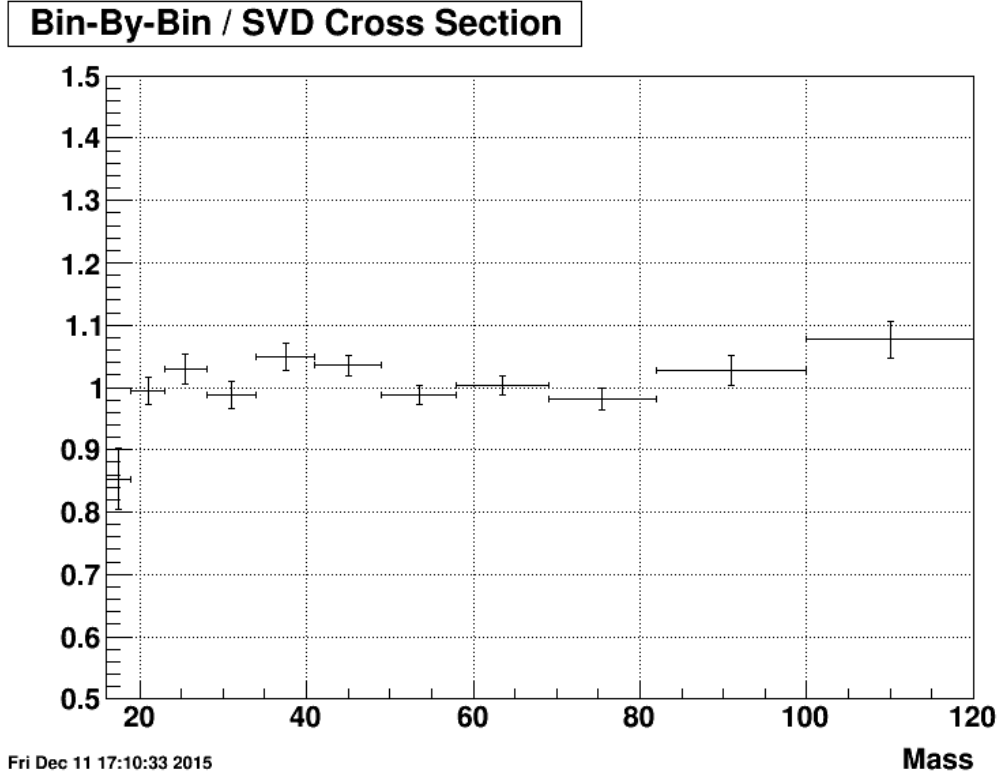


Figure 4.3: This figure shows the ratio of the di-jet cross section extracted using the Bin-by-Bin correction method to the cross section extracted using the SVD method. Note: the statistical error bars were computed using the error from the SVD method only as the Bin-by-Bin errors are not handled correctly by RooUnfold.

bin-by-bin correction to that extracted using the SVD unfolding method can be seen in figure 4.3.

The second unfolding method tested was the iterative Bayesian method which repeatedly applies Bayes' theorem to bring the raw distribution asymptotically close to the true distribution. The raw distribution serves as the input to the first application of Bayes' algorithm and that output becomes the input to the next application. This pattern continues until a set number of iterations is reached. The number of iterations used is determined by the user and is usually set at the point where subsequent iterations do not significantly change the unfolded distribution. A comparison of the cross section extracted using the iterative Bayesian method using several different numbers of iterations to the cross section extracted using the SVD method can be seen in figure 4.4.

Note that the production version of RooUnfold which was used in this analysis was version

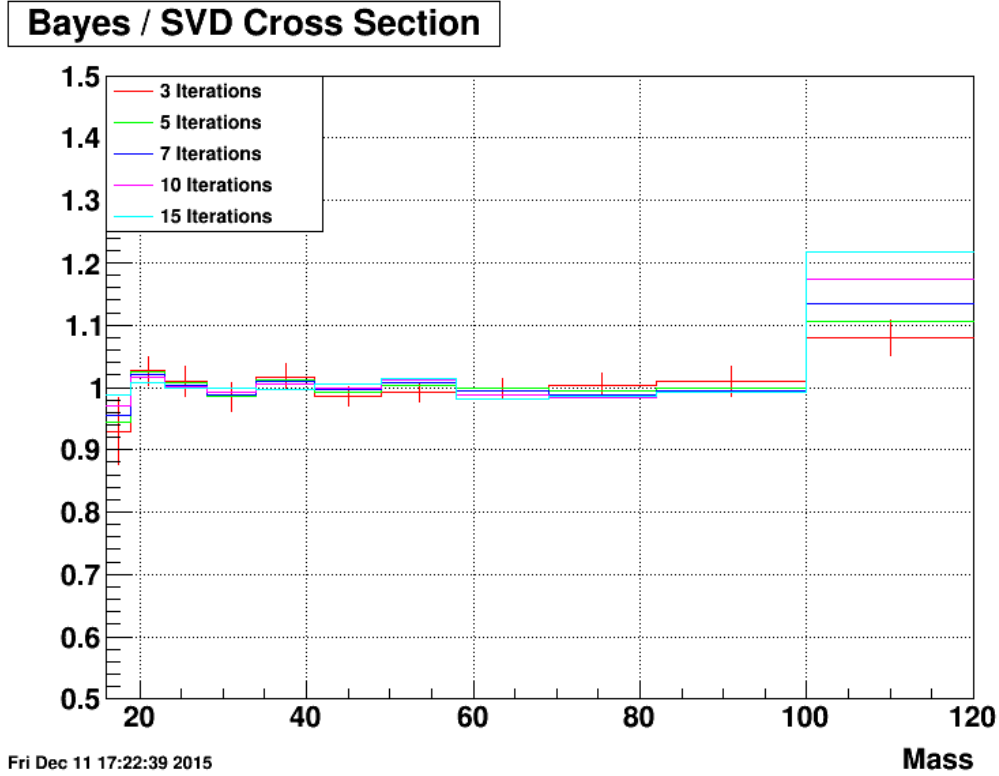


Figure 4.4: This figure shows the ratio of the di-jet cross section extracted using the iterative Bayesian method to the cross section extracted using the SVD method. Several different numbers of iterations are shown. Note: the statistical error bars were computed using the error from the SVD method only and are shown only for the iteration = 3 curve to avoid clutter.

1.1.1 . During the course of investigating the package it was found that there was an error with the way RooUnfold calculates the statistical error due to the finite statistics in the response matrix. As mentioned in section 2, the simulation was generated in 10 partonic  $p_T$  bins and each bin has an associated weight factor. In addition to these weights, an overall normalization factor can be applied which corresponds to a luminosity. This luminosity factor should not affect any physical observable, however it was found that the size of the statistical error bars on the unfolded mass spectrum depended on this overall normalization. This meant that the contribution to the statistical error from finite simulation statistics could be made arbitrarily small by changing the luminosity factor.

This issue was brought to the attention of the RooUnfold author, Tim Adye and the author of the TSVDUnfold method in ROOT (which the RooUnfold SVD method calls)

Kerstin Tackmann. With their help, the issue was identified and a fix was implemented. This fix has not been propagated through official versions of ROOT or RooUnfold yet so this analysis uses a developmental version of RooUnfold: version 346. This version can be downloaded via the RooUnfold web page [17].

## 4.3 Systematic Errors

Three main categories of systematic error on the cross section were evaluated: detector effect systematics, unfolding systematics, and systematics associated with the extraction of the integrated luminosity. The detector systematics include contributions from the TPC tracking efficiency uncertainty, the track  $p_T$  resolution, and the tower energy scale uncertainty. The uncertainty on the unfolding was evaluated by changing the SVD cutoff parameter and by adding an extra bin to the unfolding. The uncertainty on the extracted integrated luminosity is due to an apparent change in BBC response of the course of the run.

### 4.3.1 Detector Response Systematics

The systematic errors on the cross section due to uncertainty in the track finding efficiency, the track  $p_T$  scale and the BEMC tower energy are all found using the same basic procedure. Using the trigger filtered simulation sample, the detector response being evaluated is altered by a certain amount. The jet finder is then run on the altered simulation sample and di-jet events are reconstructed as described above. A response matrix is constructed and is used to unfold the raw data and obtain a cross section. The systematic error on the cross section due to the uncertainty on the particular detector response being evaluated is taken as the difference between the nominal cross section and the cross section obtained using the altered simulation.

The track finding efficiency systematic is found by randomly removing a percentage of tracks from the simulated event before the jet finder is run. The uncertainty on the track finding efficiency was taken to be 4% , so that was the maximum percentage of tracks removed. Because altering the track finding efficiency will affect the shape of the cross section, there was concern that the systematic may be underestimated in bins where the altered cross section ‘crosses over’ the nominal one. To mitigate this effect, three additional simulation samples were produced where 1, 2, and 3% of tracks were removed. The systematic in a given bin was then taken as the largest deviation between the nominal

cross section and the four cross sections constructed from the altered simulation samples. Figure 4.5 shows the size of the deviations using the four track removal percentages as well as the total systematic.

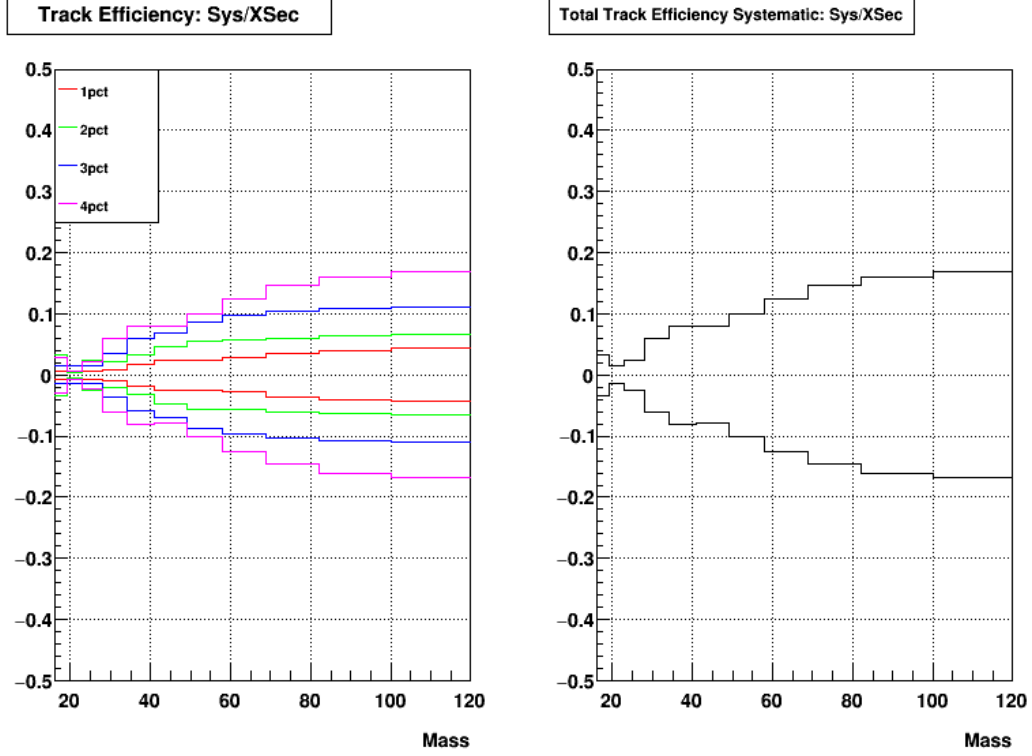


Figure 4.5: The left panel shows the size of the tracking inefficiency deviation divided by the value of the cross section for the four random track removal percentages used. The right panel shows the size of the final systematic which is the maximum deviation in each bin shown on the left.

The systematic due to the track  $p_T$  uncertainty is found by increasing and decreasing the  $p_T$  of all tracks in an event by 1%, which is a conservative estimate of the track  $p_T$  uncertainty [], before the jet finder is run. These two altered simulation samples are used to extract a plus 1% and a minus 1% cross section. The magnitudes of the differences between the nominal cross section and the two altered cross sections are found and the value of the systematic is taken as the average of these two differences. Figure 4.6 shows the size of the track  $p_T$  deviations for the  $\pm 1\%$  cases separately as well as the total systematic.

The method for determining the systematic due to the tower energy scale uncertainty

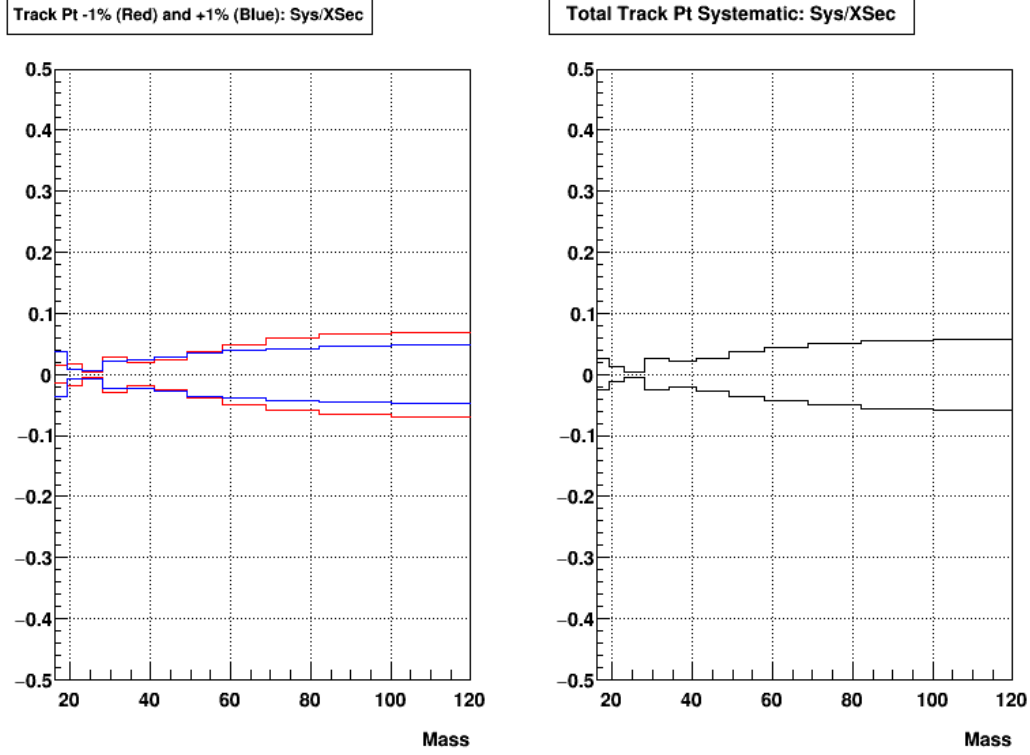


Figure 4.6: The left panel shows the size of the track  $p_T$  deviation divided by the value of the cross section for the +1% (red) and -1% (blue) cases. The right panel shows the size of the final systematic which is the average of the two deviations shown on the left.

is similar to the determination of the two systematics above. The  $E_T$  of all towers in an event are raised and lowered by a certain percent before the jet finder is run. The uncertainty on the tower energy scale is estimated to be 3.7% [] which is the maximum amount the tower  $E_T$  is raised/lowered. Values of 3.0, 2.0, and 1.0% were also used to mitigate possible spectrum ‘cross over’ effects. For each magnitude of  $E_T$  shift used, the average of the deviations from the positive and negative shifts was calculated and the largest of these averages in each bin was taken as the final systematic.

During the course of the 2012 200 GeV BEMC calibration, it was discovered that the uncertainty in the 2009 calibration was underestimated on the low side. The BEMC uncertainty should be +3.7% -4.2%. Instead of creating another simulation sample for the 4.2% deviation, a linear extrapolation was made from the 3.0% and 3.7% points was used. (In most mass bins, the progression from 1.0, 2.0, 3.0, and 3.7% is quite linear). The final  $E_T$  systematic is therefore asymmetric, with the positive value determined as described in the last paragraph and the negative value determined from the extrapolation.

Figure 4.7 shows the size of the BEMC tower energy deviations for the 1.0, 2.0, 3.0, and 3.7% values of  $E_T$  shift as well as the total systematic (this figure does not show the asymmetric systematic).

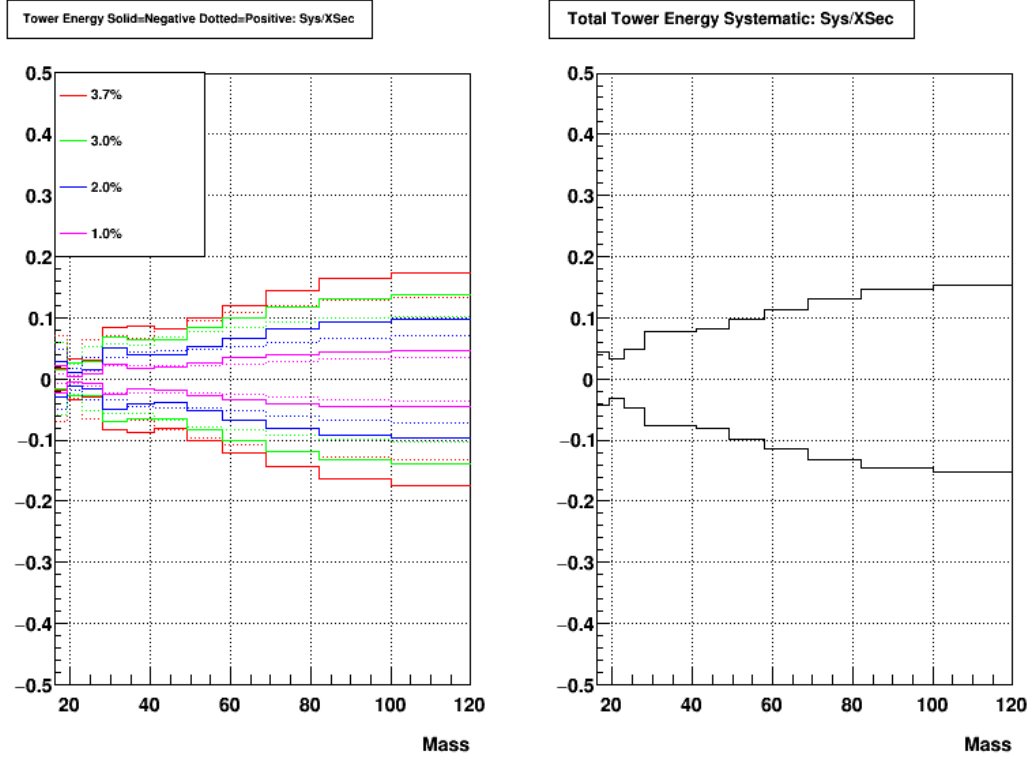


Figure 4.7: The left panel shows the sizes of the tower  $E_T$  deviations divided by the value of the cross section for all shift values. The right panel shows the size of the final tower scale systematic which is the largest of the positive/negative average deviation from a single percent change to tower  $E_T$  shown on the left.

### 4.3.2 Unfolding Systematics

This section describes the sensitivity of the unfolded cross section to the choices made in the implementation of the SVD method. The first effect looked at is the sensitivity of the cross section to the choice of cutoff parameter. As mentioned above, the nominal cross section was unfolded using a cutoff value of six. The systematic on this choice was evaluated by recalculating the cross section using cutoff values of five and seven. The value of the systematic error was taken as the average of the magnitudes of the differences between the nominal cross section and the cross section obtained using the cutoff values of five and seven. Figure 4.8 shows the size of this systematic.

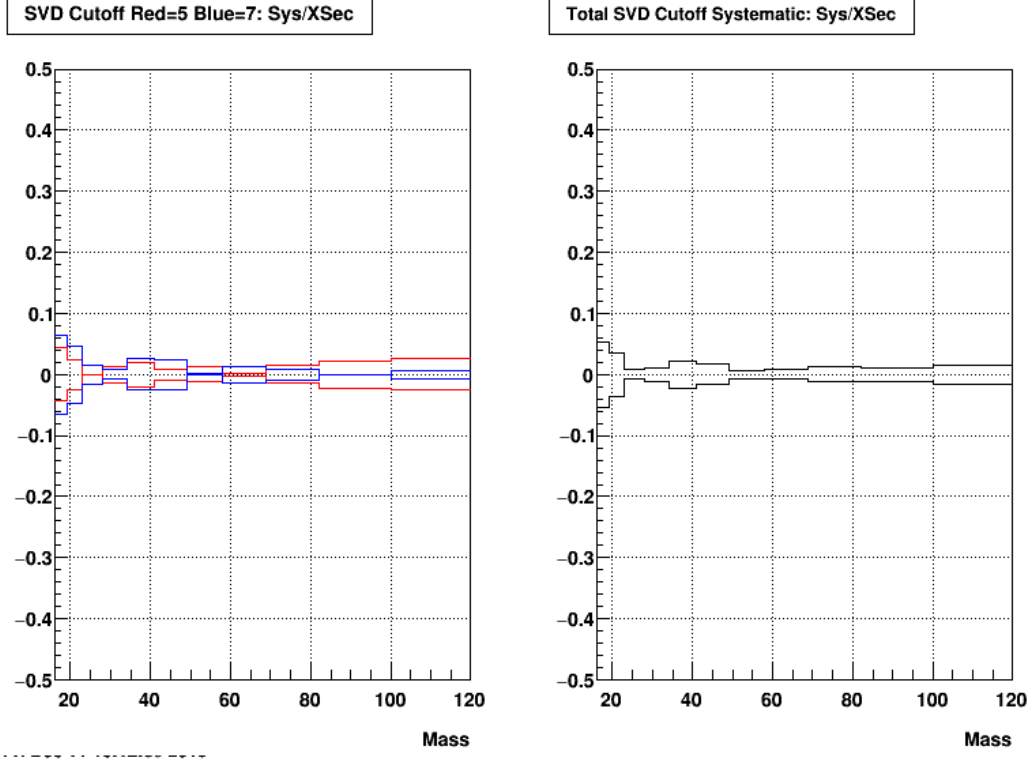


Figure 4.8: This figure shows the size of the systematic due to a change in SVD cutoff value divided by the value of the cross section for each mass bin. The red curve represents a cutoff value of 5 and the blue curve represents a cutoff value of 7.

The second systematic effect evaluated was the sensitivity of the result to the number of bins included when unfolding. For this analysis, the unfolding was done using eleven invariant mass bins but nine bins are shown in the final result. The extra two bins (one on the low mass side and one on the high mass side) are used in the unfolding to properly account for any bin migration effects into and out of the region of interest. The extra bin on the low mass side extends from 16 - 19 GeV, however, there are di-jet events which exist with invariant masses below 16 GeV. To gage the effect of the bin migration from these low mass di-jets, a bin from 13 - 16 GeV was added and the unfolding rerun. The size of the systematic was taken as the magnitude of the difference between the nominal cross section and the cross section obtained using the extra low mass bin. Figure 4.9 shows the size of this systematic.

### 4.3.3 Time Variation Systematic

The final systematic effect from data investigated in the cross section analysis was the change in cross section as a function of time throughout the run. As a consistency check,

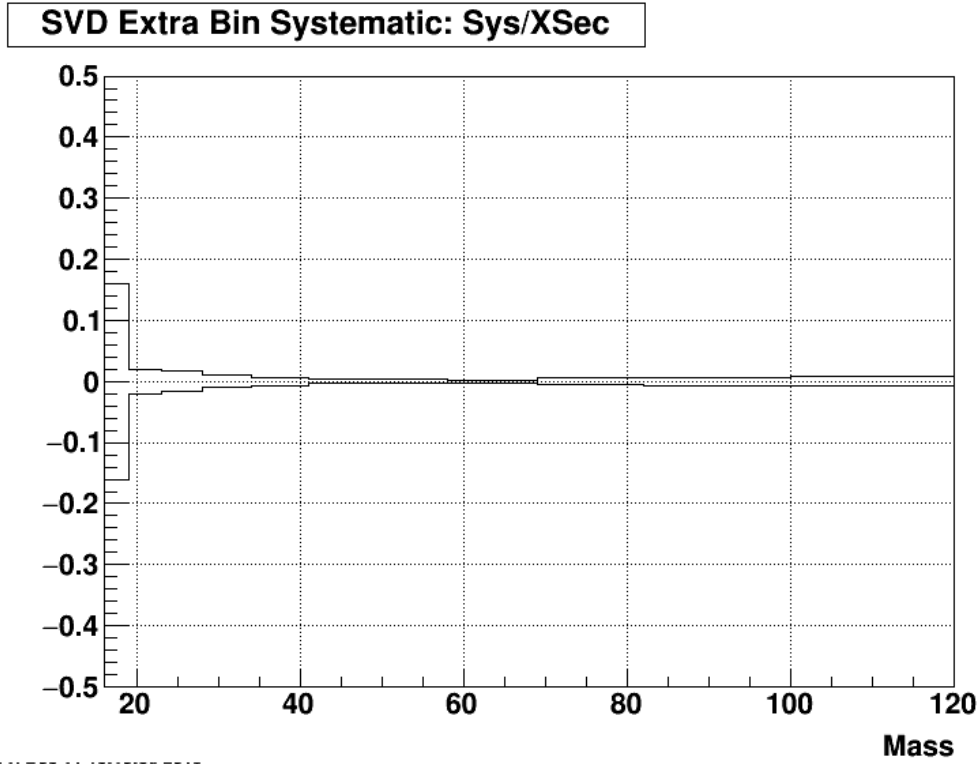


Figure 4.9: This figure shows the size of the systematic due to the addition of a low mass bin when unfolding divided by the value of the cross section for each mass bin.

the data were broken into two periods based on the magnetic field polarity and cross sections were calculated from each data set. The two cross sections were farther apart than anticipated, which led to a series of investigations to understand the change and assign an appropriate systematic.

To get a better idea of how the cross section was changing over time, the data sample was split into six periods each covering ten days (a seventh period was also included which only covered day 180). The raw di-jet yields from each period were normalized by the integrated luminosity as determined from the BBCMB-Cat2 triggers and compared to each other (see upper left-hand panel of figure 4.10). If the raw di-jet yields and integrated luminosity were constant throughout the run, all the curves should sit on top of each other. This is obviously not the case and two features are immediately obvious: first, there is a mass dependence to the ratios and second, the ratios are all offset from one-another in a roughly time ordered way (further apart in time two periods are, the more they differ).



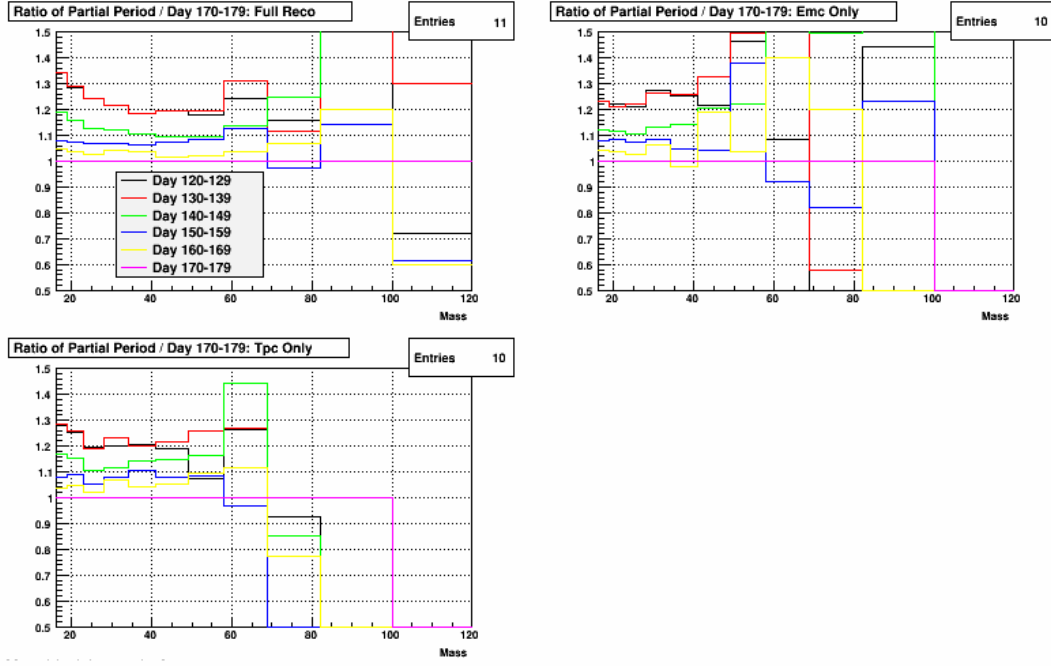


Figure 4.10: Ratios of the di-jet mass yield from a given time period to the mass yield from the end of the run. Each time period was normalized to the integrated luminosity for that period as determined from the BBCMB-Cat2 trigger. The upper left hand panel is full jet reconstruction, upper right is EMC only, and lower left is TPC only jets.

The overall differences in integrated luminosity weighted di-jet yields between different time periods was originally thought to be due to some change in the BBC response which would skew the relationship between the number of di-jets reconstructed and the number of BBCMB-Cat2 triggers fired. The observed mass dependence, however, cannot be explained by an integrated luminosity effect and must arise from some detector change which would alter the di-jet reconstruction. The two subsystems used in jet / di-jet reconstruction are the TPC and BEMC, so whatever was causing the mass change must be specific to one of those detectors. To determine which subsystem was responsible, new jet trees were created in which jets were reconstructed using only the TPC or only the BEMC (this was only done for about two-thirds of runs).

The upper right-hand and lower left-hand panels in figure 4.10 show the di-jet mass ratios for the EMC and TPC only di-jets respectively. It is seen that the mass dependence is absent for the EMC only di-jets and present for the TPC only di-jets (albeit to a lesser degree than for the full reconstruction). To better determine when the mass spectrum change occurred and correlate it with some event in the TPC, the average di-jet mass in

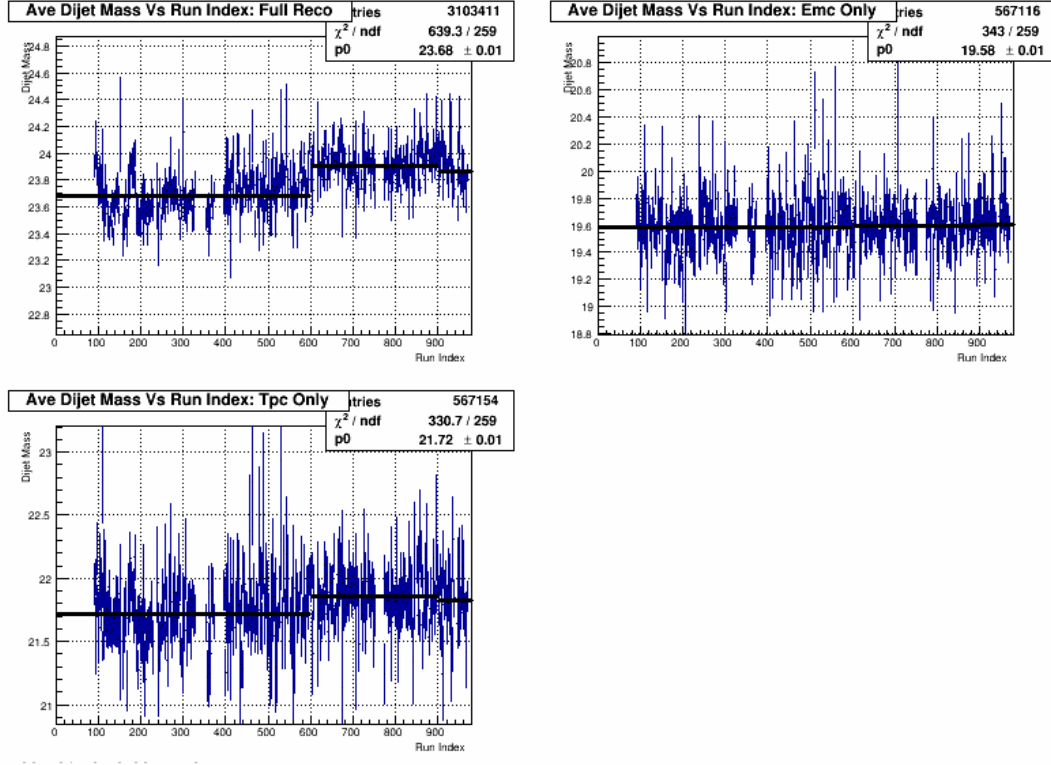


Figure 4.11: Average di-jet mass in each run as a function of run index. The upper left hand panel is full jet reconstruction, upper right is EMC only, and lower left is TPC only jets.

each run was found and plotted as a function of run index. Figure 4.11 shows the average di-jet masses for the full, EMC, and TPC only di-jets along with  $p0$  fits over three regions. A systematic change in the average di-jet mass can be seen at index 600 for the full and TPC only di-jets. Run index 600 also corresponds to the time at which RDO board 1 (which is closest to the beam line) in TPC sector 18 died and was masked out. The top panel of figure 4.12 shows the average number of TPC padrow hits used to reconstruct tracks traversing sector 18 as a function of run index and the bottom panel shows the average number of possible hit points. The loss of RDO 18-1 can clearly be seen as a drop in the average number of hits at index 600, however, there is no corresponding drop in the number of possible hit points, indicating that the loss of RDO board 18-1 was not handled properly in the offline software.

The above discussion shows that the mass dependence in the yield ratios is driven by the track contribution to the jets. In addition, the change in mass spectra corresponds to the failure and masking out of RDO board 18-1. The board failure can affect the track

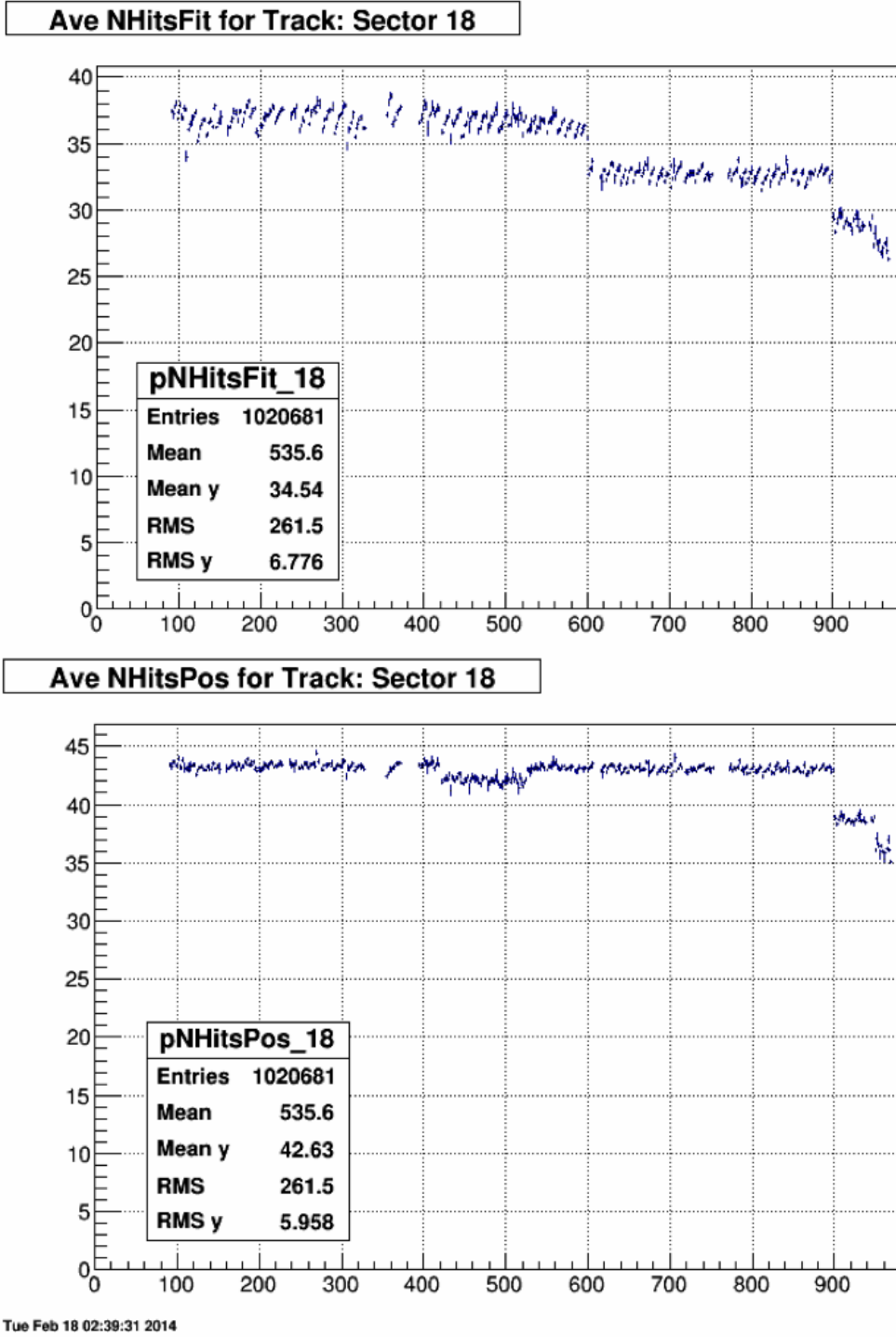


Figure 4.12: Average number of hits used in track reconstruction in TPC sector 18 in each run vs run index (top panel) and same for number of hits possible (bottom panel).

reconstruction efficiency in sector 18, so it is reasonable to check if the observed mass dependence in the time separated yields is covered by the existing tracking efficiency systematic. This was done by taking the ratio of the fully reconstructed di-jet mass spec-

tra before and after board 18-1 died, with each period normalized using the appropriate BBCMB-Cat2 trigger counts. In order to ‘subtract out’ the contribution to this ratio from the pure luminosity effect, the before and after ratio was found using the EMC only jet reconstruction. This ratio is constant as a function of mass and its value was taken as the luminosity component. The EMC only ratio was subtracted from the full ratio to give the size of the mass dependent effects without any of the luminosity differences between the two periods. Figure 4.13 shows the size of the time dependent mass shift systematic compared to the track inefficiency and  $p_T$  systematics. The efficiency systematic covers the time dependent mass systematic for all but the first and last (presented) bins, so it was decided that no additional systematic needed to be included for this effect.

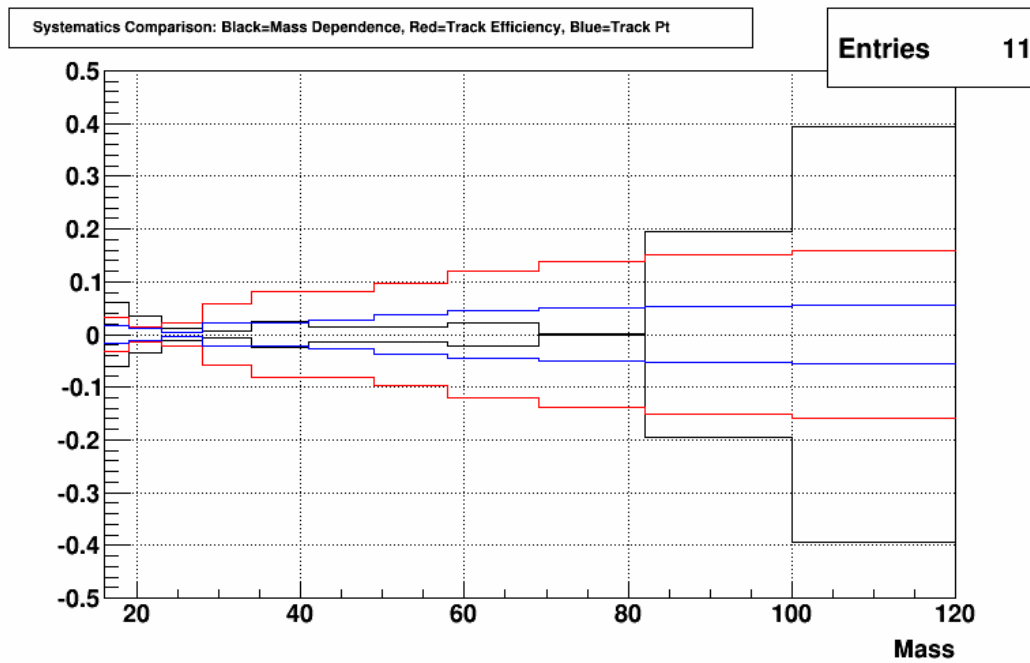


Figure 4.13: Size of the time dependent mass shift systematic (black) in relation to the existing tracking efficiency systematic (red) and track  $p_T$  systematic (blue).

The other major feature seen in figure 4.10 are the offsets between time periods. These offsets, which are constant as a function of mass, behave like they are caused by a change in the integrated luminosity determination over the course of the run. A number of studies focusing primarily on the scalar system were done to try and determine the source of these offsets. Also looked at was the BBCMB-Cat2 cross section as measured from Vernier scans.

As discussed in section 1.3, the integrated luminosity is found by counting the number of BBCMB-Cat2 triggers fired and dividing by the effective BBCMB-Cat2 cross section as determined by Vernier scans. Initially, the Vernier scans from the 200 GeV period in Run 9 were not explicitly analyzed, but Angelika Drees at CAD agreed to analyze the several scans taken over the course of the run to verify that the BBCMB-Cat2 cross section was stable throughout the run. Figure 4.14 shows the BBCMB-Cat2 cross section as extracted from 6 individual Vernier scans (the final three points were taken with the special PP2PP beam optics) and no dependence on time (low index = early in run, high index = late in run) can be seen.

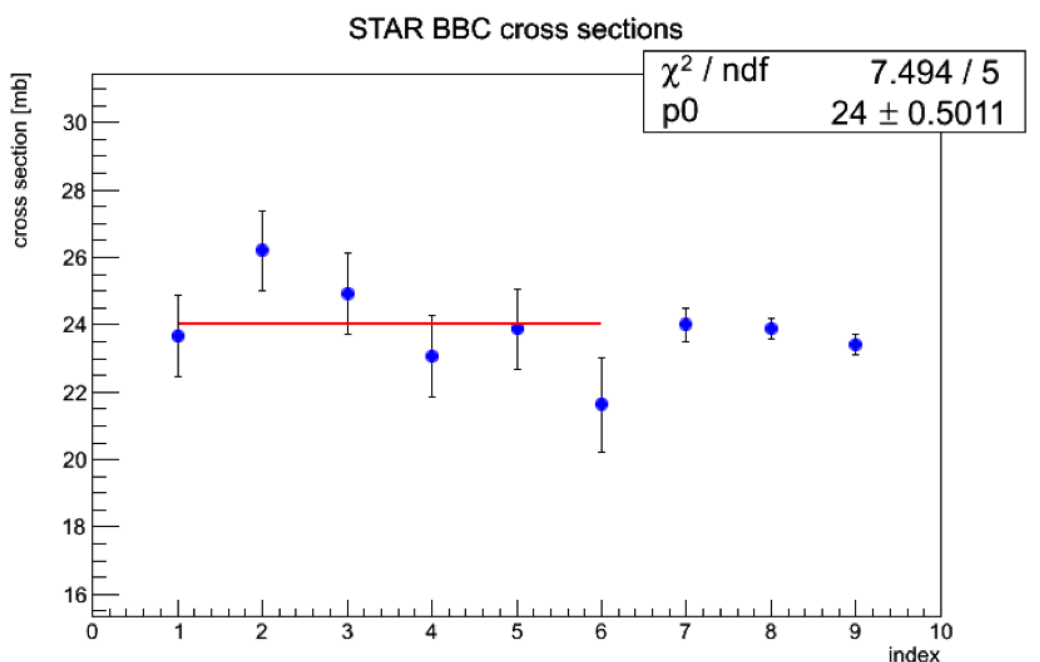


Figure 4.14: Effective BBCMB-Cat2 trigger cross section from 6 Vernier scans taken throughout the Run 9 200 GeV period.

While the effective BBCMB-Cat2 cross section seems to be stable, there still may be issues with the number of BBCMB-Cat2 triggers sent to the trigger system, or with the number of physics triggers (L2JetHigh, JP1, etc) issued. To see if there were any changes between ‘physics triggers’ and ‘luminosity triggers’ (such as BBCMB-Cat2, VPDMB, ZDCMB, etc) over the course of the run, a large number of correlations between all these quantities were investigated. An example can be seen in figure 4.15, which shows the ratio of the number of reconstructed di-jets (from the L2JetHigh trigger sample) over BBCMB-Cat2 triggers fired for each run as a function of run index. Many similar plots

were made for the VPD, and ZDCMB triggers as well as ratio involving only the number of fired L2JetHigh or JP1 triggers as opposed to reconstructed di-jets, and all show similar features. This can also be visualized as in figure 4.16, which shows for each run, the number of L2JetHigh di-jets reconstructed divided by the length of the run vs the number of BBCMB-Cat2 triggers fired vs the length of the run. In this way, information about the instantaneous luminosity dependence is also made explicit.

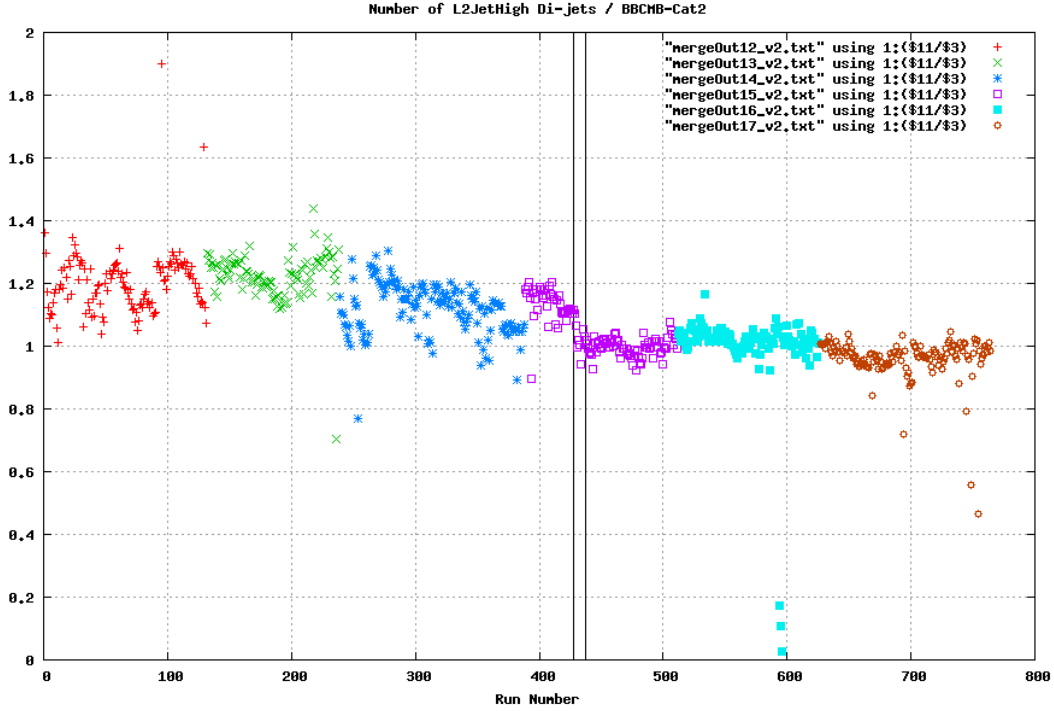


Figure 4.15: Ratio of the number of reconstructed di-jets over the number of BBCMB-Cat2 triggers fired for each run as a function of run index.

It is quite obvious from figure 4.15 that an abrupt change in behavior occurs around index 425 (left-most vertical line). In addition to the change in the average value of the ratio, the in-fill variation becomes much smaller after run index 425. Although run indices are not explicit in figure 4.16, the split between the two groupings occurs at the same time as the change seen at index 425. Index 425 corresponds to the period around days 152-154 when RHIC delivered several transverse stores and no production runs were taken. It is also very close to the time that RDO board 18-1 died in the TPC.

Despite considerable effort, no satisfactory explanation was found for the change in behavior in the di-jet to BBCMB-Cat2 ratio. Also, because of the proximity in time between

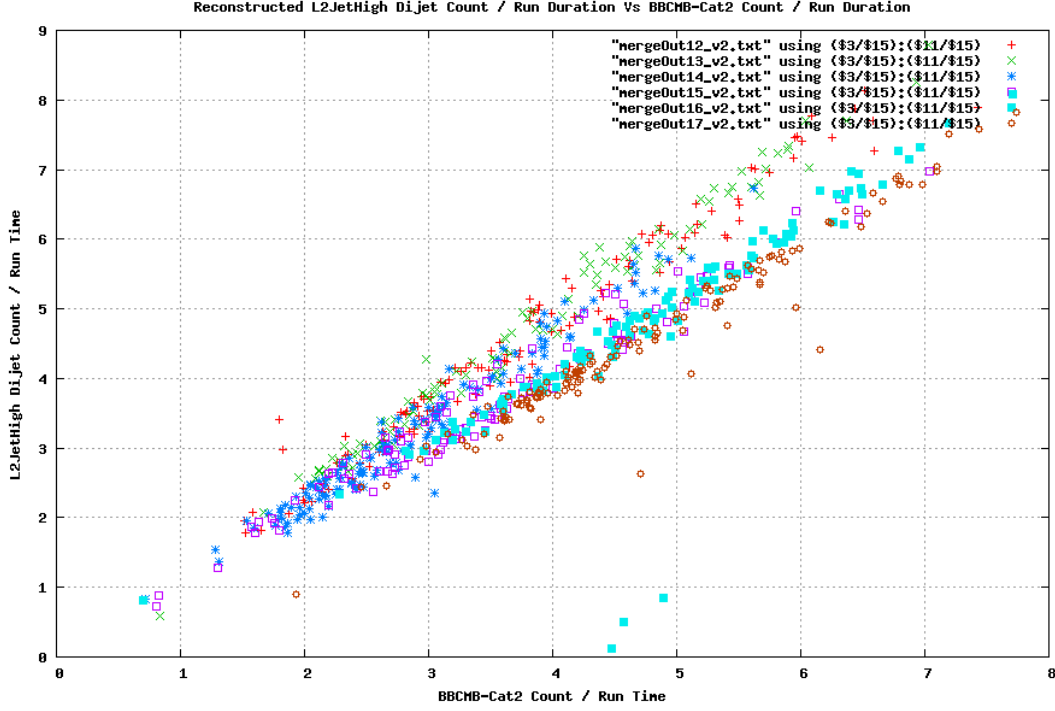


Figure 4.16: Number of reconstructed di-jets in each run divided by run length vs the number of fired BBCMB-Cat2 triggers divided by run length.

the day 152-154 split and board 18-1 dying, it is not entirely clear if board 18-1 is causing all the problems seen (mass dependence and ‘luminosity’ dependence) and the day 152-154 split is a red herring, or vice versa. It is also hard to see how all the observed features could be caused by either event.

Because of the lack of a mechanism which can explain all features, and the resulting uncertainty on which period would give the ‘correct’ cross section, it was decided that the most honest course of action would be use the entire run for the cross section analysis. The cross sections extracted from the period before and after run index 425 should bracket the cross section extracted from the entire run. Since it is believed that either the pre or post index 425 cross section is ‘correct’, it was decided to take half the difference between these cross sections as the systematic error due to our lack of knowledge about what is causing this integrated luminosity like discrepancy. The fully unfolded cross section was calculated for each period, using the embedded simulation corresponding to that time frame, and the ratio (Data-Theory)/Theory was calculated. Figure 4.17 shows the difference between these two ratios (Pre index 425 - Post index 425).

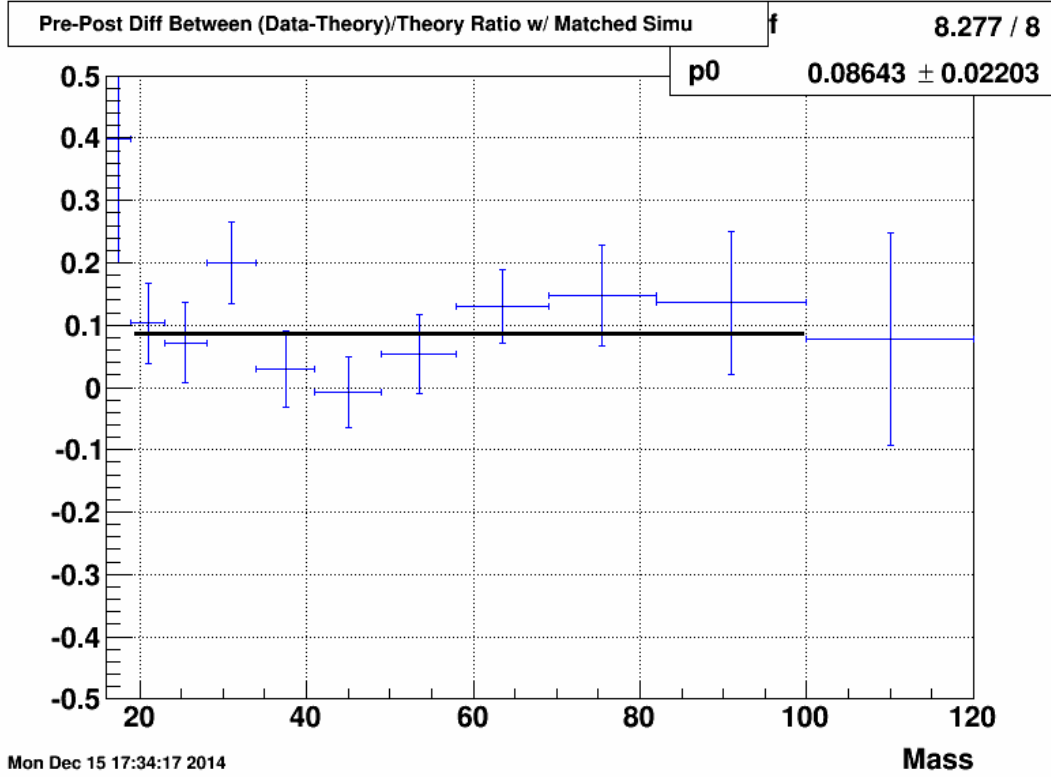


Figure 4.17: Difference between the (Data-Theory)/Theory ratios for cross sections extracted from the pre and post index 425 periods.

As can be seen in figure 4.17, the average of the differences sits around 8.6%, however, there is quite a lot of statistical scatter and some of the high mass points sit around 15%. In addition to figure 4.17, the ratio of the raw di-jet yields from before and after index 425 for the EMC only jet reconstruction was considered. This is the same ratio that was used to remove the ‘luminosity’ component when trying to determine the mass dependence caused by the loss of board 18-1 (see figure 4.13). This ratio is constant as a function of mass and has a value of 17%. So the cross section difference and EMC di-jet yield ratio give a range between 9% and 17%. To be conservative, the final determination of the difference between the two periods was taken to be 14%, yielding a final systematic due to the lack of knowledge about the time variation of the cross section of 7%. This value is added in quadrature with the error on the extracted BBC cross section from the Vernier scan (which was 5.4%) to yield a final systematic on the integrated luminosity of 8.8%.

#### 4.3.4 Total Systematic

The final overall systematic is found by adding the systematic errors from the track finding efficiency, track  $p_T$  uncertainty, BEMC tower energy scale uncertainty, SVD cutoff value



uncertainty, and SVD extra bin uncertainty together in quadrature. This final systematic error is added to and subtracted from the cross section to give the error band. Table 4.2 shows the numerical values of the individual systematics divided by the cross section for each bin. Figure 4.18 shows the size of the combined systematic error divided by the measured cross section. The effect of the asymmetric tower energy scale uncertainty is shown.

Di-jet Mass [GeV/c <sup>2</sup> ]	Track Efficiency	Track $p_T$	Hi/Lo Tower $E_T$	SVD Cutoff	SVD Bin	Hi/Lo Total
16-19	0.033	0.026	0.041 / 0.045	0.053	0.155	0.173 / 0.174
19-23	0.015	0.012	0.034 / 0.038	0.035	0.023	0.057 / 0.060
23-28	0.025	0.005	0.050 / 0.055	0.008	0.017	0.059 / 0.064
28-34	0.062	0.026	0.076 / 0.087	0.010	0.009	0.103 / 0.111
34-41	0.079	0.022	0.077 / 0.089	0.022	0.006	0.115 / 0.123
41-49	0.079	0.026	0.082 / 0.093	0.017	0.003	0.118 / 0.126
49-58	0.101	0.036	0.098 / 0.111	0.006	0.003	0.146 / 0.154
58-69	0.125	0.044	0.114 / 0.130	0.008	0.002	0.175 / 0.186
69-82	0.146	0.051	0.132 / 0.151	0.012	0.006	0.203 / 0.216
82-100	0.160	0.056	0.146 / 0.168	0.011	0.007	0.224 / 0.239
100-120	0.168	0.058	0.153 / 0.176	0.016	0.007	0.235 / 0.251

Table 4.2: Numerical values for five systematic errors on the cross section (divided by the cross section) as well as the quadrature sum for each di-jet invariant mass bin.

## 4.4 Underlying Event and Hadronization Corrections

The measured cross section is unfolded and corrected to the particle level which removes detector effects and allows the measurement to be compared to theory or to measurements from different experiments. As mentioned in section 2, the particle level of the simulation contains the effects due to underlying event and hadronization, however, the theory calculation (the calculation is next to leading order perturbative QCD implemented by Daniel deFlorian [21] using the CT10 parton distribution function set to which the extracted cross section is compared does not include these effects. In order to make a valid comparison between data and theory they must treat these effects in a consistent manner. There are two ways to ensure consistency, the first is to correct the extracted cross section to the parton level instead of the particle level as the parton level more closely reflects what is calculated in theory. The second way is to determine size of the underlying event and hadronization effects and apply that as a correction to the theory. The second method is used in this analysis.

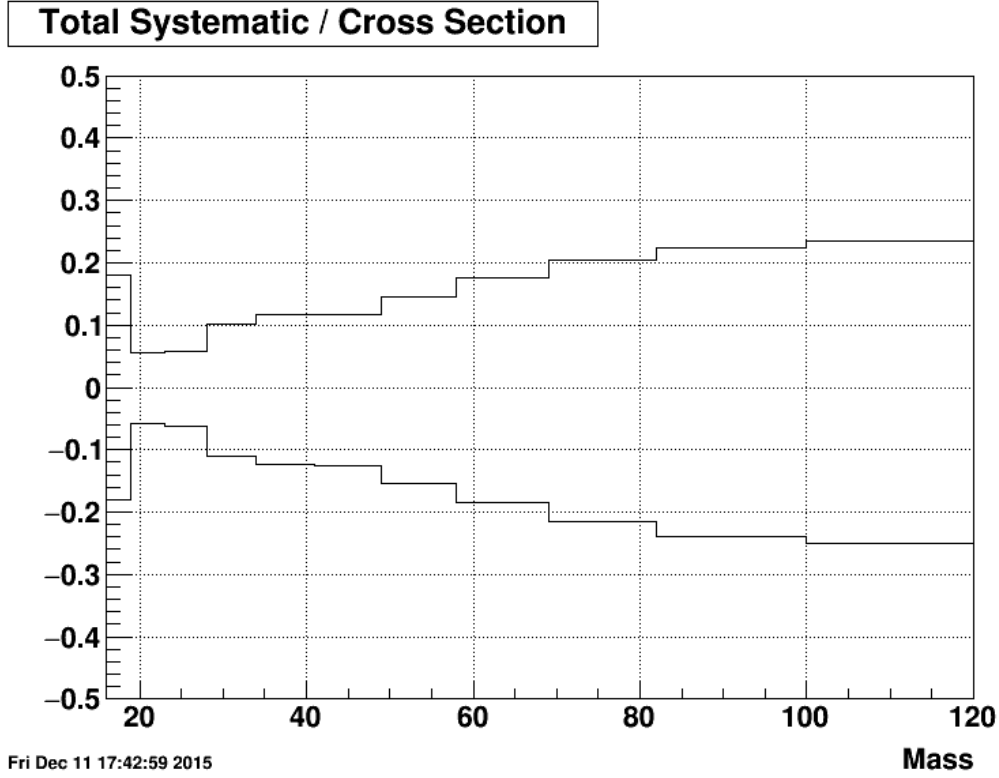


Figure 4.18: This figure shows the total systematic error divided by the cross section.

The underlying event and hadronization (UEH) correction was originally found by extracting the cross section from the raw data twice, once to particle and once to parton level. Because the particle level contains the underlying event and hadronization effects and the parton level does not, the difference between particle level and parton level cross sections should be the underlying event and hadronization contribution. This difference was then added to the theoretical values.

The draw-back of the method outlined above is that the simulation needs to be processed all the way to detector level, which is a very time consuming process. This makes it infeasible to check the sensitivity of the correction to different Pythia tune settings or PDF sets or to propagate errors from the PDF set through to an error on the correction. Because of this, it was decided to find the UEH correction by taking the ratio of particle over parton level di-jet yields (as a function of di-jet mass) from stand-alone Pythia samples which can be generated much faster than full STAR simulation samples. The theoretical cross section values are then multiplied by this ratio so a comparison to the data can be made.

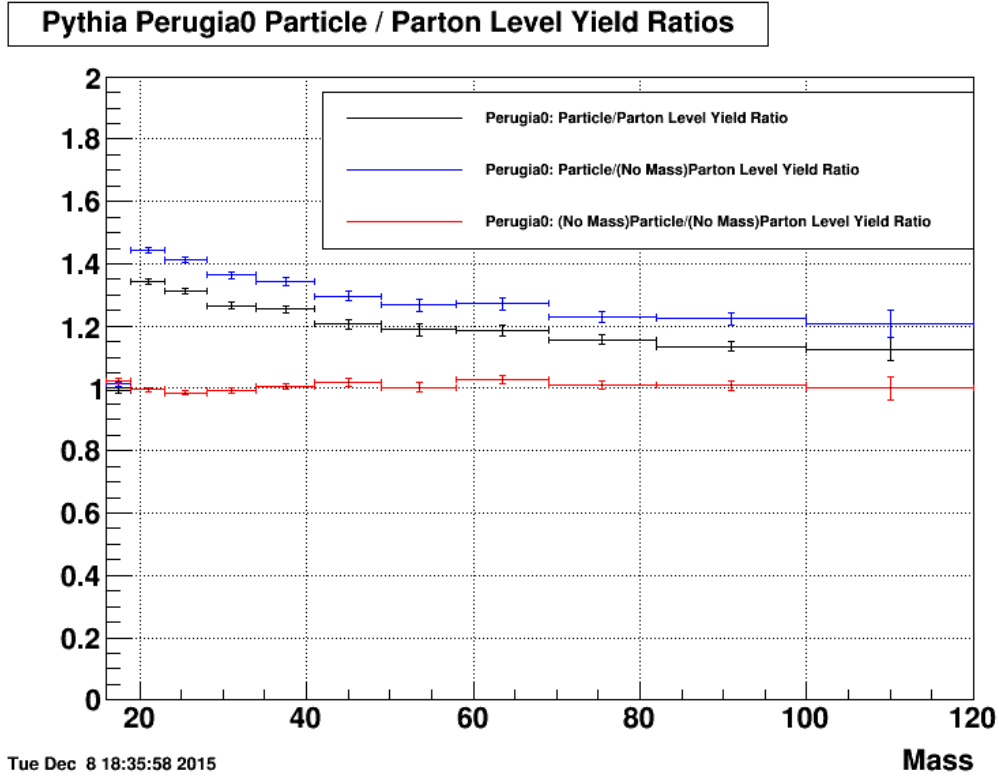


Figure 4.19: The ratios of particle over parton level di-jet yields as a function of mass. The black curve shows the ratio when both particle and parton level di-jet masses were found using the full formula. The blue curve is the UEH correction where the parton level di-jet mass was found using the no mass formula. The red curve shows the ratio when both levels used the no mass formula.

It was also decided that the denominator of the UEH correction ratio (the parton level di-jet yield) should be found using the no mass formula. The UEH correction ultimately aims to put the data and theoretical values on the same footing and since the theoretical calculations ignore the jet mass terms when calculating the di-jet mass, the no mass formula is the appropriate one to use at parton level. Figure 4.19 shows the particle over parton level yield ratio for the cases when both di-jet masses were found using the full mass formula, for the case when both di-jet masses were found using the no mass formula (this is the correction which is appropriate when extracting the cross section when using the no mass formula at data, particle, and parton levels), and for the case when the particle level mass used the full mass and the parton level used the no mass formula, which is the standard UEH correction. Figure 4.20 compares the old cross section subtraction UEH correction method with the new ratio based method. The comparison

is made by applying each correction to a CTEQ6m based NLO theory calculation and comparing the corrected theory to the data cross section. There is some tension between the two methods at low mass, but good agreement for masses above 40 GeV.

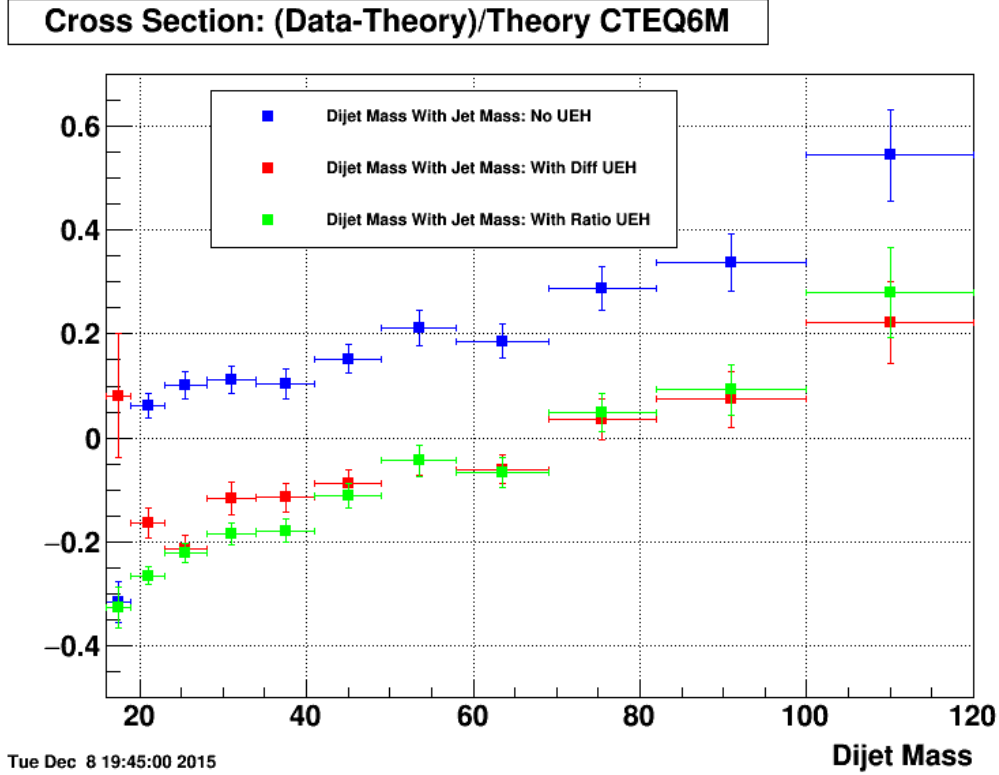


Figure 4.20: Ratio of (Data-Theory)/Theory cross sections where the theory is UEH corrected NLO using the CTEQ6M PDF set. The red curve is the ratio using the old subtraction based UEH correction and the green curve uses the new ratio based correction. In both cases, the parton level di-jet mass was calculated using the no mass formula. The blue curve shows the ratio with no UEH correction applied for comparison.

It should also be noted that the UEH correction appears to be quite sensitive, especially at low mass, to the Pythia parameter tunes and PDF sets used. The ratio of particle over no mass parton level di-jet yields was looked at for the standard Perugia0 tune, CDF TuneA, the Perugia0 tune settings using the CTEQ6 PDF set as opposed to the default CTEQ5L set, and finally, the Perugia0 tune with the intrinsic  $k_T$  changed from the default of 2.0 to a more realistic 0.5. The default Perugia0 and CDF TuneA simulations were run using Pythia version 6.4.26 while the other two simulations were run using version 6.4.28. It was found that the UEH correction is quite insensitive to this version change. Figure 4.21 shows how the corrections compare.

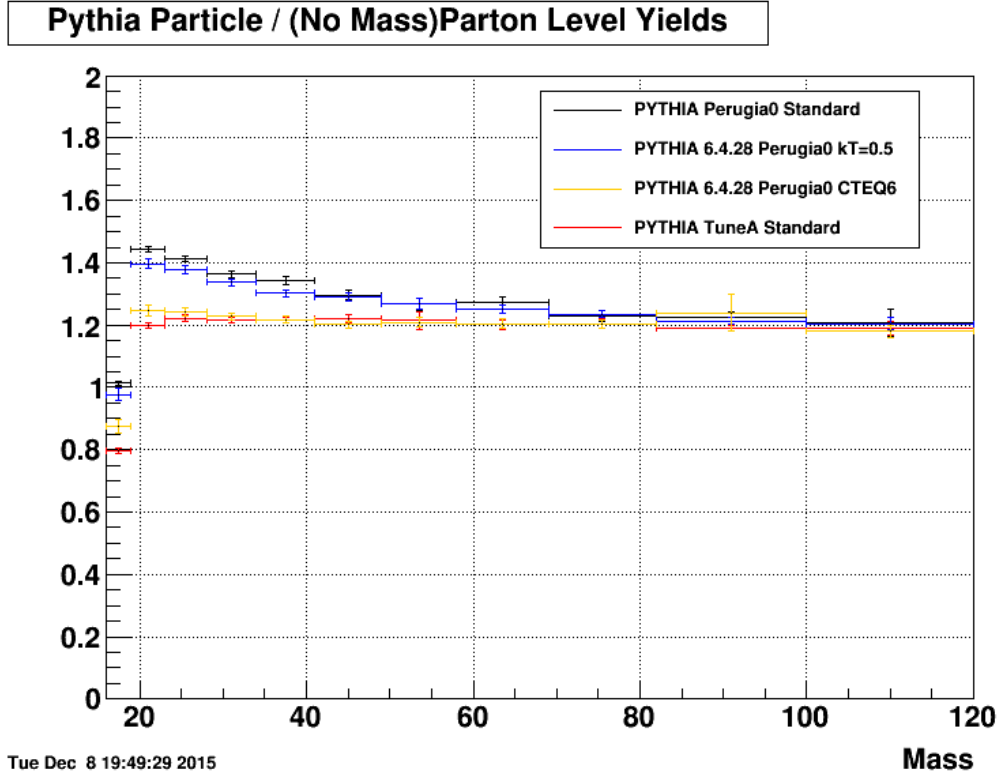


Figure 4.21: Comparison of the UEH correction factors obtained from different tune, PDF, and intrinsic  $k_T$  settings.

#### 4.4.1 UEH Correction Systematic

The systematic error on the UEH correction is taken from two sources: varying the factorization and renormalization scales by 0.5 and 2.0, and from the error sets associated with the CTEQ6 PDF set. The factorization and renormalization systematic was found by creating three small Pythia samples, one with the nominal Pythia Perugia0 settings and one each with the scale changed by a factor of 0.5 and 2.0. The scale is changed via the PARP(34) switch in Pythia which multiplies the  $2 \rightarrow 2$   $Q^2$  value by a given amount. The deviation of the 0.5 and 2.0 curves from the nominal curve is taken as the systematic. Figure 4.22 shows the particle over parton level di-jet yield ratio for the three sets.

The contribution to the UEH systematic which comes from the PDF uncertainty was found using the ‘CTEQ’ method for propagating errors from PDF sets to observables:

$$\delta X = \frac{1}{2} \sqrt{\sum_i (X[S_i^+] - X[S_i^-])^2} \quad (4.4)$$

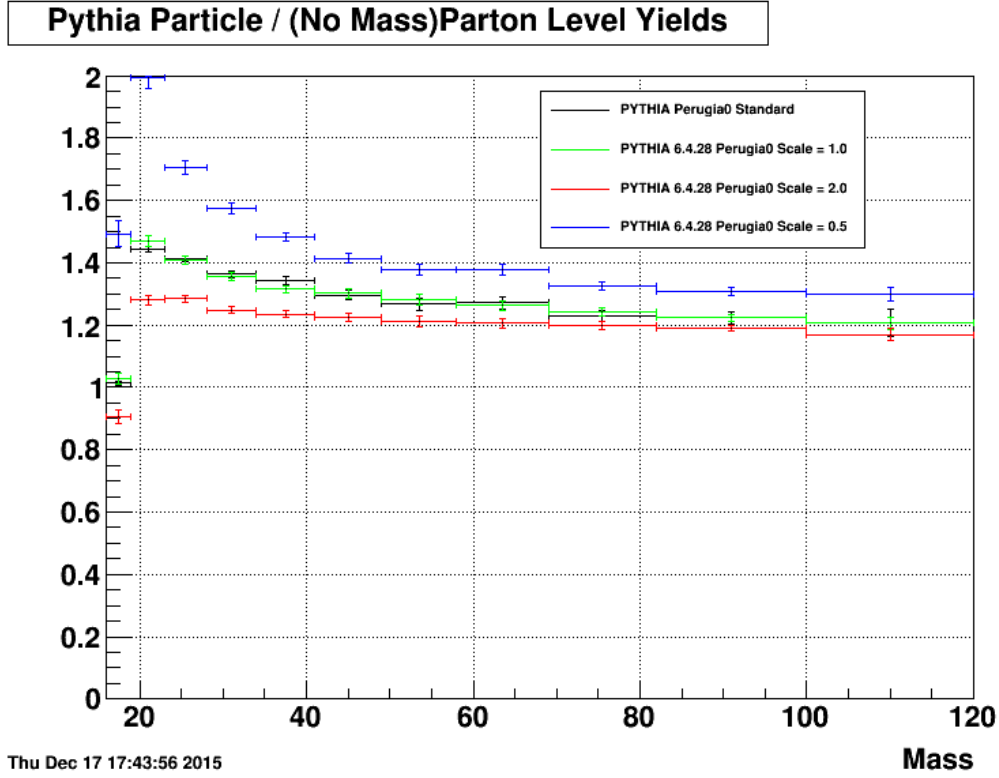


Figure 4.22: The particle over no mass parton level di-jet yield ratios for  $Q^2$  rescaled by factors of 2.0 and 0.5.

where  $X$  is the observable of interest (in this case the particle/parton ratio) and  $X[S_i^+]$  and  $X[S_i^-]$  are the values of the observable obtained using the  $i^{th}$  positive and negative error set, respectively.

A small Pythia sample was created for each error set associated with the CTEQ6M PDF set (LHAPDF ID 10000). Although the PDF associated with the Perugia0 tune is CTEQ5L, that set was not included in LHAPDF version 6.1.5 so the next available set was chosen. There are 20X2 error PDFs associated with CTEQ6M. The 20 sets correspond to 20 parameters in the PDF fitting procedure which are sensitive to current data. The error PDFs are obtained when one parameter is varied while holding all others fixed. The parameter is varied positively and negatively, resulting in two error sets for each of the 20 parameters. A comparison of the particle/parton level ratio using the nominal simulation, the CTEQ6M Perugia0 central value and the fifth positive and negative error sets can be seen in figure 4.23. The UEH correction values and associated systematic errors are shown in table 4.3.

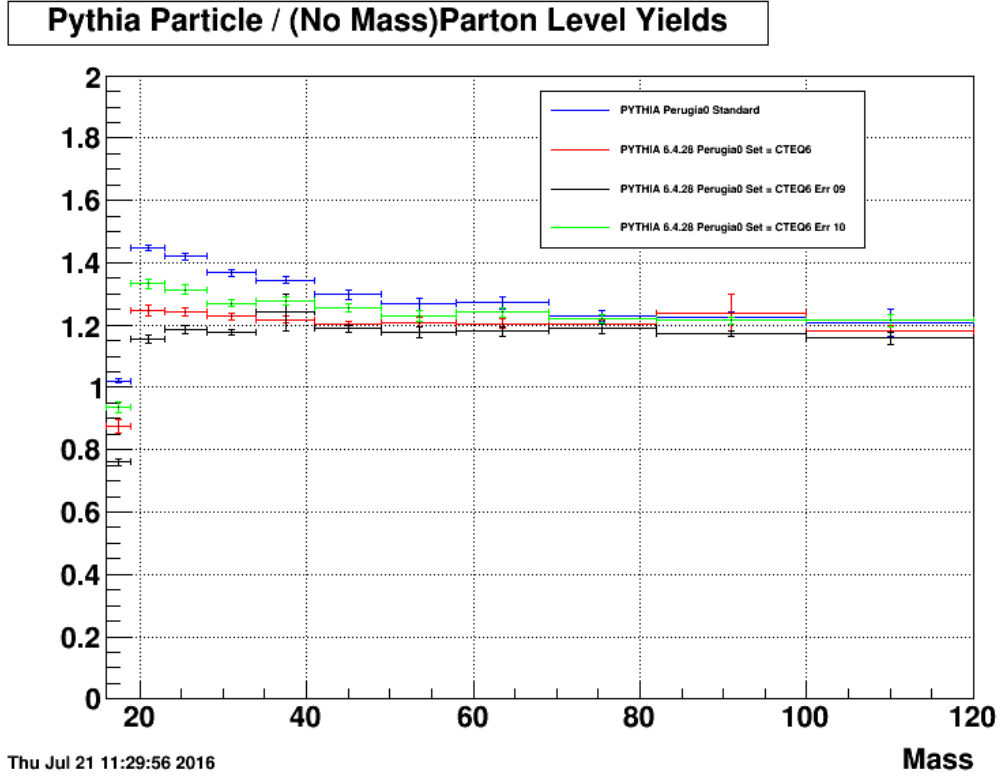


Figure 4.23: The particle over no mass parton level di-jet yield ratios for the nominal, CTEQ6M central, and CTEQ6M error set 5 samples.

Di-jet Mass [GeV/c <sup>2</sup> ]	Correction Factor	Hi/Lo Scale	PDF	Hi/Lo Total
19-23	1.447	0.378 / 0.115	0.089	0.389 / 0.145
23-28	1.420	0.201 / 0.095	0.071	0.213 / 0.119
28-34	1.368	0.152 / 0.087	0.049	0.160 / 0.100
34-41	1.345	0.104 / 0.082	0.033	0.109 / 0.088
41-49	1.298	0.090 / 0.055	0.043	0.100 / 0.070
49-58	1.268	0.087 / 0.044	0.040	0.096 / 0.059
58-69	1.271	0.085 / 0.051	0.037	0.093 / 0.063
69-82	1.228	0.079 / 0.023	0.033	0.086 / 0.040
82-100	1.223	0.070 / 0.027	0.034	0.077 / 0.043

Table 4.3: Numerical values for the UEH correction and corresponding systematics.

The Perugia0 tunes to Pythia were made assuming the CTEQ5L PDF set, so using the Perugia0 tune with CTEQ6M to estimate the PDF uncertainty contribution to the UEH systematic is not fully correct. The error will mix effects from the actual PDF differences along with various Pythia parameters. This means the uncertainties presented here are

likely slight overestimates.

## 4.5 Trigger Sample Comparison

Section 4.1 describes how di-jets are sorted into the L2JetHigh and JP1 trigger samples based on the properties of the individual jets in the di-jet. The L2JetHigh sample contains many more events than the JP1 sample (see figure 4.24) and the combination of JP1 with L2JetHigh would not significantly reduce the statistical errors versus the use of the L2JetHigh sample alone. Because of this, only the L2JetHigh sample is used for the final cross section result

Although the JP1 sample is not used in the cross section result, it is still useful as a cross check on the L2JetHigh sample. The JP1 and L2JetHigh samples have different raw di-jet invariant mass distributions, but the unfolding procedure should remove any differences due to trigger effects in the final cross sections. Any large discrepancies between the L2JetHigh and JP1 cross sections could be an indication of a problem. Figure 4.25 shows a comparison between the L2JetHigh and JP1 cross sections and good agreement is seen at low to intermediate values of invariant mass. At larger di-jet mass values the agreement becomes worse, but this is not surprising given the lack of statistics in the JP1 sample at these masses.

## 4.6 Results

The di-jet cross section is presented as a function of invariant mass in the mid-rapidity ( $|\eta| < 0.8$ ) region. The result was obtained using  $18.6\text{pb}^{-1}$  of  $pp$  data taken at  $\sqrt{s} = 200$  GeV from Run IX and used the Anti- $k_T$  jet algorithm with a radius of 0.6 to reconstruct jets. The results are compared to an NLO pQCD prediction from deFlorian using the CT10 PDF set which has been corrected for UEH effects. Systematic uncertainties on the theory were obtained by considering sensitivities to the normalization and factorization scales as well as the PDF uncertainty. The total theoretical systematic is the quadrature sum of these components with the UEH uncertainty (shown separately). The (UEH corrected) theoretical values and corresponding systematics can be seen in table 4.4. The cross section and comparison to theory are presented graphically in figure 4.26. Note that an overall scale uncertainty of 8.8% on the determination of the integrated luminosity is not shown.



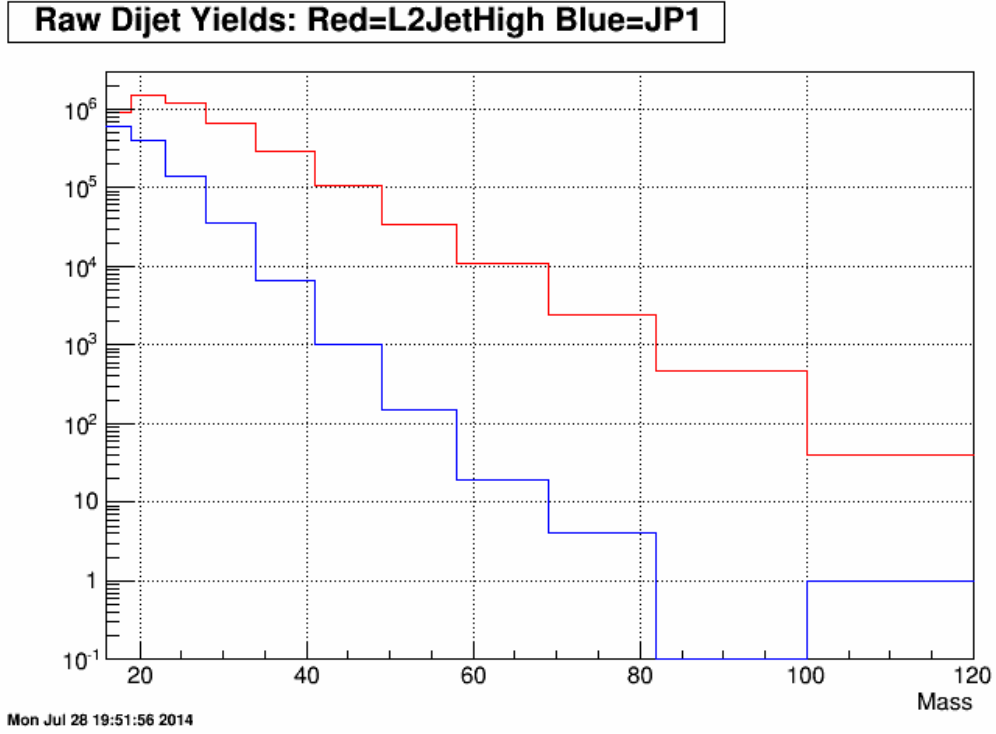


Figure 4.24: This plot shows the raw (no unfolding, normalization, or phase space factors) di-jet yields for the L2JetHigh sample (Red) and the JP1 sample (Blue).

Di-jet Mass [GeV/c <sup>2</sup> ]	Theory [ $\mu\text{b}$ ]	Hi/Lo UEH	Hi/Lo Scale	Hi/Lo PDF	Hi/Lo Total
19-23	$1.570 \times 10^{-1}$	0.389 / 0.145	0.175 / 0.188	0.073 / 0.064	0.432 / 0.246
23-28	$5.096 \times 10^{-2}$	0.213 / 0.119	0.160 / 0.183	0.078 / 0.073	0.278 / 0.230
28-34	$1.549 \times 10^{-2}$	0.160 / 0.100	0.169 / 0.191	0.119 / 0.084	0.261 / 0.231
34-41	$4.693 \times 10^{-3}$	0.109 / 0.088	0.189 / 0.195	0.101 / 0.104	0.240 / 0.238
41-49	$1.371 \times 10^{-3}$	0.100 / 0.070	0.193 / 0.200	0.117 / 0.110	0.247 / 0.238
49-58	$3.974 \times 10^{-4}$	0.096 / 0.059	0.209 / 0.212	0.137 / 0.119	0.267 / 0.250
58-69	$1.088 \times 10^{-4}$	0.093 / 0.063	0.227 / 0.219	0.167 / 0.140	0.296 / 0.268
69-82	$2.385 \times 10^{-5}$	0.086 / 0.040	0.245 / 0.231	0.223 / 0.158	0.342 / 0.283
82-100	$3.826 \times 10^{-6}$	0.077 / 0.043	0.256 / 0.245	0.341 / 0.193	0.434 / 0.315

Table 4.4: Numerical values for the UEH corrected theoretical cross section (PDF = CT10), UEH systematic, theory scale systematic, theory PDF systematic, and total systematic.

Figure 4.26 shows that the measured di-jet cross section is in good agreement with the NLO pQCD prediction. This agreement gives confidence that the di-jet observable measured by STAR is well understood and can be used in  $A_{LL}$  measurements. The size of

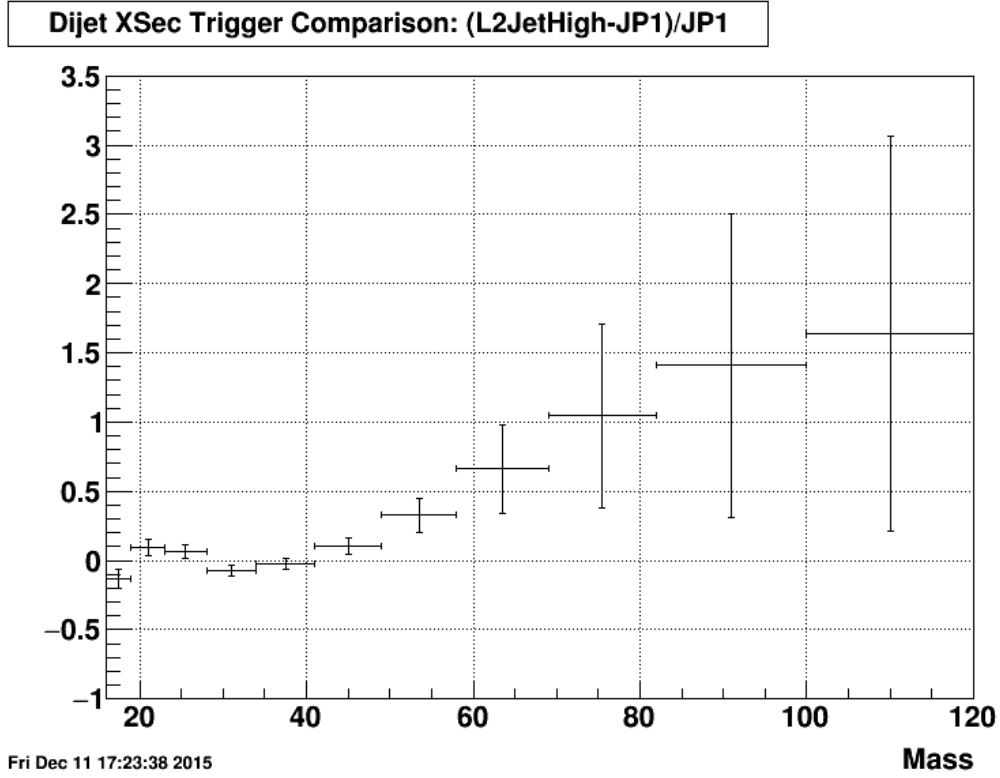


Figure 4.25: This plot shows the cross section ratio  $(\text{L2JetHigh-JP1})/\text{JP1}$ . Good agreement can be seen between the two cross sections until the JP1 counts die off.

the systematic errors on the data are smaller than the corresponding theory systematics in all bins, meaning these data have the possibility of constraining models of UEH effects at low mass and unpolarized PDFs at intermediate to high mass.

Di-jet Invariant Mass [GeV]	$\sigma \pm (\text{Stat}) \pm (\text{Sys})[\mu\text{b}]$
19-23	$(1.188 \pm 0.026^{+0.067}_{-0.071}) \times 10^{-1}$
23-28	$(4.105 \pm 0.096^{+0.242}_{-0.261}) \times 10^{-2}$
28-34	$(1.320 \pm 0.030^{+0.135}_{-0.146}) \times 10^{-2}$
34-41	$(3.923 \pm 0.080^{+0.451}_{-0.484}) \times 10^{-3}$
41-49	$(1.267 \pm 0.020^{+0.150}_{-0.160}) \times 10^{-3}$
49-58	$(4.004 \pm 0.060^{+0.583}_{-0.618}) \times 10^{-4}$
58-69	$(1.099 \pm 0.017^{+0.192}_{-0.204}) \times 10^{-4}$
69-82	$(2.578 \pm 0.048^{+0.525}_{-0.557}) \times 10^{-5}$
82-100	$(4.276 \pm 0.101^{+0.959}_{-1.022}) \times 10^{-6}$

Table 4.5: The cross section values (in  $\mu\text{b}$ ) with statistical and systematic error values for each di-jet invariant mass bin.

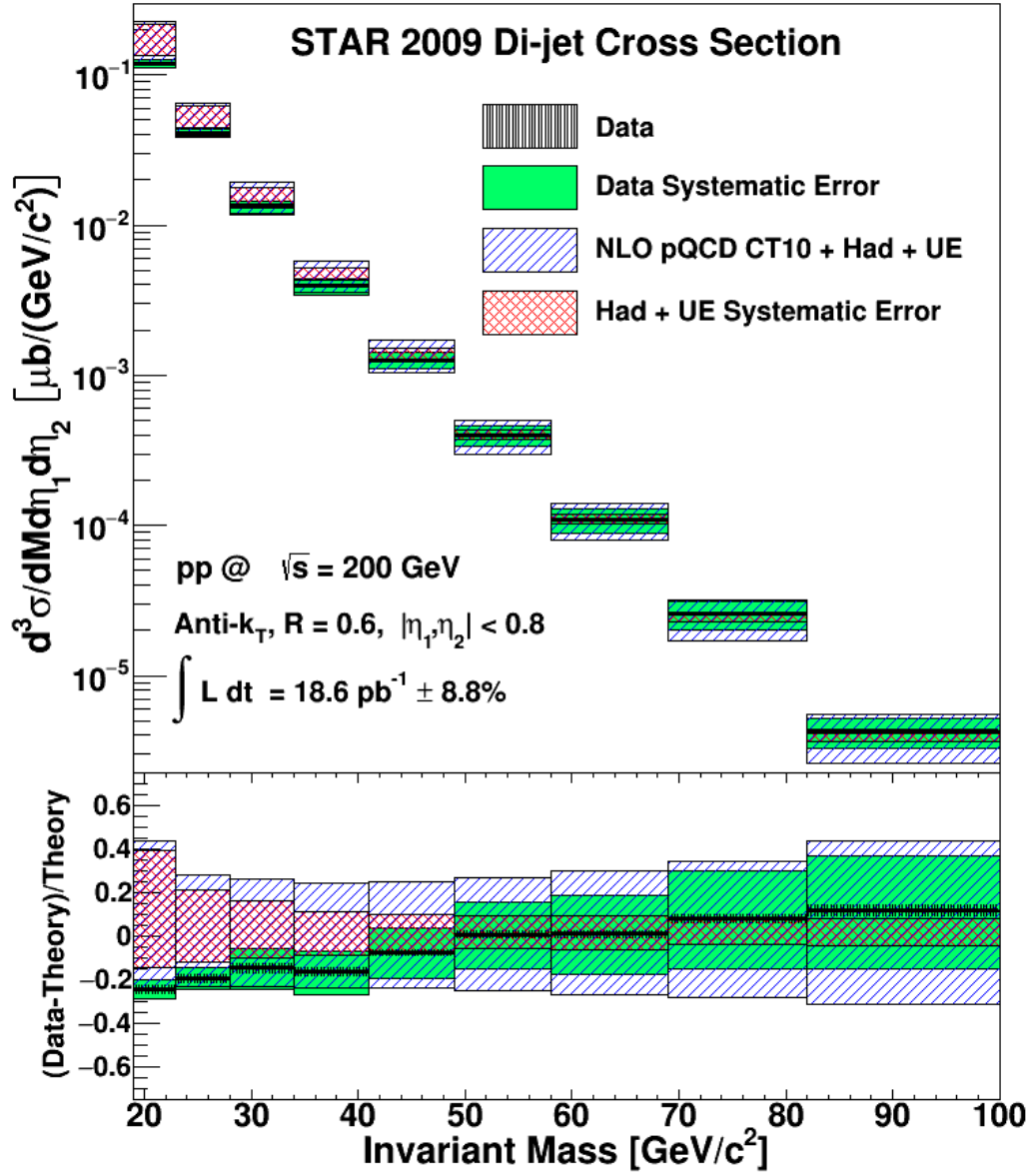


Figure 4.26: The top panel shows the measured cross section (Black lines) with the associated systematic error (green box), UEH systematic (red double-hatched box), and combined UEH and theoretical error (blue single-hatched box). The bottom panel shows the  $(\text{Data-Theory})/\text{Theory}$  ratio (where the theory has been corrected for UEH effects) with the black lines and vertical hatching representing the value and statistical error on the ratio.

# Chapter 5

## Di-jet Helicity Asymmetries

The longitudinal double-spin asymmetry,  $A_{LL}$ , is the primary observable used at RHIC to study the gluon spin contribution ( $\Delta G$ ) to the overall spin of the proton. The longitudinal double-spin asymmetry is defined in terms of helicity-dependent cross sections:

$$A_{LL} \equiv \frac{\sigma_{++} - \sigma_{+-}}{\sigma_{++} + \sigma_{+-}} \quad (5.1)$$

where  $\sigma_{++}$  ( $\sigma_{+-}$ ) is the cross section for partons of the same (opposite) helicity. In this analysis,  $A_{LL}$  is extracted directly from the helicity-dependent yields of a final state observable, so the expression for the experimentally determined  $A_{LL}$  will differ from the above equation. There are also some modifications needed when combining many independent measurements (each run can be considered an independent measurement). Equation 5.2 is the formula for  $A_{LL}$  used in this analysis. The derivation of this formula, along with discussion of the relevant assumptions made, can be found in appendix B.

$$A_{LL} = \frac{\sum_i P_{Yi} P_{Bi} [(N_i^{++} + N_i^{--}) - R_{3i} (N_i^{+-} + N_i^{-+})]}{\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i} (N_i^{+-} + N_i^{-+})]} \quad (5.2)$$

The summation is over all runs  $i$  in the data set. The factors  $P_Y$  and  $P_B$  represent the polarization values of the Yellow and Blue beams, respectively (see section 5.1 for more details).  $N^{++}, N^{--}, N^{+-}$ , and  $N^{-+}$  represent the yields for the four possible beam helicity combinations, where the first index denotes the helicity of the yellow beam bunch and the second index denotes that of the blue beam. The helicity combinations and spin patterns are detailed in section 5.2. The term  $R_3$  in equation 5.2 is the relative luminosity factor, which will be discussed in section 5.3.

The statistical error on  $A_{LL}$  can be found by standard error propagation techniques and

is given by equation 5.3. The derivation of this error is presented in appendix B.

$$\delta A_{LL} = \frac{(\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i}^2 (N_i^{+-} + N_i^{-+})])^{\frac{1}{2}}}{\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i} (N_i^{+-} + N_i^{-+})]} \quad (5.3)$$

## 5.1 Beam Polarizations

There were 103 fills used in the asymmetry measurement (see appendix A). For these fills, the average polarization values for the blue and yellow beams, respectively, were 56.38% and 56.93%. Figure 5.1 shows a plot of the fill-by-fill polarization values for the two beams.

### 5.1.1 Residual Transverse Component

The stable proton spin direction in RHIC is transverse (up-down) to the direction of motion and a series of spin rotators are used to precess the spin to a longitudinal orientation at the STAR interaction region. The spin rotators are not perfect, and there can be residual transverse components in the colliding proton bunches. The beam polarization values discussed above are for the total polarization,  $P = \sqrt{P_L^2 + P_T^2}$ . The quantity of interest for an  $A_{LL}$  measurement is  $P_L$ , so any residual transverse component will act as a dilution on the quoted beam polarization values. In this analysis, the beam polarization values are scaled by a factor  $F$  (equation 5.4) to return only the longitudinal polarization component ( $P_L = P \times F$ ).

$$F \equiv \frac{1}{\sqrt{1 + t^2}}, \quad t \equiv \frac{P_T}{P_L} \quad (5.4)$$

During most of Run IX, the beams had a somewhat large transverse component. On June 22, 2009 the rotator currents at STAR were tuned to reduce the size of these residual transverse components (see table 5.1). The values of  $t$  listed in the table were used to correct the beam polarizations used in the calculation of  $A_{LL}$ .

### 5.1.2 Polarization Decay

For each Fill, the polarimetry group provided a luminosity-weighted average polarization as well as an initial polarization value and a value for the change in polarization over time. This information was used to estimate the beam polarizations at the time of each event in the analysis.

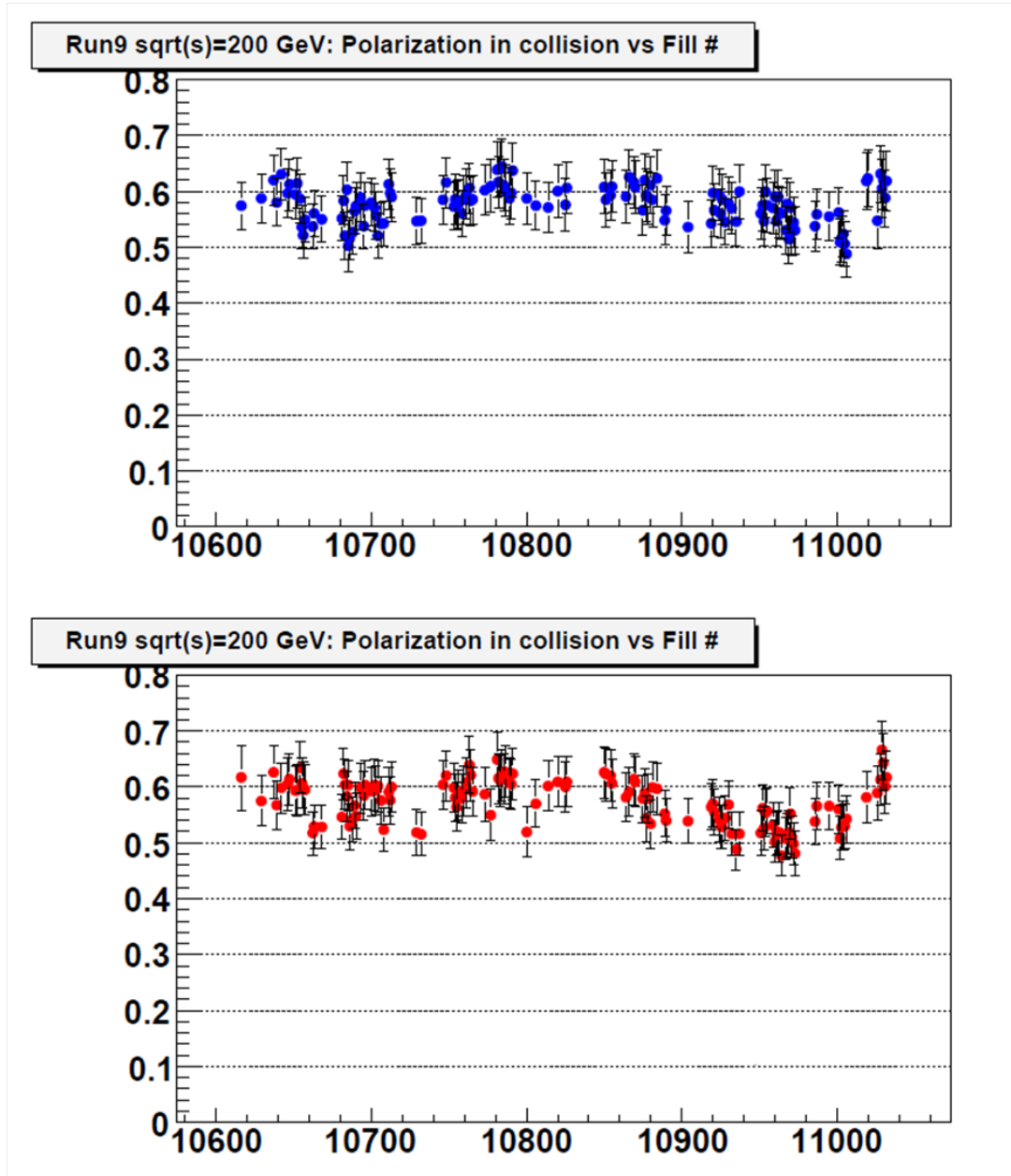


Figure 5.1: Beam polarizations by fill. The top plot shows the polarization of the Blue beam and the bottom plot shows the polarization of the Yellow beam.

The average polarization value reported by the polarimetry group for each fill was weighted by the luminosity over the course of that fill. This means that as long as the number of events sampled scales proportionally with the luminosity, the average polarization would be the correct value to use. This proportionality roughly holds for the L2JetHigh events, as that trigger was take-all throughout the run, so its event rate should scale with instantaneous luminosity. The JP1 trigger, however, was prescaled and the prescale value was

	Before Rotator Change	After Rotator Change
Yellow Beam	$t = 0.15 \pm 0.02$	$t = 0.13 \pm 0.02$
Blue Beam	$t = 0.25 \pm 0.02$	$t = 0.10 \pm 0.03$

Table 5.1: The  $t$  values ( $t \equiv P_T/P_L$ ) for the yellow and blue beams before and after the spin rotator tune, which occurred after run 1017050.

chosen in proportion to the instantaneous luminosity at the beginning of each run. This means that JP1 events are taken at a higher rate at the end of fills, when the luminosity is lower. Thus, using the average polarization value for the JP1 sample would tend to overestimate the beam polarizations appropriate for this sample. Calculating  $A_{LL}$  using the beam polarizations found as a function of event time alleviates this problem.

## 5.2 Spin Patterns

In a given fill, the polarization direction of each bunch in the two beams matched one of four spin patterns (see table 5.2). These four spin patterns were cycled through fill-by-fill over the course of Run IX.

Pattern	Blue Beam Pattern	Yellow Beam Pattern
P1	+ - + - - + - +	+ + - - + + - -
P2	- + - + + - + -	+ + - - + + - -
P3	+ - + - - + - +	- - + + - - + +
P4	- + - + + - + -	- - + + - - + +

Table 5.2: The four spin patterns used in Run IX. The ‘+’ symbol indicates positive helicity, and the ‘-’ indicates negative helicity. The patterns repeat for the remaining bunches in the fill.

At STAR, the colliding bunch helicity combination is encoded as a ‘Spin-4’ value (see table 5.3). The information on spin pattern and beam cogging is stored in an offline spin database which is accessed during an analysis, so that the Spin-4 value for any bunch crossing can be found. For a given spin pattern, each bunch crossing will have a set Spin-4 value. The Spin-4 values vs. bunch crossing number for all four spin patterns can be seen in figure 5.2.



Spin-4	Yellow Beam Helicity	Blue Beam Helicity
5	—	—
6	+	—
9	—	+
10	+	+

Table 5.3: The four ‘Spin-4’ values used to label the helicities of the colliding bunches for each event.

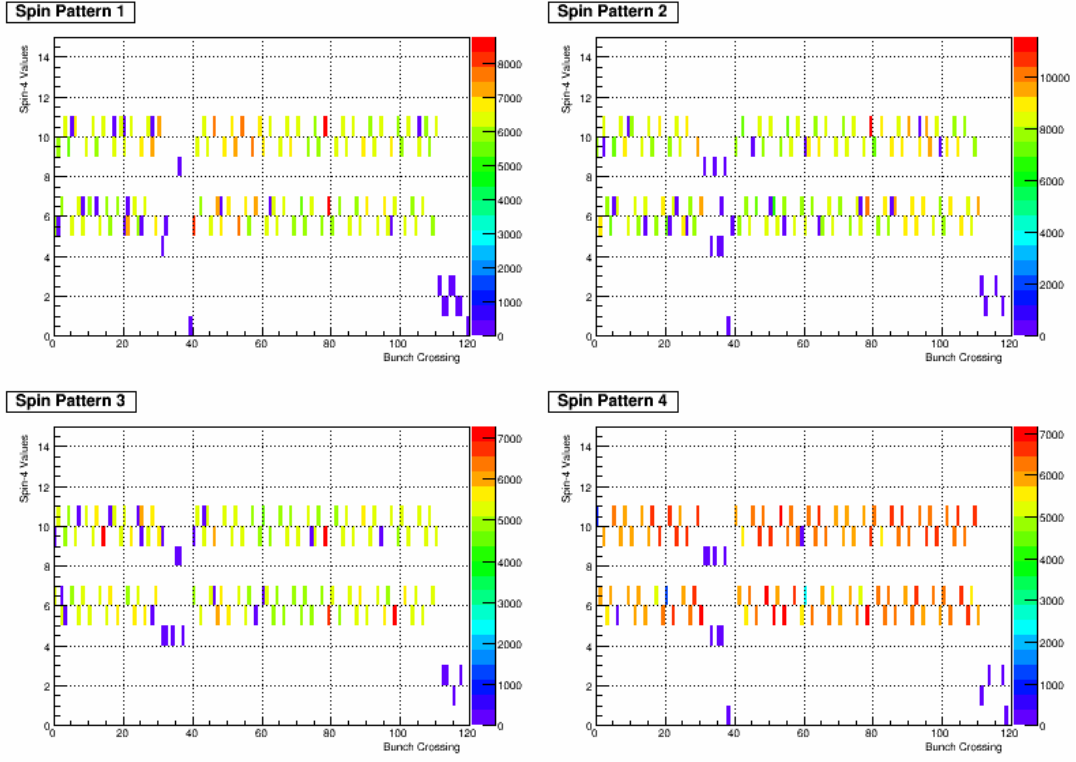


Figure 5.2: Map of which bunch crossings are tagged with which Spin-4 value for each of the four spin patterns shown in table 5.2.

### 5.3 Relative Luminosity

The final term in equation 5.2 which has not been discussed is the relative luminosity ratio,  $R_3$ .

During Run IX, the relative luminosities were measured at STAR by looking at the spin-sorted BBC coincidence rates. Relative luminosities were also calculated using the ZDC system and the differences observed between the BBC and ZDC were used to estimate

the systematic error on the quoted relative luminosity values. A detailed account of the extraction of the relative luminosities from the BBCs and ZDCs can be found here: [22].

There are six relative luminosity ratios which are germane to the asymmetry analysis, as defined in equations 5.5. The ratio  $R_3$  is obviously the quantity needed to properly normalize the helicity-sorted yields in an  $A_{LL}$  measurement. The other relative luminosities will enter into the calculation of various false asymmetries which will be discussed in section 5.5.

$$R_1 = \frac{\mathcal{L}^{--} + \mathcal{L}^{-+}}{\mathcal{L}^{++} + \mathcal{L}^{+-}} \quad (5.5a)$$

$$R_2 = \frac{\mathcal{L}^{++} + \mathcal{L}^{+-}}{\mathcal{L}^{--} + \mathcal{L}^{-+}} \quad (5.5b)$$

$$R_3 = \frac{\mathcal{L}^{++} + \mathcal{L}^{--}}{\mathcal{L}^{+-} + \mathcal{L}^{-+}} \quad (5.5c)$$

$$R_4 = \frac{\mathcal{L}^{--}}{\mathcal{L}^{++}} \quad (5.5d)$$

$$R_5 = \frac{\mathcal{L}^{-+}}{\mathcal{L}^{++}} \quad (5.5e)$$

$$R_6 = \frac{\mathcal{L}^{+-}}{\mathcal{L}^{++}} \quad (5.5f)$$

During the course of the relative luminosity investigation, it was found that certain bunch crossings had anomalous behavior. Figure 5.3 shows a typical event spectrum versus bunch crossing number. Due to this behavior, several bunches were not used when determining the relative luminosity values, and di-jet events from these bunches have also been removed from the asymmetry analysis. The bunch crossings removed were: 20, 60, 78, 79, and 80. In addition, the bunch crossings corresponding to the times when one of the abort gaps was passing through STAR have also been removed. These are bunches 31-39 and 111-119.

## 5.4 Trigger Selection and Binning

The  $A_{LL}$  values are presented as a function of the di-jet invariant mass. The mass bins retain the same boundaries as were used in the cross section analysis (see table 4.1), but several bins have been combined to reduce the statistical error on the high mass points. Table 5.4 shows the mass binning used for the  $A_{LL}$  measurement.

The di-jet events used in the asymmetry analysis were sorted into the same L2JetHigh and JP1 trigger categories as defined for the cross section analysis (see section 4.1). The JP1

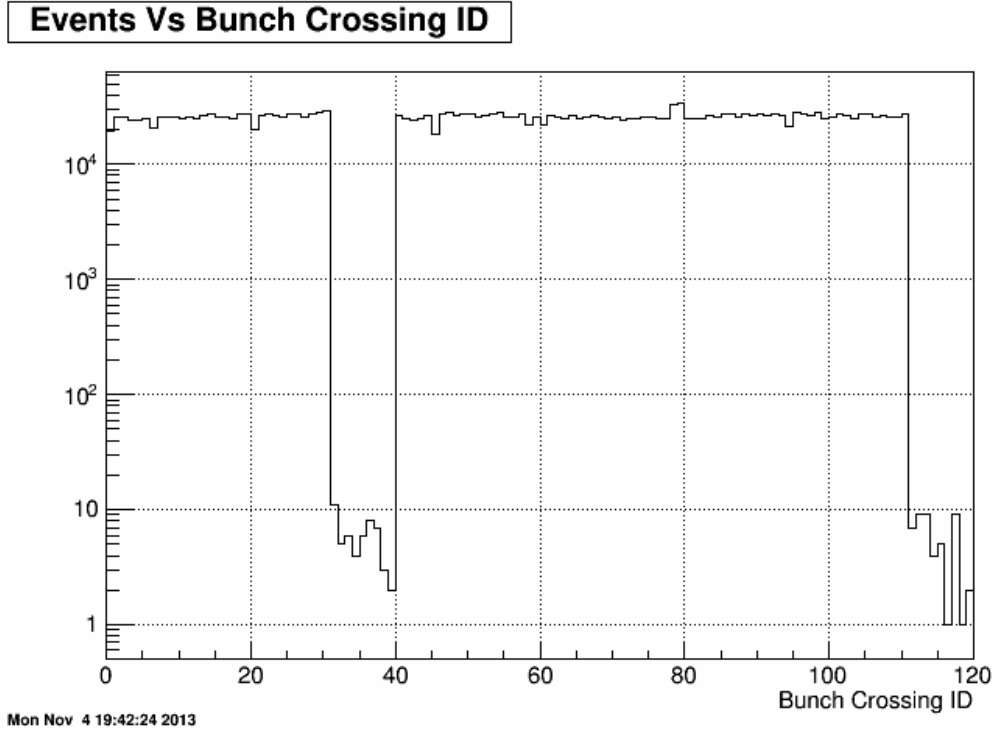


Figure 5.3: Number of events in each bunch crossing for a sample of data. The abort gaps at indices 31-39 and 111-119 are clearly seen.

Bin	Mass Range (GeV)
1	16-19
2	19-23
3	23-28
4	28-34
5	34-41
6	41-58
7	58-82
8*	82-120

Table 5.4: The mass ranges for each bin used in the  $A_{LL}$  analysis. Bin 8 was excluded from the final results due to a lack of statistics.

sample contains substantially fewer counts than the L2JetHigh sample, and was only used as a consistency check in the cross section measurement. The asymmetry measurement, unlike the cross section, is statistics limited so both the L2JetHigh and JP1 samples were used in order to get as many counts as possible. Table 5.5 shows the number of di-jet

events in each mass bin and spin configuration for each trigger category.

Bin	L2JetHigh				JP1			
	$N_5$	$N_6$	$N_9$	$N_{10}$	$N_5$	$N_6$	$N_9$	$N_{10}$
1	251632	246609	248211	250987	161290	158093	159057	159993
2	417788	406657	411783	415398	109212	107363	108676	109916
3	333795	325071	328833	332644	38590	37842	38103	38431
4	180764	175718	178158	179717	9450	9232	9216	9380
5	77086	75061	76175	76911	1755	1771	1624	1756
6	37862	36690	36560	37181	316	293	319	327
7	3600	3409	3599	3532	3	6	5	8
8	146	126	123	138	0	0	0	1

Table 5.5: Number of events in each helicity combination (see table 5.3) for the two trigger samples. The full topological sample is used.

In addition to the full data sample,  $A_{LL}$  was determined for di-jets with two jet topologies. The first sub-sample contains di-jets in which both jets are either in the East half of the BEMC ( $-0.8 \leq \eta \leq 0.0$ ) or the West half of the BEMC ( $0.0 \leq \eta \leq 0.8$ ). The second sub-sample contains di-jets with one jet in the East half of the BEMC and the other in the West. Table 5.6 shows the number of di-jet events broken down by mass bin and spin configuration for the two topological configurations.

## 5.5 False Asymmetries

In addition to the longitudinal double-spin asymmetry (equation 5.2), four other useful asymmetries can be calculated (see equation 5.6). These additional asymmetries are the yellow and blue beam single-spin asymmetries ( $A_L^Y$  and  $A_L^B$ ) and the like and unlike sign double spin asymmetries ( $A_{LL}^{ls}$  and  $A_{LL}^{us}$ ). These four quantities provide checks on the analysis scheme because they are expected to be zero in this data set. The  $A_L^Y$ ,  $A_L^B$ , and  $A_{LL}^{ls}$  asymmetries all measure parity-violating effects, and are therefore expected to be negligible. The  $A_{LL}^{us}$  asymmetry should be zero as collisions in which the yellow beam has positive helicity and the blue beam has negative helicity should be the same as the reverse case.

Bin	East-West				East-East+West-West			
	$N_5$	$N_6$	$N_9$	$N_{10}$	$N_5$	$N_6$	$N_9$	$N_{10}$
1	177487	173735	174733	175995	235435	230967	232535	234985
2	264273	257865	260592	263201	262727	256155	259867	262113
3	199764	194563	197016	198018	172621	168350	169920	173057
4	106363	102927	104378	105710	83851	82023	82996	83387
5	45378	44168	44989	45363	33463	32664	32810	33304
6	22815	22324	22160	22542	15363	14659	14719	14966
7	2345	2203	2389	2229	1258	1212	1215	1311
8	107	97	87	113	39	29	36	26

Table 5.6: Number of events in each helicity combination (see table 5.3) for the two topological configurations.

Significant deviation from zero for any of the four asymmetries could indicate a problem with the relative luminosity values, or a bug in the asymmetry calculation. Figures 5.4, 5.6, and 5.5 show the four additional asymmetries for the full sample, East-East+West-West sample, and East-West sample, respectively. The additional asymmetries agree, with reasonable  $\chi^2$  values, with constant functions and all the fits are consistent with an asymmetry of zero.

$$A_L^Y = \frac{\sum P_Y [(N^{--} + N^{-+}) - R_1 (N^{++} + N^{+-})]}{\sum P_Y^2 [(N^{--} + N^{-+}) + R_1 (N^{++} + N^{+-})]} \quad (5.6a)$$

$$A_L^B = \frac{\sum P_B [(N^{--} + N^{+-}) - R_2 (N^{++} + N^{-+})]}{\sum P_B^2 [(N^{--} + N^{+-}) + R_2 (N^{++} + N^{-+})]} \quad (5.6b)$$

$$A_{LL}^{ls} = \frac{\sum P_Y P_B (N^{--} - R_4 N^{++})}{\sum P_Y^2 P_B^2 (N^{--} + R_4 N^{++})} \quad (5.6c)$$

$$A_{LL}^{us} = \frac{\sum P_Y P_B (R_6 N^{-+} - R_5 N^{+-})}{\sum P_Y^2 P_B^2 (R_6 N^{-+} + R_5 N^{+-})} \quad (5.6d)$$

The false asymmetries were also calculated run-by-run (and fill-by-fill) to be sure there were no time dependencies. An example for the  $A_L^Y$  asymmetry can be seen in figure 5.7. The points are fit with a  $P_0$  and are consistent with zero. The behavior of the

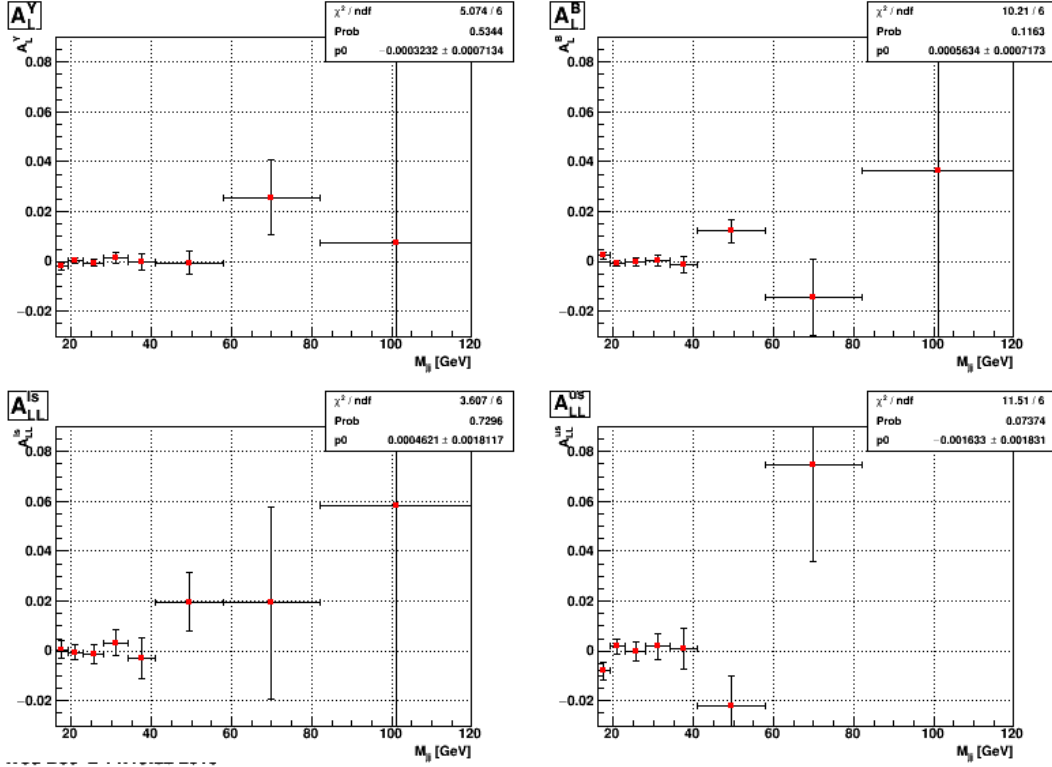


Figure 5.4: The four false asymmetries for the full topological sample. The first 7 data points are fit with a 0th order polynomial, and fit results can be seen in the legend. Note that the points are plotted at bin center and that the horizontal error bars span the width of the bin.

three remaining false asymmetries are similar. The four false asymmetries are also shown integrated over all runs and mass bins in figure 5.8.

## 5.6 Invariant Mass Shift

Unlike the cross section which drops rapidly, the di-jet  $A_{LL}$  is expected to be a slowly varying function of the invariant mass. Because of this, the full unfolding procedure used in the cross section analysis was not used in the extraction of the asymmetry. The determination of the particle or parton level mass of each data point was made by applying a simple mass shift to each point.

The mass shifts were found by comparing the detector level di-jet masses to the masses of matching particle or (no mass) parton level di-jets in simulation. The procedure for finding particle and parton level di-jets which match to detector level di-jets is described

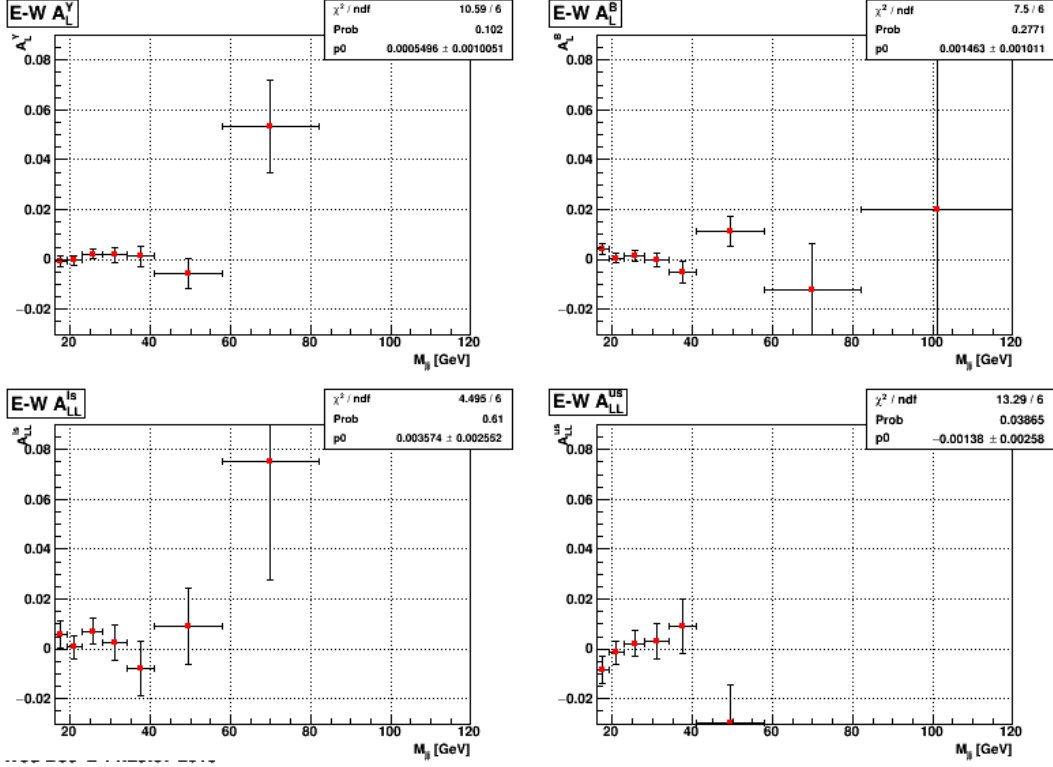


Figure 5.5: The four false asymmetries for the East-West topological sample. Details are presented in the caption of figure 5.4.

in section 4.2.1. The mass shift for a given bin was calculated by finding the difference between the particle or parton level di-jet mass and the detector level di-jet mass on an event-by-event basis for all di-jets which fall into that given mass bin. The mass shift is the mean value of this distribution.

The mass shifts were calculated separately for the L2JetHigh and JP1 trigger samples, for the East-West, East-East+West-West, and combined di-jet topologies, and for the particle and parton levels (see table 5.7 for the parton level mass shifts for each trigger and topological sample). To gage the effect that the track finding efficiency has on the mass shifts, the shifts were also calculated using the simulation samples that had 4% and 7% of the reconstructed tracks removed before the jet finder was run (see section 4.3).

The calculated mass shifts were applied to the data points as follows. First, the mass-weighted average position, as well as the number of events which are L2JetHigh and JP1 in each  $A_{LL}$  data bin, were found. The mass shift applied to the data was the sum of the L2JetHigh and JP1 mass shifts, weighted by the numbers of those types of events

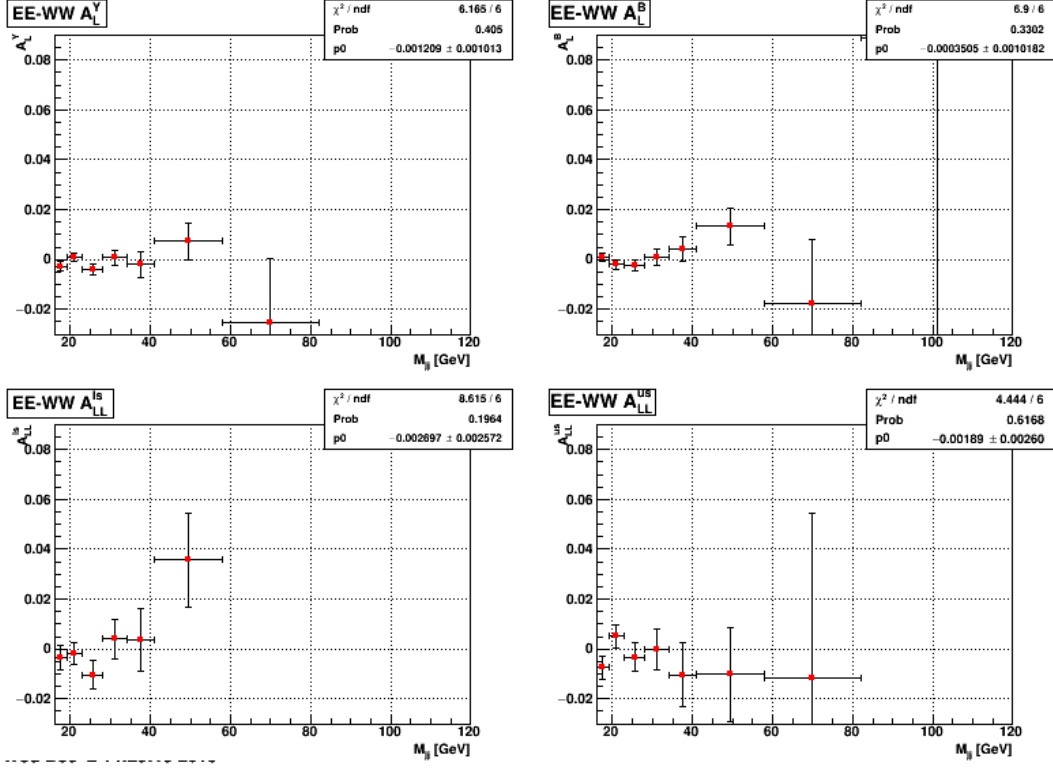


Figure 5.6: The four false asymmetries for the East-East+West-West topological sample. Details are presented in the caption of figure 5.4.

Bin	Combined			East-East+West-West			East-West		
	L2 Frac	L2 Shift	JP Shift	L2 Frac	L2 Shift	JP Shift	L2 Frac	L2 Shift	JP Shift
1	0.61	0.588	-0.356	0.64	0.438	-0.346	0.57	0.828	-0.368
2	0.79	0.813	-0.280	0.82	0.689	-0.356	0.77	0.953	-0.221
3	0.90	1.073	0.087	0.91	0.916	0.164	0.88	1.216	0.033
4	0.95	1.221	0.225	0.96	1.215	0.012	0.94	1.226	0.359
5	0.98	1.479	0.566	0.98	1.427	0.401	0.98	1.517	0.659
6	0.99	2.331	0.790	0.99	2.141	0.202	0.99	2.457	1.125
7	1.00	2.764	-	1.00	2.587	-	1.00	2.857	-
8	1.00	1.151	-	1.00	-0.060	-	1.00	1.548	-

Table 5.7: L2JetHigh event fraction from data and the mass shifts for the L2JetHigh and JP1 triggers at parton level for the three topological samples.

seen in the data. This combined mass shift was added to the mass-weighted average position to obtain the corrected position of each data point (see table 5.8). This was done independently for the three topological groupings. The parton level mass shifts with no track loss were used in the final analysis, and the difference between the no track



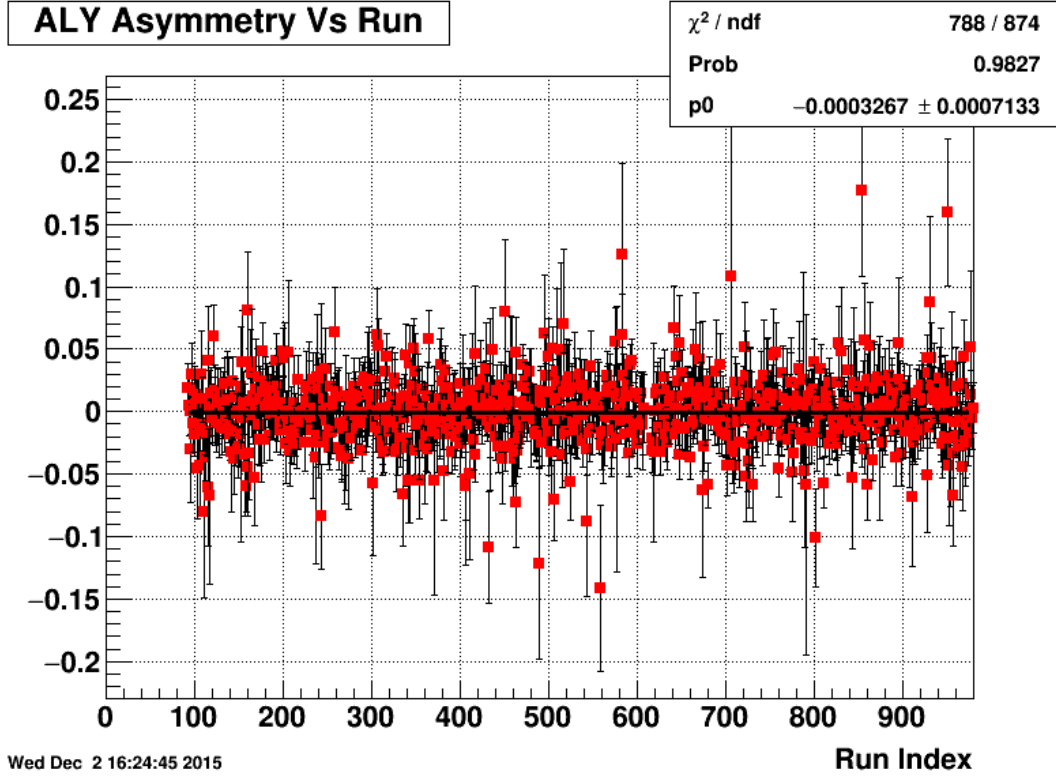


Figure 5.7: Run-by-Run yellow beam single spin asymmetry fit with a  $P_0$ .

loss and 4% track loss shifts was used in the determination of the systematic on the shift (see section 5.8.1). Figure 5.9 shows the mass shifts for the combined topological sample. The behavior for the East-West and East-East+West-West topological divisions are qualitatively similar to the combined.

## 5.7 Trigger and Reconstruction Bias Correction

The measured  $A_{LL}$  is a convolution of contributions from  $qq$ ,  $qg$ , and  $gg$  partonic scattering sub-processes. The jet-patch triggers are somewhat biased toward quark jets and this alters the sub-process fractions of the measured event sample compared to the underlying physical fractions. In addition, finite detector resolution can systematically shift the di-jet mass spectrum and thus change the sub-process fraction associated with a given mass. To address these effects, a trigger and reconstruction bias correction is applied to the extracted  $A_{LL}$  values and a corresponding systematic is evaluated.

The determination of the trigger and reconstruction bias correction and systematic follows the procedure used in the 2009 inclusive jet  $A_{LL}$  analysis [9]. The correction and

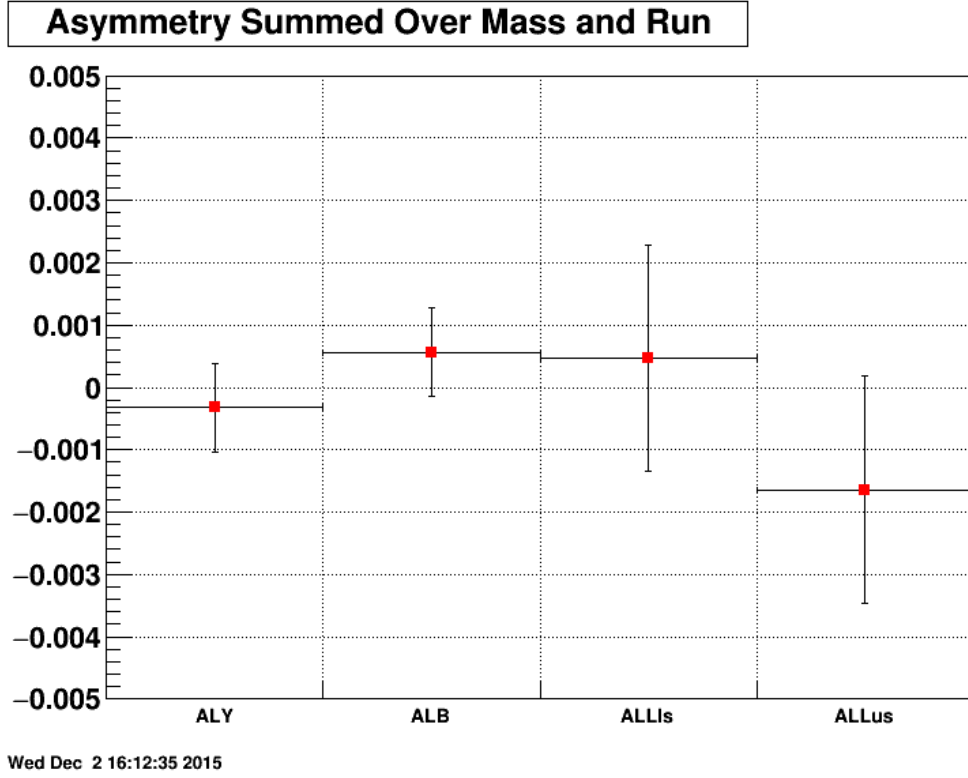


Figure 5.8: The four di-jet false asymmetries integrated over all invariant mass bins.

Bin	Combined		East-East+West-West		East-West	
	Ave Mass	Shift	Ave Mass	Shift	Ave Mass	Shift
1	17.600	0.219	17.545	0.155	17.674	0.314
2	20.873	0.585	20.843	0.497	20.904	0.678
3	25.189	0.971	25.167	0.848	25.209	1.079
4	30.524	1.171	30.499	1.163	30.544	1.179
5	36.865	1.459	36.840	1.407	36.884	1.495
6	46.270	2.318	46.158	2.127	46.345	2.444
7	64.328	2.759	64.063	2.585	64.472	2.851
8	89.087	1.146	89.057	-0.06	89.097	1.533

Table 5.8: The raw average mass values and total mass shifts at parton level for the three topological samples.

systematic were estimated by calculating the difference in  $A_{LL}$  between detector level jets and particle/parton level jets, where the  $A_{LL}$  values were obtained from the simulation by using the initial partonic kinematics from the PYTHIA record to extract values from

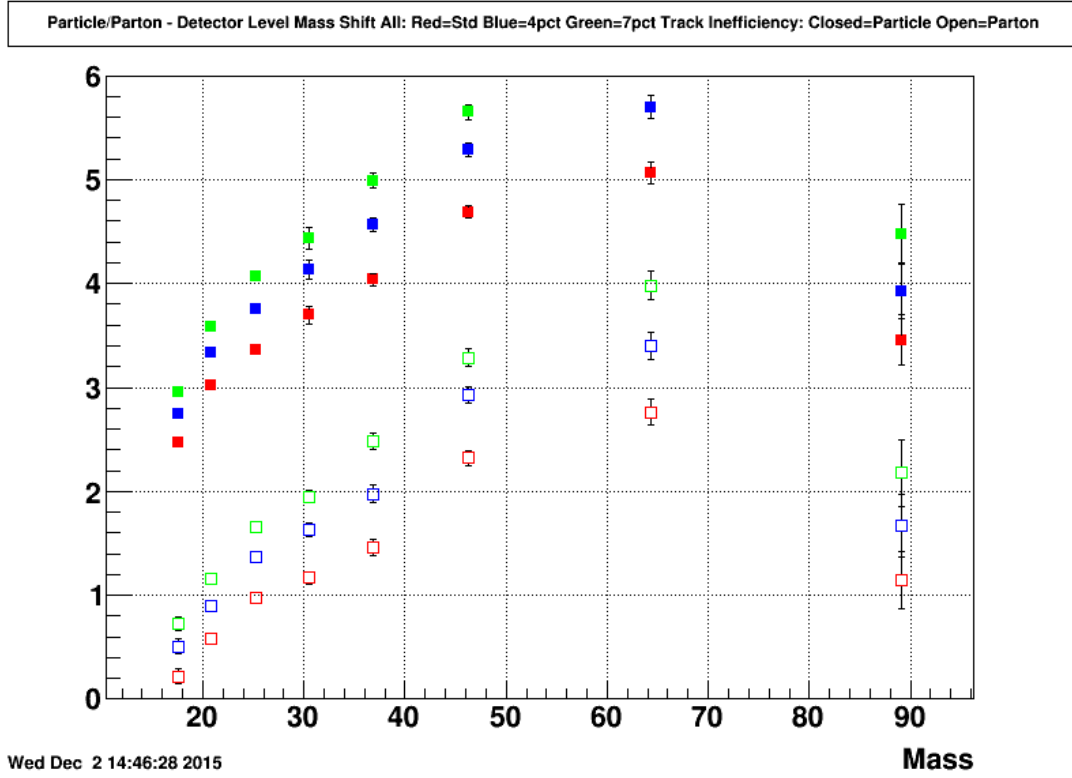


Figure 5.9: Summary of all mass shifts calculated for the full topological sample. The solid shapes represent the particle level shifts and the open shapes represent the parton level shifts. The red squares represent the 0% track-loss shifts, the blue triangles represent the 4% track-loss shifts, and the green inverted triangles represent the 7% track-loss shifts.

various polarized PDFs. Three polarized PDFs were used in the determination of the trigger and reconstruction bias correction/systematic:

**DSSV 2008** The original DSSV global analysis which incorporates DIS, SIDIS, and polarized  $pp$  data from RHIC [23], [24].

**DSSV 2009a** A private extension of the DSSV 2009 global analysis which incorporates the preliminary Run 9 STAR jet  $A_{LL}$  and PHENIX  $\pi^0$   $A_{LL}$  results.

**LSS10p** A NLO global analysis which incorporates DIS and SIDIS data. LSS10p is a fit which has a positive gluon density at the input scale [25].

The trigger and reconstruction bias corrections/systematics were calculated separately for each topology configuration using the following procedure:

1. For each event, find the particle/parton level di-jet from the full, unbiased PYTHIA sample. Use only the  $\Delta\phi$ , jet  $\eta$ , and asymmetric  $p_T$  cuts. Plot  $A_{LL}$  from the three

polarized PDFs vs the particle/parton level di-jet mass (parton level uses no mass formula). Place points at mass-weighted mean of the bin.

2. For each event, find the detector level di-jet from the simulation sample which passed the trigger filter. Place all detector level cuts but do not require detector level to particle level matching. Plot  $A_{LL}$  for the three polarized PDFs vs the detector level di-jet mass. Final  $A_{LL}$  is the trigger fraction weighted sum of the  $A_{LL}$  from the L2JetHigh and JP1 triggers. Place points at mass-weighted mean of the bin.
3. Shift the detector level masses using the factors described in section 5.6.
4. Calculate  $\Delta A_{LL}$  for each mass bin.  $\Delta A_{LL}$  is the difference between the particle/parton and detector level  $A_{LL}$  evaluated at the shifted detector level mass.

The trigger and reconstruction bias correction is taken as the average of the minimum and maximum  $\Delta A_{LL}$  values. Half the range between the minimum and maximum  $\Delta A_{LL}$  was used in the calculation of the associated systematic which will be discussed further in section 5.8.4. Figure 5.10 shows the  $\Delta A_{LL}$  for the combined topological sample. The values of the corrections are presented in table 5.9.

Bin	Combined Correction	EE+WW Correction	E-W Correction
1	0.0004	0.0002	0.0006
2	0.0009	0.0005	0.0010
3	0.0007	0.0003	0.0007
4	0.0013	0.0003	0.0018
5	0.0017	0.0015	0.0016
6	0.0026	0.0015	0.0028
7	0.0039	0.0004	0.0048

Table 5.9: Values of the trigger and reconstruction bias shifts for the three topologies.

## 5.8 Systematics

The  $A_{LL}$  systematic errors were determined in much the same way as those found in the Run 9 inclusive jet analysis [9]. The systematics were broken into two categories: systematics on the di-jet mass (x-axis uncertainty) and those on the  $A_{LL}$  value (y-axis

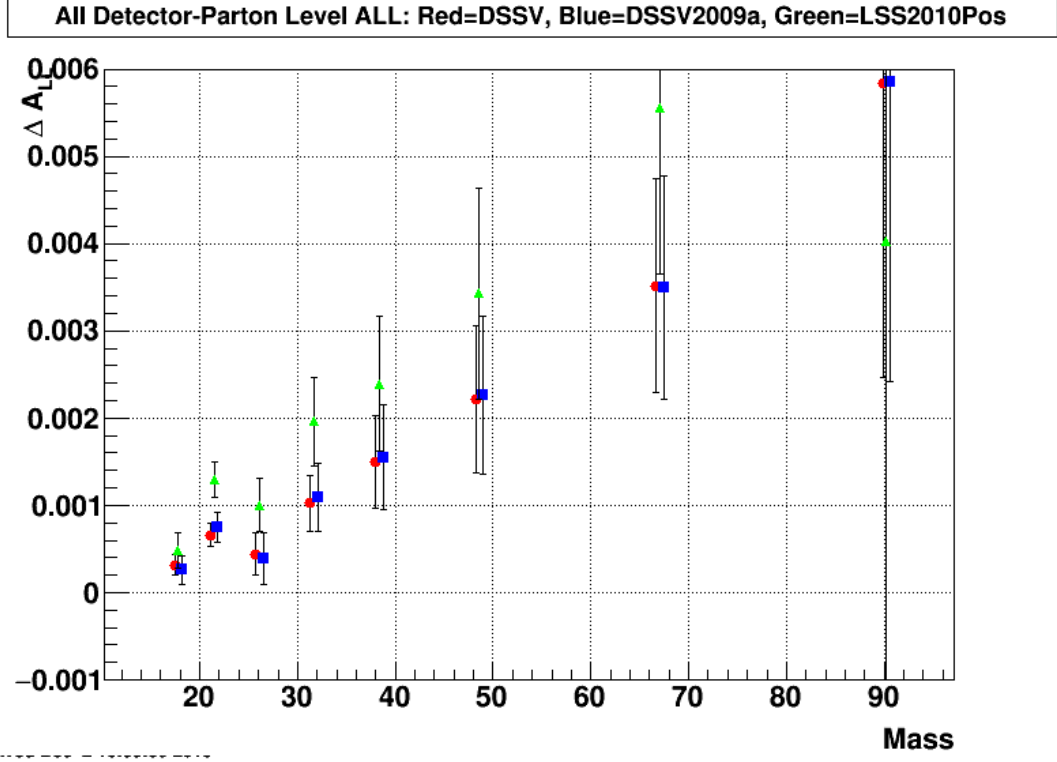


Figure 5.10:  $\Delta A_{LL}$  for the combined topological sample. Red points are from the DSSV PDF set, blue are from DSSV2009a, and green are from LSS2010Pos.

uncertainty). The di-jet mass systematic includes contributions from the mass shift systematic, the jet energy systematic, and a systematic which comes from comparing the PYTHIA and NLO cross sections. The systematic on the extracted  $A_{LL}$  includes contributions from the trigger and reconstruction bias systematic as well as a systematic due to the residual transverse polarization component of the beams. Note: A systematic due to a possible contribution to  $A_{LL}$  from background events was evaluated for the inclusive jet sample and was found to be less than 2% of the statistical uncertainty of the measurement for all bins and deemed negligible. Because of the correlation requirement, the di-jet sample is expected to be less sensitive to background contributions and the background  $A_{LL}$  systematic was not evaluated in this analysis.

In addition to the point-to-point systematics mentioned above and described in more detail below, there are two scale uncertainties which affect all points equally: the polarization uncertainty and the relative luminosity systematic uncertainty. The polarization uncertainty was determined by the polarimetry group and has a value of 6.5%. The relative luminosity systematic used is the same as from the inclusive jet analysis which

evaluated BBC/ZDC differences and false asymmetry magnitudes to arrive at a value of  $\pm 0.0005$ .

### 5.8.1 Mass Shift Systematic

The systematic error on the mass shift incorporates both the uncertainty which arises from the simulation statistics and the difference in shift between the nominal and 4% track loss branch. The statistical error was determined by adding in quadrature the errors from the two trigger samples weighted by the trigger fractions. In each bin, the maximum statistical error between the particle and parton levels was added in quadrature to the maximum nominal to 4% difference to get the total systematic.

### 5.8.2 Jet Energy Scale Systematic

The systematic error on the reconstructed di-jet mass comes solely from the jet energy scale uncertainty. Contributions from the  $\eta - \phi$  position uncertainties for individual jets are not considered. The determination of the jet energy scale uncertainty follows the procedure used in the Run 9 inclusive jet analysis and contains contributions from the track BEMC scale uncertainty, the track  $p_T$  uncertainty, and the response of the calorimeter to hadrons.

As mentioned in section 1.5, the BEMC scale uncertainty was estimated to be 3.7% and a BEMC efficiency uncertainty, due mostly to uncertainty in the status tables, was taken to be 1%. These uncertainties are added in quadrature and scaled by the average neutral fraction of all jets in the di-jet mass bin.

The average BEMC response to hadrons was taken to be 30% and the track response uncertainty was taken to be 9% (both taken from Fig. 4 in [26]). The tracking efficiency was estimated to be 81%. The fraction of the EMC response in the projected tower was estimated as 50%. Finally, a scale factor from charged to total hadrons of  $1/0.86$  (see [26]) was used. Using these values, the BEMC hadron response uncertainty was calculated as  $(1/0.86 - 0.5 \times 0.81)/0.81.3 \times 0.09 = 0.025$ . This value was scaled by the average charged fraction of all jets in the di-jet mass bin.

The total jet energy scale systematic is then taken as the quadrature sum of the BEMC scale uncertainty which has been scaled by the neutral fraction, the track  $p_T$  uncertainty,

and the hadron response uncertainty, both of which have been scaled by the average charged fraction.

### 5.8.3 Pythia to NLO Systematic

The final contribution to the systematic uncertainty on the x-position of the  $A_{LL}$  points is the mass shift needed to bring the NLO and parton level PYTHIA cross sections into agreement. This systematic is only found for the parton level. The systematic is determined as follows:

1. Calculate the NLO di-jet cross section in 1 GeV bins using the CTEQ6M PDF set and the di-jet theory code. Extract the PYTHIA di-jet cross section also in 1 GeV bins at parton level from the full simulation sample.
2. Minimize statistical bin-to-bin fluctuations by smoothing using a moving sum technique where the smoothed value is given by:  $\sigma_i^{\text{smooth}} = (\sigma_{i-1} + \sigma_i + \sigma_{i+1})/3$
3. Using exponential extrapolation, find the NLO and PYTHIA cross sections at the average parton level masses (with mass shift applied) in each bin.
4. Determine the change in mass needed to make the NLO cross section at a given mass equal the PYTHIA cross section at that mass.

The change in mass needed to bring the cross sections into agreement is taken as the systematic. The systematic is determined for the combined, EEWW, and EW samples individually. Figure 5.11 shows the PYTHIA to NLO cross section ratio.

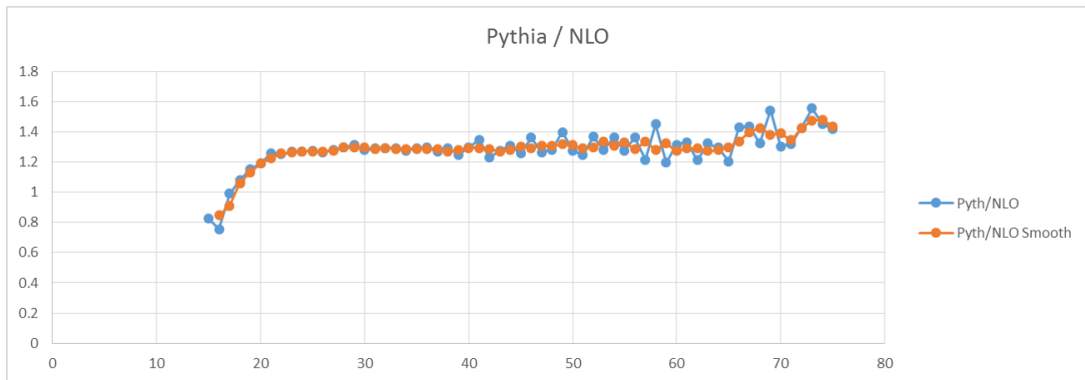


Figure 5.11: Ratio of the di-jet Pythia cross section at parton level to the NLO cross section as a function of invariant mass for the full topological sample. The orange curve is just the ratio after both cross sections had the smoothing procedure applied.

### 5.8.4 Trigger and Reconstruction Bias Systematic

The systematic error on the trigger and reconstruction bias correction includes contributions from both the observed range of  $\Delta A_{LL}$  values and the finite simulation statistics. At both the particle and parton levels, half the difference between the largest and smallest  $\Delta A_{LL}$  value was found and the largest of these was added in quadrature to the largest of the statistical errors from the three  $\Delta A_{LL}$  points from either the particle or parton levels to obtain the final systematic. This procedure was carried out for the EW, EEWW, and combined topology samples independently.

### 5.8.5 Residual Transverse Component Systematic

The determination of the systematic error on  $A_{LL}$  due to residual transverse beam components follows the procedure used in the inclusive jet paper [ ] which in turn followed that from the Run 6 paper. The systematic is evaluated as:

$$\delta A_{LL} = |\tan \theta_Y \tan \theta_B \cos(\phi_Y - \phi_B) A_\Sigma| \quad (5.7)$$

where  $\theta$  and  $\phi$  are the polar and azimuthal angles of the polarization vectors, respectively. The  $\tan \theta$  terms were calculated for the periods before and after the rotator tuning and then waited by the fraction of the events collected during that period. The  $\cos(\phi_Y - \phi_B)$  term was conservatively set to 1. Because there was no transverse running during Run 9, the  $A_\Sigma$  values used were taken from Run 6. These values were all consistent with zero, so to be conservative, the statistical error on each  $A_\Sigma$  measurement was taken as the value used in the calculation of the systematic. The  $A_\Sigma$  were measured as a function of jet  $p_T$  so for a given di-jet mass point, the  $p_T$  value at which  $A_\Sigma$  was evaluated was half that mass.

### 5.8.6 Total Systematic Error

The above systematic errors were calculated for each mass bin and each of the three topological configurations separately. Tables 5.10, 5.11, and 5.12 summarize the systematic errors for the combined, EEWW, and EW topological samples, respectively.



Bin	$\delta\text{Mass}$				$\delta A_{LL}$		
	Mass Shift	JES	Pythia $\rightarrow$ NLO	Total	TRB	Trans Pol	Total
1	0.296	0.408	0.028	0.51	0.0003	0.00028	0.00041
2	0.318	0.479	0.526	0.78	0.0004	0.00028	0.00049
3	0.404	0.571	0.775	1.04	0.0004	0.00028	0.00049
4	0.468	0.685	0.976	1.28	0.0007	0.00032	0.00077
5	0.528	0.823	1.024	1.42	0.0009	0.00065	0.00111
6	0.610	1.031	1.239	1.72	0.0013	0.00132	0.00185
7	0.654	1.432	2.022	2.56	0.0022	0.00253	0.00335
8	0.594	1.978	-	-	0.0058	0.00405	0.00707

Table 5.10: The systematic errors on the di-jet mass and  $A_{LL}$  for the combined topological sample.

Bin	$\delta\text{Mass}$				$\delta A_{LL}$		
	Mass Shift	JES	Pythia $\rightarrow$ NLO	Total	TRB	Trans Pol	Total
1	0.306	0.407	0.235	0.56	0.0003	0.00028	0.00041
2	0.308	0.477	0.904	1.07	0.0004	0.00028	0.00049
3	0.396	0.569	1.141	1.33	0.0005	0.00028	0.00057
4	0.486	0.684	1.105	1.39	0.0009	0.00032	0.00096
5	0.530	0.822	1.497	1.79	0.0014	0.00064	0.00154
6	0.582	1.028	1.815	2.17	0.0020	0.00130	0.00239
7	0.573	1.426	2.045	2.56	0.0032	0.00251	0.00407
8	0.509	1.977	-	-	0.0133	0.00397	0.01388

Table 5.11: The systematic errors on the di-jet mass and  $A_{LL}$  for the East-East+West-West topological sample.

## 5.9 Results

The di-jet  $A_{LL}$  is presented as a function of invariant mass for two independent topological configurations, as well as the combination of the two. The result was obtained using 19.5 pb<sup>-1</sup> of polarized  $pp$  data taken at  $\sqrt{s} = 200$  GeV from Run IX. The anti- $k_t$  jet algorithm with a radius of 0.6 was used to reconstruct jets. The two topological configurations are East-West and East-East+West-West. The East-West configuration contains those di-

Bin	$\delta\text{Mass}$				$\delta A_{LL}$		
	Mass Shift	JES	Pythia $\rightarrow$ NLO	Total	TRB	Trans Pol	Total
1	0.298	0.411	0.192	0.54	0.0004	0.00028	0.00049
2	0.339	0.480	0.759	0.96	0.0005	0.00028	0.00057
3	0.418	0.572	1.117	1.32	0.0006	0.00028	0.00066
4	0.471	0.686	1.502	1.72	0.0008	0.00032	0.00086
5	0.535	0.824	1.388	1.70	0.0011	0.00065	0.00128
6	0.636	1.032	1.672	2.07	0.0017	0.00133	0.00216
7	0.718	1.435	2.943	3.35	0.0026	0.00255	0.00364
8	0.754	1.978	-	-	0.0061	0.00408	0.00734

Table 5.12: The systematic errors on the di-jet mass and  $A_{LL}$  for the East-West topological sample.

jets which have one jet in the East half of the BEMC and one jet in the West, while the East-East+West-West configuration contains those di-jets which have both jets in either the East or West halves of the BEMC. The full topological configuration is simply the combination of the two independent topological samples.

The  $A_{LL}$  points were plotted at the mass-weighted average position of each bin. A mass shift was then applied to these average positions, which was determined by taking the event-weighted average of the L2JetHigh and JP1 mass shifts listed in Table 5.8. The original average bin position and the mass shifts were calculated independently for the three topological samples. The  $A_{LL}$  results can be found in Tables 5.13, 5.14, and 5.15 for the full, East-East+West-West, and East-West topological configurations, respectively.

The  $A_{LL}$  results for the three topological groups are shown in Figure 5.12. Note that an overall scale uncertainty of 6.8% due to uncertainty in the beam polarization measurements is not shown. As an illustration of the size of  $A_{LL}$  versus the false asymmetries (and zero), figure 5.13 shows  $A_{LL}$  and the four false asymmetry values integrated over all mass bins. Note that no trigger and reconstruction bias correction is applied to the  $A_{LL}$  value.

Bin	Mass $\pm$ (Sys) [GeV/c <sup>2</sup> ]	$A_{LL} \pm$ (Stat) $\pm$ (Sys)
1	17.82 $\pm$ 0.51	0.0063 $\pm$ 0.0025 $\pm$ 0.0004
2	21.46 $\pm$ 0.78	0.0091 $\pm$ 0.0022 $\pm$ 0.0005
3	26.16 $\pm$ 1.04	0.0110 $\pm$ 0.0027 $\pm$ 0.0005
4	31.70 $\pm$ 1.28	0.0093 $\pm$ 0.0038 $\pm$ 0.0008
5	38.32 $\pm$ 1.42	0.0102 $\pm$ 0.0058 $\pm$ 0.0011
6	48.59 $\pm$ 1.72	0.0186 $\pm$ 0.0084 $\pm$ 0.0019
7	67.09 $\pm$ 2.56	0.0077 $\pm$ 0.0273 $\pm$ 0.0034

Table 5.13:  $A_{LL}$  results for the combined topological sample.

Bin	Mass $\pm$ (Sys) [GeV/c <sup>2</sup> ]	$A_{LL} \pm$ (Stat) $\pm$ (Sys)
1	17.70 $\pm$ 0.56	0.0067 $\pm$ 0.0034 $\pm$ 0.0004
2	21.34 $\pm$ 1.07	0.0088 $\pm$ 0.0032 $\pm$ 0.0005
3	26.02 $\pm$ 1.33	0.0162 $\pm$ 0.0039 $\pm$ 0.0006
4	31.66 $\pm$ 1.39	0.0024 $\pm$ 0.0056 $\pm$ 0.0010
5	38.25 $\pm$ 1.79	0.0130 $\pm$ 0.0089 $\pm$ 0.0015
6	48.28 $\pm$ 2.17	0.0336 $\pm$ 0.0133 $\pm$ 0.0024
7	66.65 $\pm$ 2.56	0.0755 $\pm$ 0.0460 $\pm$ 0.0041

Table 5.14:  $A_{LL}$  results for the East-East+West-West topological sample.

Bin	Mass $\pm$ (Sys) [GeV/c <sup>2</sup> ]	$A_{LL} \pm$ (Stat) $\pm$ (Sys)
1	17.99 $\pm$ 0.54	0.0059 $\pm$ 0.0039 $\pm$ 0.0005
2	21.58 $\pm$ 0.96	0.0096 $\pm$ 0.0032 $\pm$ 0.0006
3	26.29 $\pm$ 1.32	0.0068 $\pm$ 0.0037 $\pm$ 0.0007
4	31.72 $\pm$ 1.72	0.0151 $\pm$ 0.0050 $\pm$ 0.0009
5	38.38 $\pm$ 1.70	0.0083 $\pm$ 0.0077 $\pm$ 0.0013
6	48.79 $\pm$ 2.07	0.0092 $\pm$ 0.0109 $\pm$ 0.0022
7	67.32 $\pm$ 3.35	-0.0282 $\pm$ 0.0340 $\pm$ 0.0036

Table 5.15:  $A_{LL}$  results for the East-West topological sample.

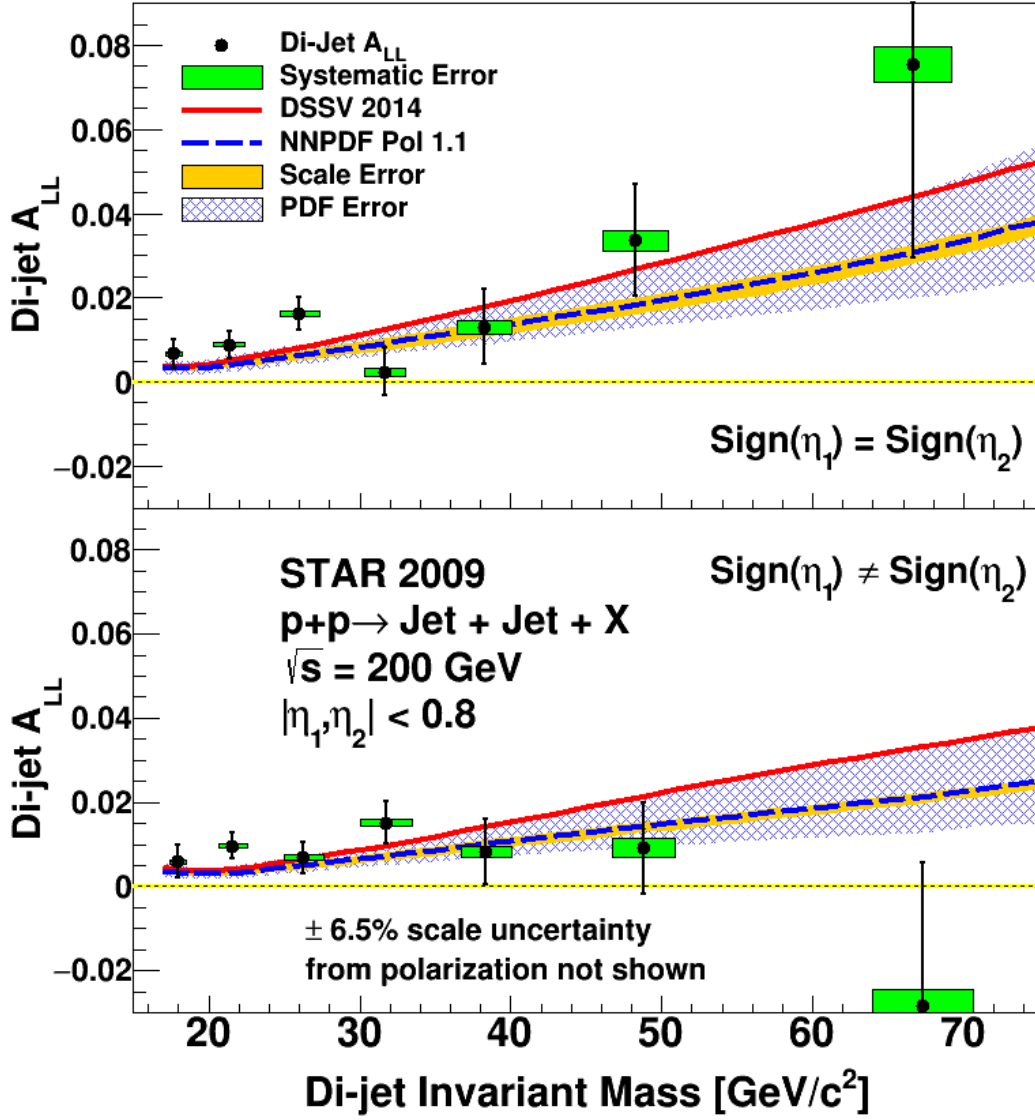


Figure 5.12: The di-jet  $A_{LL}$  for the EE+WW (top) and EW (bottom) topological samples.

### 5.9.1 Comparison to Theory

The asymmetry results are compared to theoretical curves generated using code originally developed by Frixione et al and modified by Dmitry Kalinkin as described in chapter 6. Two curves were produced using the DSSV 2014 and NNPDFpol1.1 polarized PDF sets. The unpolarized PDF sets used in the denominator of the asymmetry were MRST2008 and NNPDF2.3, respectively.

The  $\chi^2$  between the data and theory was calculated using the ‘CTEQ formula’ which accounts for the presence of correlated systematics (here I consider two sources of correlated

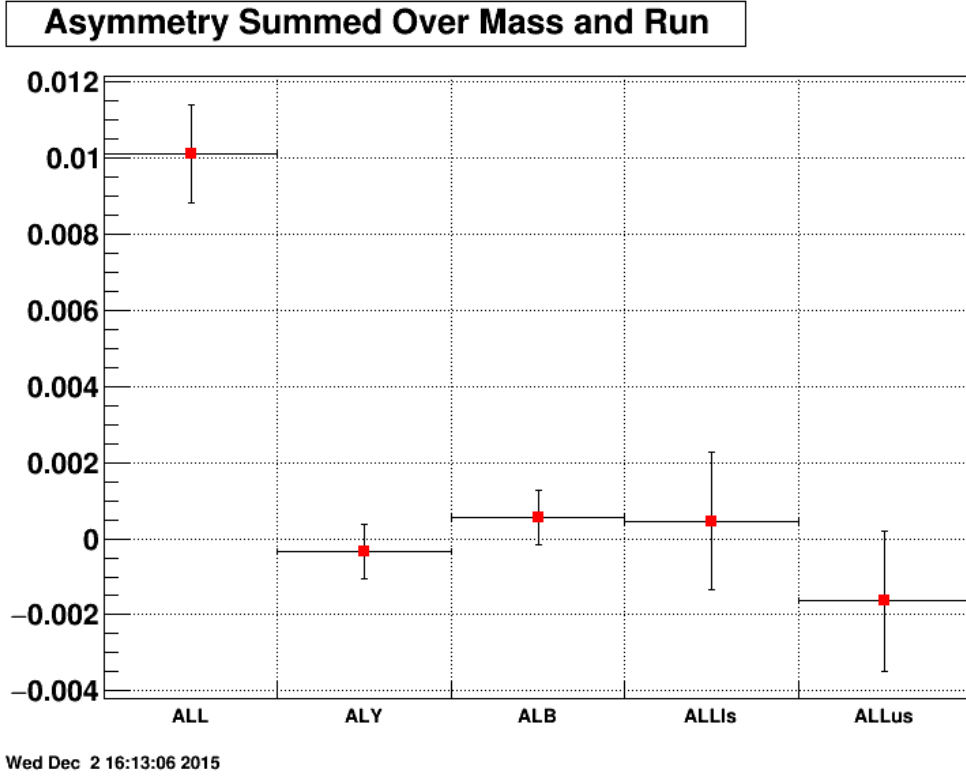


Figure 5.13: The di-jet  $A_{LL}$  and false asymmetries integrated over all mass bins.

error):

$$\chi^2 = r_1^2 + r_2^2 + \sum_{ij} \left[ \frac{D_i - \beta_{1i}r_1 - \beta_{2i}r_2 - T_i}{\sigma_i} M_{ij} \frac{D_j - \beta_{1j}r_1 - \beta_{2j}r_2 - T_j}{\sigma_j} \right] \quad (5.8)$$

where  $D_i$ ,  $T_i$ , and  $\sigma_i$  are the values of the data asymmetry, theoretical asymmetry, and the quadrature sum of the statistical and (uncorrelated) systematic error on the data asymmetry, respectively of the  $i^{th}$  point.  $M_{ij}$  is the covariance matrix. Here  $\beta_1 = \sigma_{RL}$  is the correlated error contribution from the relative luminosity and  $\beta_2 = \frac{\partial A_{LL}}{\partial M} \times \sigma_M$  is the correlated error contribution from the jet energy scale. The  $r_1$  and  $r_2$  factors are free parameters associated with the relative luminosity and jet energy scale, respectively. Optimizing for the values of  $r_1$  and  $r_2$  one can write the above equation as (here we assume the covariance matrix is the identity matrix):

$$\chi^2 = \sum_i \left( \frac{D_i - T_i}{\sigma_i} \right)^2 - B^T A^{-1} B \quad (5.9)$$

where the matrix  $A$  is defined as:

$$A = \begin{bmatrix} 1 + \sum_i \left( \frac{\beta_{1i}}{\sigma_i} \right)^2 & \sum_i \left( \frac{\beta_{1i}\beta_{2i}}{\sigma_i^2} \right) \\ \sum_i \left( \frac{\beta_{1i}\beta_{2i}}{\sigma_i^2} \right) & 1 + \sum_i \left( \frac{\beta_{2i}}{\sigma_i} \right)^2 \end{bmatrix} \quad (5.10)$$

and the column vector  $B$  is defined as:

$$B = \begin{bmatrix} \sum_i \left( \frac{D_i - T_i}{\sigma_i} \times \frac{\beta_{1i}}{\sigma_i} \right) \\ \sum_i \left( \frac{D_i - T_i}{\sigma_i} \times \frac{\beta_{2i}}{\sigma_i} \right) \end{bmatrix} \quad (5.11)$$

The values for  $\chi^2$ ,  $r_1$ , and  $r_2$  for both topologies and theory curves can be found in table 5.16.

Polarized PDF	Topology	$\chi^2$	$r_1$	$r_2$
DSSV 14	EE+WW	9.9	0.37	0.32
	E+W	9.2	0.29	0.06
NNPDFpol1.1	EE+WW	12.0	0.64	0.77
	E+W	8.8	0.50	0.35

Table 5.16:  $\chi^2$  and  $r$  values for data / theory comparison for both polarized models and topologies.

# Chapter 6

## Theory Code

The cross section and  $A_{LL}$  theory curves were obtained using code provided by the DSSV group. The code has been modified by Dmitry Kalinkin to increase ease of use, integrate LHAPDF, and slightly change the di-jet mass definition. This chapter documents these changes and provides instructions on running the code.

### 6.1 Calculation

The code implements a calculation of the jet cross sections in polarized proton-proton collisions. Based on the unpolarized case calculation by S. Frixione et al. [27][28][29] this has an additional option[30] to calculate the polarized cross section  $\sigma_{++} - \sigma_{+-}$  which is sensitive to the helicity of the initial state partons. The ratio of polarized and unpolarized cross sections gives the asymmetry:

$$A_{LL} \equiv \frac{\sigma_{++} - \sigma_{+-}}{\sigma_{++} + \sigma_{+-}}$$

The code uses VEGAS to integrate the cross section over the phase space, the calculation defines both inclusive jet and di-jet cross-sections at NLO.

Recent PDF fits[31][32] including RHIC data prefer higher values of the gluon polarization (Fig. 6.1), so it is natural to compare newest results from STAR jet analyses against the theory for these newer PDF sets. This document describes modifications made to the theory code that was used to produce the updated curves.

### 6.2 Modifications

The version used for curve production contains following modifications:

- *LHAPDF backend (SET=50)*: The original code supports only predefined list of PDF sets. Interfacing to LHAPDF allows to get results for a wide variety of modern PDF sets such as CT10, NNPDFpol1.1. This is also the first interface that allows to access PDF error vectors.
- *DSSV2014 backend (SET=51)*: DSSV2014 PDF set is not distributed for LHAPDF, so it was added as an additional Fortran driver.
- *Fix minor Fortran compilation warnings*: Previously the code would compile with warnings which was polluting the compiler output.
- *Separate parameters for the factorization and renormalization scales*: Previously the factorization and renormalization scale factors could only be set to a common value by a single parameter in the steering. New version provides a separate options for these two parameters to allow for more detailed evaluation.
- *Output weighted parton kinematic configurations as HepMC events to Rivet[33]*: Builtin histogramming package doesn't allow to book variable bin sized histograms, it also forces user to define cuts manually in Fortran. The new interface feeds weights into the Rivet analysis that user have to define in C++. The advantages is that Rivet provides flexible library to define custom jet and custom observable definitions, flexible histogramming, and tools for merging results from parallel jobs. Rivet analysis can later be reused without modifications with major MC's such as Sherpa, Herwig and Pythia. Rivet also provides tools to compare and merge different runs.
- *Disable VEGAS grid optimisation*: The original code runs multiple VEGAS calls where phase space sampling grid is adjusted to minimize numerical  $\chi^2$  of the total cross section of the given order. In practice having to define the number of such iterations in the steering is confusing for the user and technique doesn't seem to significantly contribute to the statistical precision per CPU time. It is also not clear whether optimising grid for the total cross section value is beneficial for the uncertainties on the higher  $p_T$  bins. The number of VEGAS calls was hardcoded to 1 and the user now only sets the number of iterations.
- *Accept input parameters as command arguments rather than from stdio*: The original code would take its arguments from the standard input which made it cumbersome to experiment with parameters, run the code in HTCondor jobs and to add new options into the code.



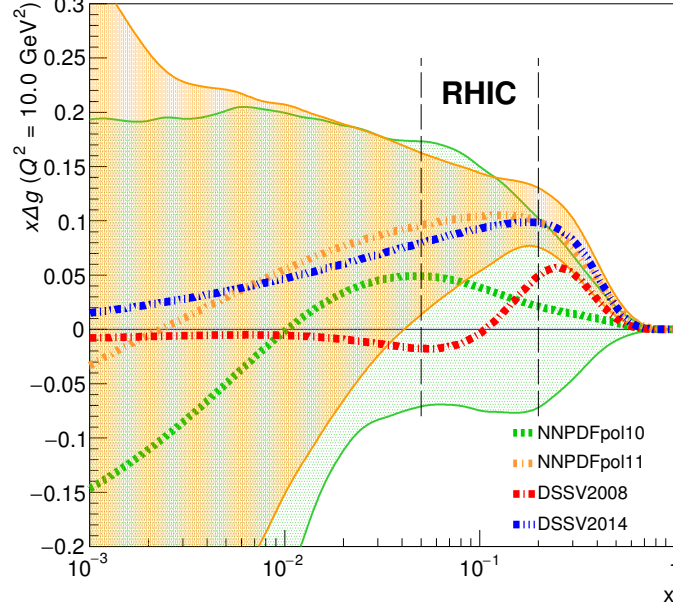


Figure 6.1: The gluon component of the polarized PDF sets fit with RHIC data (NNPDFpol1.1, DSSV2014) and without (NNPDFpol1.0, DSSV2008).

- *Compilation with CMake build system:* The original Fortran code was compiled with a single `gfortran` call. With addition of C++ bits it is more natural to have some sort of Makefile to keep the build process simple.
- *Change of the di-jet mass definition:* Original code uses an unusual definition for the di-jet mass. It was replaced with a more conventional one (see Sec. 6.4.2)

Version with only the first three modifications exists in the git repository under *release2* tag.

## 6.3 Results

Results were produced using the described code linked to Rivet 2.4.0 with Rivet analysis for run 9 200 GeV di-jets implemented in [34]. The result for unpolarized cross-section produced by the code was compared against a very similar calculation implemented by NLOJet++ code[35], which showed an agreement to a sub-percent level (Fig. 6.2). The same result in comparison to the STAR data is given on Fig. 6.3

Predictions for  $A_{LL}$  shown on Fig. 6.4 were calculated using pairs of polarized and unpolarized PDF sets in accordance with Table 6.5. Scale uncertainties were obtained as envelope of cross sections at six combination of renormalization and factorization scale

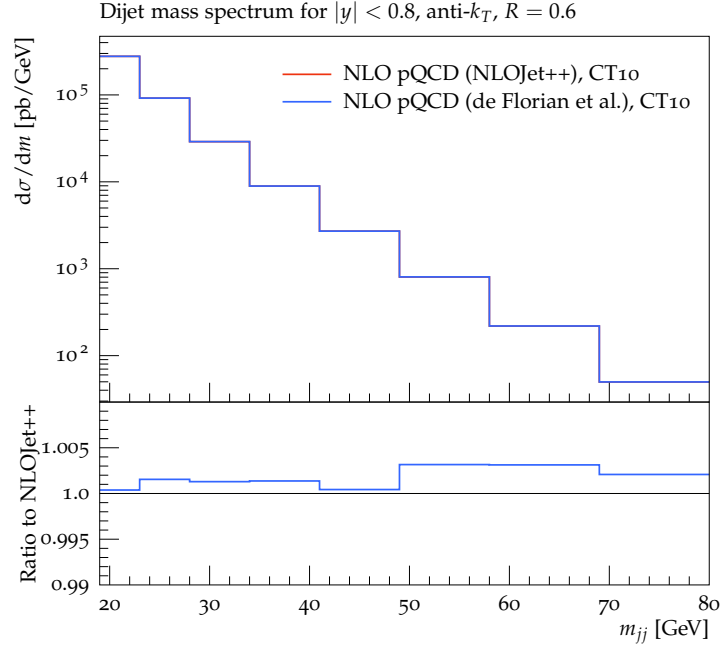


Figure 6.2: Comparison of the two pQCD calculations of the unpolarized di-jet cross section at  $\sqrt{s} = 200$  GeV for CT10 PDF set.

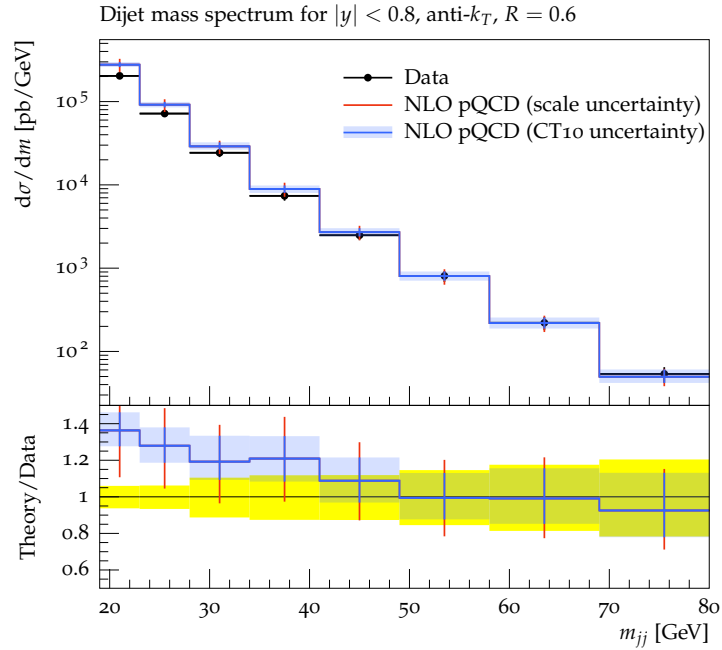


Figure 6.3: Theory prediction for unpolarized di-jet cross section for CT10 PDF set compared to the STAR data corrected to the parton level at  $\sqrt{s} = 200$  GeV.

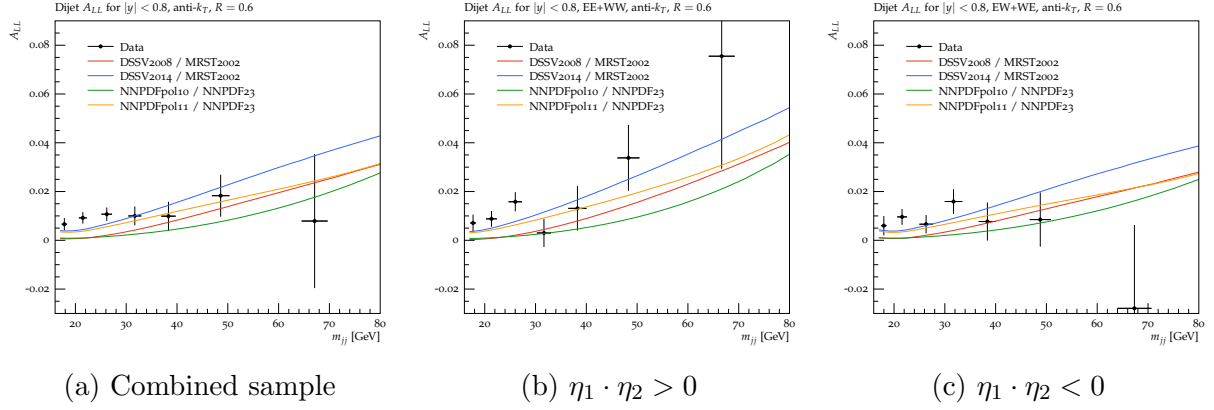


Figure 6.4: Di-jet  $A_{LL}$  theory predictions against the STAR data at  $\sqrt{s} = 200$  GeV. This plot needs an update to show DSSV2014/MRST2008 (small difference).

Unpolarized PDF set	Polarized PDF set
DSSV2008	MRST2002
DSSV2014	MRST2008
NNPDFpol1.0	NNPDF2.3
NNPDFpol1.1	NNPDF2.3

Figure 6.5: Correspondence between the used PDF sets

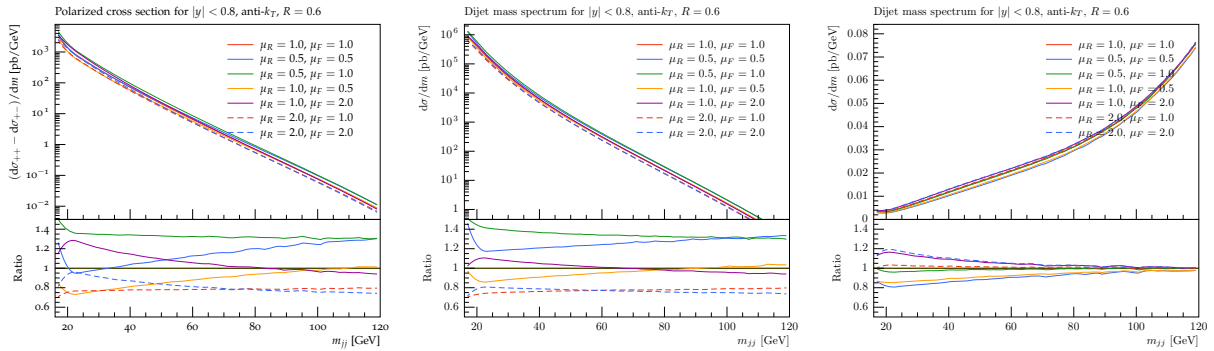


Figure 6.6: Polarized, unpolarized di-jet cross sections and di-jet  $A_{LL}$  for NNPDFpol1.1/NNPDF2.3 for various renormalization and factorization scale factors at  $\sqrt{s} = 200$  GeV.

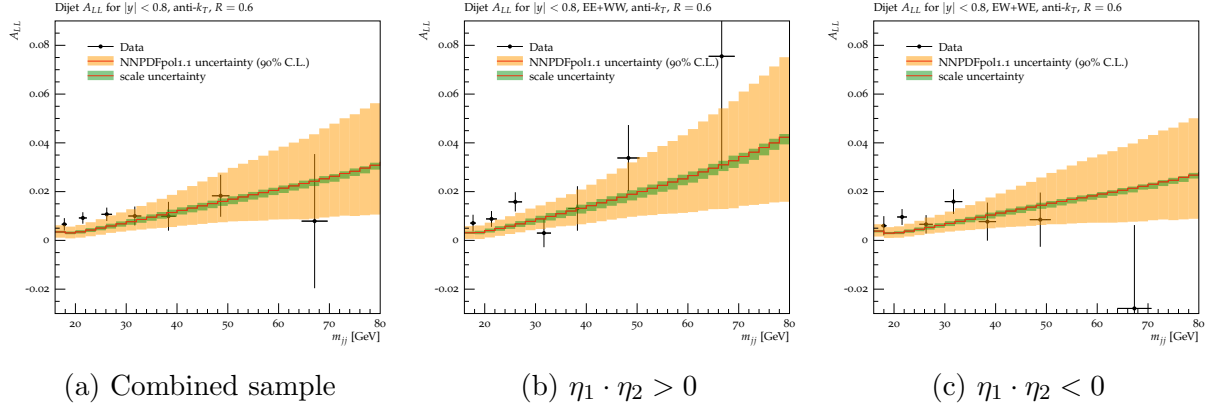


Figure 6.7: Di-jet  $A_{LL}$  theory predictions for NNPDFpol1.1 with PDF and scale uncertainties against STAR data at  $\sqrt{s} = 200$  GeV.

factors (Fig. 6.6):

$$\mu_R, \mu_F = (1.0, 2.0), (1.0, 0.5), (2.0, 1.0), (0.5, 1.0), (2.0, 2.0), (0.5, 0.5)$$

PDF uncertainties were obtained using a standard "master" formula:

$$\delta_+ \sigma = \sqrt{\sum_k [\max \{ \sigma(S_0) - \sigma(S^{+(k)}), \sigma(S_0) - \sigma(S^{-(k)}), 0 \}]^2}$$

$$\delta_- \sigma = \sqrt{\sum_k [\min \{ \sigma(S_0) - \sigma(S^{+(k)}), \sigma(S_0) - \sigma(S^{-(k)}), 0 \}]^2}$$

where  $\sigma$  is the observable (in this case the cross section in the bin),  $S_0$  is the central PDF set and  $S^{\pm(k)}$  is the  $k$ -th odd(even) error vector. Only uncertainty of the polarized PDF went into the calculation for  $A_{LL}$  uncertainty as it is the one that dominates and is of interest to us. Results for uncertainties for NNPDFpol1.1 are presented on Fig. 6.7.

## 6.4 Di-jet mass definition

### 6.4.1 Jet merging scheme

The scheme used to merge partons into a jet is defined in the original user.f in the xmerge2 subroutine:

```
subroutine xmerge2
#      (xkt1,xkt2,xkt3,xeta1,xeta2,xeta3,xphi1,xphi2,xphi3)
```

c

```

c Defines the jet variables (xkt3,xeta3,xphi3) in terms of the
c partonic variables when two partons are merged
c Recombination in E-scheme for merging
c xeta3= true rapidity (not pseudorapidity)
c
      implicit real * 8 (a-h,o-z)
      parameter (pi=3.14159265358979312D0)
c
c energy and momentum of parton 1
      c1 = (exp(2*xeta1)-1)/(exp(2*xeta1)+1)
      e1 = xkt1/(1-c1**2)**0.5
      px1 = xkt1 * Cos(xphi1)
      py1 = xkt1 * sin(xphi1)
      pz1 = e1 * c1
c energy and momentum of parton 2
      c2 = (exp(2*xeta2)-1)/(exp(2*xeta2)+1)
      e2 = xkt2/(1-c2**2)**0.5
      px2 = xkt2 * Cos(xphi2)
      py2 = xkt2 * sin(xphi2)
      pz2 = e2 * c2
c energy and momentum of parton 1+2
      e3 = e1 + e2
      px3 = px1 + px2
      py3 = py1 + py2
      pz3 = pz1 + pz2
c phi,eta and Et for merged jet
      xphi3 = atan2(py3,px3)
      xeta3 = 0.5d0 * log( (e3+pz3)/(e3-pz3)) !y = rapidity
      pmod=(pz3**2 + px3**2 + py3**2)**0.5
c      xeta3= 0.5d0 * log( (pmod+pz3)/(pmod-pz3)) ! eta pseudorapidity
      xkt3=(px3**2+py3**2)**0.5
c
      return
end

```

The comment claims that E-scheme merge is performed. E-scheme by definitions does a simple four vector sum[14]. Adding two on-shell four vectors doesn't generally give on-shell four vector, yet the result of the merge is stored in just three variables.

## 6.4.2 Di-jet mass

The definition of dijet mass in the original code can be found in user.f:

C Define kinematical variables

```
cs12=cos(zphi(n1)-zphi(n2))
dphi=dacos(cs12)
pl1=zkt(n1)*sinh(zeta(n1))
pl2=zkt(n2)*sinh(zeta(n2))
e1=zkt(n1)*cosh(zeta(n1))
e2=zkt(n2)*cosh(zeta(n2))
```

c invariant mass

```
xm2=2*(e1*e2-pl1*pl2-zkt(n1)*zkt(n2)*cs12)
xmjj=sqrt(xm2)
```

This dijet mass is used as renormalization and factorization scales as well as for mass in the cross section plot. One should remember that the `zeta(i)` is not pseudorapidity  $\eta$  but rapidity  $y$ . To evaluate the expression for `xm2` we use following expressions:

$$\eta \equiv \frac{1}{2} \log \left( \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right), \quad \cosh(\eta) = \frac{|\vec{p}|}{k_T}, \quad \sinh(\eta) = \frac{p_z}{k_T}$$

$$y \equiv \frac{1}{2} \log \left( \frac{E + p_z}{E - p_z} \right), \quad \cosh(y) = \frac{E}{\sqrt{m^2 + k_T^2}}, \quad \sinh(y) = \frac{p_z}{\sqrt{m^2 + k_T^2}}$$

So the code defines:

$$\begin{aligned} xm2 &= 2k_{T1}k_{T2}(\cosh(y_1)\cosh(y_2) - \sinh(y_1)\sinh(y_2) - \cos(\Delta\varphi)) = \\ &= 2(E_1E_2 - p_{z1}p_{z2}) \frac{k_{T1}k_{T2}}{\sqrt{m_1^2 + k_{T1}^2}\sqrt{m_2^2 + k_{T2}^2}} - 2k_{T1}k_{T2}\cos(\Delta\varphi) = \\ &= 2(E_1E_2 - p_{z1}p_{z2}) \left( 1 - \frac{1}{2} \frac{m_1^2}{k_{T1}^2} + \mathcal{O}\left(\frac{m_1^4}{k_{T1}^4}\right) \right) \left( 1 - \frac{1}{2} \frac{m_2^2}{k_{T2}^2} + \mathcal{O}\left(\frac{m_2^4}{k_{T2}^4}\right) \right) - 2k_{T1}k_{T2}\cos(\Delta\varphi) = \\ &= -(E_1E_2 - p_{z1}p_{z2}) \left( \frac{m_1^2}{k_{T1}^2} + \frac{m_2^2}{k_{T2}^2} + \mathcal{O}\left(\frac{m^4}{k_T^4}\right) \right) + 2(E_1E_2 - p_{z1}p_{z2} - k_{T1}k_{T2}\cos(\Delta\varphi)) = \\ &= -\cosh(\eta_1 - \eta_2) \left( m_1^2 \frac{k_{T2}}{k_{T1}} + m_2^2 \frac{k_{T1}}{k_{T2}} \right) + \mathcal{O}\left(\frac{m^4}{k_T^4}\right) + 2(p_1)_\mu(p_2)^\mu \end{aligned}$$

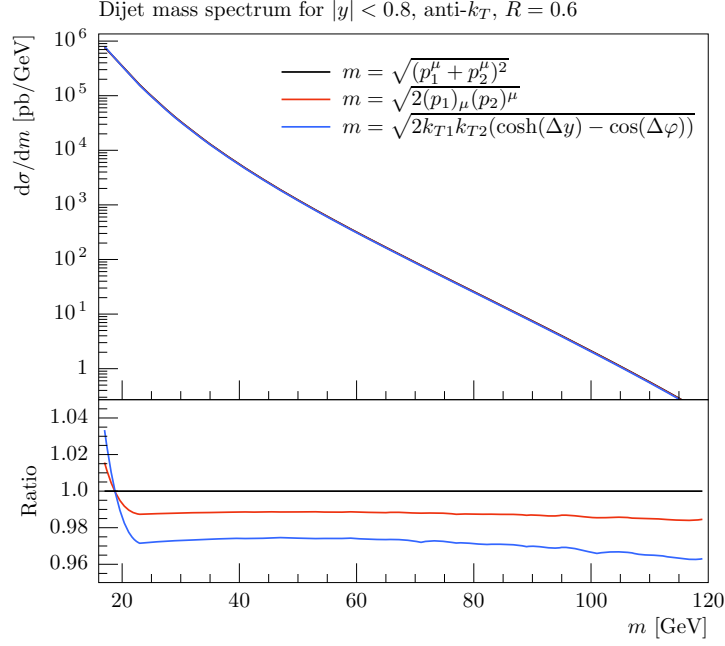


Figure 6.8: NLO pQCD theory prediction for unpolarized di-jet cross section spectrum for different definitions of mass (x axis). The renormalization and factorization scales are taken to be  $\mu_R = \mu_F = \sqrt{(p_1^\nu + p_2^\nu)^2}$  for all three curves.

Impact of this definition in comparison to standard invariant mass definition  $m = \sqrt{(p_1^\nu + p_2^\nu)^2}$  and a “reduced”  $m = \sqrt{2(p_1)_\nu(p_2)^\nu}$  definition. The code was modified to use invariant mass definition for the renormalization and factorization scales. Mass variable of the plot is defined by user in Rivet analysis, for the purposes of this study was also set according to the same standard invariant mass definition.

## 6.5 Di-jet definition cut ordering

One can come up with two possible ways to select a di-jet event:

### Definition 1

1. Find all jets in the region of interest  $|\eta| < 0.8$
2. Find two  $p_T$  leading jets
3. Apply cuts on  $p_T$  and  $\Delta\varphi$

### Defintion 2

1. Find all jets in the full phase space

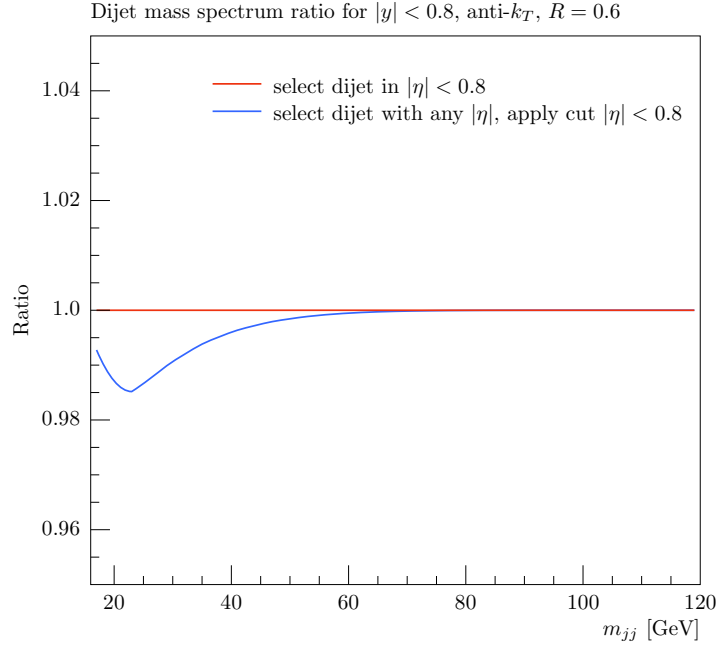


Figure 6.9: Comparison of NLO pQCD theory predictions for different di-jet definition cuts.

2. Find two  $p_T$  leading jets
3. Apply  $|\eta| < 0.8$  cut on both jets as well as cuts on  $p_T$  and  $\Delta\varphi$

Relative effect of utilizing the Definition 2 in NLO calculation is presented on Fig. 6.9.

## 6.6 Using the code

This section describes usage of Rivet version of the code. This assumes that you are running on RACF in STAR environment (i.e. with csh shell)

### 6.6.1 Installation

```
# Obtaining the code
git clone /star/u/veprbl/public/poljet/pol-jet-code-new.git
cd pol-jet-code-new
git checkout rivet-release1
# Compilation
mkdir build
cd build
```



```

sh -c "CMAKE_PREFIX_PATH=/star/u/veprbl/public/poljet/rivet cmake28 .."
make
# Obtain a sample Rivet analysis
cp /star/u/veprbl/public/poljet/STAR_DIJET_EXAMPLE.cc .
# Set proper environment variables
source /star/u/veprbl/public/poljet/rivet/rivetenv.csh
setenv RIVET_ANALYSIS_PATH `pwd`
# Compile the analysis
rivet-buildplugin STAR_DIJET_EXAMPLE.cc
# Run
./r.e -a STAR_DIJET_EXAMPLE --sqrts=500 --pdf-backend=50 --lhpdf-pdfname=CT10nlo --p
# Plot the histograms
setenv PATH /afs/cern.ch/sw/XML/texlive/latest/bin/x86_64-linux/:$PATH
rivet-mkhtml Rivet.yoda

```

A following warning may appear when you are running the calculation:

```
Rivet.Analysis.Handler: WARN Analysis 'STAR_DIJET_EXAMPLE' is unvalidated: be carefu
```

You can safely ignore it. However if you see

```
Rivet.Analysis.Handler: WARN Analysis 'STAR_DIJET_EXAMPLE' not found
```

that means that Rivet couldn't find your analysis and will not generate any useful histograms. You should then check `RIVET_ANALYSIS_PATH` environment variable is set to point to the directory with your analysis.

If you want to submit parallel jobs you can utilize `genjob.py` script that provides basic support for generation of the HTCondor job files. YODA histogramming package has a script `yodamerge` that can be used for merging the different runs.

## 6.6.2 Old version

Version with LHAPDF support only is available with instruction at:

<https://drupal.star.bnl.gov/STAR/blog/veprbl/nlo-dijet-code-installation-guide>

# Appendix A

## List of Runs and Fills

### A.1 Cross Section Runs

This is the list of runs used in the cross section analysis.

10120063	10122007	10122067	10124050	10126003	10128043	10128098	10134024
10120065	10122010	10122071	10124053	10126004	10128046	10128099	10134025
10120078	10122013	10122086	10124110	10126005	10128047	10128100	10134026
10120079	10122014	10122087	10124111	10126012	10128048	10128101	10134027
10120082	10122015	10122095	10124113	10126017	10128049	10128102	10134028
10120085	10122016	10122099	10125001	10126018	10128050	10129003	10134030
10120086	10122017	10123004	10125008	10126019	10128052	10129005	10134035
10120093	10122019	10123007	10125009	10126024	10128053	10129006	10134036
10120097	10122022	10123010	10125010	10126025	10128054	10129007	10134037
10120100	10122023	10123086	10125014	10126026	10128055	10129008	10134044
10121001	10122024	10123087	10125015	10126083	10128056	10129011	10134085
10121017	10122047	10123090	10125016	10126084	10128059	10131039	10134086
10121020	10122048	10124013	10125017	10126087	10128060	10131040	10134101
10121022	10122049	10124014	10125022	10126088	10128061	10131041	10134102
10121029	10122050	10124037	10125023	10126089	10128063	10131042	10134103
10121039	10122054	10124038	10125075	10126090	10128065	10131043	10135001
10121040	10122055	10124044	10125076	10127008	10128066	10131045	10135002
10121043	10122060	10124045	10125080	10127009	10128070	10131047	10135005
10121044	10122061	10124046	10125083	10127011	10128072	10131052	10135006
10122006	10122065	10124049	10125091	10128041	10128094	10134021	10135007

10135008	10136037	10139007	10143014	10143106	10145030	10148026	10151002
10135009	10136061	10139008	10143015	10144001	10145032	10148027	10151003
10135011	10136063	10139009	10143018	10144002	10145034	10148028	10151004
10135016	10136069	10139010	10143023	10144003	10145036	10148033	10151005
10135017	10136070	10139014	10143025	10144022	10145038	10148034	10151006
10135018	10136071	10139015	10143026	10144026	10145042	10148035	10151034
10135058	10136073	10139017	10143027	10144027	10145046	10149008	10151035
10135059	10136074	10139018	10143029	10144028	10145047	10149012	10151039
10135063	10136077	10141018	10143043	10144029	10145070	10149023	10151040
10135064	10136078	10141019	10143044	10144030	10145072	10149024	10151041
10135065	10136079	10141020	10143045	10144034	10145073	10149025	10151042
10135066	10136092	10141023	10143047	10144035	10145076	10149026	10151043
10135070	10136096	10141025	10143051	10144036	10145078	10149028	10151044
10135072	10136097	10141026	10143052	10144037	10145079	10149031	10151045
10135076	10136099	10141027	10143053	10144044	10145081	10149032	10151047
10135077	10136100	10141030	10143054	10144045	10145082	10149033	10152001
10135081	10137003	10141031	10143058	10144046	10146040	10149034	10152004
10135082	10137004	10141032	10143062	10144072	10146047	10149035	10152005
10135083	10137006	10142029	10143063	10144074	10146048	10150005	10152006
10136001	10137008	10142031	10143064	10144075	10146049	10150008	10152007
10136011	10137045	10142034	10143065	10144076	10146050	10150009	10152008
10136012	10137046	10142035	10143076	10144083	10146051	10150010	10152009
10136017	10137048	10142036	10143077	10144085	10146052	10150011	10152010
10136019	10137049	10142041	10143078	10144087	10146054	10150012	10154060
10136020	10138047	10142042	10143082	10144090	10146055	10150013	10154061
10136021	10138049	10142043	10143083	10144091	10146073	10150018	10154064
10136024	10138052	10142044	10143086	10144092	10146084	10150021	10154065
10136025	10138053	10142047	10143090	10144093	10146086	10150022	10154066
10136026	10138054	10142050	10143092	10144098	10146087	10150024	10154067
10136027	10138055	10142056	10143095	10145011	10146091	10150025	10154083
10136028	10138098	10142057	10143098	10145012	10148002	10150052	10155001
10136030	10138099	10142058	10143099	10145013	10148005	10150053	10155010
10136031	10138100	10143007	10143102	10145016	10148006	10150056	10155014
10136035	10139002	10143008	10143103	10145018	10148021	10150057	10155015
10136036	10139003	10143009	10143104	10145027	10148025	10151001	10155016

10155019	10158010	10159049	10162029	10167007	10169047	10171011	10173007
10155020	10158013	10160005	10162030	10167008	10169048	10171014	10173008
10155022	10158014	10160006	10162031	10167009	10169049	10171015	10173009
10156031	10158015	10160009	10162032	10167012	10169065	10171016	10173012
10156034	10158016	10160010	10162033	10167013	10169070	10171019	10173013
10156037	10158017	10160011	10162034	10167014	10169074	10171021	10173015
10156038	10158018	10160012	10162035	10167015	10169075	10171022	10173016
10156039	10158021	10160013	10162036	10167016	10169076	10171025	10173017
10156040	10158042	10160014	10162037	10167017	10169077	10171034	10173031
10156058	10158043	10160016	10162038	10167020	10169078	10171036	10173032
10156086	10158047	10160017	10162040	10167048	10169080	10171041	10173033
10156087	10158048	10160071	10163048	10167049	10170003	10171042	10173039
10156090	10158049	10160072	10163051	10167050	10170011	10171043	10173048
10156092	10158050	10160075	10163052	10167053	10170012	10171044	10173051
10156093	10158051	10160077	10163053	10167054	10170013	10171045	10173053
10156095	10158054	10160078	10163054	10167056	10170016	10171048	10173055
10156096	10158055	10160079	10163055	10167057	10170017	10171060	10174012
10157001	10158074	10160081	10163056	10167058	10170018	10171061	10174013
10157004	10158075	10160084	10163059	10167059	10170019	10171069	10174016
10157005	10158076	10161005	10164002	10169005	10170023	10171070	10174023
10157015	10158079	10161006	10164009	10169006	10170024	10171071	10174025
10157016	10158080	10161010	10164010	10169009	10170025	10171078	10174026
10157019	10158082	10161011	10164011	10169010	10170026	10172001	10174027
10157020	10158083	10161014	10164013	10169012	10170029	10172002	10174028
10157021	10158086	10161015	10164016	10169013	10170045	10172003	10174031
10157022	10158087	10161016	10164017	10169014	10170046	10172007	10174044
10157023	10158089	10161019	10164018	10169021	10170047	10172060	10174045
10157027	10158090	10161020	10164025	10169030	10170050	10172061	10174048
10157051	10159006	10161021	10164026	10169031	10170052	10172064	10174049
10157052	10159039	10161025	10164029	10169032	10170053	10172082	10174050
10157056	10159040	10161026	10164030	10169033	10170054	10172083	10174051
10157057	10159044	10161027	10164031	10169041	10170060	10172085	10174052
10157058	10159045	10161030	10164034	10169042	10170061	10172089	10174094
10158001	10159046	10162024	10166061	10169043	10170064	10172090	10175005
10158004	10159048	10162025	10166067	10169044	10170065	10172094	10175008

10175009	10175038	10176022	10178036	10179018	10179043	10179097	10180027
10175010	10176001	10176025	10179005	10179019	10179044	10179098	10180028
10175011	10176008	10176028	10179006	10179022	10179045	10180003	10180029
10175012	10176016	10178022	10179007	10179031	10179085	10180004	10180030
10175013	10176017	10178023	10179008	10179032	10179086	10180007	
10175014	10176018	10178026	10179009	10179033	10179088	10180021	
10175019	10176020	10178029	10179010	10179042	10179096	10180022	

## A.2 $A_{LL}$ Runs

The additional runs that were included for the  $A_{LL}$  analysis are listed here.

10124062	10131075	10138011	10139040	10140007	10156004	10164079	10170081
10124066	10131076	10138012	10139044	10140010	10156007	10164082	10170089
10124071	10135030	10138013	10139067	10140011	10156008	10165007	10171037
10124072	10137051	10138014	10139068	10141010	10156009	10165008	10172077
10124075	10137052	10138016	10139069	10141013	10156011	10165015	10172079
10124076	10137055	10138021	10139070	10142086	10156013	10165016	10177055
10129048	10137059	10138022	10139073	10142093	10162006	10165017	10177056
10129050	10137060	10138023	10139074	10142098	10162007	10165018	10177057
10130011	10137061	10138024	10139076	10143085	10162010	10165019	10177060
10130012	10137063	10138025	10139077	10149087	10164057	10165022	10177061
10130014	10137064	10138026	10139102	10150001	10164060	10165023	
10130015	10137065	10138027	10139107	10151046	10164061	10165024	
10131009	10137066	10138030	10140002	10154068	10164062	10165027	
10131012	10137067	10138032	10140005	10155095	10164067	10170075	
10131029	10138008	10139038	10140006	10155097	10164078	10170078	

## A.3 Cross Section Fills

The runs used in the cross section analysis come from the following fills.

10682	10684	10688	10690	10696	10703	10706	10712
10683	10685	10689	10695	10700	10704	10708	10713

10729	10763	10789	10854	10880	10932	10960	10986
10746	10773	10790	10855	10881	10935	10961	10987
10748	10777	10791	10866	10884	10937	10963	11001
10753	10781	10800	10869	10889	10951	10964	11002
10754	10782	10806	10870	10890	10952	10967	11003
10755	10783	10814	10875	10904	10953	10968	11005
10756	10784	10820	10876	10919	10954	10970	11006
10758	10785	10825	10877	10920	10955	10971	
10761	10786	10826	10878	10921	10959	10973	

## A.4 $A_{LL}$ Fills

The additional runs used in the  $A_{LL}$  analysis come from the following unique fills.

10717	10730	10764	10813	10924	10957
10720	10749	10765	10864	10925	10998
10727	10759	10778	10902	10956	

# Appendix B

## $A_{LL}$ and Mass Formula Derivation

The first section of this appendix details the derivation of the  $A_{LL}$  formula presented in equation 5.2 which is appropriate for the combination of many individual measurements. The second section outlines the derivation of the statistical error formula for  $A_{LL}$ . In both cases, the assumptions inherent in the formulas will be made explicit. The derivation of the  $A_{LL}$  formula is closely based on arguments laid out by Scott Wissink in private communication and the derivation of the  $A_{LL}$  error formula is based on standard error propagation techniques which can be found in most books on statistical methods, such as [36].

### B.1 Asymmetry Formula Derivation

The final  $A_{LL}$  values shown in this dissertation are the result of the combination of many individual measurements (each run is considered an independent measurement). The asymmetry for an individual measurement can be expressed schematically as:

$$A_i = \frac{1}{P_i} \frac{N_i^+ - R_i N_i^+}{N_i^+ + R_i N_i^+} \quad (\text{B.1})$$

where  $P$ ,  $N^{+/-}$ , and  $R$  are the appropriate polarizations, spin-sorted yields, and relative luminosity factors, respectively. The best value for the asymmetry (denoted  $\bar{A}$ ) based on multiple measurements of  $A$  can be found using the maximum likelihood method, which requires the minimization of the function:

$$\chi^2 \equiv \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{\sigma_i} \right)^2 \quad (\text{B.2})$$

with respect to  $\bar{x}$ . Here,  $\sigma_i$  is the error on an individual measurement  $x_i$ . Doing the

minimization yields the following expression for  $\bar{x}$ :

$$\bar{x} = \frac{\sum_i (x_i / \sigma_i^2)}{\sum_i (1 / \sigma_i^2)} \quad (\text{B.3})$$

Plugging equation B.1 into B.3 gives an expression for the best value of the asymmetry using a number of independent measurements:

$$\bar{A} = \frac{\sum_i (A_i / \sigma_{A_i}^2)}{\sum_i (1 / \sigma_{A_i}^2)} \quad (\text{B.4})$$

The exact uncertainty on  $A_i$  is a messy function and finding the uncertainty can be difficult when the yields are small. Things can be simplified immensely if it is assumed that the fractional errors on the polarization and relative luminosity terms are small compared to the fractional error on the yield asymmetry itself. With these assumptions, it is easy to show, using standard error propagation techniques, that the uncertainty on an  $A_i$  can be expressed (with the run index  $i$  dropped for clarity):

$$\sigma_A^2 = \frac{1}{P^2 (N^+ + RN^-)} [(1 - \epsilon)^2 \sigma_{N^+}^2 + (1 + \epsilon)^2 R^2 \sigma_{N^-}^2] \quad (\text{B.5})$$

where  $\epsilon \equiv P \times A$ . So as long as the size of the yield asymmetry is small, the  $\epsilon$  term in the above equation can be neglected. In addition, if the relative luminosity  $R$  is close to unity, the numerator and denominator have equivalent form. Finally, the error on the yield is assumed to be the square root of the yield, such that  $\sigma_N^2 = N$ . Using the above assumptions, equation B.5 for the error on the asymmetry can be expressed as:

$$\sigma_A^2 \approx \frac{1}{P^2 (N^+ + RN^-)} \quad (\text{B.6})$$

Plugging equation B.6 into the expression for the asymmetry, equation B.4, yields:

$$\bar{A} = \frac{\sum_i P_i (N_i^+ - R_i N_i^-)}{\sum_i P_i^2 (N_i^+ + R_i N_i^-)} \quad (\text{B.7})$$

Substituting the polarization, yield, and relative luminosity terms appropriate for  $A_{LL}$  into the above equation yields the expression for  $A_{LL}$  shown in section 5:

$$A_{LL} = \frac{\sum_i P_{Yi} P_{Bi} [(N_i^{++} + N_i^{--}) - R_{3i} (N_i^{+-} + N_i^{-+})]}{\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i} (N_i^{+-} + N_i^{-+})]} \quad (\text{B.8})$$

So it is seen that the above expression for the  $A_{LL}$  is valid so long as the fractional uncertainties on the polarizations and relative luminosities are small compared to the fractional uncertainty on the yield asymmetry, the relative luminosity values are close to unity, and finally that the asymmetry itself is small. All of these conditions are satisfied to a good degree by the data used in the asymmetry measurements.



## B.2 Asymmetry Error Formula Derivation

With the expression for  $A_{LL}$  determined, the next step is to find the expression for the statistical uncertainty. As in the derivation of the asymmetry formula, it is assumed that the errors on the polarization and relative luminosity can be neglected. The asymmetry formula has the following basic form:

$$X = \frac{\sum_i [P_i (A_i - R_i B_i)]}{\sum_i [P_i^2 (A_i + R_i B_i)]} = \frac{P_1 (A_1 - R_1 B_1) + P_2 (A_2 - R_2 B_2) + \dots}{P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2) + \dots} \quad (\text{B.9})$$

where the summation has been explicitly expanded for to two places. The  $P$ ,  $R$ ,  $A$ , and  $B$  terms represent the beam polarization, relative luminosity, and spin-dependent yields, respectively. The statistical uncertainty on the above formula is given by:

$$\sigma_X^2 = \sigma_{A_1}^2 \left( \frac{\partial X}{\partial A_1} \right)^2 + \sigma_{A_2}^2 \left( \frac{\partial X}{\partial A_2} \right)^2 + \sigma_{B_1}^2 \left( \frac{\partial X}{\partial B_1} \right)^2 + \sigma_{B_2}^2 \left( \frac{\partial X}{\partial B_2} \right)^2 \quad (\text{B.10})$$

where the  $\sigma$  terms are the statistical errors on the indicated yields ( $\sigma_N^2 = N$ ). After some algebra, the partial derivative terms can be expressed as:

$$\frac{\partial X}{\partial A_1} = \frac{2P_1^3 R_1 B_1 + P_1 P_2^2 (A_2 + R_2 B_2) - P_1^2 P_2 (A_2 - R_2 B_2)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.11a})$$

$$\frac{\partial X}{\partial A_2} = \frac{2P_2^3 R_2 B_2 + P_2 P_1^2 (A_1 + R_1 B_1) - P_2^2 P_1 (A_1 - R_1 B_1)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.11b})$$

$$\frac{\partial X}{\partial B_1} = \frac{-2P_1^3 R_1 A_1 - P_1^2 P_2 R_1 (A_2 - R_2 B_2) - P_1 P_2^2 R_1 (A_2 + R_2 B_2)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.11c})$$

$$\frac{\partial X}{\partial B_2} = \frac{-2P_2^3 R_2 A_2 - P_2^2 P_1 R_2 (A_1 - R_1 B_1) - P_2 P_1^2 R_2 (A_1 + R_1 B_1)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.11d})$$

The above expressions for the partial derivative terms look messy, but they can be written quite succinctly in terms of the asymmetry itself (equation B.9). Using the asymmetry

expression, the partial derivative terms above can be written as:

$$\frac{\partial X}{\partial A_1} = \frac{P_1 (1 - P_1 X)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.12a})$$

$$\frac{\partial X}{\partial A_2} = \frac{P_2 (1 - P_2 X)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.12b})$$

$$\frac{\partial X}{\partial B_1} = \frac{-P_1 R_1 (1 + P_1 X)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.12c})$$

$$\frac{\partial X}{\partial B_2} = \frac{-P_2 R_2 (1 + P_2 X)}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.12d})$$

Again, if the asymmetry  $X$  itself is small, that term can be neglected in the above expressions. Plugging equations B.12 into the general expression for the uncertainty on the asymmetry (equation B.10), and remembering that the squared uncertainty on a yield is just the yield itself, gives the following formula for the statistical error on the uncertainty:

$$\sigma_X^2 = \frac{P_1^2 A_1 + P_2^2 A_2 + P_1^2 R_1^2 B_1 + P_2^2 R_2^2 B_2}{[P_1^2 (A_1 + R_1 B_1) + P_2^2 (A_2 + R_2 B_2)]^2} \quad (\text{B.13})$$

The formula above was worked out explicitly for the case of two measurements, but can easily be extended to  $N$ . The form of the error expression is evident and the summation can be reintroduced to give:

$$\sigma_X^2 = \frac{\sum_i [P_i^2 (A_i + R_i^2 B_i)]}{[\sum_i P_i^2 (A_i + R_i B_i)]^2} \quad (\text{B.14})$$

Placing the proper beam polarization, relative luminosity, and yield factors into the above equation (and using  $\delta$  to signify the uncertainty), the expression for the statistical error on  $A_{LL}$  presented in section 5 is recovered:

$$\delta A_{LL} = \frac{(\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i}^2 (N_i^{+-} + N_i^{-+})])^{\frac{1}{2}}}{\sum_i P_{Yi}^2 P_{Bi}^2 [(N_i^{++} + N_i^{--}) + R_{3i} (N_i^{+-} + N_i^{-+})]} \quad (\text{B.15})$$

As a final exercise, a very simple expression for the statistical error can be found by assuming the beam polarizations are constant and that the relative luminosity is unity. With these assumptions, it is easy to see that the statistical uncertainty can be approximated as:

$$\delta A_{LL} = \frac{1}{P_Y P_B} \frac{1}{\sqrt{N}} \quad (\text{B.16})$$

where  $P_Y$  and  $P_B$  are the average yellow and blue beam polarizations, respectively, and  $N$  is the total di-jet yield. Large discrepancies between this approximation and the errors given by equation B.15 are an indication of a mistake in the calculation of the statistical uncertainties.

### B.3 Invariant Mass Formula

The invariant mass formula used in the cross section and asymmetry analyses is simply the invariant mass of a relativistic system represented by two 4-vectors:  $M^2 = (\mathcal{P}_3 + \mathcal{P}_4)^2$ . Here  $\mathcal{P}_3$  and  $\mathcal{P}_4$  are the momentum 4-vectors of the two outgoing jets. For our purposes, it is more convenient to express the 4-vector in terms of the transverse mass, rapidity, transverse momentum, and azimuthal angle as in equation B.17.

$$\mathcal{P} = \begin{bmatrix} m_T \cosh y \\ p_T \cos \phi \\ p_T \sin \phi \\ m_T \sinh y \end{bmatrix} \quad (\text{B.17})$$

Here  $m_T \equiv \sqrt{m^2 + p_x^2 + p_y^2}$  is the transverse mass and  $y \equiv \frac{1}{2} \ln \left( \frac{E+p_z}{E-p_z} \right)$  is the rapidity. Using the definitions for the transverse mass and rapidity, it is easy to show that  $m_T \cosh y = E$  and  $m_T \sinh y = p_z$ .

There are three terms which need to be evaluated when calculating the mass:  $\mathcal{P}_3^2$ ,  $\mathcal{P}_4^2$ , and  $2\mathcal{P}_3 \cdot \mathcal{P}_4$ . The first two terms are the masses of the two outgoing jets:

$$\begin{aligned} \mathcal{P}^2 &= m_T^2 \cosh^2 y - p_T^2 \cos^2 \phi - p_T^2 \sin^2 \phi - m_T^2 \sinh^2 y \\ &= m_T^2 (\cosh^2 y - \sinh^2 y) - p_T^2 (\cos^2 \phi + \sin^2 \phi) \\ &= m^2 + p_x^2 + p_y^2 - p_x^2 - p_y^2 \\ &= m^2 \end{aligned}$$

The last term is the combination of  $\mathcal{P}_3$  and  $\mathcal{P}_4$ :

$$\begin{aligned} 2\mathcal{P}_3 \cdot \mathcal{P}_4 &= 2 [m_{T3} m_{T4} (\cosh y_3 \cosh y_4 - \sinh y_3 \sinh y_4)] \\ &\quad - 2 [p_{T3} p_{T4} (\cos \phi_3 \cos \phi_4 + \sin \phi_3 \sin \phi_4)] \\ &= 2 \left[ \sqrt{m_3^2 + p_{T3}^2} \sqrt{m_4^2 + p_{T4}^2} \cosh (y_3 - y_4) - p_{T3} p_{T4} \cos (\phi_3 - \phi_4) \right] \end{aligned}$$

Putting the above expressions together, we get the invariant mass formula:

$$M = \left[ m_3^2 + m_4^2 + 2\sqrt{m_3^2 + p_{T3}^2}\sqrt{m_4^2 + p_{T4}^2} \cosh(y_3 - y_4) - 2p_{T3}p_{T4} \cos(\phi_3 - \phi_4) \right]^{\frac{1}{2}} \quad (\text{B.20})$$

# Appendix C

## Code Location

The Run 9 di-jet cross section and  $A_{LL}$  analysis codes are stored in CVS and can be checked out with the following command: `cvs co offline/paper/psn0649`

An overview of the code structure and running instructions can be found below, but more detailed explanations can be found in the README files contained in each subdirectory.

### C.1 Di-jet Trees

The Run 9 di-jet analyses utilize a set of personalized di-jet trees which are based off the standard Run 9 jet trees. Details on the jet trees can be found in Appendix B of [9].

Currently, the Run 9 data jet trees are on disk at:

`/gpfs/mnt/gpfs01/star/pwg_tasks/spin01/` . They are permanently stored on HPSS at: `/home/pibero/2009jets.P11id` . The simulation jet trees were created specifically for these analyses using Pibero's 2009 embedding sample stored at PDSF. These simulation jet trees are currently on disk at: `/star/institutions/iucf/pagebs/embedding/SL11d/` in the sub-directories:

`pythiajets.SL11d.embed.(R)FF.fixed`,

`pythia.SL11d.embed.(R)FF.fixed`, `/systematics/jets.SL11d.embed.(R)FF.sys2` . All these simulation trees are backed up on HPSS at: `/home/pagebs/embedding_SL11d/` .

As opposed to the jet trees which store all jets from all vertices for several choices of jet algorithm, the di-jet trees only store the two highest jets from the best vertex of the event and only use the anti-k algorithm with a radius of 0.6. In addition, it is required that

at least one jet satisfy one of the trigger conditions. These selection criteria significantly reduce the disk footprint of the di-jet trees and the time needed to process all runs. The di-jet trees also contain all auxiliary information needed to perform an analysis such as the jet track, tower, particle, and pythia information as well as event level information such as fill, polarization, relative luminosity, spin state, etc.

The structure of the di-jet trees is defined in `StForwardDijetEvent` located at: `...createDijetTrees/StRoot/StForwardDijetEvent/`

All di-jet trees needed for the analysis can be found at `/gpfs/mnt/gpfs01/star/pwg/pagebs/` or more permanently on HPSS at `/home/pagebs/diJetTrees/2009XSecALL/v6/`. If there is a need to regenerate any di-jet trees, the code is located in `.../createDijetTrees/` and the submission scripts are located in the sub-directory `xmlSubmit/`.

Before the trees can be created, you need to run ‘cons’ and then ‘compile’ from the `createDijetTrees/` directory. Before running compile, make sure all paths match the directory you have installed the code in. You can then submit jobs from the `xmlSubmit/` sub-directory. Data trees are created by executing the ‘`loopSubmitScript.sh`’ script while the various simulation trees are created by directly submitting the appropriate xml file using ‘`star-submit`’.

## C.2 Cross Section Code

The code and files needed to perform the cross section analysis are located in `.../crossSection/`. Again, before anything can be done, you must run ‘cons’ and ‘compile’ (remembering to change paths as needed) from the `.../crossSection/` directory. In addition, the RooUnfold package is needed, so you will need to ‘cd’ to the RooUnfold directory and run ‘make’ in order to compile the RooUnfold code.

The cross section calculation is done using several root histogram files as input. These files are stored in the `output/` sub-directory. The necessary histogram files are generated by the programs: `analyzeData.v1.exe`, `analyzeData.v1a.exe`, `responseMatrixParticle.v1.exe`, `responseMatrixParticle.v1a.exe`, `responseMatrixParton.v1.exe`, `responseMatrixParton.v1a.exe`, and `pythia.v1.exe`.

The above programs are run by the following scripts: `runDataComp.sh`, `runResponseMatrix.sh`, and `runPythia.sh`.

Once the resultant histograms have been created and placed in the `output/` directory, the cross section can be calculated using the root script: `plotCalcXSec.C`, which is run by typing: `root -l plotCalcXSec.C'(#)'`, where `#` is an integer flag which determines whether the nominal cross section is calculated, or one of those used in systematic evaluations. This script will print out a series of values for the cross section and the corresponding statistical error.

The values printed out from `plotCalcXSec.C` are entered into the `plotXSec.C` script. This script plots a number of useful comparisons as well as the systematics associated with the cross section.

The final cross section values, systematic errors, and theory values / systematics are entered into the script: `plotPresentation.C`. This script generates the final cross section figure as well as the data / simulation comparison figure.

### C.3 Asymmetry Code

The code and files needed to perform the asymmetry analysis are located in `.../asymmetry/`. As with the cross section analysis, you need to run ‘cons’ and then ‘compile’ before anything can be run.

The asymmetry analysis is done using several root histogram files as input. These files are stored in the `output/` sub-directory. The files are generated by the programs: `analyzeDataAsym_v1.exe`, `asymMassShiftParticle_v1.exe`, `asymMassShiftParton_v1.exe`, `trigAndRecoBiasFilter_v1.exe`, and `trigAndRecoBiasNoFilter_v1.exe`. These programs are run from the scripts: `runAsym.sh`, `runMassShift.sh`, and `runTrigRecoBias.sh`.

The root script: `plotAsym_v1.C` has a number of functions which plot the raw  $A_{LL}$  and false asymmetries in several different ways. The script: `plotMassShiftAndTrigReco_v1.C` calculates and displays the mass shifts which are applied to each asymmetry point as well as the trigger and reconstruction bias. Note that the mass shift function needs to be run before the trigger and reconstruction bias function. The final asymmetry figure is

generated using the plotPresentation.C script.



# Bibliography

- [1] J. Beringer et al., “Review of Particle Physics,” *Phys. Rev. D.*, vol. 86, p. 010001, (2012).
- [2] A. Drees, Z. Xu in *Proceedings of the Particle Accelerator Conference, Chicago*, p. 3120, (2001).
- [3] J. Adams et al. (STAR Collaboration), “Transverse-momentum and collision-energy dependence of high- $p_T$  hadron suppression in au+au collisions at ultrarelativistic energies,” *Phys. Rev. Lett.*, vol. 91, p. 172302, (2003).
- [4] G. Van Buren et al., “Correcting for distortions due to ionization in the STAR TPC.” arXiv:physics/0512157, (2005).
- [5] Michael Betancourt et al., “2006 BEMC tower calibration report.” Internal STAR Document: [https://drupal.star.bnl.gov/STAR/system/files/2006-Calibration-Report\\_3.pdf](https://drupal.star.bnl.gov/STAR/system/files/2006-Calibration-Report_3.pdf), (2009).
- [6] Matthew Walker and Michael Betancourt, “2009 BEMC tower calibration report.” Internal STAR Document: <https://drupal.star.bnl.gov/STAR/system/files/2009-Calibration-Report.pdf>, (2010).
- [7] M. Cormier et al., “STAR Barrel Electromagnetic Calorimeter Absolute Calibration Using Minimum Ionizing Particles from Collisions at RHIC.” Internal STAR Note: <https://drupal.star.bnl.gov/STAR/starnotes/public/sn0436>, (2001).
- [8] J. Stevens, “Measurement of the  $W$  and  $Z$  production Cross Sections at Mid-rapidity in Proton-Proton Collisions at  $\sqrt{s} = 500$  GeV in Run 9.” Internal STAR Note: <https://drupal.star.bnl.gov/STAR/starnotes/private/psn0546>, (2011).
- [9] P. Dhawotho, C. Gagliardi, A. Cudd, “2009 Inclusive Jet Double-Helicity Asymmetry  $A_{LL}$  in Polarized p+p Collisions at  $\sqrt{s} = 200$  GeV Analysis Note.” Internal STAR Note: <https://drupal.star.bnl.gov/STAR/starnotes/private/psn0573>, (2013).

- [10] T. Sjostrand, S. Mrenna, P. Skands, “PYTHIA 6.4 physics and manual,” *JHEP*, vol. 05, p. 026, (2006).
- [11] P. Skands, “Tuning monte carlo generators: The perugia tunes,” *Phys. Rev. D*, vol. 82, p. 074018, (2010).
- [12] “GEANT 3.21, CERN Program Library.”
- [13] M. Cacciari, G.P. Salam, G. Soyez, “The anti-kt jet clustering algorithm,” *J. High Energy Phys.*, vol. 2008, no. 04, p. 063, (2008).
- [14] M. Cacciari, G.P. Salam, G. Soyez, “FastJet user manual,” (2011).
- [15] S.D. Ellis, J. Huston, K. Hatakeyama, et al., “Jets in hadron-hadron collisions,” *Prog. Part. Nucl. Phys.*, vol. 60, p. 484, (2008).
- [16] G.P. Salam, “Towards jetography,” *Eur. Phys. J. C*, vol. 67, p. 637, (2010).
- [17] T. Adye in *Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, CERN, Geneva, pp. 313–318, (2011). [hepunix.rl.ac.uk/~adye/software/unfold/RooUnfold.html](http://hepunix.rl.ac.uk/~adye/software/unfold/RooUnfold.html).
- [18] “<http://root.cern.ch/drupal>.”
- [19] G. D’Agostini, “A multidimensional unfolding method based on bayes’ theorem,” *Nucl. Instrum. Meth A*, vol. 362, pp. 487–498, (1995).
- [20] A. Hoecker, V. Kartvelishvili, “SVD Approach to Data Unfolding,” *Nucl. Instrum. Meth A*, vol. 372, pp. 469–481, (1996).
- [21] A. Mukherjee and W. Vogelsang, “Jet production in (un)polarized  $pp$  collisions: Dependence on jet algorithm,” *Phys. Rev. D*, vol. 86, p. 094009, (2012).
- [22] J. Hayes-Wehle, J. Seele, H. Spinka, B. Surrow, “Relative luminosity analysis for run9  $pp$  200 GeV running.” Internal STAR Note: <http://drupal.star.bnl.gov/STAR/starnotes/private/psn0570>, (2012).
- [23] D. de Florian, R. Sassot, M. Stratmann, W. Vogelsang, “Global analysis of helicity parton densities and their uncertainties,” *Phys. Rev. Lett.*, vol. 101, p. 072001, (2008).

- [24] D. de Florian, R. Sassot, M. Stratmann, W. Vogelsang, “Extraction of spin-dependent parton densities and their uncertainties,” *Phys. Rev. D*, vol. 80, p. 034030, (2009).
- [25] E. Leader, A. Sidorov, and D. Stamenov, “Determination of polarized parton densities from a QCD analysis of inclusive and semi-inclusive deep inelastic scattering data,” *Phys. Rev. D*, vol. 82, p. 114018, (2010).
- [26] J. Adams et al.(STAR Collaboration), “Measurements of transverse energy distributions in au+au collisions at  $\sqrt{s_{NN}} = 200$  gev,” *Phys. Rev. C*, vol. 70, p. 054907, (2004).
- [27] S. Frixione, Z. Kunszt, and A. Signer, “Three jet cross-sections to next-to-leading order,” *Nucl. Phys.*, vol. B467, pp. 399–442, 1996.
- [28] S. Frixione, “A General approach to jet cross-sections in QCD,” *Nucl. Phys.*, vol. B507, pp. 295–314, 1997.
- [29] S. Frixione and G. Ridolfi, “Jet photoproduction at HERA,” *Nucl. Phys.*, vol. B507, pp. 315–333, 1997.
- [30] D. de Florian, S. Frixione, A. Signer, and W. Vogelsang, “Next-to-leading order jet cross-sections in polarized hadronic collisions,” *Nucl. Phys.*, vol. B539, pp. 455–476, 1999.
- [31] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, “Evidence for polarization of gluons in the proton,” *Phys. Rev. Lett.*, vol. 113, no. 1, p. 012001, 2014.
- [32] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo, “A first unbiased global determination of polarized PDFs and their uncertainties,” *Nucl. Phys.*, vol. B887, pp. 276–308, 2014.
- [33] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, “Rivet user manual,” *Comput. Phys. Commun.*, vol. 184, pp. 2803–2819, 2013.
- [34] STAR\_DIJET analysis <https://github.com/veprbl/star-rivet-analyses/> commit 79d340c8c515.
- [35] Z. Nagy, “Next-to-leading order calculation of three jet observables in hadron hadron collision,” *Phys. Rev.*, vol. D68, p. 094002, 2003.

- [36] P.R. Bevington and D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, 3 ed., (2003).