

BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**A SEARCH FOR NEW RESONANCES WITH THE DIJET ANGULAR
RATIO USING THE COMPACT MUON SOLENOID EXPERIMENT**

by

JASON MICHAEL ST. JOHN

A.B., Harvard University, 2001

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2012

Approved by

First Reader

James Rohlf, Ph.D.
Professor of Physics

Second Reader

Robert Carey, Ph.D.
Professor of Physics

Acknowledgments

We are the stories we tell ourselves, and increasingly there is evidence for this from neuroscience. Surely we learn this habit of self-construction from our loved ones. They tell us the story of ourselves before we even have words in our heads, when the world is a whirlwind of impressions, figures without captions. They are there, telling our story alongside us throughout life, helping to shape who we are.

I have been very fortunate to have many loving narrators and excellent storytellers in my story, making this dissertation possible.

My parents' love, support, and sense of humor encouraged a small and cautious boy on ever-grander explorations, from the bugs under a rotting log to the largest science experiment in the world. Anne, my little big sister, and David, my big little brother, are my closest allies and my first friends. I love you all very much.

In Boston, I must thank Kevin Smith and his international cadre of students: Alex 'Two-Legs' 'Night Train' 'Hey Quit Calling Me That' DeMasi, Leyla Colakerol, and Yufeng Zhang. Together we were Yacko, Smacko, Boffo, and Yob, and no trip across the Long Island Sound, no long summer at BNL, would have the same relish without you. (Alex, stop that.)

Ed Kearns, Jennifer Raaf, Fanny Dufour, Mike Litos, and p-Dan Gastler, my adopted HEP family, you welcomed me into your barrio before I even had an experiment. I continue to check my office for mischief each time I return to it. Thank you. Eric Hazen, thanks to you I can read in hex, and even have nice memories of learning to do so.

Tibor Antal, Claudio Castelnovo, Anna neé Izsak, Merco Mazza, Sylvia Viola-Kuminsky-etc, Andrea Velenich, Sameet Sreenivasan, Vishal Sood, and the whole Lower Allston/BU Physics dinner party circuit which gathered in a former whorehouse on Everett; Rachele Dominguez, Kaca Bradonic, Claudia DeGrandi, Kevin Stokely, Rebecca Bartorsky, Andrew Inglis, Sebastian Remi, and all the BU physics crowd I have been fortunate enough to know; David and Michelle Lapidus, Elizabeth Engelhardt, David Camden, Shelby Condray, Elissa Krakauer, and all the housemates who have become the friends (no doubt my imperfect recall has left out important names):

The purpose of life is dinner parties. I don't think anyone I really know will be surprised to hear me say as much, or rather that dinner parties encapsulate the purpose of human life. There may be preparation and anticipation, but there can often be spontaneity; there is some bringing-to-bear of resources and effort; convergence of old friends and new acquaintances; learning new information, opinions, and skills; tensions which arise to resolve or set aside; necessary chores; the satisfaction of overcoming challenges; people brought closer in those challenges; and there is time to savor existence. Thank you. Please let me know when you're available to make

dinner together again.

KathyCopic, D. Spoons Spoonhauer, Aras Papadelis, Ellie Twedt, Mama Perez, Toyoko Orimoto, Seth Caughron, Jimmy Degenhardt, the real Christian Ohm, Max Scherzer, Mark Cooke, Rachel Yohay, Efe Yazgan, Luis Lebolo, Jeff Temple, Tullio Grassi, Zach Marshall, Daniel and Katrine Whiteson, Ted Laird, Dick Kellogg, Arjan Heering, Tina Vernon, and many more: If it takes a village to raise a child, it takes a good chunk of the planet to raise a physicist. Thank you all for extending me your friendship, in a land where successfully ordering a coffee without resorting English is a major victory.

Dragoslav Lasic, Laza: my face, my living arrangements, my manners, and my time at CERN were all vastly better for knowing you. I hope someday to introduce you to the cat.

Dean Hidas, Carley Kopecky, Louise Suter, Warren Clarida, Kristi Clarida, Bo Jayatilaka, Kirby, John C. Freeman, Homer Wolfe, James Agnew, Mika Vestrienen, Mark & Michelle Mathis, Jason Manour, Scott Braithwaite, Teppei Katori, Sarah neé Croft of CBM, John neé Backus-Mayes of CBM, Carrie McGivern, Zeynep Isvan, Tim Head, Joe Zennamo, Joe Walding, Joe Grange, Ranjan Dharmapalan, Alex Radovic, the great 21st century experimentalist Josh Spitz, Vasundhara Chetluru, and many, many others, including many of the names above this paragraph: My life on the Science Prairie has been made sweet and livable by your unhesitating friendship and excellent company. Thank you.

Amnon Harel, Marek Zielinski, as well as the rest of the 2010-2011 dijet ratio team, I am deeply grateful for your hard work, and your support of my work. None of this work would have been possible without you. Taylan Yetkin, for getting the ball rolling on this analysis, thank you very much.

My thesis committee, Rob Carey who gives such thorough and thoughtful feedback, Ken Lane, Claudio Chamon, and Rama Bansil; thank you for your time and attention, your questions, and for that sweetest of first words each doctor hears: "Congratulations."

Jim Rohlf, my major advisor: Thank you for opening the gateway to the LHC for me, for many long conversations not only about physics, but about the realities of doing science, culinary arts, and whatever else came to our minds.

Jim Hirschauer, thank you for being a sane mind with as much attention to a good explanation as to good friendship. Let's work on something together soon.

Rob Harris, I have learned an enormous amount thanks to your guidance and example. I am better at dealing both with science and with scientists for having worked closely with you. I don't know how to thank you enough, except to be the best colleague in research I can be.

It has been suggested that the only reasonable reaction to the fact of one's own existence is gratitude. How much greater is the gratitude I have, for having found all these others whose existence enhances the quality of my own. Thank you all, for being as you have become, and for everything good I have become with you around.

**A SEARCH FOR NEW RESONANCES WITH THE DIJET
ANGULAR RATIO USING THE COMPACT MUON SOLENOID
EXPERIMENT**

(Order No.)

JASON ST. JOHN

Boston University Graduate School of Arts and Sciences, 2012

Major Professor: James Rohlf, Professor of Physics

ABSTRACT

A search for dijet resonances is performed using 2.2 fb^{-1} of proton-proton collision data at $\sqrt{s} = 7 \text{ TeV}$ recorded by the CMS detector at CERN. The study is based on the dijet angular ratio, the ratio of the number of events with the two leading jets having pseudorapidity difference $|\Delta\eta| < 1.3$ to the number of events with $1.3 < |\Delta\eta| < 3.0$. Models of new resonances which decay into two jets typically predict dijet angular distributions and hence, values of the dijet angular ratio which differ from standard model processes. We thus use the measurement of the angular ratio as a function of mass to set limits on the cross sections of new spin- $\frac{1}{2}$ quark-gluon resonances. We exclude excited quarks of mass less than 3.2 TeV at 95% confidence level, where a limit of 2.8 TeV is expected.

Contents

Acknowledgments	iii
Abstract	vii
Table of Contents	viii
List of Tables	xii
List of Figures	xiii
List of Abbreviations	xvii
1 Introduction	1
2 Theory and Simulation	3
2.1 Quantum Chromodynamics and the Standard Model	3
2.2 Parton Distribution Functions	5
2.3 QCD Interactions	5
2.4 Beyond the Matrix Element	10
2.4.1 Parton Showers	12
2.4.2 Multiple Interactions	13
2.4.3 Hadronization	14
2.5 Reconstruction of Jets: Clustering Algorithms	15
2.6 Dijet Resonances	17
3 Experimental Setup	20
3.1 LHC	20
3.2 CMS	22
3.3 Coordinate System	23
3.4 Silicon Tracking	24
3.5 Crystal Calorimeter	26
3.6 Hadronic Calorimeter	28
3.7 Muon Systems	30

3.8	Trigger System	31
3.8.1	Hardware Trigger	32
3.8.2	Software Trigger	33
4	Data Analysis	34
4.1	The Dijet Angular Ratio	35
4.2	Data and Monte Carlo Inputs	36
4.2.1	Common Tools	36
4.2.2	Collision Data	36
4.2.3	Monte Carlo Datasets	37
4.3	Jet Reconstruction	38
4.4	Event Selection	39
4.4.1	Optimization of Dijet Angular Ratio Definition	41
4.4.2	Dijet Mass Bins	44
4.4.3	Dijet Mass Requirement and Trigger Efficiency	44
4.4.4	Data Quality	50
4.5	Dijet Angular Ratio Predictions	52
4.5.1	QCD Prediction from PYTHIA	53
4.5.2	Dijet Resonance Predictions from PYTHIA	57
4.6	Dijet angular ratio in Data	62
4.6.1	Choice of Reconstruction Method	62
4.6.2	Measured Dijet angular ratio	64
4.6.3	Highest-Mass Events	65
4.6.4	Choice of Sideband Region	65
4.7	Statistical Method	72
4.7.1	Likelihood Ratio Definiton	72
4.7.2	Setting Limits	73
4.8	Systematic Uncertainties	75
4.8.1	Pileup (PU)	75
4.8.2	Additive Offset to Prediction from Sideband	77
4.8.3	Non-Perturbative Corrections	78
4.8.4	Jet Energy Correction (JEC)	79
4.8.5	General Detector Effects	80
4.8.6	QCD production rate normalization	81
4.8.7	Signal Model Uncertainty	82
4.8.8	Total Systematic Uncertainty	83
5	Conclusion	87
	Appendices	88
A	Highest Mass Dijet Events	89

B	HCAL Readout Electronics System	99
C	Monitoring	103
C.1	DCC Crate Monitoring	103
C.1.1	Monitored Registers	104
C.1.2	The hcalDCCMonitoring Class	104
C.2	DQM: Data Quality Monitoring	105
C.2.1	Problem rates	107
C.2.2	Corruption in the Raw Data	108
C.2.3	Monitoring Data Flow	112
C.2.4	Expert Diagnostics	113
	Curriculum Vitae	120

List of Tables

4.1	Standard CMS jet identification requirements.	40
4.2	Trigger fully efficient points by dijet mass.	45
4.3	Physical characteristics of the ten highest inner events.	67
4.4	Choices of the sideband region, excluded from contributing to the signal discrimination of the analysis, but required to determine the best fit value of the additive shift applied to our QCD MC prediction for $R_{ \Delta\eta }$	69
C.1	Abbreviations used in the plot "07 LRB Data Corruption Indicators".	110
C.2	Abbreviations used in the plot "08 Half-HTR Data Corruption". . . .	111
C.3	Abbreviations used in the plot "Channel Integrity" plots, both the summary and the breakouts by DCC.	111

List of Figures

2.1	Illustration of the Mandelstam variables.	6
2.2	Lowest-order Feynman diagrams for $qq' \rightarrow qq'$	7
2.3	Lowest-order Feynman diagrams for $q\bar{q} \rightarrow q'\bar{q}'$	7
2.4	Lowest-order Feynman diagrams for $q\bar{q} \rightarrow q\bar{q}$	7
2.5	Lowest-order Feynman diagrams for $q\bar{q} \rightarrow q'\bar{q}'$	8
2.6	Lowest-order Feynman diagrams for $q\bar{q} \rightarrow gg$	9
2.7	Lowest-order Feynman diagrams for $qq \rightarrow qq$	9
2.8	Lowest-order Feynman diagrams for $gg \rightarrow gg$	10
2.9	The highest-mass dijet event in this dissertation.	11
2.10	Feynman diagram for a generic dijet resonance.	17
3.1	Diagram of the LHC accelerator complex.	21
3.2	Exploded view of the CMS detector	22
3.3	The all-silicon CMS particle tracking system.	24
3.4	Lorentz angle and charge sharing in the Pixel tracker.	26
3.5	Schematic of the CMS electromagnetic crystal calorimeter.	27
3.6	Upper right quarter of a CMS cross sectional view showing HCAL.	28
4.1	PYTHIA predictions for scatterplot distributions of η_1 and η_2 in leading dijets.	41
4.2	FOM value showing a maximum at $\Delta\eta_{inner} = 1.3$, for all values of $\Delta\eta_{outer}$	43
4.3	Outer dijet trigger turn-on curves for PF dijets. (Corrected jet triggers.)	46
4.4	Outer dijet trigger turn-on curves for PF dijets. (Uncorrected jet triggers.)	47
4.5	Outer dijet trigger turn-on curves for calo dijets. (Corrected jet triggers.)	48
4.6	Outer dijet trigger turn-on curves for Calo dijets. (Uncorrected jet triggers.)	49
4.7	Distribution of the jet electromagnetic fraction.	50
4.8	Distribution of the jet ID variable $n90$	51
4.9	Distribution of the jet ID variable $fHPD$	51
4.10	The distribution of the number of tracks within jets at the associated vertex.	52

4.11	The distribution of pseudorapidity for jets comprising inner or outer dijets.	52
4.12	The distribution of transverse momentum for jets used in this analysis.	53
4.13	The distribution of pseudorapidity for jets comprising dijets.	54
4.14	The distribution of pseudorapidity differences between jets comprising dijets.	55
4.15	The fitted inner differential mass spectrum.	57
4.16	The differential mass spectra of inner and outer dijets.	58
4.17	The lineshapes for excited quarks.	61
4.18	The event yield of inner and outer events for calo and PF dijets.	62
4.19	The measured dijet angular ratio in calo and PF dijets.	63
4.20	The quotient of the dijet angular ratio in PF and Calo jets.	64
4.21	The total dijet event yield, and the differential dijet mass spectra.	66
4.22	The PYTHIA prediction before it is offset.	68
4.23	The measured dijet angular ratio in calo dijets, with MC expectations.	69
4.24	Pulls: Data minus MC over statistical uncertainty.	70
4.25	The measured dijet angular ratio in PF dijets.	71
4.26	The PF pulls.	71
4.27	Limit summary plots for qg dijet resonances without systematic uncertainties.	75
4.28	The measured value of $R_{ \Delta\eta }$ with one or more than vertex.	76
4.29	The measured value of $R_{ \Delta\eta }$ in events with above- or below-average counts of reconstructed vertices	77
4.30	The PYTHIA6 $R_{ \Delta\eta }$ non-perturbative effects.	79
4.31	The $R_{ \Delta\eta }$ in PYTHIA6 and the NLO prediction.	80
4.32	Systematic due to JEC uncertainty.	81
4.33	The JEC systematic uncertainty envelope.	82
4.34	The $R_{ \Delta\eta }$ in genjets and calojets.	83
4.35	The normalized PYTHIA inner dijet rate prediction.	84
4.36	Final lower exclusion limit plot with systematics.	85
4.37	Effect of systematic uncertainties on limit.	86
A.1	The second pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??	90
A.2	The second pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??	91
A.3	The third pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??	92
A.4	The fourth pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??	93
A.5	The last pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??	94

A.6	The first pair of the eight highest-mass outer dijet events.	95
A.7	The second pair of the eight highest-mass outer dijet events.	96
A.8	The third pair of the eight highest-mass outer dijet events.	97
A.9	The last pair of the eight highest-mass outer dijet events.	98
B.1	Cutaway diagram of the HPD.	100
B.2	Block diagram of the DCC.	101
C.1	An example plot from the electronics monitoring system MonVis. . .	106
C.2	The format specification for DCC events at the level of raw bits. . . .	107
C.3	The HCAL Raw Data Monitor’s Problem Plots.	108
C.4	The corruption folder.	109
C.5	Keys to Faux-LED plots.	110
C.6	The Data Flow folder.	112
C.7	The key for ”01 Data Flow Indicators.”	112

List of Abbreviations

CMS	Compact Muon Solenoid
APD	Avalanche Photodiode
API	Application Programming Interface
CL	Confidence Level
CMS	Compact Muon Solenoid
CRC	Cyclical Redundancy Check
CSC	Cathode Strip Chamber
CTEQ	Coordinated Theoretical-Experimental Project on QCD
DAQ	Data Acquisition
DCC	Data Concentrator Card
DIS	Deep Inelastic Scattering
DQM	Data Quality Monitoring
DT	Drift Tube
EB	ECAL Barrel
ECAL	Electromagnetic Calorimeter
EE	ECAL Endcap
FED	Front-End Device
FIFO	First-In, First-Out buffer
FPGA	Field-Programmable Gate Array
FSR	Final-State Radiation
GCT	Global Calorimetric Trigger
GOL	Gigabit Optical Link
GT	Global Trigger
HAD	Hadronization
HAL	Hardware Access Library
HB	HCAL Barrel
HCAL	Hadronic Calorimeter
HE	HCAL Endcap
HF	HCAL Forward
HLT	High-Level Trigger
HO	HCAL Outer
HPD	Hybrid Photodiode
HTR	HCAL Trigger and Readout
ISR	Initial-State Radiation
JEC	Jet Energy Correction
JSON	JavaScript Object Notation
L1A	Level-1 Accept
LLR	Log Likelihood Ratio
LO	Leading Order
LRB	Link Receiver Board

MET	Missing Transverse Energy
MPI	Multiple-Parton Interactions
NLO	Next-to-Leading Order
P5	Point 5 of the LHC
PCI	Peripheral Component Interconnect
PDF	Parton Momentum Function
PDS	Pseudo Data Set
PKAM	Previously Known As Monster
PMT	Photo-Multiplier Tube
PS	Parton Shower
PU	Pileup
RCT	Regional Calorimetric Trigger
RPC	Resistive Plate Chamber
SM	Standard Model
SPS	Super Proton Synchrotron
TEC	Tracker Endcap
TIB	Tracker Inner Barrel
TID	Tracker Inner Disk
TLA	Three-Letter Acronym
TOB	Tracker Outer Barrel
TTC	Trigger Timing & Control
TTS	Trigger Timing System
USC55	Underground Service Cavern at P5
VME	VERSAmodule Eurocard bus
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
WLS	Wavelength Shifting

Chapter 1

Introduction

“Thou cannot see an isolated quark, no matter how hard thou tryest.” -Sheldon Glashow, *Interactions*

This dissertation presents a study of the angular distributions of inclusive dijets at the Compact Muon Solenoid experiment [?], which is fed colliding protons by CERN’s Large Hadron Collider. The most common hard collision outcome features a pair of jets, a dijet. By searching in these events for signs of a new resonance, we are playing to the strengths of the events nature is willing to give us. And by focusing on the angular distributions, which are unaffected by the usual complications of energy measurements of jets, we are playing to our own strengths.

Particle physics has a long, proud history of measuring angular scattering distributions. Ernest Rutherford’s discovery of the nucleus within the atom a century ago is such an example, and essentially founded the field of sub-atomic physics [?]. Robert Hofstadter’s resolution of the nucleon size [?] was the next big step. Nowadays we have probed to such small scales that a fixed target in a particle beam just won’t do. We have graduated to vast synchrotrons colliding ultra-relativistic particles, and surrounded the collision points with detectors the size of small buildings. The science of the very small has grown quite large. The Standard Model has been expanded and then re-unified as humanity’s reach has extended, and still we know that it doesn’t explain everything. We are probing to the smallest scales and the highest energies in search of some clue as to how the universe makes sense.

This thesis begins with an overview of theoretical expectations in Chapter ??, tracing the path from two partons exchanging momentum in a pp collision to the emergence of two jets. Chapter ?? explains the CMS experimental apparatus which detects and measures these jets. Finally in Chapter ?? we discuss the analysis of the data themselves. This chapter covers the definition of the dijet angular ratio, expectations for the standard model and for possible signals which new resonances would yield, the findings in data, and the resulting conclusions which we draw.

Chapter 2

Theory and Simulation

2.1 Quantum Chromodynamics and the Standard Model

Quantum Chromodynamics (QCD) is the quantum field theory of those forces which quarks feel and which leptons such as the electron do not. QCD describes a force-carrying 'color' field whose quantum is the gluon. Like Quantum Electrodynamics (QED), which describes the electric field whose quantum is the photon, QCD describes forces between color-charged particles. Where the photon is exchanged by particles which are charged positively or negatively, the gluon is exchanged by particles with "red," "green," or "blue" charges, described in group theory as the symmetry $SU(3)$.

Just as a positive and a negative charge compose a QED-neutral object, such as a hydrogen atom, so a red quark, a green quark, and a blue quark compose a QCD-neutral object, such as the proton.

The Dirac equation [?] makes clear that the fermions composing normal matter can be described in keeping with relativity *if* they have antimatter partners. The "positive electron" was the first such antiparticle known, discovered by Anderson [?]. A particle whose charges are positive (under QED) and red (under QCD) would have an antiparticle whose charges are negative and anti-red. Negative charges are present in normal matter, but anti-red, anti-blue, and anti-green are unique to antimatter. This means that three quarks, one of each color, can be bound together in a color-neutral state (called a baryon), as can three anti-quarks, or one quark and one anti-quark (called a meson). In a final twist, the gluons whose exchange mediates the color force themselves carry two color charges, and from this much mischief is born, as we discuss below.

We must note that QCD is part of the larger Standard Model of particle physics, the other half of which is the electroweak theory which enlarges QED by describing its connection to the weak sector. The matter content of the Standard Model is comprised of six quarks and six leptons, arranged in three successively more mas-

sive generations of $SU(2)_Y$ doublets ¹. While all twelve of these fermionic fields participate in the weak interaction through left-handed matter (or right-handed antimatter) components, only the quarks participate in the strong interaction of QCD. The Standard Model Lagrangian is globally invariant under the Poincaré symmetry, which corresponds to conservation of energy, momentum, and angular momentum. It is also locally invariant under the gauge symmetry $SU(3) \times SU(2) \times U(1)$, whose gauge bosons are the force-carrying gluon of QCD and the photon, W^\pm , and Z which mediate electroweak interactions.

The large cross sections of QCD processes in the experiment we describe in these pages dominate other standard model processes. They are several orders of magnitude greater. As we will detail in Ch. ??, high-momentum-transfer electroweak physics processes are not included in the models we use, yet these all-QCD models describe our inclusive event sample extremely well.

2.2 Parton Distribution Functions

The quarks and gluons within the proton share its momentum among them, with gluons carrying about half of the momentum. Gluon splitting gives a 'sea' of virtual quarks and antiquarks, which also share in the momentum. We express this momentum sharing as a probability for a parton of a given species to be found in the proton as a function of the longitudinal momentum fraction x . These parton distribution functions (PDFs) determine the probability for an initial pair of partons to have a certain combination of momenta, and they dictate a large spread in the collisions' center-of-mass energy below that of the protons.

Measuring PDFs is a challenge being met by parameterizing collected scattering experiment results with continuous functions. Widely used PDFs are published by the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [?]. There is a complication here in that the set of proton PDFs is slightly different depending upon the invariant mass of the four-momentum exchanged by probe particle and parton. We discuss this in section ??.

2.3 QCD Interactions

The following discussion follows the lead in [?]. For reasons we discuss in ?? and ??, it is difficult to distinguish between a gluon and a quark emerging from a collision. It is even more difficult to know which two partons initiated the collision. Thus, even our first approximation from QCD for dijet production must cover all the bases.

¹We can be more exact. The left-handed fields are in the doublets, while the right-handed components of the charged leptons and quarks live in $SU(2)$ singlets. Nature has not yet revealed right-handed neutrinos.

In this dissertation, the convention is that Feynman diagrams are read left-to-right, like English, and not bottom-to-top, like the comment threads on YouTube. We adopt the usual Mandelstam variables [?] to express the cross sections:

$$\begin{aligned} s^2 &\equiv (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ t^2 &\equiv (p_1 - p_3)^2 = (p_2 - p_4)^2 \\ u^2 &\equiv (p_1 - p_4)^2 = (p_2 - p_3)^2 \end{aligned}$$

where p_1, p_2, p_3 , and p_4 refer to the four-vectors (E, \vec{p}) of the external legs in Fig. ???. The above definitions follow from conservation of energy. For s^2 this is the statement that the initial energy must be the same as the final energy. For t^2 and u^2 , we are reminded that the energy gained by one incoming particle must be balanced by the energy lost from the other incoming particle.

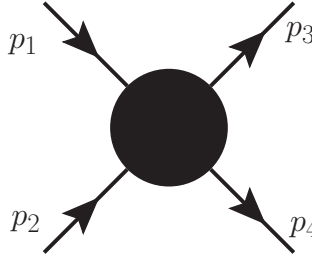


Figure 2.1: The four generic four-momenta of two-to-two scattering, useful for expressing Mandelstam variables s, t, u . The arrows indicate an idea of flow, not the fermionic character of these four fields, and the blob obscuring the vertex may be any legal interaction.

One simple place to start is to calculate all the two-quarks-in, two-quarks-out cross sections. There are four such processes at lowest order (so-called tree level), ??-??, and their corresponding Feynman graphs are shown in Figs. ?? to ??. In practice these must be calculated taking into account the PDFs of the proton. The PDFs are different for every quark flavor, are different again for the antiquarks, and the gluons have a proton PDF, too, to complete the set.

$$\frac{d\sigma}{d\hat{t}}(qq' \rightarrow qq') = \frac{4\pi\alpha_s^2}{9\hat{s}^2} \left[\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} \right] \quad (2.1)$$

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow q'\bar{q}') = \frac{4\pi\alpha_s^2}{9\hat{s}^2} \left[\frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right] \quad (2.2)$$

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow q\bar{q}) = \frac{4\pi\alpha_s^2}{9\hat{s}^2} \left[\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} - \frac{2}{3} \frac{\hat{u}^2}{\hat{s}\hat{t}} \right] \quad (2.3)$$

$$\frac{d\sigma}{d\hat{t}}(qq \rightarrow qq) = \frac{4\pi\alpha_s^2}{9\hat{s}^2} \left[\frac{\hat{u}^2 + \hat{s}^2}{\hat{t}^2} + \frac{\hat{t}^2 + \hat{s}^2}{\hat{u}^2} - \frac{2}{3} \frac{\hat{s}^2}{\hat{u}\hat{t}} \right] \quad (2.4)$$

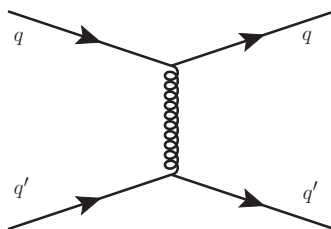


Figure 2.2: $(qq' \rightarrow qq')$ Any pairing quarks or antiquarks of dissimilar flavor exchange momentum at tree level.

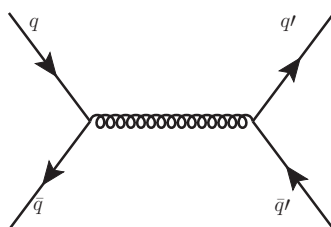


Figure 2.3: $(q\bar{q} \rightarrow q'\bar{q}')$ A quark-antiquark pair annihilate and the resulting gluon pair-produces a quark and antiquark of dissimilar flavor.

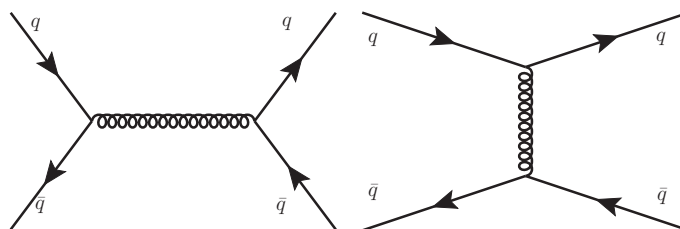


Figure 2.4: $(q\bar{q} \rightarrow q\bar{q})$ A quark-antiquark pair annihilate and the resulting gluon pair-produces a quark and antiquark of the same flavor, or they exchange a gluon.

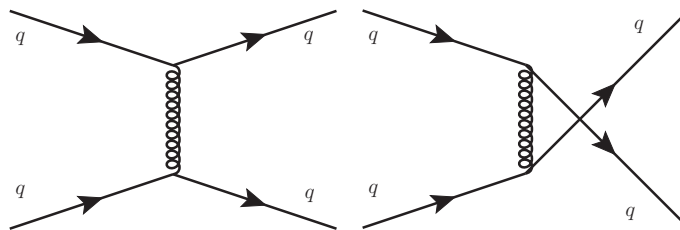


Figure 2.5: $(qq \rightarrow q'q')$ Two quarks (or two antiquarks) of the same flavor exchange momentum.

Next we take on the two-to-two QCD interactions with two external gluon legs. There are three diagrams at tree level which take a quark-antiquark pair to a pair of gluons, Fig. ???. These same diagrams time-reversed, although not shown, give the same cross section up to a factor reflecting the different number of initial states for a gluon pair; this is the old saw “sum-over-initial, average-over-final” beloved of the quantum mechanic. There are also three tree-level diagrams for $qg \rightarrow qg$ scattering, shown in Fig. ???.

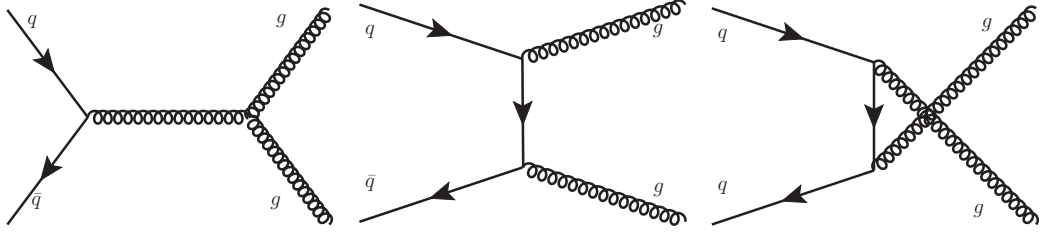


Figure 2.6: $(q\bar{q} \rightarrow gg)$ Two quarks (or two antiquarks) annihilate into two gluons.

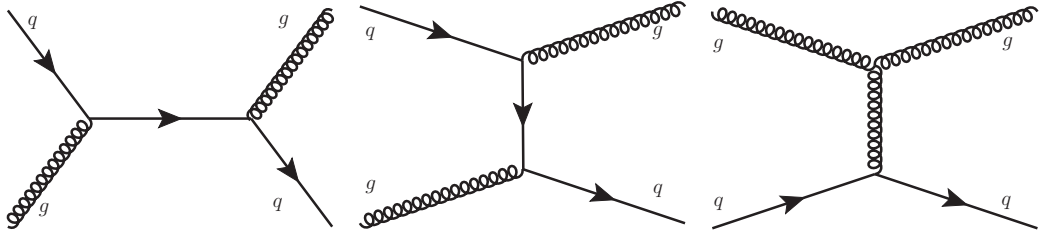


Figure 2.7: $(gq \rightarrow gq)$ A gluon is absorbed only to be re-emitted, or it exchanges momentum with a quark of any flavor.

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow gg) = \frac{32\pi\alpha_s^2}{27\hat{s}^2} \left[\frac{\hat{u}}{\hat{t}} + \frac{\hat{t}}{\hat{u}} - \frac{9}{4} \left(\frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right) \right] \quad (2.5)$$

$$\frac{d\sigma}{d\hat{t}}(gg \rightarrow q\bar{q}) = \frac{\pi\alpha_s^2}{6\hat{s}^2} \left[\frac{\hat{u}}{\hat{t}} + \frac{\hat{t}}{\hat{u}} - \frac{9}{4} \left(\frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right) \right] \quad (2.6)$$

$$\frac{d\sigma}{d\hat{t}}(qg \rightarrow qg) = \frac{4\pi\alpha_s^2}{9\hat{s}^2} \left[-\frac{\hat{u}}{\hat{s}} - \frac{\hat{s}}{\hat{u}} + \frac{9}{4} \left(\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} \right) \right] \quad (2.7)$$

Finally, because QCD is a non-Abelian gauge theory, there are four processes by which two gluons may scatter into two gluons, shown in Fig. ???. Their cross section is given in equation ???.

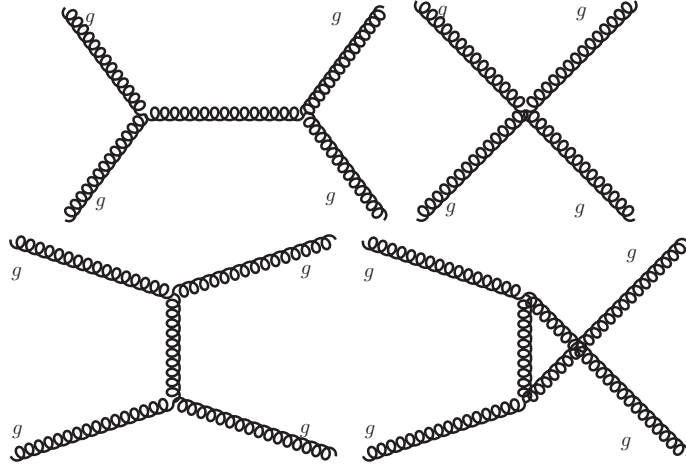


Figure 2.8: $(gg \rightarrow gg)$ Two gluons in, two gluons out. The quartic coupling of gluons contributes its scattering amplitude here.

$$\frac{d\sigma}{dt}(gg \rightarrow gg) = \frac{9\pi\alpha_s^2}{2\hat{s}^2} \left[3 - \frac{\hat{t}\hat{u}}{\hat{s}^2} - \frac{\hat{s}\hat{u}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2} \right] \quad (2.8)$$

All of these processes are simulated by the PYTHIA 6.4 Monte Carlo simulation [?] using CTEQ6 parton distribution functions [?], as described in section ??.

2.4 Beyond the Matrix Element

High-energy collisions of one or more systems of bound quarks, *hadrons*, reveal two extremes of behavior in color-charged objects. Within the hadrons they make up, quarks and gluons (collectively, *partons*) are remarkably independent. One parton may interact with a high-energy probe particle, without an apparent effect from its fellow hadron constituents. On the other hand, partons are never seen outside of hadrons. Frank Wilczek’s 2004 Nobel prize was for work relating to this *asymptotic freedom*, a coin whose other side is infrared slavery. Asymptotic freedom means that at short distance scales (high energies), quarks behave as if they’re free. The practical consequence of this state of affairs is the observation of hadronic jets exiting collisions more or less where one would expect naked partons, as if the partons were free and alone during the collision, but on the way out the door became suddenly self-conscious. It is clear that the matrix element description is not enough to completely describe jet phenomenology, as a glance at Fig.?? confirms.

Naked partons must give up some of their energy to the creation of new partons which join up into color-neutral hadrons. Often this process creates dozens of hadrons

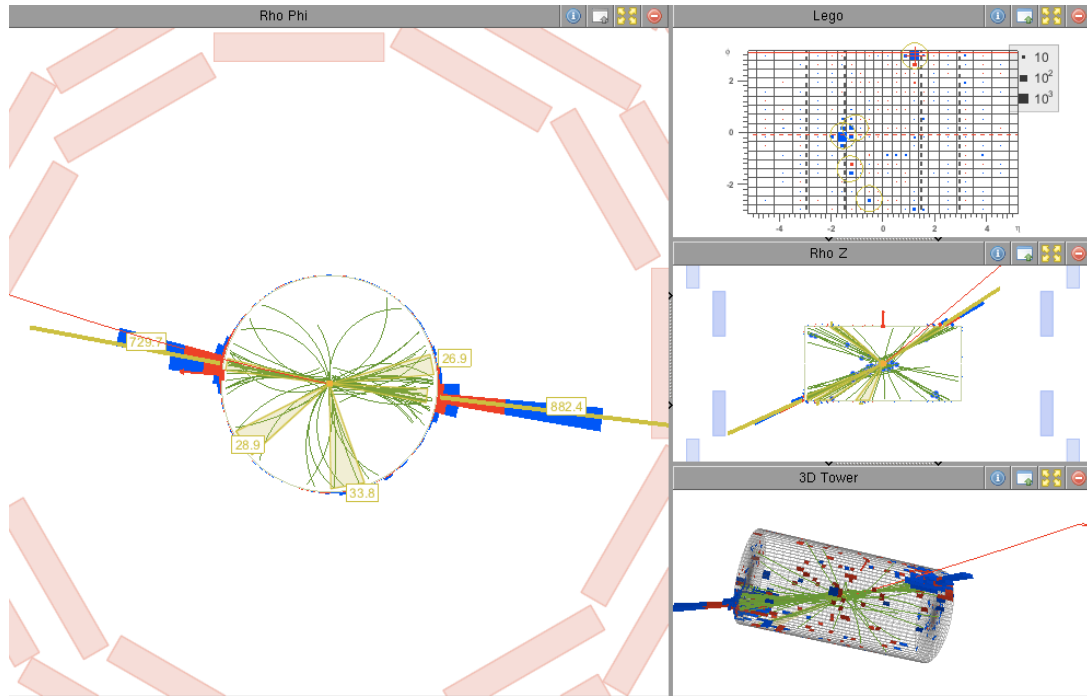


Figure 2.9: The highest-mass dijet event in this dissertation. Notice the rich activity in the (green) tracks, and the presence of energy deposits away from the jets (red electromagnetic, blue hadronic). Details of this event are given in Table ?? along with the other high-mass dijets.

headed away from a particle collision, mostly in tight clumps around the scattered partons' original directions. When we collide two hadrons at high enough energies, we may pretend to ourselves that we collide two partons thanks to asymptotic freedom, but thanks to infrared slavery it is jets we will see flying out of the collisions.

Starting from the approximation of free partons in two-to-two scattering, the usual approach to simulation is to begin adding in successive effects until the resulting events resemble actual observations as closely as needed. These effects include initial- and final-state radiation, an underlying event, multiple parton interactions, and hadronization. They are all included in the PYTHIA event simulations used in this analysis.

2.4.1 Parton Showers

Initial- and final-state radiation are consequences of the fact that quarks and gluons can radiate gluons. Also gluons can split into quark-antiquark pairs, quarks can radiate photons, and indeed the entire Standard Model is invited to the party, in principle. Probed with low momentum exchange, the parton distribution functions

seem quite immune to changes in Q^2 , the square of four-momentum exchanged by the scattered partons. As Q^2 increases (or the parton momentum fraction becomes small) we find that the PDFs are not quite immune to the changes. The DGLAP evolution equations [? ? ?] successfully describe the logarithmic Q^2 dependence of PDFs in terms of splitting functions. The splitting functions reflect the modification to each $1 \rightarrow 2$ probability due to the energy sharing fraction of the mother particle and one of the daughter particles. The evolution equations allow measurements of PDFs made at one Q^2 to be evolved into a prediction of the PDFs at another Q^2 .

The splitting functions introduced by DGLAP for evolving PDFs to new Q^2 values are also put to work in leading-log parton showers. This is a Monte Carlo technique using probabilities which depend upon the momentum scale and the energy sharing of the splitting. The goal is to recreate the topology and kinematics of a believable parton shower. To achieve this, the PYTHIA monte carlo generator used in this thesis proceeds from the hard scatter, the main event, downward in momentum scale to a cutoff scale near 1 GeV. It follows color flow as well as conserving energy and momentum at each branching. The probability \mathcal{P}_a for each particle a to split is found from its differential

$$d\mathcal{P}_a = \sum_{b,c} \frac{\alpha_{abc}}{2\pi} P_{a \rightarrow bc}(z) dt dz \quad (2.9)$$

where $t \equiv \ln\left(\frac{Q^2}{m_a^2}\right)$ and $E_b = zE_a$. Loosely speaking, t plays the role of time in the shower development in that it is restricted to decrease monotonically from the hard scatter in final-state showers, and to build monotonically to meet the hard scatter in initial-state showers. Thus the probability for a branching at a given t is found by integrating $\mathcal{P}_a(z)$ across the kinematically allowed z range (there are different definitions of this range available in PYTHIA), and multiplying by the Sudakov form factor $S_a(t)$, which is the exponentiated integral over all earlier times that some branching has not already occurred.

If mother partons are originally massless, their subsequent branchings may assign them a mass retroactively. Occasionally this causes conflict with the masses and z values already assigned, and arbitration takes place to repair or reject the branchings in question.

When the dust has settled and complete parton showers are generated, it's time to face the music of hadronization.

2.4.2 Multiple Interactions

Having settled on the primary interaction, there is still the problem of possible secondary interactions among the other partons present. We set aside the possibility of simultaneous collisions by other protons, known as pileup, as a separate problem. The proton remnants may be color-connected to the partons entering the primary

interaction, must recoil as dictated by momentum and energy conservation, and they may have interactions of their own—multiple interactions. One simplifying assumption is the independence of the primary and any secondary interactions, although we want to be careful that the secondary interactions do not rival the primary interaction for hardness. The energy and momentum available to partons in the proton remnants must be less by the amount taken up in the primary interaction. The impact parameter of the collision should affect the chance for secondary interactions, by changing the effective overlap of the original protons. Finally, a parton already taken by another interaction, such as a valence quark, must not be used again by another interaction. All of these concerns are addressed to some approximation by the implementation of multiple parton interactions in our PYTHIA samples.

2.4.3 Hadronization

Hadronization is the process of partons from the shower becoming hadrons. In the Lund string model [10, 11], strings of self-interacting strong field lines are treated as physical objects. Their decay into color-triplet quarks and antiquarks which are paired into color-singlet hadrons, predicts hadronization observations quite successfully. It is for this success that PYTHIA, which models hadronization thus, is best known. The Lund model in its modern form reveals that between beam remnants and the jets of completely showered partons, where color connections must stretch, additional hadrons are found. This has been confirmed by dedicated studies in different types of events selected for their color topologies [12, 13, 14]. The completely hadronized event, with all partons and color strings resolved into hadrons, is a bunch of particles (mostly pions, the lightest hadrons) headed away from the collision point with their assigned four momenta. At last our scattered partons have graduated to become hadronic jets.

2.5 Reconstruction of Jets: Clustering Algorithms

Many physics objects may be reconstructed from the signals and noise found for each event CMS records. These objects include muons, electrons, and photons, but for the purpose of this dissertation, the physics object of interest is the hadronic jet.

Now we have a jet-like group of particles long lived enough to escape the interaction point and pass the inner wall of the beampipe into the Compact Muon Solenoid particle detector. Electrically charged particles leave tracks, executing helices in the central solenoidal field. Nearly all of the particles deposit their energy in the calorimeters.

Clumps of particle tracks and energy deposits form the observable evidence of hadronic jets. Of course, it is impossible to review each such clump by hand to assign its single energy-momentum four-vector. As a practical measure, algorithms have been developed which can automate this process.

Historically these algorithms searched for cones of fixed size [? ?] which contained the jets as well as possible. The modern successors of cone jets are successive combination algorithms. Arguably the simplest of these is the the Cambridge-Aachen algorithm [?], which merges the closest pair of jet constituents, removing them from the list of energy deposits, tracks, or simulated four-vectors. In their place goes the merged object, and the new closest pair on the list is merged. The algorithm stops when the smallest angular separation² of any two objects is greater than some maximum, set as a parameter. All objects left when the dust settles are jets.

The anti-kT algorithm [?] enlarges slightly upon the angular separation criterion, weighting it by the lower p_T^{-1} in each pair of constituents. For constituents i and j , the distance measure is

$$d_{ij} = \min(p_{T,i}^{-1}, p_{T,j}^{-1}) \frac{(\Delta R_{ij})^2}{(\Delta D)^2} \quad (2.10)$$

and an analogous 'distance' measure is added for the individual constituents,

$$d_i = p_{T,i}^{-1}. \quad (2.11)$$

(When the power of p_T is +1 rather than -1, one has the kT algorithm.) The d_i are the generalization of the single angular cutoff now that a nonzero power of p_T is part of the comparison.

The value of ΔD is specified as a parameter, normalizing the $(\Delta R_{ij})^2$ by the angular cutoff. A single comparison step is iterated:

- ◊ If the smallest distance measure is a d_i , that object is declared a jet and moved to the list of complete jets.
- ◊ If the smallest distance measure is one of the d_{ij} , those two objects are merged.

After either operation, all the affected distance measures referring to i (or i and j) are recomputed. The list of jet constituents rapidly drops to zero, leaving only the list of jets. As is widely noted, anti-kT jets are built up around high- p_T cores, and take on perfectly round shapes as they incorporate lower- p_T constituents. When nearby jets abut, the jet with the greater p_T wins proportionally greater conicality. For two overlapping jets with p_{T1} and p_{T2} , a common border b emerges such that $\frac{\Delta R_{1b}}{p_{T1}} = \frac{\Delta R_{2b}}{p_{T2}}$ [?]. In the end the anti-kT algorithm, for all its elegance, behaves like an ideal cone algorithm.

It is very important that the jets built by anti-kT are perfectly unchanged even if a jet constituent is replaced by two 4-vectors, each with half the magnitude of the original. This immunity is known as collinear safety. Anti-kT jets are also infrared safe, because adding infinitely soft jet constituents will have no effect. Collinear and

²The angular coordinates in this thesis are those explained in section ??, the azimuthal angle ϕ about the beamlines and the pseudorapidity η . Thus for any angular separation $\Delta R^2 = \Delta\phi^2 + \Delta\eta^2$.

infrared safety historically dogged the early cone algorithms, although ultimately solutions have been found. Both properties are necessary in order to compare jet observations meaningfully with theory.

2.6 Dijet Resonances

In the following pages we search for processes producing a narrow resonance X that decays to dijets: $pp \rightarrow X \rightarrow 2 \text{ jets}$, inclusive, as pictured in Fig. ??.

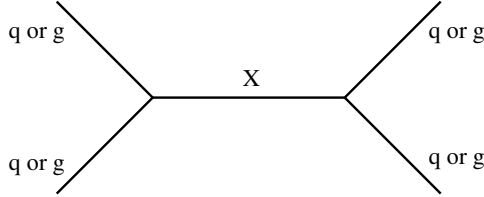


Figure 2.10: Feynman diagram for a generic dijet resonance. The initial state and final state both contain two partons (quarks, antiquarks, or gluons) and the intermediate state contains an s -channel resonance X .

In our benchmark resonance model [?] and [?], the fermions of the standard model are composite. In this framework, excited quarks have nonzero masses even before electroweak symmetry breaking and any possible Yukawa couplings to massive scalars with nonzero vacuum expectation values³. Excited quarks would couple to the gauge bosons, and through the gauge bosons, also couple to the groundstate quarks. The gauge-mediated transition between f^* excited states and the more familiar fields f appears in the Lagrangian density as a term such as ([?] eqns. (2-4))

$$\mathcal{L}_{trans} = \frac{1}{2\Lambda} \bar{f}_R^* \sigma^{\mu\nu} \left[g_s f_s \frac{\lambda^a}{2} G_{\mu\nu}^a \right] f_L + h.c. \quad (2.12)$$

where $G_{\mu\nu}^a$ is the field-strength tensor of the gluon, f_s modifies the coupling strength (and may be a structure function $f_s(q^2)$), and Λ is the compositeness scale.

The predicted decay width of u^* or d^* into ordinary quarks and gauge bosons takes the form

$$\Gamma(q^* \rightarrow qg) = \frac{1}{3} \alpha_s f_s^2 \frac{m^{*3}}{\Lambda^2}. \quad (2.13)$$

³Protons, familiar composite objects, get their mass this way. Quark rest masses are very small compared to the proton mass; most of the mass in the visible universe comes about thanks to the strong force!

Taking the excited quark mass as the compositeness scale, this gives 3.9×10^{-2} as a fraction of the excited quark mass. With such narrow resonances, experimental dijet mass resolution broadens the measured peak significantly. Initial- and final-state radiation further soften the peak. In this model, 85% of all decays are of the form $q^* \rightarrow qg$, giving a pair of energetic jets, and the remaining decays are electroweak rather than strong.

The resulting dijet decay signature, wider than any inherent width and with significant tails, is just what we are looking for, and would be the same for any narrow, spin- $\frac{1}{2}qg$ resonance.

This is a a generic resonance search. Our experimental motivation is that the LHC is a parton-parton collider, and resonances of partons must decay to a parton pair, giving two jets in the final state. The theoretical interest is broad since there are many models predicting narrow dijet resonances. The analysis is performed in a general way that does not limit the search to any specific dijet resonance model. We consider the excited composite quark model of [?] and [?] an example of new physics to which this search has sensitivity. The cross section limit which we set applies to any spin- $\frac{1}{2}$ quark-gluon resonance, and allows us to set a mass limit on the excited quark model specifically.

Chapter 3

Experimental Setup

3.1 LHC

The world's largest and highest-energy synchrotron, the Large Hadron Collider straddles the French border at the far western tip of Switzerland. Its tunnel is 175 m to 50 m under the ground, the former LEP tunnel whose 1.4% gradient slopes downward toward Lake Geneva for reasons of costs and geo-engineering [?]. It accepts beams of 450 GeV protons from CERN's Super Proton Synchrotron (SPS) (See Fig. ??) and circulates them in opposite directions around its 26.659 km circumference, accelerating them by successive passes through radio frequency cavities to an energy of 3,500 GeV. The protons are bent to this course while passing through 1232 two-aperture, superconducting dipole magnets [?].

At four of the LHC's eight interaction points, the two separate beams can be made to collide at low- β^* ¹. At the interaction point designated Point 5 lives CMS.

The instantaneous proton-proton luminosity of the LHC is unprecedented. There are several ways to estimate the instantaneous luminosity, such as

$$\mathcal{L}_{bb} = \frac{f_{rev} N^2 \cos^2\left(\frac{\alpha}{2}\right) F}{2\pi \Sigma_x \Sigma_y} = \frac{f_{rev} N^2 \cos^2\left(\frac{\alpha}{2}\right) SF}{4\pi \sigma_x \sigma_y} \quad (3.1)$$

where \mathcal{L}_{bb} is the bunch-bunch instantaneous luminosity, f_{rev} is the frequency of bunch revolution or orbit, N is the number of protons per bunch, α is the beams' crossing angle, F is the bunch separation factor, and Σ_x and Σ_y are the effective beam areas in the two transverse directions at the interaction point. If the geometric factor S is used, these last can be replaced by the transverse beam sizes σ_x and σ_y with modified coefficients as shown.

The LHC achieved its nominal instantaneous luminosity of 3×10^{33} Hz/cm² in early September of 2011, just a little after the data analyzed in these pages had been collected.

¹ β^* is the distance after the nominal interaction at which a beam's width has doubled.

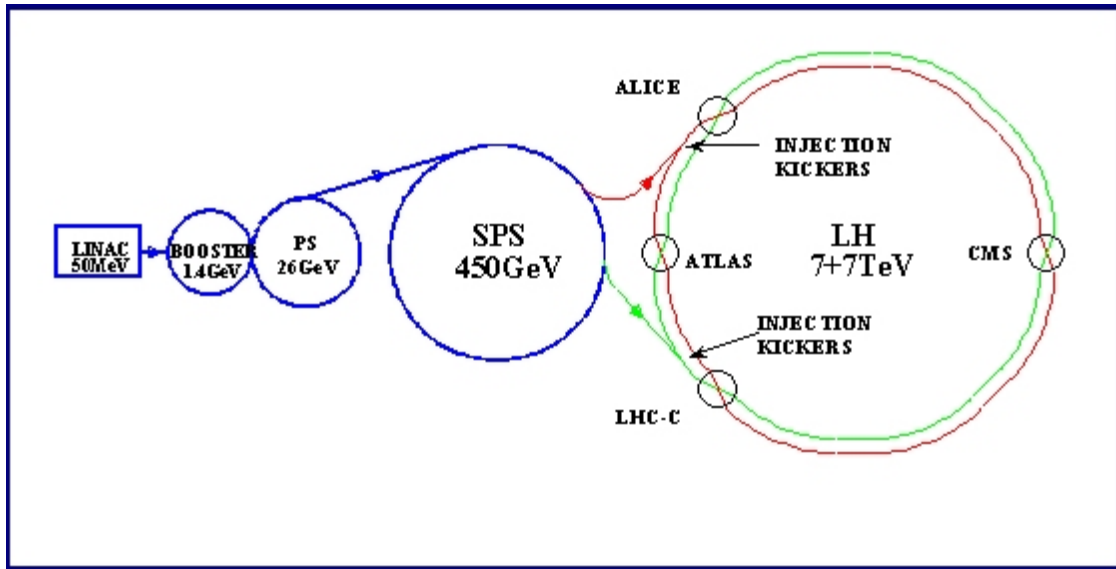


Figure 3.1: Diagram of the LHC accelerator complex.

3.2 CMS

The CMS detector, Fig. ??, is like a 12.5 kilotonne onion only it's a cylindrical, general-purpose particle detector. Its axis lies horizontally, coinciding with the beampipe of the Large Hadron Collider at Interaction Point 5 (P5). Down this pipe from both ends fly the LHC's proton bunches, colliding at the heart of the detector. From inner layer to outer, the debris of particle collisions traverse the silicon trackers, the crystal electromagnetic calorimeter, the brass hadronic calorimeter, the world's largest superconducting solenoid, and finally the magnet's iron return yoke which is instrumented with three independent systems of gas ionization muon trackers. If their trajectory makes a sufficiently small angle with the beamline, particles originating in collisions may miss all of these, to traverse only the steel of the forward calorimeter. These systems are describe below in much, much more detail. Much of the following detector system detail reflects the same information found in [?].

3.3 Coordinate System

The CMS coordinate system takes the z -axis at the center of the detector, aligned with the counter-clockwise orbiting beam. The transverse plane is vertical, with the y -axis oriented upward and the x -axis pointing into the center of the ring. The azimuthal angle ϕ is zero at the positive x -axis and runs $-\pi$ to $+\pi$. The pseudorapidity

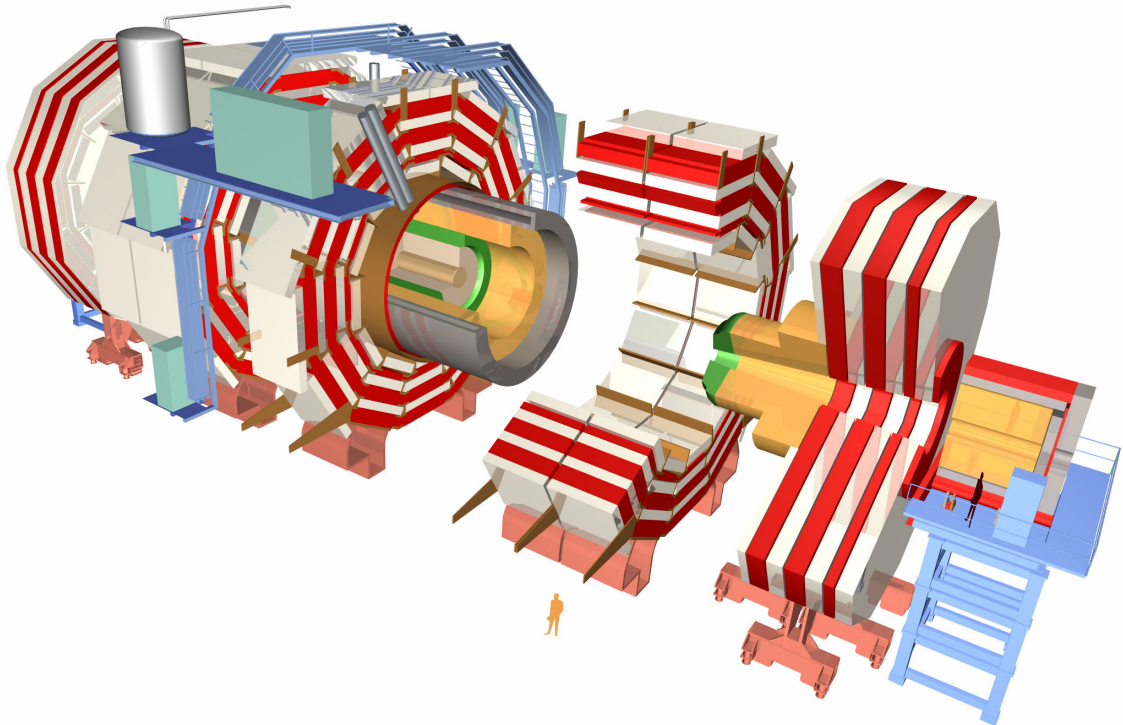


Figure 3.2: Exploded view of the CMS detector, showing inner tracking systems, electromagnetic and hadronic calorimeters, the world’s largest superconducting solenoid, and muon-tracking instrumentation in the iron of the magnetic return yoke. Also shown is the design of the endcap and far forward calorimeter. A CERN standard-sized technician is shown in the foreground for scale.

η is defined as

$$\eta \equiv -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (3.2)$$

where θ is the polar angle. The pseudorapidity η is often used at colliders because it may be measured in the lab frame and can also be written in terms of the momentum p

$$\eta = \frac{1}{2} \ln \left[\frac{|\vec{p}| + p_L}{|\vec{p}| - p_L} \right] \quad (3.3)$$

which in the massless, relativistic limit approximates the rapidity

$$y = \frac{1}{2} \ln \left[\frac{E + p_L}{E - p_L} \right]. \quad (3.4)$$

Hence like y , η in the massless, relativistic limit is additive under longitudinal boosts, very handy for describing relativistic particles emerging from collisions whose center of mass frame is longitudinally boosted by parton distribution functions. η has two further advantages: It is independent of particle mass, which is not always known. Because the detector has nonuniformities (dead regions, hot cells, etc.), it is also helpful that η is a direction in the detector. Finally, particle trajectories are more evenly distributed in η than in θ . η is steeply sloped with respect to θ and extremal as it approaches the very active regions near the beamlines.

3.4 Silicon Tracking

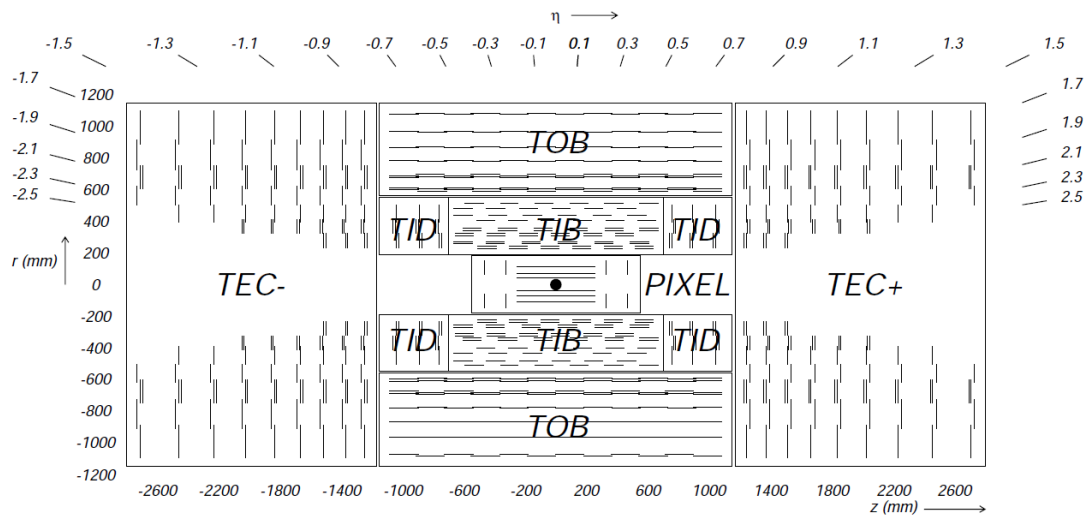


Figure 3.3: The all-silicon CMS particle tracking system loves its TLA's. Shown here in a transected view are the Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB), and Tracker End Caps (TEC).

CMS employs two concentric silicon-based systems for detecting the tracks of charged particles. The inner is the Silicon Pixel Detector (Pixel), nested just within the Silicon Strip Tracker (Tracker), and both lie fully within the axial part of the experiment's 3.8 T solenoidal magnetic field. They are designed as high-resolution vertex-determining systems which operate even in a high multiplicity environment.

The Pixel's barrel comprises three layers starting at an inner radius of 4.4 cm and surrounded by ten layers of the Tracker barrel, which reach an outer radius of 1.1 m. The tracking coverage of the barrel is improved by endcap disks, two in the Pixels and nine in the endcaps. Altogether these systems constitute about two

hundred square meters of active silicon tracking. Their acceptance for singly charged particles having transverse momentum at least 1 GeV extends up to $|\eta| < 2.5$.

Charged particles traversing the silicon liberate electron-hole pairs, which then drift apart under the influence of the applied bias voltage. The electrons are collected by anodes for processing as analog signals. In the inner tracker, these take the form of dots on a square lattice, effectively dividing the sensitive bulk into pixels. In the outer strip tracker, the anodes are linear electrodes. The holes drift to the cathode on the back surface of the silicon, where they are absorbed.

In both designs, the barrel sections drift electrons at a significant angle to the 3.8 T magnetic field. This gives rise to large Lorentz angle of 32 between the electric field and the drift direction of the electrons [?], spreading their charge among two or three anodes. This is sketched in Fig. ???. This charge sharing leads to improved spatial resolution of the hit locations. In the endcaps, where the sensors are hosted on fan-blade shaped modules, the tilt of the blades about their radial axis also provides a crossed field geometry in order to take advantage of this effect. Without this tilt, the magnetic field would be nearly parallel with the electric field in the silicon, and very little charge sharing would take place in the endcaps.

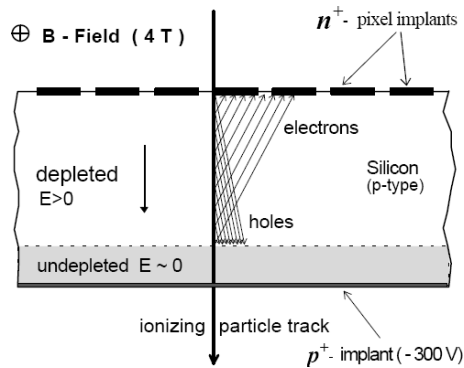


Figure 3.4: Lorentz angle and charge sharing in the Pixel tracker. As depicted in [?]Fig. 2.2

As illustrated in Fig. ??, the depth of the depleted, sensitive bulk from the undepleted region to the anode face affects the proportion of charge sharing from Lorentz angle effects. Ionization only liberates electron-hole pairs in the sensitive bulk, and the farther this happens from the pixels, the more spread the charges will be at the anode. The depletion depth is in turn tied to (limited by) the sensor thickness of 250 μm . There are other factors which informed the optimization of pixel size and geometry. Pixels which are too large would be unable to distinguish hits from closely spaced tracks. Estimating from simulation, a hit occupancy of 1% was used as a guiding requirement to minimize this type of error [?].

3.5 Crystal Calorimeter

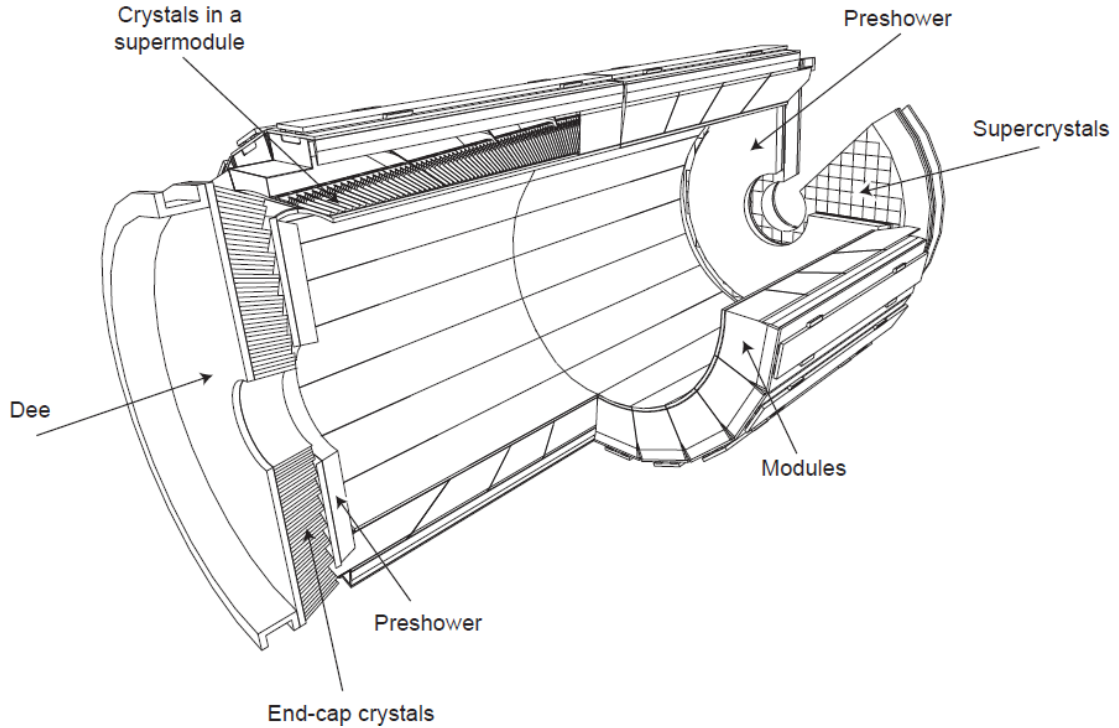


Figure 3.5: Schematic of the CMS electromagnetic crystal calorimeter. Illustrated here are the cylindrical 'barrel' and circular 'endcap' sections, in cutaway.

The electromagnetic calorimeter (ECAL) was designed an important goal in mind: detecting Higgs bosons decays to photon pairs. The ECAL employs scintillating crystals of lead tungstate (PbWO_4) because of their fast scintillation decay time (80% drop within 25 ns), radiation resistance, and small Molière radius [?]—only 2.2 cm. These crystals allow fast measurements of signals which might be as frequent as 80 MHz, a long useful life in an intense radiation environment, and good spatial resolution of electromagnetic showers in a highly active environment.

The ECAL's scintillating crystals are oriented just 3° (both azimuthally and in the polar direction) from pointing their long axis directly at the nominal center of the detector, to preserve directional information even while eliminating cracks which would otherwise give direct paths for high-energy particles to escape detection. The high density (8.28 g/cm^3) and short radiation length (0.89 cm) mean that the 230 mm of crystal length packs in $28.5 \chi_0$, for excellent shower containment. The ECAL barrel (EB) and endcap (EE) calorimeters subtend pseudorapidity ranges $|\eta| < 1.479$ and $1.479 < |\eta| < 3.0$, as well as the complete azimuthal angle ϕ . There are 61,200 crystals in the barrel and 7,342 in each of the endcaps.

Lead tungstate poses notable challenges to electromagnetic calorimetry. Radiation damage induces optical absorption by localized damage which heals on a short timescale, giving a performance equilibrium which depends upon the dose rate. Laser illumination of the crystals allows in-situ calibration for this effect.

Temperature dependence of the signal is another challenge met by the ECAL. The scintillation yield of lead tungstate is temperature dependent, losing $2.1\%C^{-1}$ at 18C [?]. In the EB, Avalanche PhotoDiodes (APDs) are mounted directly to the crystals to provide photodetection, and the negative temperature dependence of the APD signal brings the overall response variation to $-3.8\pm 0.4\%C^{-1}$ [?]. An extensive water cooling manifold stabilizes the temperature against the cooling influence of the silicon tracker and the heat load of the ECAL on-detector readout electronics.

3.6 Hadronic Calorimeter

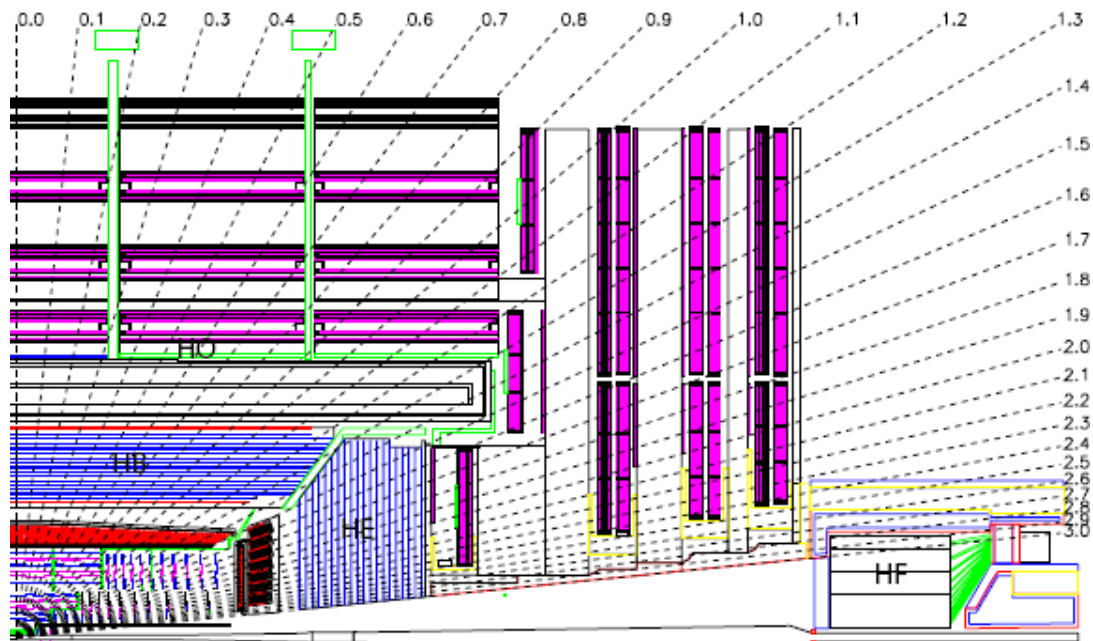


Figure 3.6: Upper right quarter of a CMS cross sectional view, where the interaction point is at the lower left. HCAL is shown in blue, its four sections labeled HB (barrel), HE (endcap), HF (forward), and HO (outer). Black dashed lines label cones of constant η .

The Hadronic Calorimeter (HCAL) is designed to measure the energy of the LHC collisions' hadronic effluence. It achieves this by the same method that a 500-pound gorilla gets the seat it wants in the theater—by being massive enough. The HCAL's density and thickness ensure that it crams in an enormous number of nucleons along

the paths taken by hadrons coming from the proton collisions. The first of two important consequences is that these hadrons have as many chances as possible to initiate a hadronic shower, so that they begin to deposit their energy where it can be measured. Naturally the particles constituting the shower are subject to the same forces, and therefore themselves are also likely to deposit their energy before traveling far. Thus the second consequence of HCAL's absorber material density is that the shower will be contained within the HCAL for measurement.

Hadronic showers are notorious for their individual variations of such properties as longitudinal and transverse profile, and forms taken by energy deposited. Nonetheless the notion of an average hadronic shower provides a useful length scale from which to design. Accordingly the HCAL segmentation in the polar angular direction η and the azimuthal direction ϕ is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$. In the barrel these dimensions correspond closely to the typical size of a hadronic shower.

The HCAL has its barrel (HB) and endcap (HE) sections within the solenoid's interior, surrounding and supporting the ECAL. They are made of brass alloy absorber, cartridge brass # 260 with 70% copper and 30% zinc [?], interleaved with plastic scintillator tiles which sample energy deposited by the particles. Wavelength shifting (WLS) fibers are laid into grooves in the scintillator and spliced to colorless optical fibers, which bring the light yield to on-detector photoconverters.

The HB extends from an inner radius of 1.77 m to the inner wall of the solenoid at 2.95 m, occupying a pseudorapidity of $|\eta| < 1.3$. The tiles of HB are laid in 72 sections azimuthally, giving an angular width of 0.087 (5). The HE shares this segmentation scheme at the overlap region and continues it forward to $\eta < 1.6$. After $\eta > 1.6$ and up to the forward edge of the HE at $|\eta| = 3.0$, the azimuthal granularity is doubled to 10.

On the outside of the solenoid, instrumented scintillator tiles comprise the "outer" hadronic calorimeter (HO) meant to sample hadronic energy deposited after HB. Massive steel plugs form the absorber of the forward calorimeter (HF), hosting embedded quartz fibers which run parallel to the beam pipe. Their Cherenkov light output is used to detect particle activity in the absorber material. Within each tower of HF some of the fibers do not run all the way to the front face but stop short, helping to differentiate the early-deposited electromagnetic shower energy from hadronic showers.

Much more about the back-end electronics of HCAL can be found in Appendix ???. A description of the monitoring systems which report their good performance and assure high data quality can be found in Appendix ???.

3.7 Muon Systems

CMS makes use of three independent muon tracking systems. They are all gas-under-voltage detectors with specialized geometries. Drift Tubes (DTs) are ganged

in 'chambers' of parallel tubes with carefully controlled cross sectional geometry approximating a rectangle. Chambers are found together in 'stations' at four longitudinal depths of the magnetized iron yoke, crossed with respect to other chambers in their station to measure both the $r - \phi$ and z position of through-going muons. The DTs cover $|\eta| < 1.2$, which is the CMS barrel section.

The endcap disks of the return yoke are home to the Cathode Strip Chambers (CSCs). These muon trackers are notably radiation resistant, which is important because they cover the region $0.9 < |\eta| < 2.4$ making them subject to high fluence during proton collisions. Cathode strips are oriented radially, and the anode wires run across them in the tangent direction. Six layers are ganged in each CSC, and there are CSCs at each of the four endcap stations within the two endcap disks.

These two systems are complemented by the fast-response Resistive Plate Chambers (RPCs) in the region $|\eta| < 1.6$. These thin-gapped particle detectors use pixelated electrodes and pulse timing to locate particle tracks. Their additional information is useful for improving track-finding efficiencies, but the RPCs are best known for their star role in the CMS muon triggers, for which they are well-suited.

The whole of the CMS muon system, together with the magnet yoke on which they are installed, is carefully surveyed by a laser alignment system. The sensitivity varies but is only a few microns in most of the system. Distortions of the iron yoke change the geometry both within and between the muon tracking components, and this effect is measured so as to be taken into account. For a dramatic example, the CMS solenoid at its full magnetic field of 3.8 T draws in the centers of the multiton endcap iron plates about 16 mm.

3.8 Trigger System

At LHC design luminosity, it would be impractical and prohibitively expensive for CMS to record all that it detects with each filled bunch crossing. Indeed, this is true even for filling schemes and instantaneous luminosities substantially below nominal. The solution is to record events use an automated, carefully prepared selection trigger, allocating a small amount of recording bandwidth for events triggered with minimum bias and using the rest to record only events containing some signals of particular interest. CMS achieves this using a novel two-stage trigger system.

A 'counting house' of electronic hardware makes a real-time decision as to what signals of interest are present, if any. When programmed criteria are met, these Level-1 Accepts (L1A's) are broadcast to all CMS subdetectors by the Trigger Timing and Control system (TTC), prompting the readout hardware to send the associated event to the central CMS data acquisition (DAQ), which both assembles the event fragments from the various subsystems and tests the complete event with a software filter called the High Level Trigger (HLT). The HLT has access to the entirety of the recorded event, reconstructed as quickly as possible, and makes a rapid decision to reject the event or to forward it on for recording.

3.8.1 Hardware Trigger

The first stage in a trigger decision is the hardware trigger, or “Level-1”. The hardware trigger has a hierarchical design. Local triggers, systems of independent units at a lower level, feed into less numerous units at a higher level. Higher levels summarize the most significant signals present in the lower levels, until a final decision on the event can be reached based on global knowledge.

This hardware comprises the Regional Calorimetric Trigger (RCT), the Global Calorimetric Trigger (GCT), and the Global Trigger (GT). The GT is also fed by systems whose inputs come from the muon chambers. For purposes of this dissertation, they are irrelevant but for the finite triggering bandwidth, which they share. The hardware trigger as a whole must arrive at an L1A rate no higher than 100 kHz (although this has varied in practice), on an input rate as high as 80 MHz.

3.8.2 Software Trigger

The second and final stage in a trigger is the HLT. It is realized as a Filter Farm, a system of hundreds of CPUs being fed complete CMS events and performing a nearly complete reconstruction on their raw data. A builder network is responsible for routing event fragments to the builder unit which is assembling each event, and which then forwards the event along to one of its associated Filter Units for evaluation against the full menu of software trigger algorithms. When an event passes a filtering algorithm (and the prescale factor count for that algorithm is satisfied), the event is forwarded to the Storage Manager. Both raw data and the HLT decisions are included. The overall rate of stored events cannot exceed about 100 Hz, although again in practice this has varied with the size of the event, the duration of temporary storage request rate excursions, and related factors.

Chapter 4

Data Analysis

As discussed in chapter ??, the dominant high-momentum-transfer events in our LHC pp collisions feature at least two energetic jets, *dijets* arising from momentum exchange between two partons. They are predominantly forward- and backward-scattered jets, reflecting the momentum of the incoming protons. In its center-of-mass frame the associated dijet tends to have large opening angles, with one jet close to each beamline of the incoming protons, in a way which does not depend strongly on the total energy involved. This pattern of scattering is predicted by QCD.

By contrast consider two-to-two parton scattering mediated by a massive resonance. The decay of the resonance into the original parton combination is nearly isotropic in the rest frame of the resonance, and the interaction is greatly enhanced when the total energy matches the mass of the resonant state. Thus resonances compared to QCD dijets have greatly enhanced counts of dijets with small pseudo-rapidity difference. The dijet masses from resonance decays also cluster in mass near the resonant state.

When searching for such dijet resonances, the initial analysis may put aside the angular information and examine the dijet mass spectrum directly for significant bumps. This comes down to making a careful study of the number of dijets counted within some ranges of dijet invariant mass, per range size. One makes many such ranges, non-overlapping bins of dijet mass, and chooses them to match the mass resolution at the dijet mass. The CMS collaboration has published the results of just such a search [?], as has the ATLAS collaboration [?]. These are traditional 'bump-hunt' studies of the dijet differential cross section which use 1 fb^{-1} of data. They set limits on the excited quark model which is used in this thesis as a benchmark; CMS set a lower limit at 2.49 TeV (2.68 TeV expected), and ATLAS set their at 2.99 TeV (2.81 TeV expected). All are limits at the 95% confidence level.

4.1 The Dijet Angular Ratio

The search for dijet resonances using only the dijet mass spectrum need not stand alone. An important cross check takes advantage of the more isotropic, less forward scattering which typifies the decay of massive particles. This angular information is brought to the fore by making two independent counts of dijet events, both binned in dijet mass as above, one for dijets with small pseudorapidity difference, *inner* dijets, and one with larger pseudorapidity difference, *outer* dijets, then taking the ratio of inner to outer event counts. In the steeply falling dijet cross section, this angular ratio approach has a major advantage when compared to bump hunting; the uncertainty in the luminosity divides away exactly.

Making the ratio of two dijet mass spectra from the same detector has another advantage. The dijet angular ratio divides away the largest systematic uncertainty, the jet energy scale. To the extent that they are under-measured by the same amount, an inner dijet whose mass is systematically under-measured will still wind up in the numerator of the ratio with an outer dijet of the same true mass. As we show in ??, this cancellation is nearly perfect. It leaves a small systematic uncertainty on the dijet angular ratio due to the uncertainty on the jet energy scale, which has some angular dependence.

4.2 Data and Monte Carlo Inputs

4.2.1 Common Tools

The joy of collaborating is not needing to reinvent wheels. We make use of the standard data processing and resulting ntuple of 2.2 fb^{-1} shared by our CMS colleagues conducting other physics analyses of high- p_T jets. The ntuple [?] is a list of events and contains only the event information needed for high- p_T jet analyses. We also use CMS-standard Monte Carlo datasets for fully reconstructed pseudo-events, after processing by the same ntuple-maker.

4.2.2 Collision Data

JavaScript Object Notation (JSON) files [?], despite their distracting name, serve the purpose of restricting the consulted data to high-quality run sections. They are the product of an automated data certification [?] process with considerable human oversight. Using them to mask out rough spots identified in the data, we improve the trustworthiness of the data which enter the measurement. We consider 2.2 fb^{-1} of integrated luminosity in runs 136033 to 173692 using these official JSON files:

```
===Jet dataset===
(136033-149442)
```

```

/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions10
Cert_136033149442_7TeV_Apr21ReReco_Collisions10_JSON.txt
(160404-163869)
/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions11
Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt
(165088-173692)
/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions11
Cert_160404-173692_7TeV_PromptReco_Collisions11_JSON.txt

```

With the JSON files in place as a filter, we run our ntuple-maker on the following datasets, listed by their run range:

```

===Calo, PF Jet===
(136033-141949) /JetMETTau/Run2010A-Apr21ReReco-v1/AOD
(141950-145761) /JetMET/Run2010A-Apr21ReReco-v1/AOD
(145762-149442) /Jet/Run2010B-Apr21ReReco-v1/AOD
(160404-163869) /Jet/Run2011A-May10ReReco-v1/AOD
(165088-168437) /Jet/Run2011A-PromptReco-v4/AOD
(170053-172619) /Jet/Run2011A-PromptReco-v5/AOD
(172620-173692) /Jet/Run2011A-PromptReco-v6/AOD

```

These datasets are centrally produced for use in CMS physics analyses.

4.2.3 Monte Carlo Datasets

We use CMS Summer11-PU_S3_START42_V11-v2/AODSIM MC samples generated with PYTHIA [?]:

For PYTHIA predictions for the dijet angular ratio in fully reconstructed QCD, we use samples of 20k-2M events generated in 20 bins of \hat{p}_T :

```

/QCD_Pt-XtoY_TuneZ2_7TeV_pythia6/

```

For predictions of the dijet angular ratio in the excited quark events, we use 100k events for each resonance with $MASS = 700, 1200, 2000, 3500$ TeV:

```

/QstarToJJ_M-MASS_TuneZ2_7TeV_pythia6/

```

Although our acceptance is different, the line shape of resonances is essentially unchanged with respect to that used in [?]. This constancy is due to the ratio of inner and outer dijets from resonance decays, which is almost perfectly flat across dijet masses. In this analysis we can explicitly make the approximation of an identical lineshape in inner and outer events for resonances. See section ??.

4.3 Jet Reconstruction

We reconstruct jets using the anti-kT algorithm [?] with size parameter $\Delta D \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.7$. Calorimeter towers are the input taken by anti-kT as jet constituents. They are built from reconstructed energy deposits in one HCAL cell and in the ECAL crystals before it, in a projective geometry. The four-vector of a calo tower has an energy which is the sum of its hits. The directional component of a calo tower four-vector points from the interaction vertex to the energy-averaged center of the calorimeter cells in the tower. Four-vectors of calo towers are taken to be massless.

The reconstructed jet energy E is defined as the scalar sum of the calorimeter tower energies clustered into the jet. The jet momentum \vec{p} is the corresponding vector sum: $\vec{p} = \sum E_i \hat{u}_i$ with \hat{u}_i being the unit vector pointing from the interaction vertex to the energy deposition E_i inside the jet. The jet transverse momentum p_T is the component of \vec{p} in the transverse plane. The E and \vec{p} of a reconstructed jet are corrected for the non-linear response of the calorimeter to a generated jet, as explained below.

Monte Carlo jets reconstructed without detector simulation (*genjets*) are the result of applying the same jet algorithm to the Lorentz vectors of stable generated particles' four-vectors. The jet corrections are chosen so that, on average, the p_T of a simulated, corrected jet is equal to the p_T of the corresponding genjet. The corrections used for this analysis are the CMS standard pileup (L1), relative (L2), and absolute (L3) jet corrections for η and p_T variation of the jet response. These corrections are based on MC and validated with data [?].

The dijet we are interested in is simply the object comprised of the two jets in an event with the greatest transverse momenta. Other objects in the event can be ignored. It is possible to conduct the analysis presented in this thesis with such an inclusive definition because the two-jet event topology is overwhelmingly dominant at the Large Hadron Collider. What we ignore is a negligible fraction of what we study.

The dijet mass is given by $m_{jj} = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2}$.

4.4 Event Selection

Following standard CMS practice we require at least one 'good' Primary Vertex in each event, placing requirements on the minimum number of tracks with a minimum quality associated to the vertex. To cut down on contributions from satellite bunch crossings and even beam-gas events from beyond the edges of the detector's nominal interaction volume, we also restrict our vertices to $|z| < 24$ cm.

We run our CMSSW 4.2.8 ntuple-producing EDAnalyzer on fully reconstructed datasets to produce a single Processed ROOT ntuple. In this step we select anti-KT 0.7 jets (both Calo jets and, for use in studies, Particle Flow jets) and apply their

jet corrections. We select events in data that have passed jet- p_T high-level trigger paths. With these processed ntuples we perform the final analysis.

Taking the ntuple described above as input, we select events for which each of the two leading jets lies within the barrel and endcap calorimeters ($|\eta| < 2.5$). This cut simplifies the analysis in several ways. Surviving dijets are measured with very little uncertainty on their energy and transverse momentum. Together with their central barrel sections, the ECAL and HCAL endcap calorimeters extend to $|\eta| = 3.0$ and $|\eta| = 2.5$. The cut on maximal $|\eta|$ limits the need to consider the separate systematic errors of the forward calorimeters. The two leading jets must each meet standard jet identification criteria as listed in Table ?? [?].

Table 4.1: Standard CMS jet identification requirements, applied separately so as to arrive at an independent event selection for calojets and PFjets.

Calo jets 'loose' ID
fHPD < 0.98
n90hits > 1
EMF > 0.01
PF jets 'tight' ID
NeutralHadronFraction < 0.90
NeutralEMF < 0.90
Constituents > 1
For $ \eta < 2.4$:
ChargedHadronFraction > 0
ChargedEMFraction < 0.99
ChargedHadronMultiplicity > 0

Further cuts are needed to classify events as inner or outer events, and the choice of these cut values directly influences our sensitivity to new physics and our ability to use the triggers available. In section ?? we explain how these cuts were chosen and then their ultimate effect on the use of data from our single-jet- p_T triggers is given in subsection ??.

4.4.1 Optimization of Dijet Angular Ratio Definition

This analysis consists of counting two types of events, binning them by dijet mass, and taking the ratio. We conducted an optimization of three parameters which together define the numerator and denominator of the dijet angular ratio.

Three likely schemata were considered for dividing the $\eta_1 - \eta_2$ phase space into signal-rich and background-enriched regions (where η_1 and η_2 refer to the pseudorapidities of the leading and next-to-leading jet in p_T .)

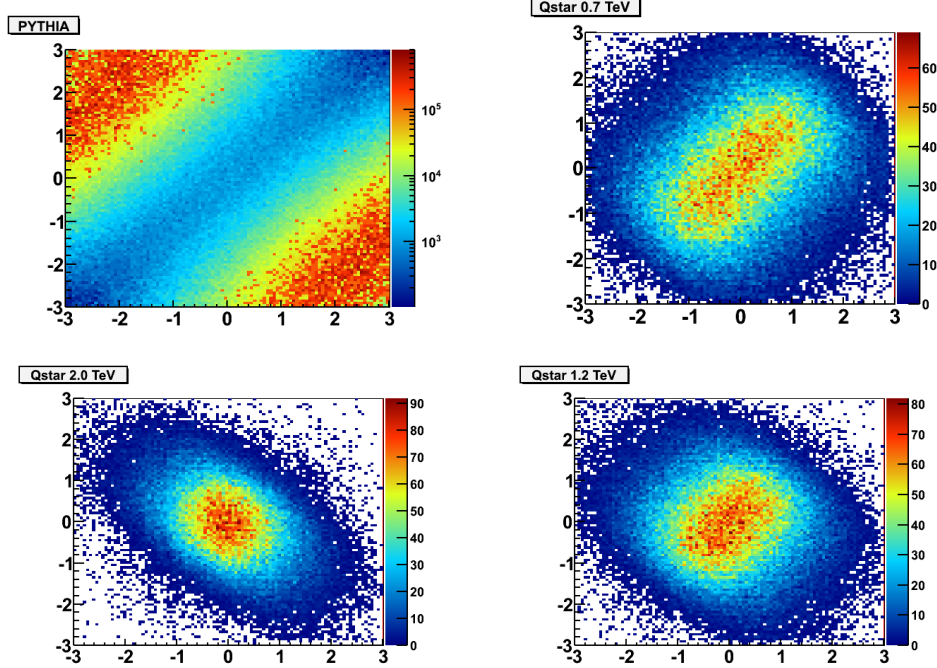


Figure 4.1: PYTHIA predictions for distributions of η_1 and η_2 in leading dijets. In the top left plot, QCD scattering is shown for dijet masses 500 to 6000 GeV. Clockwise around them from top right, massive, excited quarks are formed and decay. Shown are 0.7, 1.2, and 2.0 TeV excited quarks.

The first two schemata impose maximum values of $|\eta|$ on both jets of an inner dijet ($|\eta| < \eta_{inner}$), and for outer dijets require that both jets are in a more forward region than this ($\eta_{inner} < |\eta| < \eta_{outer}$). In the second schema we also disallow same- η -sign dijets in the outer region. Such same-sign events are rare in t-channel scattering such as QCD dijets, as Fig. ?? shows.

Indeed, Fig. ?? inspired our third schema, which retains a maximum $|\eta|$ for all jets, but defines an inner region for dijets whose constituent jets are separated by less than $\Delta\eta_{inner}$, and an outer region where $\Delta\eta_{inner} < |\Delta\eta| < \Delta\eta_{outer}$.

It would be computationally expensive to arrive at predicted limit-setting ability for points densely spanning the three-dimensional parameter spaces $\eta_{max}-\eta_{inner}-\eta_{outer}$ for the first two schemata and $\eta_{max}-\Delta\eta_{inner}-\Delta\eta_{outer}$ for the third schema. Instead, we

carried out a sensitivity optimization using a figure of merit described in Equation ??:

$$\text{FOM} \equiv \sum_{\text{bins}} \frac{R_{\text{bin}}^{\text{S+B}} - R_{\text{bin}}^{\text{B}}}{\sigma_{\text{bin}}^{\text{B}}} \quad (4.1)$$

where $R_{\text{bin}}^{\text{S+B}}$ is the value of the dijet angular ratio in the bin when both QCD and the resonance are present, $R_{\text{bin}}^{\text{B}}$ is the same without the resonance, and $\sigma_{\text{bin}}^{\text{B}}$ is the size of the error bar on $R_{\text{bin}}^{\text{B}}$ in the bin. A weighted statistic might be able to improve slightly on this selection, but is not investigated here.

This figure of merit was consistently higher by about a percent when vetoing same-side outer dijets than when including them, indicating improved sensitivity for such an $R_{|\eta|}$ search. Yet $R_{|\Delta\eta|}$ gave higher sensitivity still, and so it was chosen as the basis for this dijet resonance search.

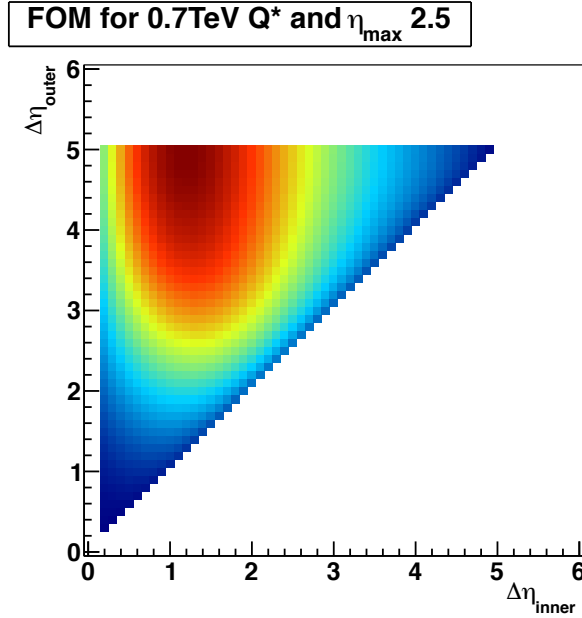


Figure 4.2: FOM value showing a maximum at $\Delta\eta_{\text{inner}} = 1.3$, for all values of $\Delta\eta_{\text{outer}}$.

We restrict our jets to be well contained within the CMS barrel and endcap, $|\eta| < 2.5$. The resulting sensitivity optimization study is summarized by Fig ??. For signals of 700 GeV excited quarks, our FOM is best for $\Delta\eta_{\text{inner}} = 1.2$ and $\Delta\eta_{\text{outer}} = 5.0$, that is, $\Delta\eta_{\text{outer}}$ so large that it places no constraint at all. As the mass of the resonance rises, the peak FOM for $|\eta| < 2.5$ moves to higher values of $\Delta\eta_{\text{inner}}$. At a resonance mass of 2.0 TeV, it has moved to 1.4 and the peak value of the FOM has dropped precipitously. We take this $\Delta\eta_{\text{inner}}$ parameter to be 1.3.

Trigger efficiency concerns drove the consideration of a non-maximal value $\Delta\eta_{outer}$. Because the $\Delta\eta_{inner}$ which maximizes our sensitivity FOM does not depend upon the choice of $\Delta\eta_{inner}$, we chose $\Delta\eta_{outer} = 3.0$, and accepted the trigger efficiencies which resulted. See section ??.

A still more sensitive search could be made by finely binning in both dijet mass and dijet $|\Delta\eta|$, and optimizing the cuts in this space for each resonance mass. We are interested here only in showing the power of the dijet angular ratio as defined above.

4.4.2 Dijet Mass Bins

The data are binned according to dijet mass. The bin boundaries are listed below in units of GeV. This is the same binning as that used in the dijet mass spectrum analysis [?], and we list only the dijet mass bins relevant to the present analysis.

The width of each dijet mass bin is chosen a priori to be approximately equal to the dijet mass resolution at that mass, based on MC studies, the same dijet mass binning used by CMS in [?]. The dijet mass resolution is parameterized with a fit to the difference of dijet mass found in position-matched genjets and reconstructed jets. Starting from a first bin with low edge of 1 GeV, we used this parameterization to obtain the bin edges for our analysis. The bin edges we ultimately used in this analysis are these, in GeV:

489, 526, 565, 606, 649, 693, 740, 788, 838, 890, 944, 1000, 1058,
 1118, 1181, 1246, 1313, 1383, 1455, 1530, 1607, 1687, 1770, 1856,
 1945, 2037, 2132, 2231, 2332, 2438, 2546, 2659, 2775, 2895, 3019,
 3147, 3279, 3416, 3558, 3704, 3854, 4010, 4171, 4337, 4509, 4686,
 4869, 5058, 5253, 5455, 5663, 5877, 6099, 6328, 6564, 6808, 7060

Consequences of trigger prescales and the difficulties of modeling at low dijet mass (see below) drive the choice of the bin edge lists' beginning. Its end point reflects the low statistics of very high-mass dijets.

4.4.3 Dijet Mass Requirement and Trigger Efficiency

Trigger inefficiencies have an impact on the dijet ratio because the jet triggers are more efficient for the inner region than the outer region, at a given dijet mass. Thus trigger inefficiency can bias the dijet ratio upward. We minimize this effect by placing a minimum on the dijet mass accepted from each trigger. In doing so we require that the outer dijets are triggered with more than 99.9% efficiency by the firing trigger. The efficiency of triggering on inner dijets is still higher. The resulting wiggle room for possible trigger bias in the measured dijet angular ratio is therefore less than 0.1%.

Table ?? summarizes Figs. ??-??, showing the jet triggers' turn-on curves for outer dijets. We include an event in the final dataset if the dijet mass of the event is above the fully efficient mass of any firing trigger. In this way the trigger is prevented from generating an upward bias in the angular ratio, which could give a false signal.

This procedure was carried out independently for two CMS standard jet reconstructions, calojets and Particle Flow jets [?]. Each has their charms, but only one wins our hearts, as we discuss in section ??. Calo jets were used in the trigger throughout data taking. There are inevitably discrepancies between calo and PF jets, as the slightly softer PF turn-on curves attest.

Table 4.2: Dijet mass bin at which Jet P_T triggers first exceed 99.9% efficiency.

Trigger Name Scale (GeV)	Fully Efficient Bin (GeV) PF jets	Fully Efficient Bin (GeV) Calojets
HLT _{jet} 370	2231	2132
HLT _{jet} 300	1770	1770
HLT _{jet} 240	1455	1313
HLT _{jet} 190	1118	1058
HLT _{jet} 150	838	788
HLT _{jet} 110	740	606
HLT _{jet} 180U	1455	1313
HLT _{jet} 140U	1118	1058
HLT _{jet} 100U	890	788
HLT _{jet} 70U	693	606
HLT _{jet} 50U	526	453

4.4.4 Data Quality

We present here the distributions of several key quantities in our final data. Figs. ??-?? show distributions of properties of jets passing all cuts, and give us confidence that they represent physical jets rather than detector noise.

The distribution of jet η and transverse momentum are shown in Fig.?? and Fig.?.?. These are also free from unexpected features, and indicate a healthy dataset.

Detailed Pseudorapidity Distributions

We examine the data in fine bins of η and $\Delta\eta$. In order to keep statistical power uniform, we merge contiguous mass bins to make several coarse dijet mass bins of

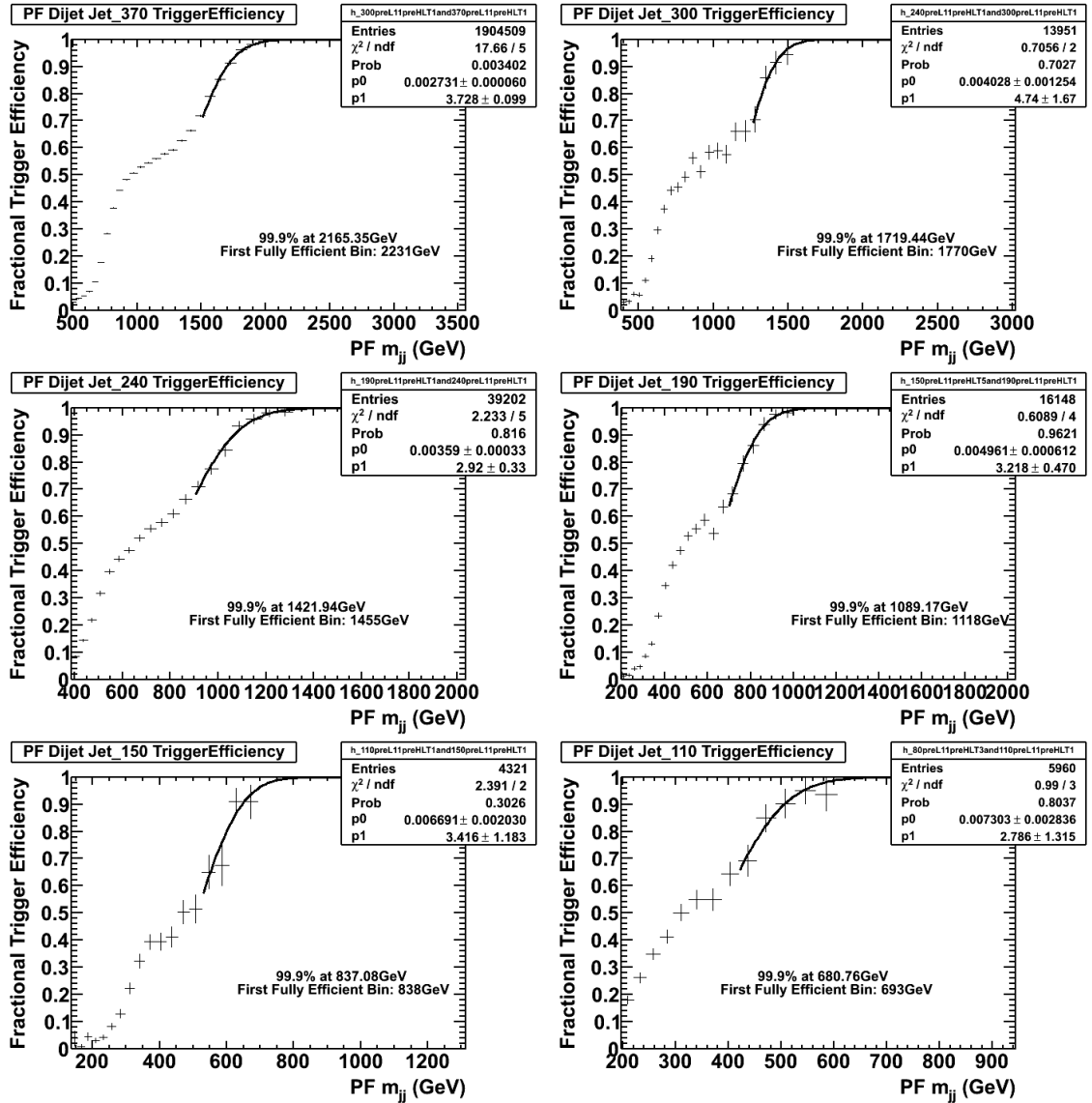


Figure 4.3: Efficiencies of the single-jet p_T trigger versus dijet mass for the outer regions of pseudorapidity $1.3 < |\Delta\eta| < 3.0$. 99.9% efficiency, and the dijet mass where the next bin begins, are given on the plots.

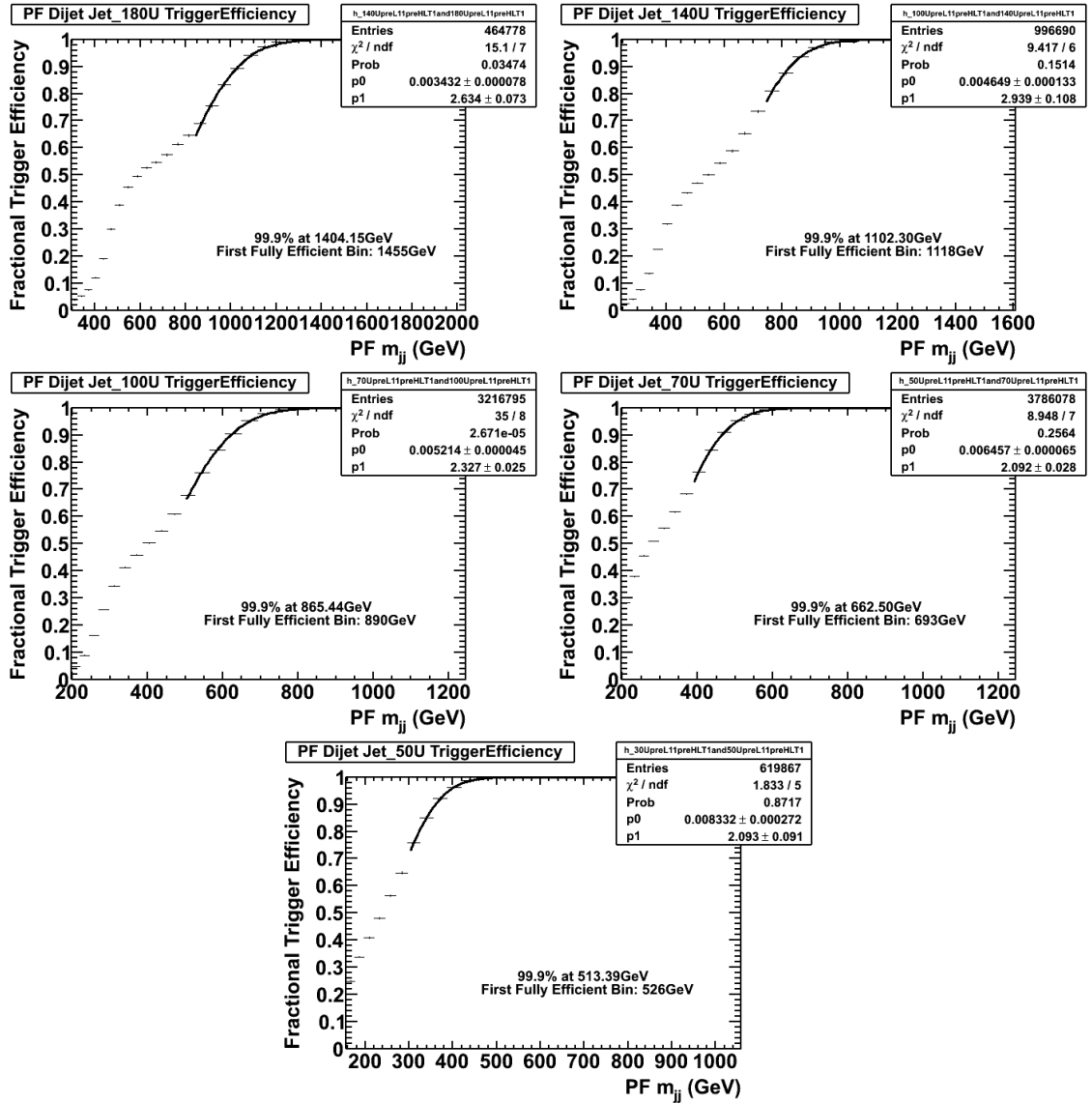


Figure 4.4: Efficiencies of the uncorrected single-jet p_T trigger versus dijet mass for the outer regions of pseudorapidity $1.3 < |\Delta\eta| < 3.0$. 99.9% efficiency, and the dijet mass where the next bin begins, are given on the plots.

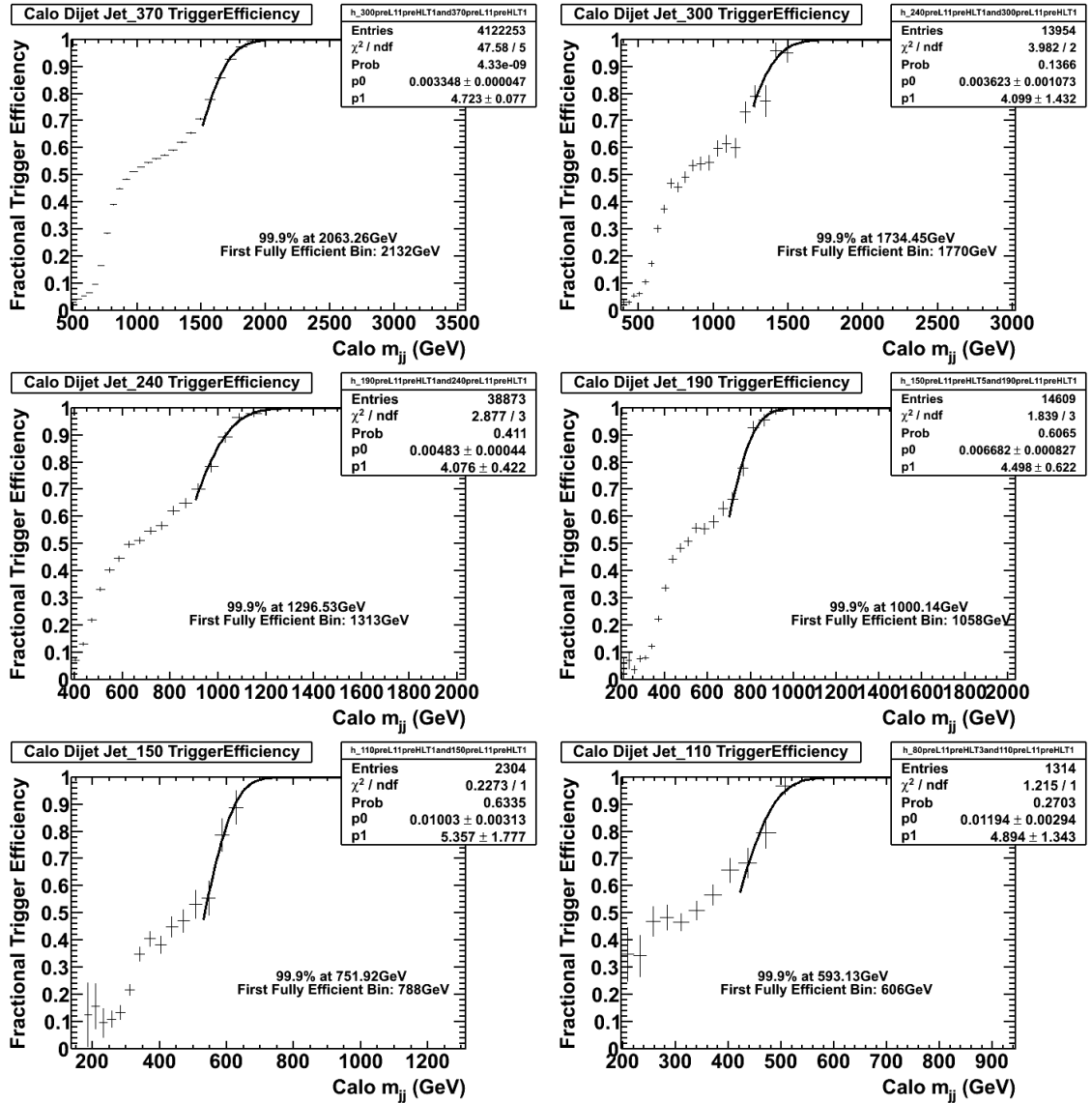


Figure 4.5: Efficiencies of the single-jet p_T trigger versus calo dijet mass for the outer regions of pseudorapidity $1.3 < |\Delta\eta| < 3.0$. 99.9% efficiency, and the dijet mass where the next bin begins, are given on the plots.

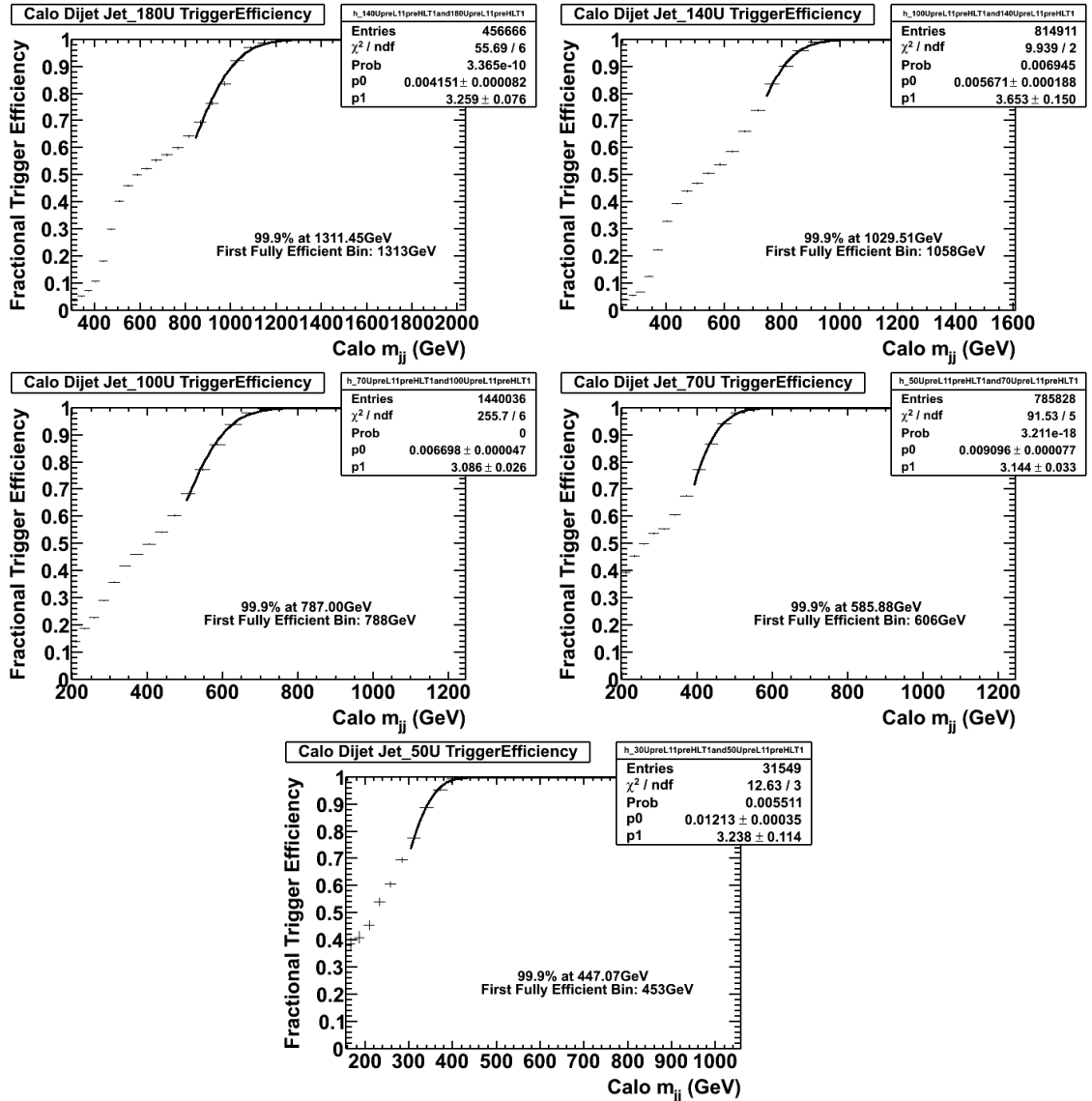


Figure 4.6: Efficiencies of the uncorrected single-jet p_T trigger versus calo dijet mass for the outer regions of pseudorapidity $1.3 < |\Delta\eta| < 3.0$. 99.9% efficiency, and the dijet mass where the next bin begins, are given on the plots.

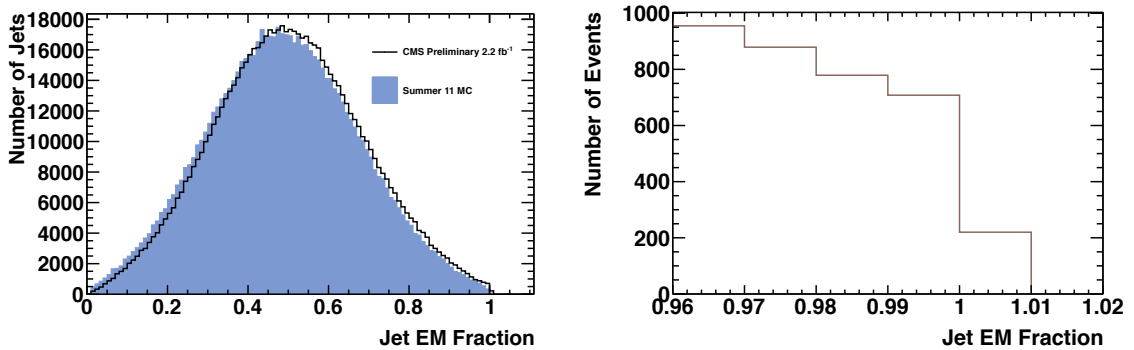


Figure 4.7: Distribution of the jet electromagnetic fraction, among jets comprising dijets used in this analysis. The distribution is smooth and reveals no noise-like features. Photons reconstructed as jets, and jets which have fluctuated into mostly π^0 s, contribute to the bin at $\text{EMF}=1$.

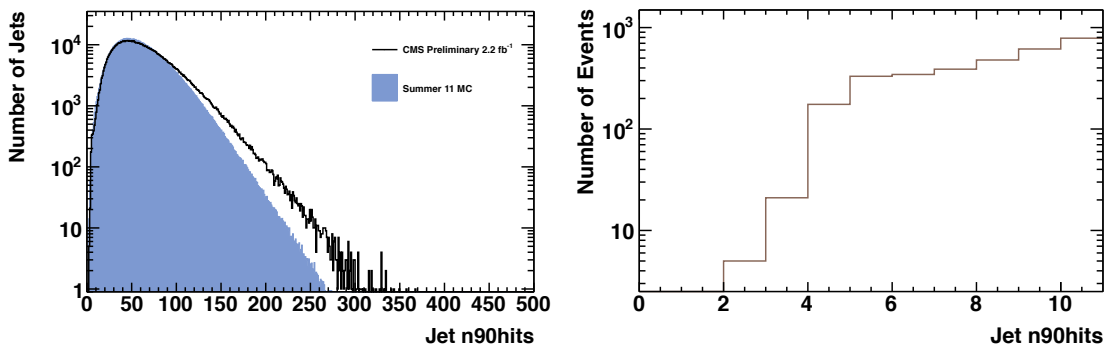


Figure 4.8: Distribution of the jet ID variable $n90$, the minimum number of calorimeter towers which contain 90% of the jet's total energy. The distribution is smooth and reveals no noise-like features, even near the cut at 1.

approximately equal event counts. One of these just so happens to be the low-mass sideband discussed in ??.

While the jet η distributions compare well to Monte Carlo results (see Fig. ??), there is a slight under-prediction of small- $|\Delta\eta|$ with respect to large- $|\Delta\eta|$ in PYTHIA6 when comparing to data, present in all mass ranges. This is shown in Fig. ?. In the dijet angular ratio, this is manifest as a systematic under-prediction of the ratio in PYTHIA which we handle by shifting the PYTHIA prediction to fit a low-mass sideband. The discussion is found in ?. Leading-order simulation such as this lacks hard gluon emission, and does not fully reproduce the angular distributions of data. Next-to-leading order predictions almost completely fix the disagreement, as

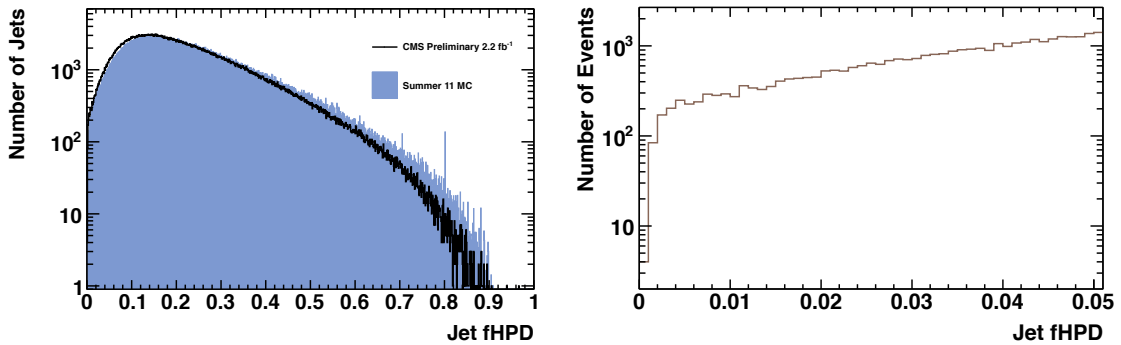


Figure 4.9: Distribution of the jet ID variable $fHPD$, the maximum fraction of a jet's energy which is contained within any single photodetector of the HCAL. The distribution is smooth and reveals no noise-like features.

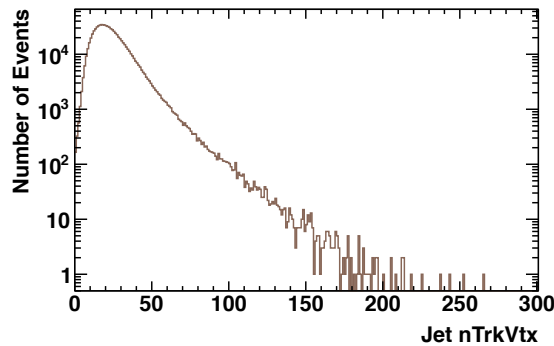


Figure 4.10: The distribution of the number of tracks within jets at the associated vertex. The distribution is smooth and reveals no noise-like features, such as a spike at very low numbers of tracks.

we discuss in ??.

4.5 Dijet Angular Ratio Predictions

Figure ?? gives a pictorial motivation for using a dijet angular ratio to search for dijet resonances. With only standard model processes in effect, the dominant two-to-two parton scattering gives rise to jets with high absolute pseudorapidities. The absolute difference $|\Delta\eta|$ is typically large, clustering away from the line $\eta_1 = \eta_2$. By contrast, if the parton scattering is mediated by a massive resonance, the individual pseudorapidities have smaller absolute values, and this is all the truer for their

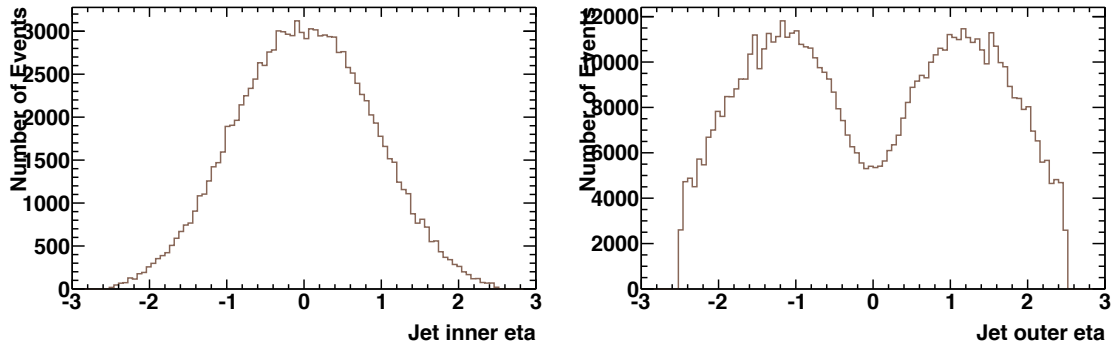


Figure 4.11: The distribution of pseudorapidity for jets comprising inner or outer dijets.

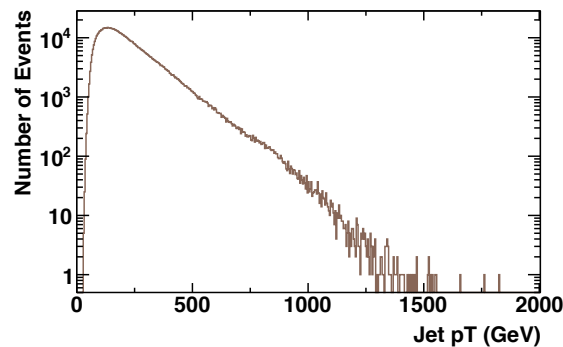


Figure 4.12: The distribution of transverse momentum for jets used in this analysis. The histogram is smooth and steeply falling, indicating physics-dominated events.

difference.

We obtain predictions for the dijet angular ratio for the SM alone and with new, resonant physics using PYTHIA plus full detector simulation. We find that the dijet ratio in generated jets (*genjets*) is compatible with that of fully simulated jets (*calojets*). We use *genjets* to make our parameterized background expectation, since we are able to generate larger samples of these than of *calojets*. The good level of agreement between *genjets* and *calojets* is described in detail in ??.

4.5.1 QCD Prediction from Pythia

Two-to-two parton scattering events were simulated in PYTHIA6 using CTEQ6 PDFs [?]. In addition to PYTHIA's lowest-order matrix element calculations, the effects of parton showering, initial- and final-state radiation, multiple parton

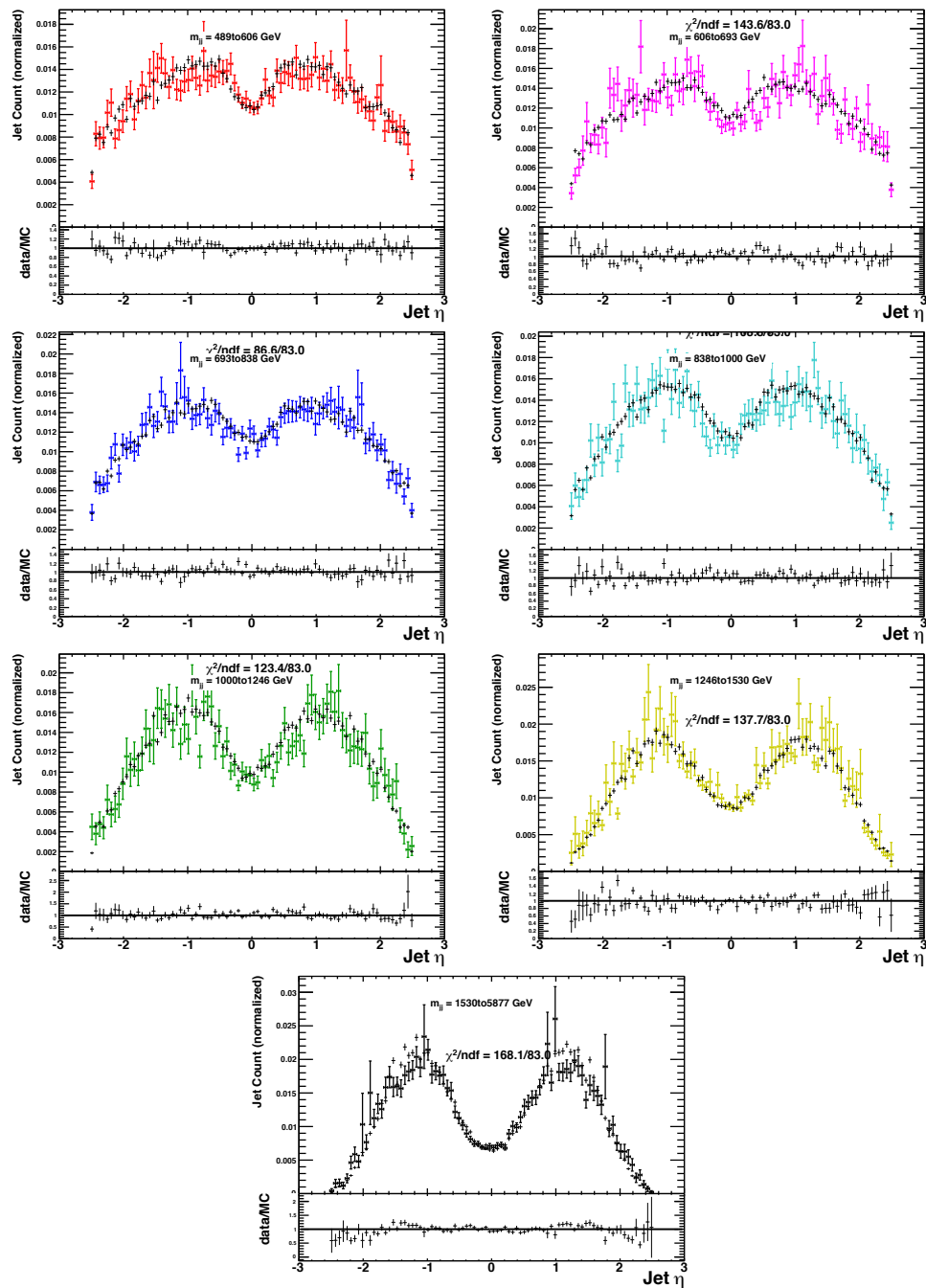


Figure 4.13: The area-normalized distribution of pseudorapidity for jets comprising dijets. Generator-level PYTHIA6 results are overlaid in fine black markers, and the ratio of data/MC presented below.

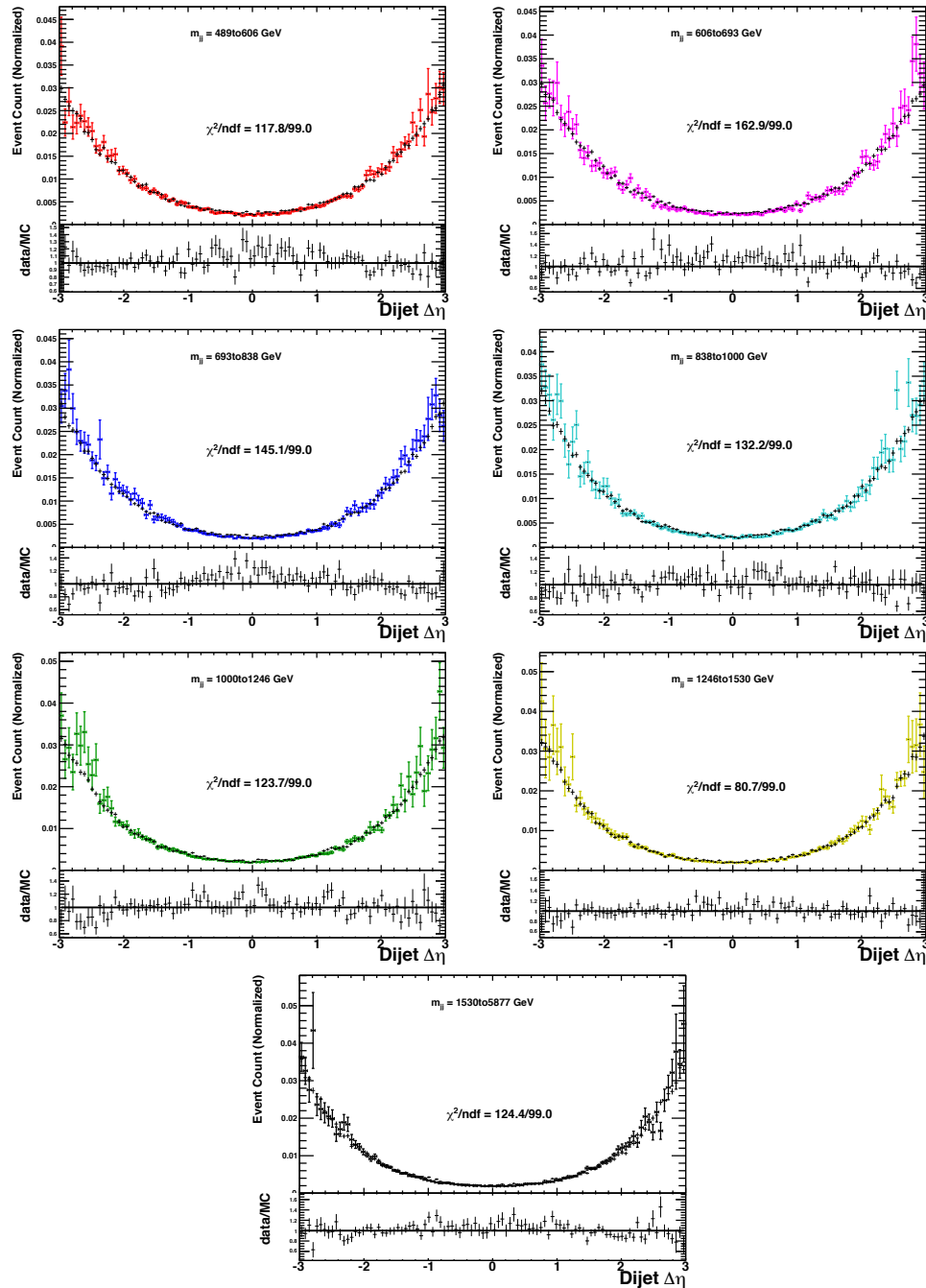


Figure 4.14: The area-normalized distribution of pseudorapidity differences between jets comprising dijets. Generator-level PYTHIA6 results are overlaid in fine black markers, and the ratio of data/MC presented below. There is a slight over-prediction of small- $|\Delta\eta|$ over large- $|\Delta\eta|$ in PYTHIA6 with respect to observation.

interactions, and hadronization were simulated. The underlying event simulation used the parameter tune known as Z2. The resulting differential dijet mass spectra, together with their ratio, are show in Fig. ???. The form of the fit function used is

$$\frac{d\sigma}{dm} = \frac{p_0(1 - y + p_3y^2)^{p_1}}{y^{-p_2-p_4 \ln y}}(p_5 + p_6m_{jj} + p_7m_{jj}^2) \quad (4.2)$$

where $y \equiv \frac{m_{jj}}{\sqrt{s}}$. The form of this fit has an illustrious ancestry [?] [?] beginning with the UA2 collaboration. There the power law in mass was introduced, $p_0/m_{jj}^{p_1}$, inspired by the matrix element of the hard scatter [? ?]. At CDF polynomial mass dependence $(1 - m_{jj}/\sqrt{s})^{p_2}$ was introduced [? ?], borrowed from the PDFs' mass dependence at their average momentum fraction. The term p_3y^2 was brought in to help agreement to the data in the second of these papers. Good agreement with still larger dijet mass ranges has necessitated more terms being used, starting an exponentiated power series in $\ln(m_{jj}/\sqrt{s})$.

Here we have continued that tradition with our polynomial in the dijet mass $p_3y^2)^{p_1}$. This gives a $\chi^2/ndf = 67.35/53$ to our leading-order prediction. We note that the form

$$\frac{d\sigma}{dm} = \frac{p_0(1 - y + p_3y^2)^{p_1}}{y^{-p_2-p_4 \ln y + p_5 \ln^2 y + p_6 \ln^3 y + p_7 \ln^4 y}} \quad (4.3)$$

can also be used, attaining $\chi^2/ndf = 62.06/53$ over the same mass range 270-6500 GeV. The two fit functions have the same parameter count. The difference between this fit and the one used in this analysis is less than one or two thousandths of the fit value everywhere. This difference is within the scatter of the monte carlo prediction about the fit.

Parameterized fits to the inner dijet and dijet angular ratio predictions are used by the limit-setting procedure discussed in ???. The rate of QCD events is determined by the fit to PYTHIA6 predictions in Fig. ???. It extends through fifteen order of magnitude, describing the leading order Monte Carlo results continuously.

4.5.2 Dijet Resonance Predictions from Pythia

We consider dijet resonances from models of excited quarks (q^*) [?]. The choice of parameterization for the dijet angular ratio in the dijet resonances is driven by the parameter for which we will set limits: the cross section for $|\Delta\eta| < 3.0, |\eta_{1,2}| < 2.5$ for each resonance $\sigma_{|\Delta\eta|<3.0}^R$. We write the inner and outer spectra for the sum of the resonance and QCD in terms of this observable:

$$\frac{d}{dm}\sigma_{\text{out}}^{\text{sum}}(m) = \frac{d}{dm}\sigma_{\text{out}}^{\text{QCD}}(m) + \frac{A(M_R)}{R(M_R)} \cdot \sigma_{|\Delta\eta|<3.0}^R \cdot \frac{d}{dm}\mathcal{P}(m), \quad (4.4)$$

$$\frac{d}{dm}\sigma_{\text{in}}^{\text{sum}}(m) = \frac{d}{dm}\sigma_{\text{in}}^{\text{QCD}}(m) + A(M_R) \cdot \sigma_{|\Delta\eta|<3.0}^R \cdot \frac{d}{dm}\mathcal{P}(m), \quad (4.5)$$

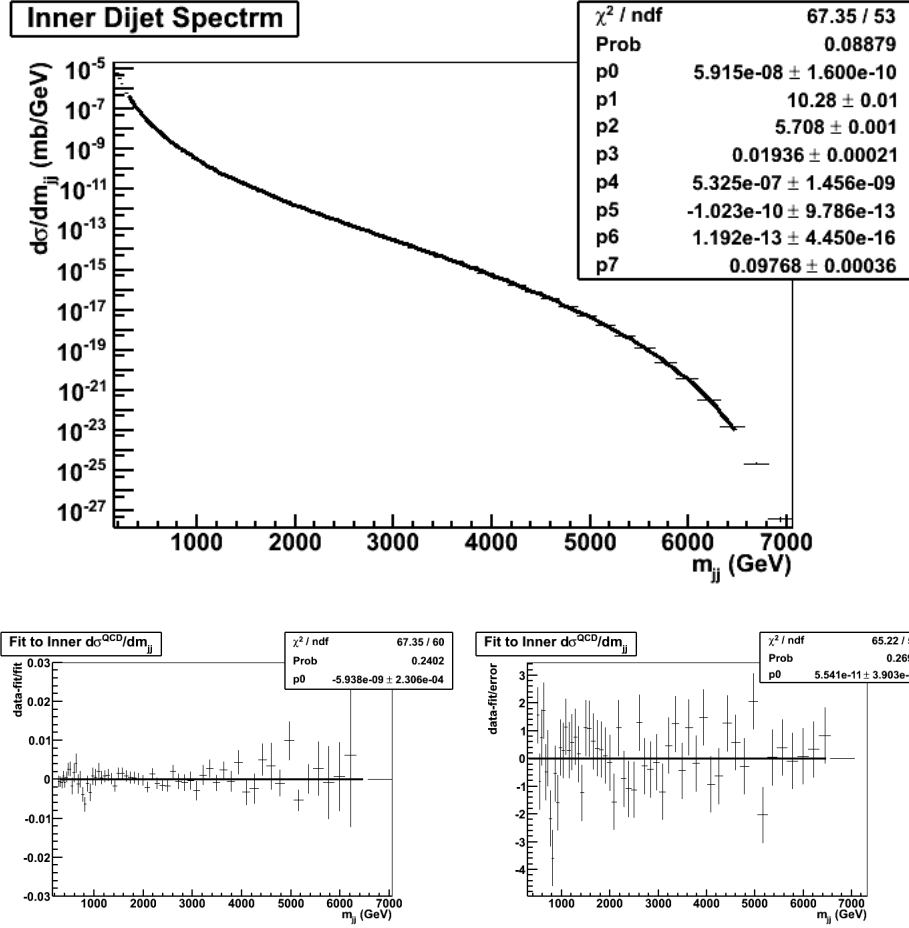


Figure 4.15: The fitted inner differential mass spectrum. Below it, the fractional difference of the histogram and the fit, and the same difference divided by the uncertainty for each bin.

where m is the dijet mass, M_R is the resonance mass, $R(M_R)$ is the average ratio for resonance R of mass M_R for the dijet resonance:

$$R(M_R) = \frac{\sigma_{|\Delta\eta|<1.3}^R}{\sigma_{1.3<|\Delta\eta|<3.0}^R}, \quad (4.6)$$

$A(M_R)$ is the acceptance of the inner $\Delta\eta$ region divided by the total inner plus outer $|\Delta\eta| < 3.0$ region for resonance R of mass M_R :

$$A(M_R) = \frac{\sigma_{|\Delta\eta|<1.3}^R}{\sigma_{|\Delta\eta|<3.0}^R}, \quad (4.7)$$

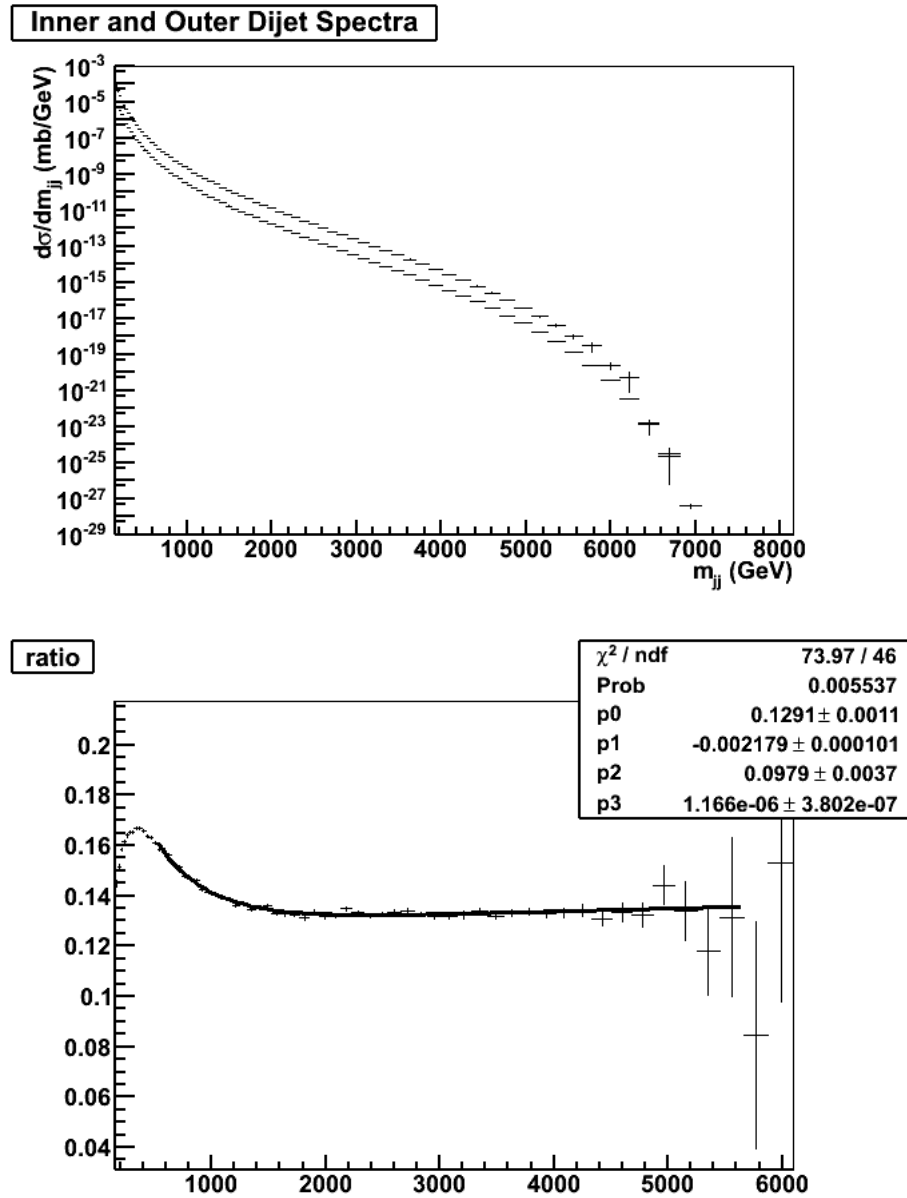


Figure 4.16: The differential mass spectra of inner (lower) and outer (upper) dijets. Also shown is their ratio, which is the basis for our background expectation in this search, used in the likelihood definition.

and $\frac{d}{dm}\mathcal{P}(m)$ is the lineshape for each resonance. The lineshapes come from PYTHIA MC plus full CMS simulation and in general depend primarily on resonance mass, width, and parton content. The inherent width is taken to be small in relation to the dijet mass resolution, leaving only the mass and parton content to have an impact on the lineshape of resonances.

We allow the QCD-plus- M_R -resonance model to differ from the QCD model only for $0.3 < \frac{m}{M_R} < 1.3$. On the high side of this range, the tail of the resonance falls more slowly than the QCD spectrum causing a significant enhancement of the dijet ratio. The shape of the ratio in this tail region is affected by interference between the resonance and QCD amplitudes, but this interference is not included in the PYTHIA treatment of these events. To be conservative we exclude the tail regions. They include less than 1% of the resonance signal, yet would yield substantive discriminatory power based on poorly modeled signal shapes. We show $R(M_R)$ values for q^* at several masses in Figs. ??

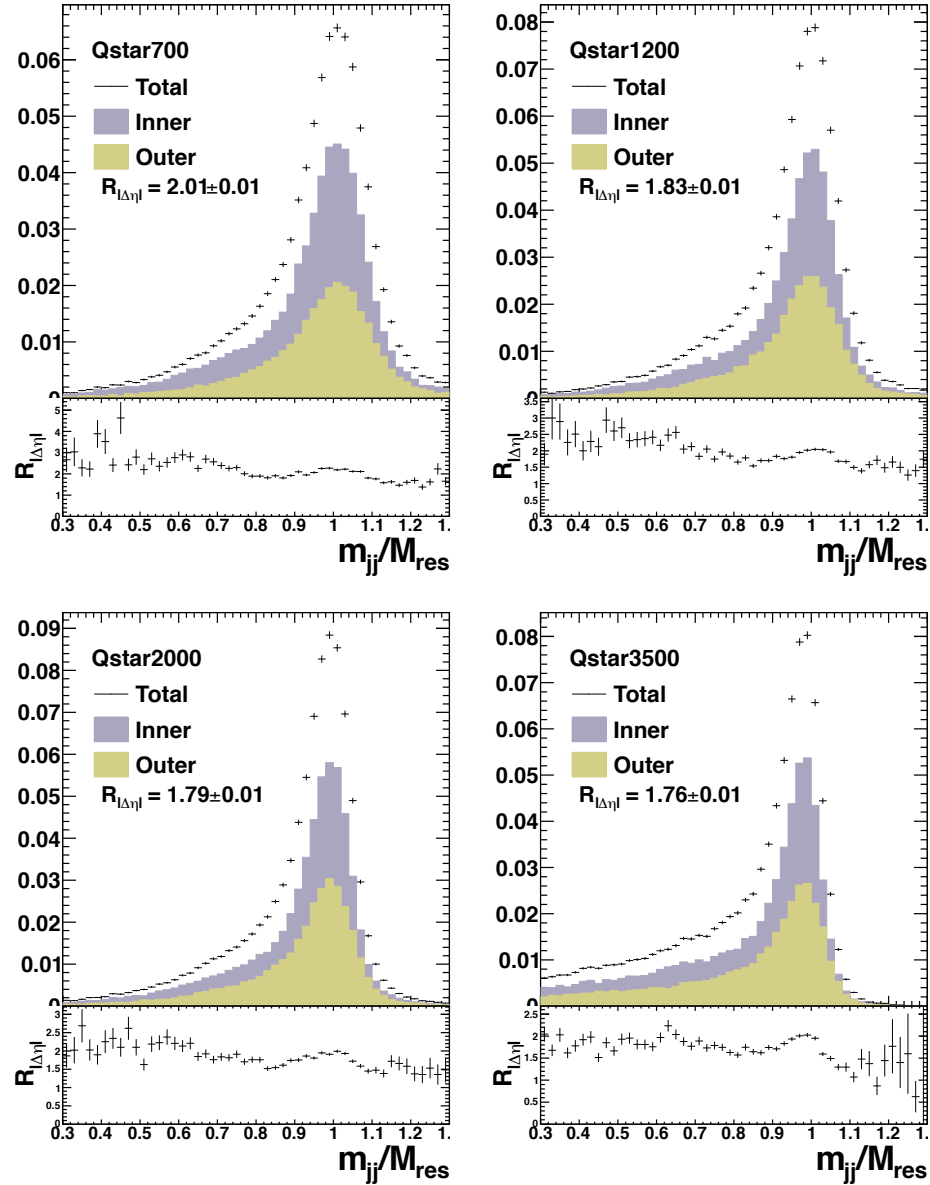


Figure 4.17: The lineshapes for 700 GeV, 1200 GeV, 2000 GeV, and 3500 GeV excited quarks in dijet mass as a fraction of resonance mass. Shown with crosses is the total lineshape, and filled histograms are overlaid separately for the inner and outer events. Inner events outnumber outer events. Below each is the dijet angular ratio, and a fit of the dijet angular ratio to a constant line in these signal-only events is given in the legend.

4.6 Dijet angular ratio in Data

4.6.1 Choice of Reconstruction Method

We considered both calojet and PFjet reconstructions. These two reconstructions give consistent values of the $R_{|\Delta\eta|}$, as shown in Fig.?? and Fig.?. With their steeper trigger turnons and lower-mass fully efficient points, calojets increase the analysis's statistical power. We choose to proceed with calojets. The event yield from both calo and PF jet reconstructions after taking into account these slight differences of trigger efficiency are show in Fig. ??.

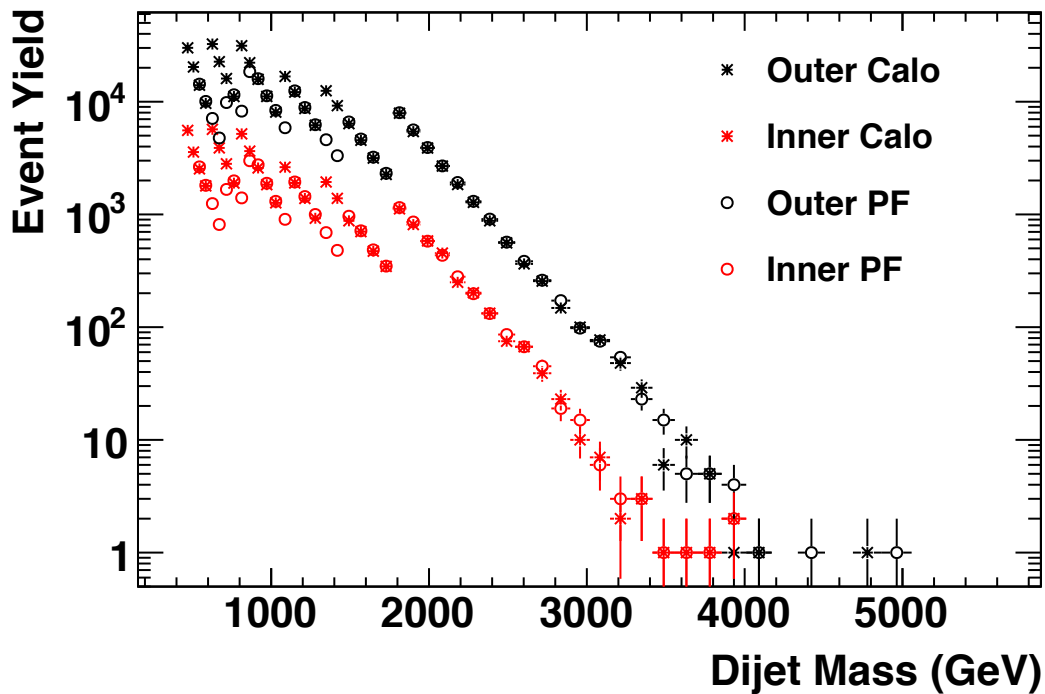


Figure 4.18: The event yield of inner and outer events for calo and PF dijets. In the low-mass region, jet triggers are reaching their fully efficient (and hence, unbiased) dijet masses, determined independently for calo and PF dijets. These come later for PF dijets, lessening their statistical power.

4.6.2 Measured Dijet angular ratio

The event yield from calojets is shown in Fig. ??, along with the smooth spectrum after accounting for the effect of trigger prescales. The ratio $R_{|\Delta\eta|}$ in our calojet

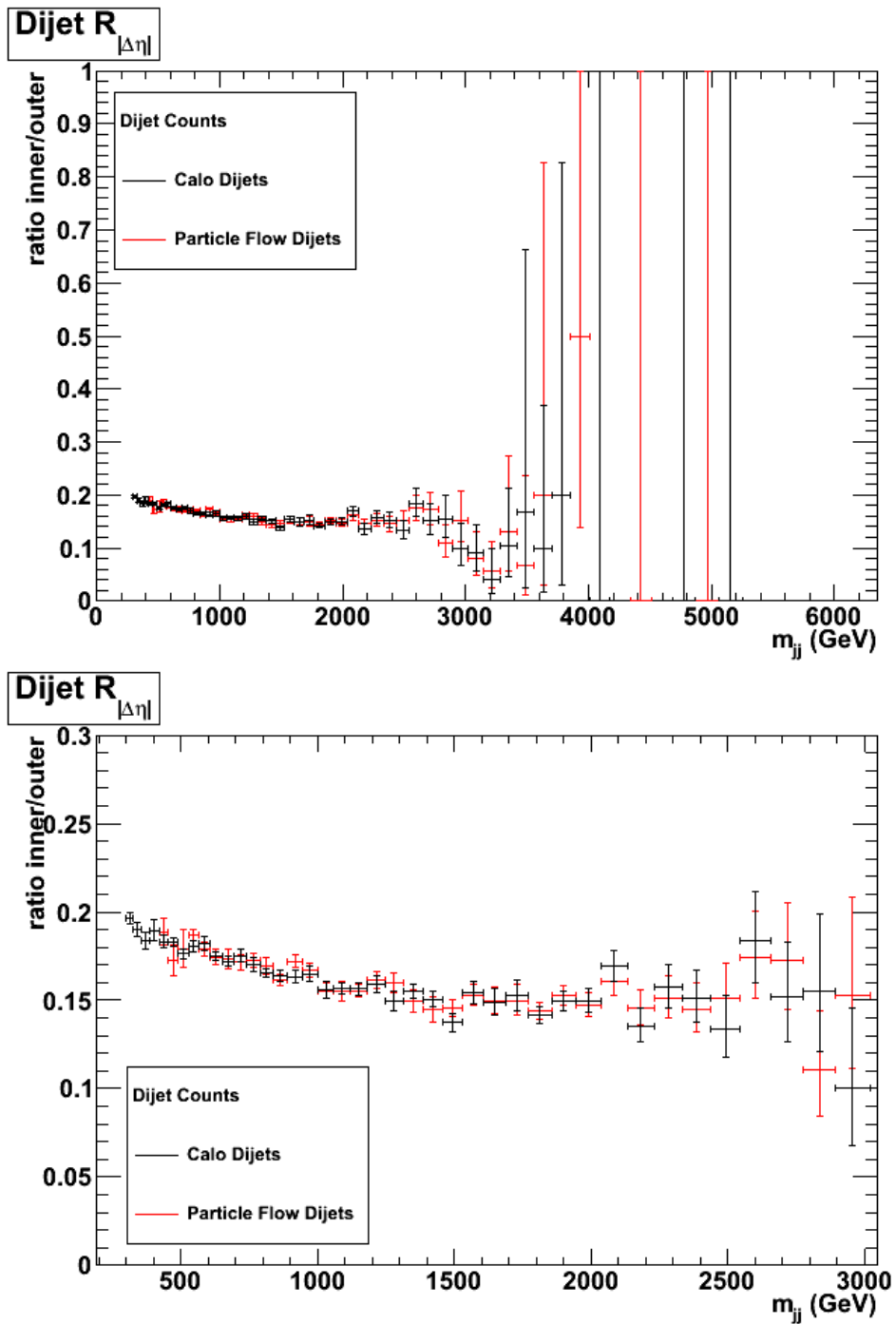


Figure 4.19: The measured dijet angular ratio in calo and PF dijets. Shown are the asymmetric error bars for the ratio of Poisson means. The measured $R_{|\Delta\eta|}$ in these two is compatible, although the calojet reconstruction offers better statistical power.

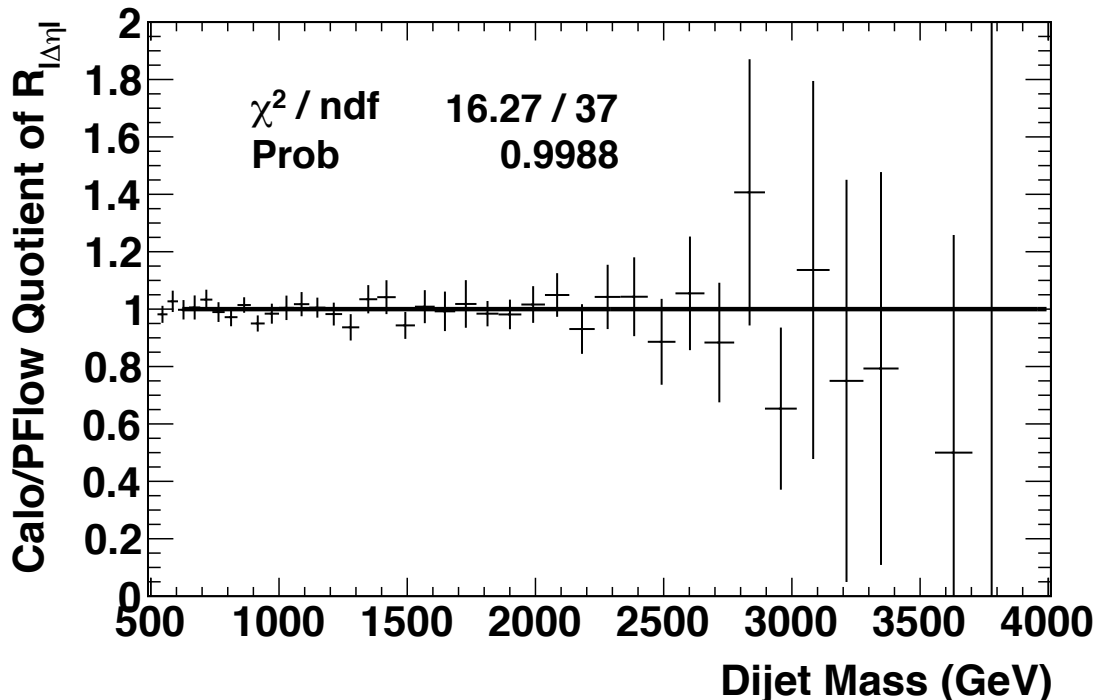


Figure 4.20: The quotient of the dijet angular ratio in PF and Calo reconstructions. The same dataset is used in both.

sample gives the overlaid data in Fig. ???. The 68% confidence intervals are shown by the error bars. These follow the Clopper-Pearson prescription [?] for the uncertainty of measurements of the ratio of two counted quantities.

The QCD expectation, derived from MC as described in section ??, is given an additive offset of 0.0188 ± 0.00186 which gives the best fit in the low-mass sideband region. The MC predictions are only considered after this additive offset is applied. For masses above the sideband, the data scatter about the MC, as detailed in Fig.???. The same treatment is shown for PF dijets, in Figs. ??-??, for completeness.

We proceed with our analysis using calojets only, for the reasons given above.

4.6.3 Highest-Mass Events

We examined the ten inner events with the highest dijet masses and all the outer events having dijet masses at or above the least of these masses. These all appear to contain good-quality dijets. Displays of these events may be found in Appendix ???. Details of the ten highest-mass inner dijet events are shown in Table ???. The jet p_{TS} are well balanced, their azimuthal separation $|\Delta\phi|$ is roughly π , and their jet-ID

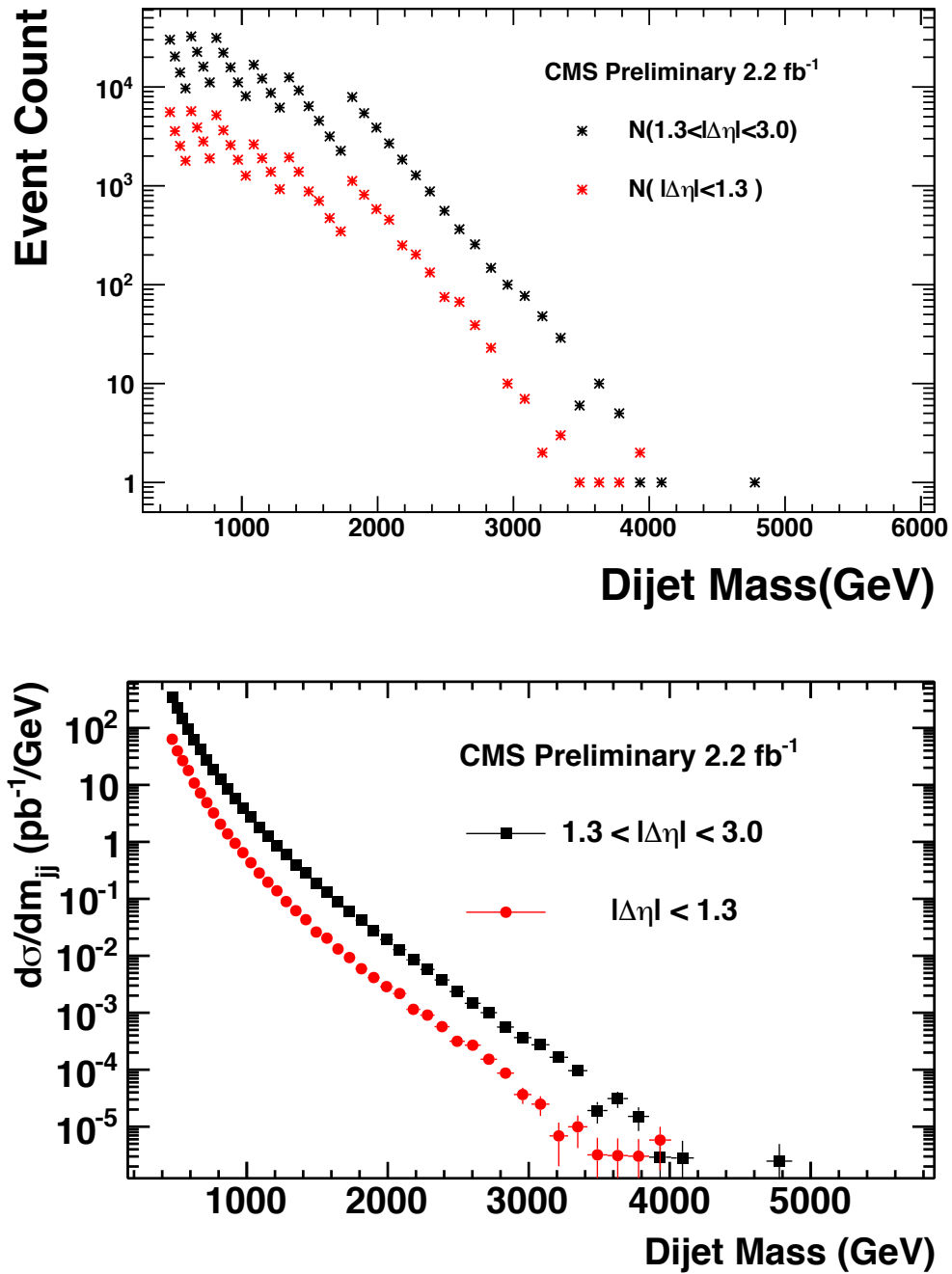


Figure 4.21: Top: The total event yield of numerator and denominator dijet events by dijet mass. The effect of prescales on the triggers listed in Table ?? is clearly present in the sawtooth appearance of these raw counts. Bottom: The differential dijet mass spectra given by these data.

quantities emf , $fHPD$, and $n90hits$ are typical of the distributions in Figs. ??-??. The missing transverse energy (MET) is a small fraction of the total transverse energy in each event. If large contributions from detector noise were present, these statements would no be possible.

4.6.4 Choice of Sideband Region

As we discuss in ??, our leading order QCD Monte Carlo is able to describe the shape of the $R_{|\Delta\eta|}$ in data very well, for dijets of 500 GeV and above. We take that mass as the lower bound on a sideband region which is not used as part of the signal discrimination. Rather, these few sideband bins are used to find the best fit for a constant, additive shift which we apply to our MC prediction at all dijet masses. (Scaling rather than shifting gives essentially the same result.) While a larger sideband region will better constrain the value of this additive shift, a smaller one will leave more dijet mass bins available for our search.

We present the resulting best-fit, additive shift values for several choices of sideband in Table ??. We choose to work with the three-bin sideband henceforward. This choice permits us to set cross section limits for the widest range of resonance masses allowed by the data. It does not appreciably affect the value of the offset, nor of its uncertainty.

These data are well described by the PYTHIA6 model after applying the additive shift taken from the sideband. In Fig. ?? we make a simple statistical test of the goodness of this comparison, showing the difference between data and the normalized QCD model.

Having found our low-mass sideband region, we can consider the data in comparison to our improved model of QCD. We can also add signal as from spin- $\frac{1}{2}$ qg resonances to see how readily those could be accomodated by our data. This is shown in Fig. ??.

Table 4.3: Physical characteristics of the ten inner events having the highest reconstructed dijet masses. There are no worrying indications of noise.

Run Event Lumi	m_{jj} (TeV) $p_{T,1}$ $p_{T,2}$	$ \Delta\eta $ η_1 η_2	$ \Delta\phi $ ϕ_1 ϕ_2	emf ₁ emf ₂	fHPD ₁ fHPD ₂	n90Hits ₁ n90Hits ₂	$\frac{MET}{\sum E_T}$	Good Vertex Count
167746	3.987	0.567	3.137					
385009283	1.825	0.38	-0.98	0.54	0.32	22	0.02	7
314	1.762	-0.19	2.16	0.33	0.19	95		
172208	3.868	1.286	3.134					
58850658	1.531	-0.63	2.00	0.20	0.64	20	0.02	7
71	1.409	0.66	-1.13	0.52	0.24	24		
171578	3.745	1.039	3.111					
487293629	1.536	-0.58	-2.53	0.74	0.13	32	0.02	5
479	1.521	0.46	0.58	0.53	0.31	24		
166895	3.641	1.166	3.126					
367873378	1.521	0.55	2.06	0.36	0.48	32	0.04	5
402	1.337	-0.62	-1.07	0.17	0.56	21		
170901	3.420	0.962	3.130					
22809456	1.469	-0.27	-1.53	0.44	0.47	15	0.02	5
29	1.358	0.69	1.62	0.63	0.27	25		
166781	3.359	0.607	3.112					
420884464	1.470	-0.33	-2.77	0.48	0.26	130	0.02	8
354	1.455	0.28	0.34	0.73	0.11	68		
173389	3.335	0.914	3.108					
373649287	1.445	0.48	1.20	0.42	0.32	43	0.05	10
291	1.345	-0.43	-1.91	0.36	0.33	37		
173198	3.332	0.148	3.105					
45863530	1.552	-0.37	-0.18	0.63	0.18	24	0.01	6
58	1.550	-0.22	2.92	0.50	0.34	26		
166033	3.217	1.180	3.133					
1441098465	1.281	-0.67	-0.05	0.17	0.72	11	0.03	6
1092	1.212	0.51	3.09	0.85	0.05	51		
172949	3.163	0.724	3.139					
1172865747	1.399	-0.65	-0.64	0.35	0.41	27	0.02	5
828	1.347	0.07	2.50	0.48	0.37	43		

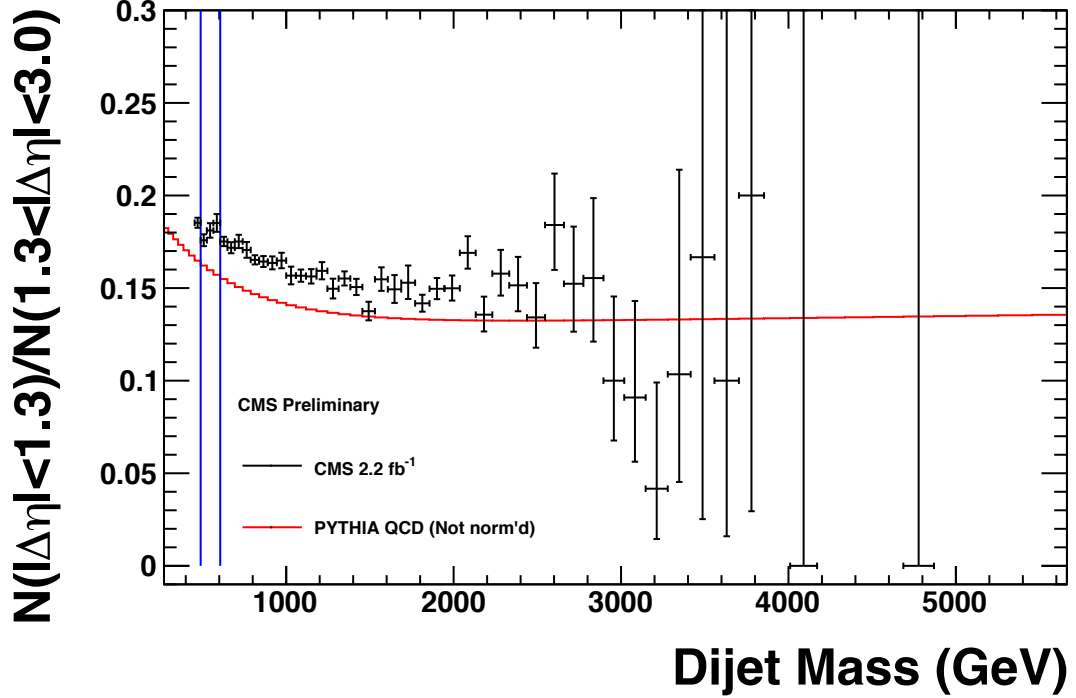


Figure 4.22: The measured dijet angular ratio in calo dijets. Shown are the asymmetric Clopper-Pearson error bars. The data are overlaid on the expectation from PYTHIA6 QCD simulation at generator level. The agreement of shape motivates us to apply an additive offset to the MC taken from the sideband region. The sideband region is bounded by blue, vertical lines. A single data point of $2.0^{+14.7}_{-1.66}$ in the bin below $4TeV$ is outside this plot.

Table 4.4: Choices of the sideband region, excluded from contributing to the signal discrimination of the analysis, but required to determine the best fit value of the additive shift applied to our QCD MC prediction for $R_{|\Delta\eta|}$.

No. Bins	Range m_{jj} (GeV)	Additive Shift
5	489 – 693	0.0194 ± 0.00147
4	489 – 649	0.0194 ± 0.00147
3	489 – 606	0.0188 ± 0.00186

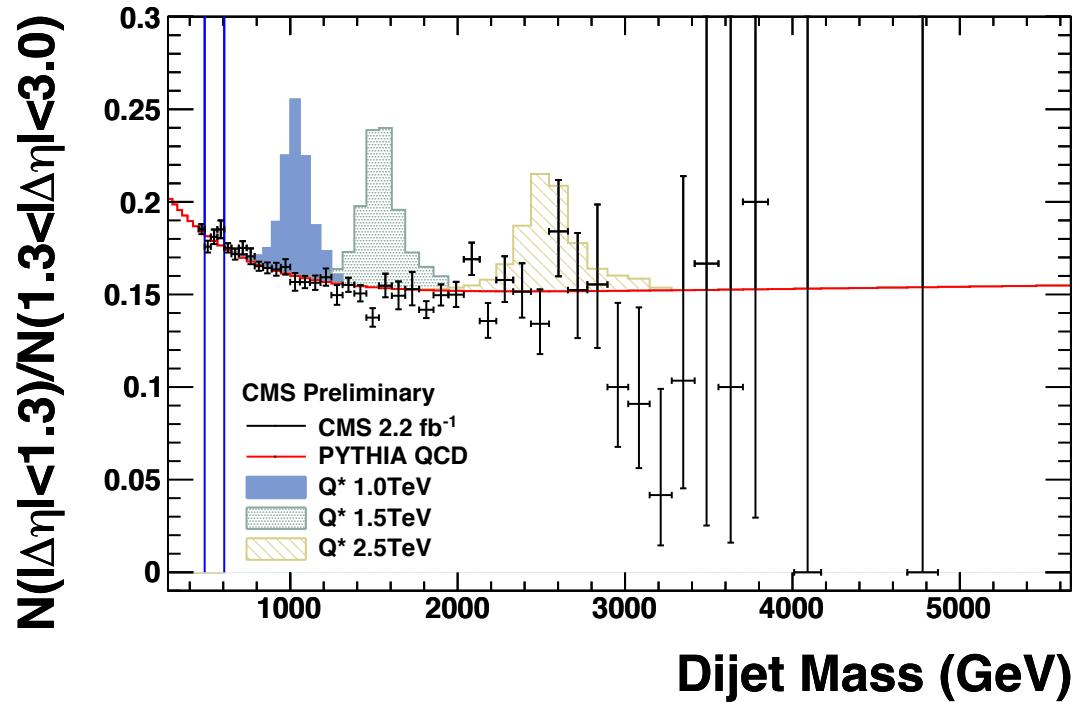


Figure 4.23: The measured dijet angular ratio in calo dijets. Shown are the asymmetric Clopper-Pearson error bars. The data are overlaid on the expectations for the scenarios QCD-only, and QCD plus excited quarks at three reference masses. A single data point of $2.0^{+14.7}_{-1.66}$ in the bin below 4 TeV is outside this plot. The sideband region is bounded by blue, vertical lines.

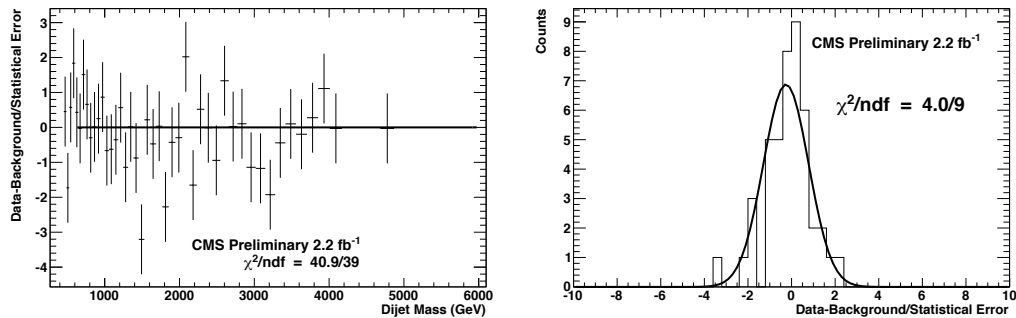


Figure 4.24: The “pulls”, data minus MC over statistical uncertainty, of the measured dijet angular ratio in calo dijets, with respect to the background expectation. Shown are the pulls fitted with a flat line; also shown is the distribution of the pulls fitted with a Gaussian. The mean is close to zero and the width is close to one.

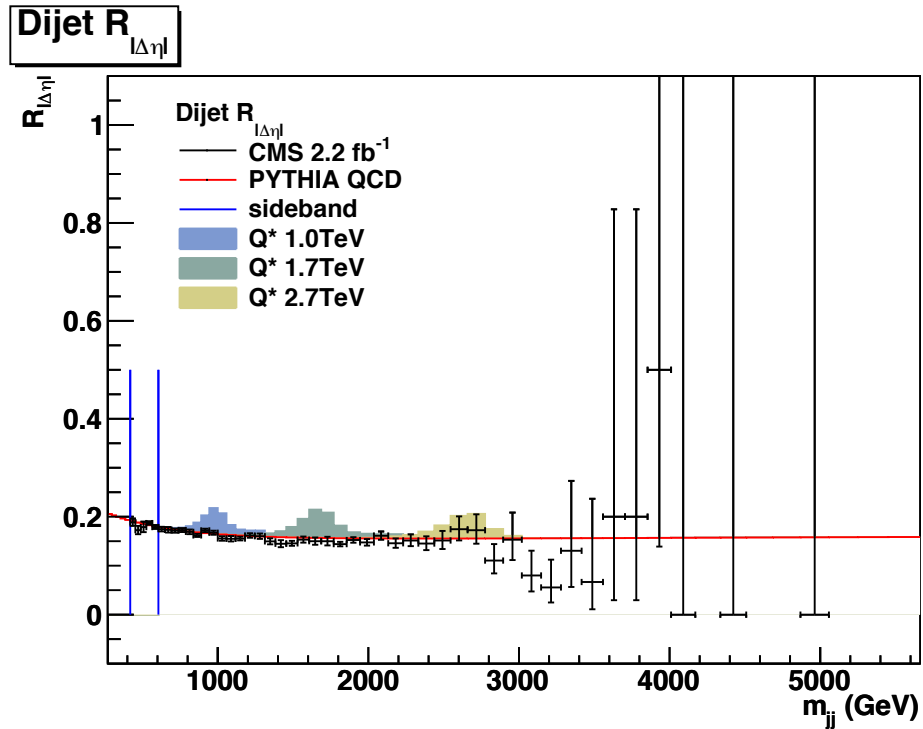


Figure 4.25: The measured dijet angular ratio in PF dijets. Shown are the asymmetric Clopper-Pearson error bars, which describe the uncertainty of a ratio of two Poisson means. The data are overlaid on the expectations for the QCD-only scenario as well as QCD-plus-excited-quarks at three excited quark masses.

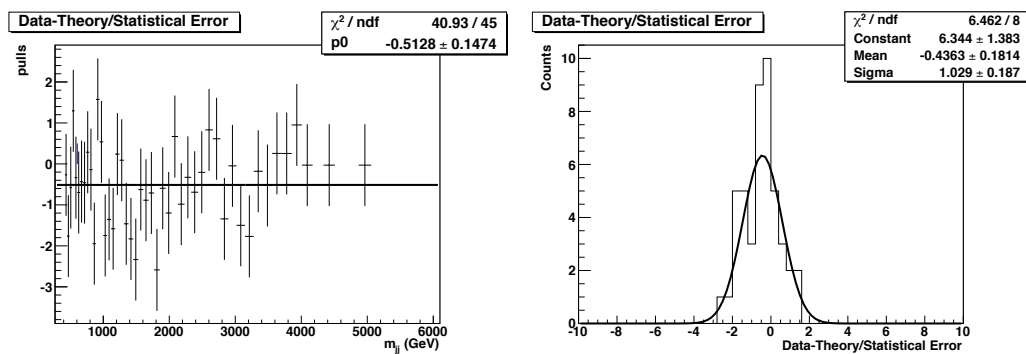


Figure 4.26: The “pulls”, data minus MC over statistical uncertainty, of the measured dijet angular ratio in PF jets, with respect to the background expectation. Shown are the pulls fitted with a flat line; also shown is the distribution of the pulls fitted with a Gaussian. The mean is close to zero and the width is close to one.

4.7 Statistical Method

Given that the data are in good agreement with the Standard Model, we set limits on the cross section of dijet resonances at a wide range of masses using a CL_s [?] criterion taking a log likelihood ratio (LLR) for its test statistic. This LLR does not intrinsically include the systematic uncertainties, rather they are included in ensemble testing with the Cousins-Highland method [?], as we describe.

4.7.1 Likelihood Ratio Definiton

We define the likelihood \mathcal{L}_j for each dijet mass bin j as the probability to observe n_{in} and n_{out} events independently, with a Poisson expectation of observing μ_{in} and μ_{out} :

$$\mathcal{L}_j = \mathcal{P}(n_{\text{in}}, n_{\text{out}} | \mu_{\text{in}}, \mu_{\text{out}}) \quad (4.8)$$

$$= \mathcal{P}_{\text{Pois.}}(n_{\text{in}} | \mu_{\text{in}}) \mathcal{P}_{\text{Pois.}}(n_{\text{out}} | \mu_{\text{out}}). \quad (4.9)$$

This product of Poisson functions in the number of inner and outer dijets can be written as the product of Poisson and binomial functions in the observed number of inner dijets, the observed number of total dijets ($n_{\text{tot}} = n_{\text{in}} + n_{\text{out}}$), and the ratio ρ of the expected values:

$$\mathcal{L}_j = \mathcal{P}_{\text{Pois.}}(n_{\text{tot}} | \mu_{\text{tot}}) \mathcal{P}_{\text{Binom.}}(n_{\text{in}} | n_{\text{tot}}, \rho), \quad (4.10)$$

where $\rho \equiv \mu_{\text{in}} / (\mu_{\text{in}} + \mu_{\text{out}})$. Finally, since all the information on the ratio is contained in the binomial factor, the per-bin likelihood can be reduced to simply

$$\mathcal{L}_j = \mathcal{P}_{\text{Binom.}}(n_{\text{in}} | n_{\text{tot}}, \rho). \quad (4.11)$$

Precisely speaking, we use the total count as an ancillary statistic, so that the likelihood of the ratio is conditioned by the observed total count. This binomial probability is computed using the Gaussian approximation for $n_{\text{tot}} > 10^8$ and the incomplete beta function otherwise. We would like to thank the authors of [?] for pointing out the numerical difficulties in evaluating the incomplete beta function for large numbers ($N > 10^3$) and providing their code which resolves this for $N < 10^8$.

The total likelihood for all dijet mass bins is the product of all \mathcal{L}_j . In practice, since the natural logarithm is a monotonic operator, it is equivalent and convenient to sum $\ln \mathcal{L}_j$ such that

$$\ln \mathcal{L} = \sum_{\text{bins}} \ln \mathcal{L}_j. \quad (4.12)$$

The log likelihood ratio (LLR) for QCD and each alternate hypothesis (QCD plus dijet resonance of a particular mass and cross section) is then written as

$$\text{LLR} = \ln \mathcal{L}_{\text{alt}} - \ln \mathcal{L}_{\text{QCD}}. \quad (4.13)$$

It is important to note that the LLR in any dataset is a function of the background expectations and signal-plus-background expectations being considered. Changing either of these models changes the LLR of any given dataset or pseudo-dataset.

4.7.2 Setting Limits

For every resonance hypothesis (mass and cross section), we generate pseudodatasets (PDSs) of binomially fluctuated $R_{|\Delta\eta|}$ bin values. Ensembles of one hundred thousand PDSs are made from the background-only and signal-plus-background hypotheses, each. The binomial distribution governing each bin's probability distribution is predicated on the total number of events in each bin in the data. In this way our pseudodatasets represent the two scenarios (realistic QCD with or without signal) and yet have the same statistical power as our collected data.

Under any pair of signal and background hypotheses, each PDS take its value of LLR, the test statistic. The measured data likewise take a value of LLR given the background-only and signal-plus-background hypotheses being tested.

We use the distribution of background-assumption (B) and signal-plus-background assumption (S+B) LLRs to find our predicted lowest limit at 95% confidence. Our measured limit comes about by comparing these distributions to the LLR of data. We integrate the B and S+B PDS LLR distributions from their lowest points to the 95% LLR_{CLs} for which

$$\frac{\int_{-\infty}^{LLR_{CLs}} PDS_{S+B}}{\int_{-\infty}^{LLR_{CLs}} PDS_B} = 0.05. \quad (4.14)$$

When the LLR of data lies below a 95% LLR_{CLs} , we exclude the signal scenario with 95% confidence. When the mean (best expectation) LLR of QCD lies below a 95% LLR_{CLs} , we *expect* to set a 95% exclusion limit with this CLs method.

In Fig. ?? we show the 95% CLs exclusion limits, both the expected limit (the dashed line with 1- and 2-sigma bands) and the observed limit. Here the systematics are not yet taken into consideration. At the dijet masses where the data differ from the expectation, we can and do observe corresponding features in the measured $R_{|\Delta\eta|}$ with respect to the QCD expectation, and in the appropriate direction.

4.8 Systematic Uncertainties

As mentioned above, systematic uncertainties on the dijet centrality ratio are not included in the calculation of the LLR itself. They enter the limit setting procedure through the ensemble generation. When generating the pseudo-datasets, the value of a quantity for which we assign a systematic uncertainty is drawn from a Gaussian distribution of width equal to the systematic uncertainty.

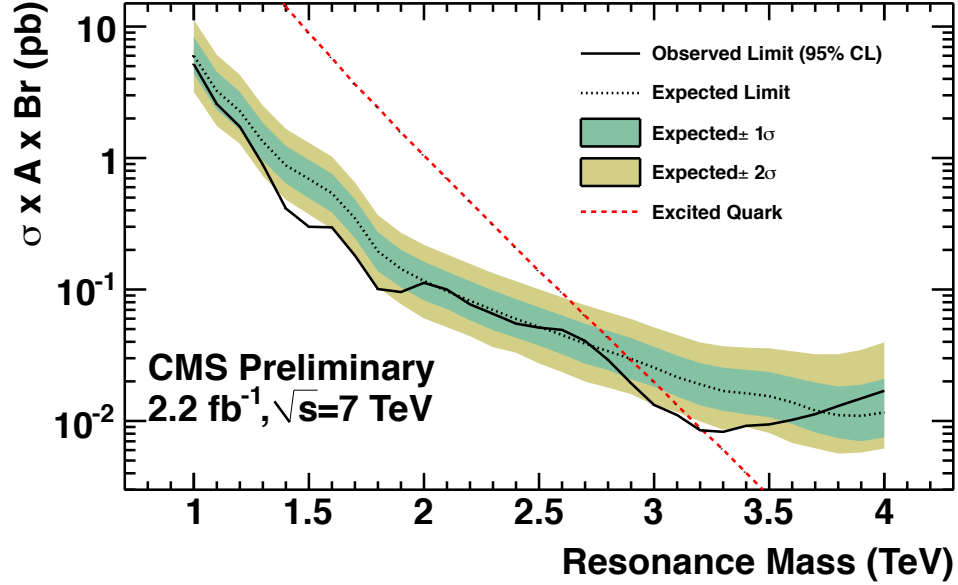


Figure 4.27: Limit summary plots for qg dijet resonances without systematic uncertainties. Within the acceptance, we show the lowest cross section excluded at 95% for the data (*solid black*), the SM expectation (*dashed black*) with 1 and 2 σ bands (*green/yellow*) versus M_{res} . For reference we show the theoretically predicted cross section for a massive excited quark decaying into this final state *dashed red*.

4.8.1 Pileup (PU)

Instantaneous luminosity at the LHC is so high that more than one proton-proton interaction can be expected at every crossing of proton bunches. The hardest collision in a given crossing will be the most likely source of jets with high transverse momentum, and as such should be the interaction which is most likely to fire the trigger. We demonstrate in Fig. ?? and ?? that the difference in the dijet angular ratio is consistent with zero by comparing events with only one 'good' primary vertex to all events with more than one such vertex.

4.8.2 Additive Offset to Prediction from Sideband

The detailed shape of $R_{|\Delta\eta|}(m_{jj})$ matches well between LO QCD predictions from PYTHIA and the low-mass dijet bins where our data have strong statistical power. For this reason we take an additive offset, constant in m_{jj} , from a sideband region, as a correction to the prediction. We apply this offset to the prediction at all dijet masses. This is discussed in ?? and presented in Table ?. The uncertainty on the

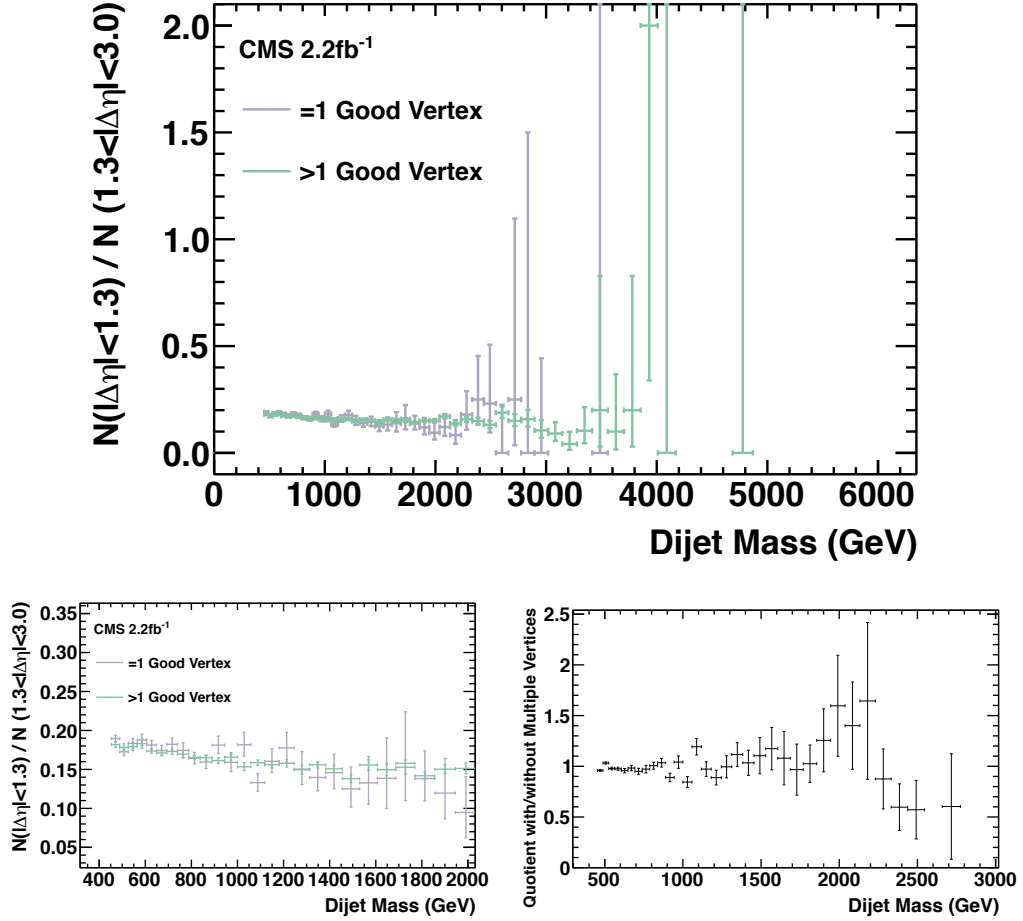


Figure 4.28: The measured value of $R_{|\Delta\eta|}$ in events with only one good vertex, and in events with more than one such vertex. Also shown are the low-mass, high-statistics region on an expanded scale and the quotient of these two distributions.

value of this fit is 0.0019 units of ratio, the uncertainty of the fit to find the offset. Thus it is a percent-level uncertainty in the modeling of the $R_{|\Delta\eta|}$.

4.8.3 Non-Perturbative Corrections

Here we examine the effect of nonperturbative effects applied in PYTHIA6. In Fig. ?? we show the $R_{|\Delta\eta|}$ data in the thick black crosses, and leading order PYTHIA6 simulation in green (*LO*). Switching on parton showering (*PS*) has negligible impact, but initial-state radiation (*ISR*) and *ISR* with electromagnetic component (*ISR2*) systematically drop the $R_{|\Delta\eta|}$ prediction at low mass. The inclusion of final-state radiation (together with *ISR*, marked as *IFSR*) also has a small effect in the same

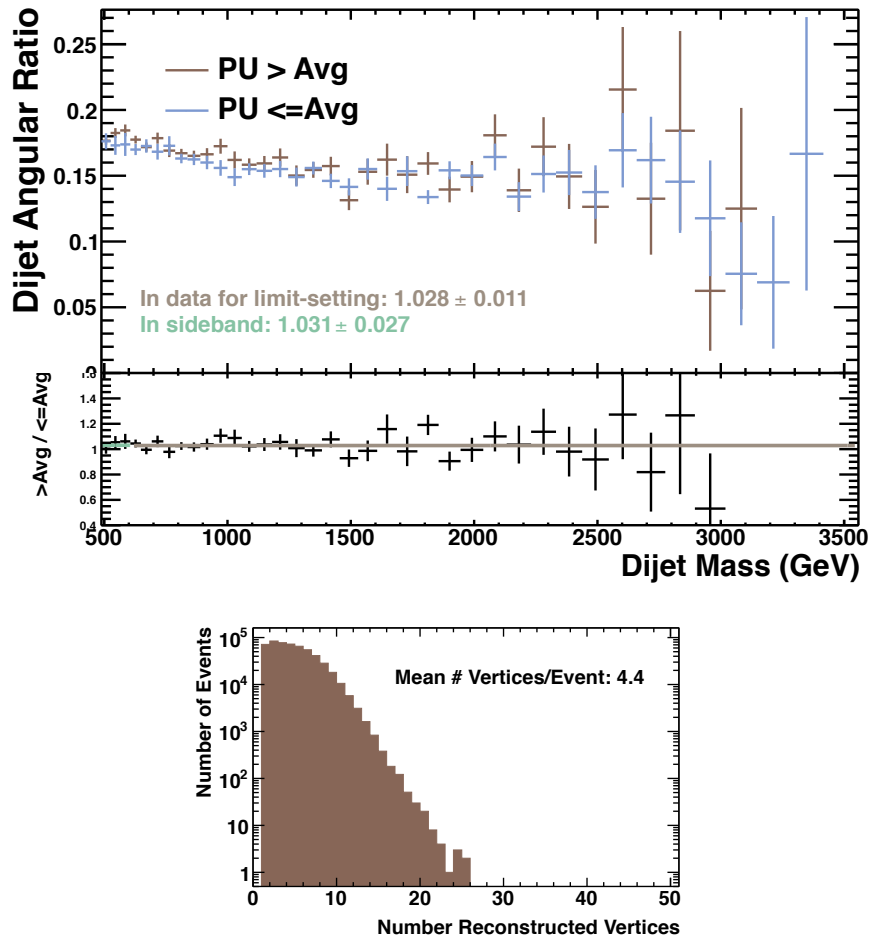


Figure 4.29: The measured value of $R_{|\Delta\eta|}$ in events with above- or below-average counts of reconstructed vertices, as well as the distribution of vertices reconstructed per event. The $R_{|\Delta\eta|}$ is quite compatible in the two populations of events.

direction. To this point the slope of the thick black crosses does not match the simulation well.

By including multiple parton interactions and hadronization (MPI, HAD), the shape of black and red points agree for dijet masses 500 GeV and above. Because of the turn-over in the fully corrected LO prediction, not present in our data, we chose to exclude dijet masses below 500 GeV. Above this mass, the PYTHIA6 prediction for the ratio with all non-perturbative effects agrees with CMS data up to the constant offset which we take from data.

We also present a study at next-to-leading order, in Fig. ???. We extract the fractional change in the PYTHIA prediction due to the non-perturbative effects of multiple parton interactions, hadronization, and the underlying event. We apply

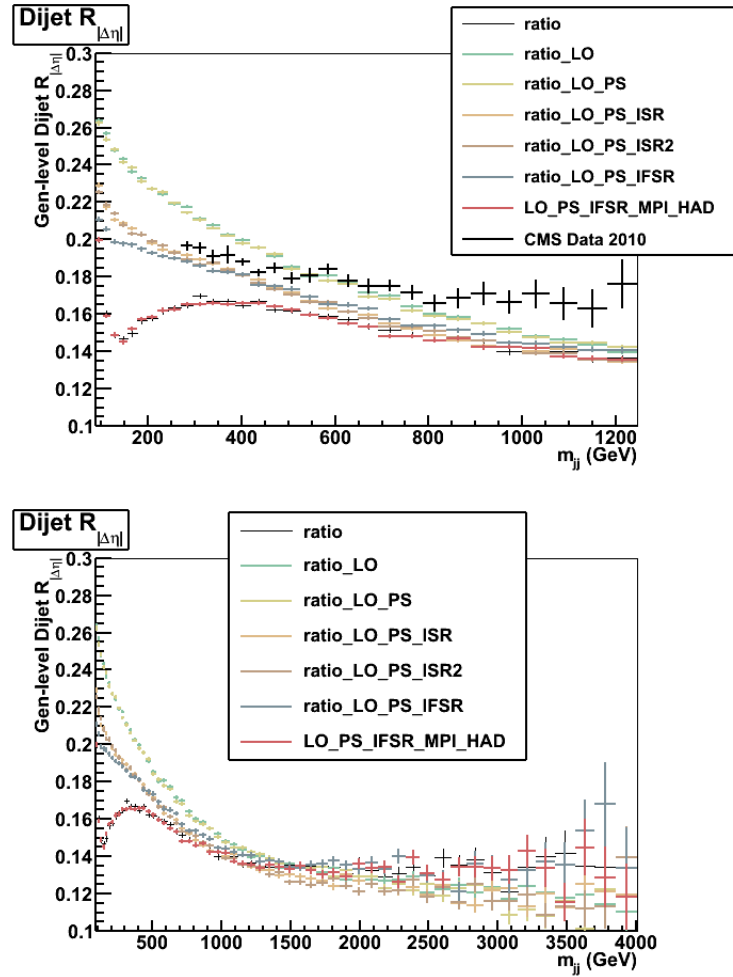


Figure 4.30: The $R_{|\Delta\eta|}$ in PYTHIA6 with accumulating effect from non-perturbative effects. Top, a zoomed view of the low dijet mass bins, with data overlaid on prediction. Bottom, the same theory predictions at a larger scale, and data removed.

this to the NLO prediction. The resulting “corrected NLO” prediction for the dijet angular ratio is shown in this figure, lying remarkably close to the shifted PYTHIA6 prediction used as our background expectation. The difference between these two curves is taken as an uncertainty on the shape of the background expectation, shown in Fig. ?? with the shape uncertainty due to the uncertainty on the jet energy correction, discussed in the next subsection.

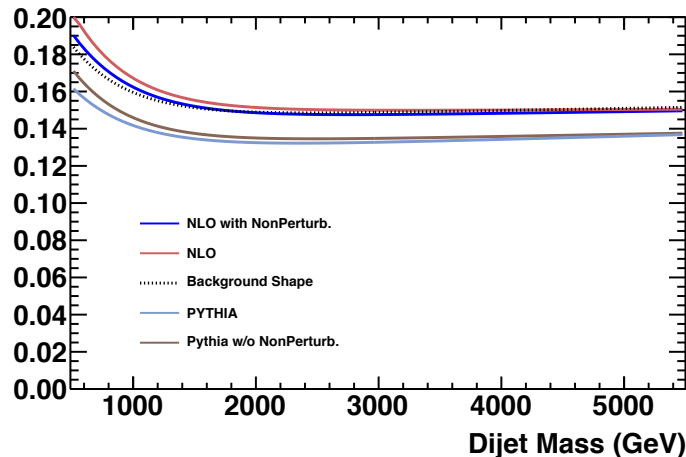


Figure 4.31: The $R_{|\Delta\eta|}$ in PYTHIA6 (pale blue) and PYTHIA6 without non-perturbative effects (brown). Also shown is the NLO prediction from fastNLOv2 and NLOJet++ 4.1.3. (red line at top) and the same after scaling by the effect of the non-perturbative effects (blue). This last prediction is very close to the dashed line, which is the background prediction used in this analysis.

4.8.4 Jet Energy Correction (JEC)

The CMS Jet Energy Correction has been measured from data, and has a pseudorapidity dependence. The jet corrections were applied by CMSSW_4.2.8 in September of 2011, and the global tag used to get our ak7CaloJetsL2L3 corrections was GR_R_42_V19::All. Misassigning the correct value of the JEC within its uncertainty would bias the $R_{|\Delta\eta|}$. Shifting the JEC up and down by the amount of this uncertainty allows us to measure the systematic uncertainty on the $R_{|\Delta\eta|}$ due to the JEC uncertainty.

We apply this procedure to our QCD genjets MC sample, and take the difference between the fits we obtain as a systematic uncertainty on the predicted $R_{|\Delta\eta|}$. See Fig.???. Together with the background shape systematic discussed above, the difference between these curves and the nominal prediction are shown in Fig. ??.

4.8.5 General Detector Effects

Using monte carlo simulated data, the jets used in this analysis have been corrected in their energy and momentum so as to match, on average, the original particle-level jet. By comparing the $R_{|\Delta\eta|}$ found in particle-level monte carlo to that found in fully reconstructed jets we measure any general detector effect on our quantity of interest. We observe no strong effect, and a flat-line fit to the quotient of gen-

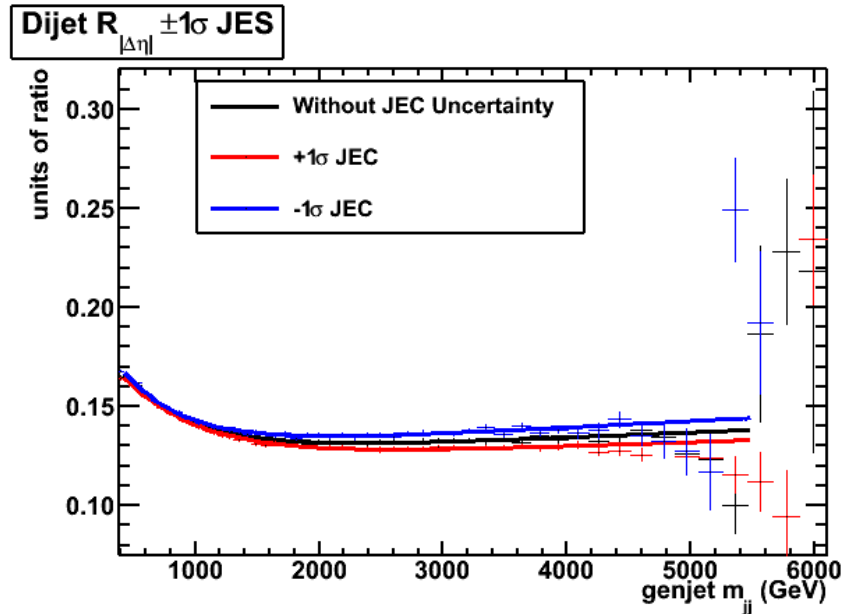


Figure 4.32: The $R_{|\Delta\eta|}$ having scaled the four-vectors of all jets up or down by one sigma. The vertical scale is in units of the dijet angular ratio.

level and reconstructed dijet angular ratio indicates they are compatible to within $0.9\% \pm 1.7\%$. See Fig. ??.

4.8.6 QCD production rate normalization

PYTHIA6 predicts a differential dijet spectrum $d\sigma/dm_{jj} vsm_{jj}$ whose shape matches data well, but whose overall rate is low [?]. Comparing the data from unrescaled triggers in the full $2.2 fb^{-1}$, we find the best fit normalization comes from multiplying the PYTHIA6 prediction by 1.577 ± 0.026 . See Fig. ?. The statistical uncertainty of the normalization fit translates into a systematic uncertainty on the background rate of 1.6%.

4.8.7 Signal Model Uncertainty

While the uncertainty on the absolute jet energy scale does not much affect the measurement of the ratio, it does affect the measurement of the dijet mass. For the nearly flat dijet ratio predicted by QCD, a 2.2% uncertainty on the dijet mass is not very important, but for a resonant bump the uncertainty on the dijet mass has a 2.2% effect on its mass. We take this 2.2% dijet mass uncertainty as a systematic uncertainty on the signal mass.

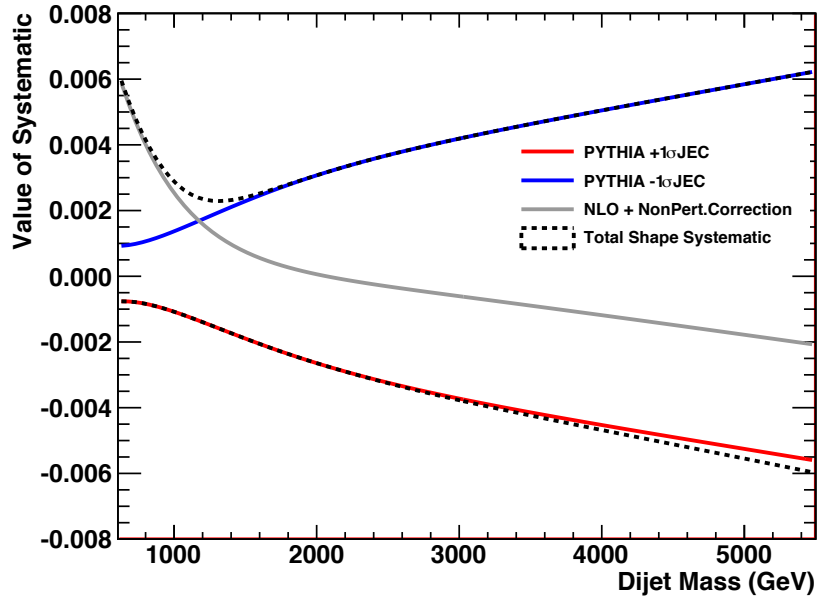


Figure 4.33: The envelope of the shape uncertainty from the JEC uncertainty and the consideration of NLO, together. The vertical scale for both plots is in units of the dijet angular ratio.

We also take into account the present 4.5% uncertainty in quoted luminosity, as this affects the cross section limits which may be quoted for new physics processes.

4.8.8 Total Systematic Uncertainty

We implement all our measured systematic uncertainties as independent variations when generating pseudodatasets for use in determining the limit. The result is to broaden appropriately the bin-by-bin binomial variations in our PDS's. These variations also refine upward the cross section excluded by the data at 95% CL. Refer to Section ?? for more on this procedure.

In Fig. ?? we show the ratio of the exclusion limit obtained with and without systematics, for both the observed and the expected limit. The result confirms that systematic uncertainties have a clear effect on the limit-setting power at low mass where statistical uncertainty is small but become unimportant at high mass where statistical uncertainty dominates.

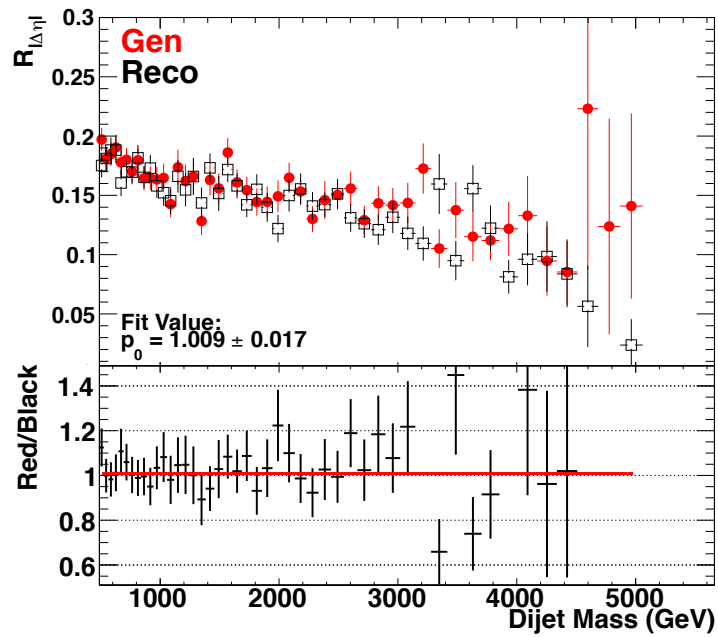


Figure 4.34: The measured value of the $R_{|\Delta\eta|}$ in fully simulated QCD monte carlo in the absence of pileup (to which we are not sensitive; see ??). Fully reconstructed jets are matched as well as possible to their genjets by requiring $R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.05$ and taking the genjet with the smallest R available. Reconstructed jets give compatible $R_{|\Delta\eta|}$ to within a percent or so.

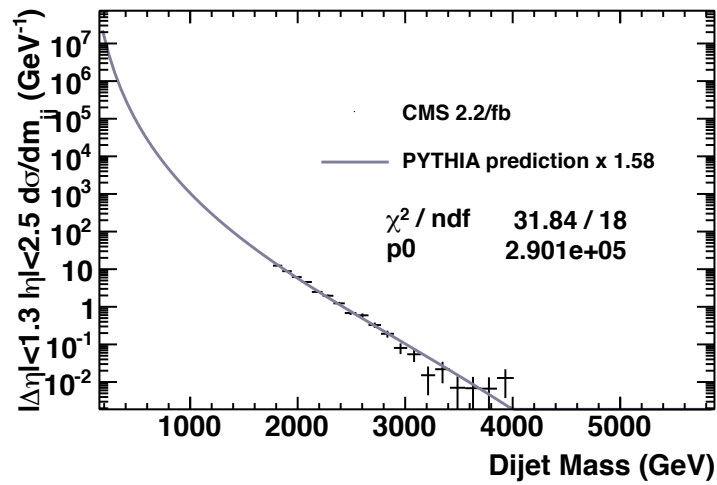


Figure 4.35: The comparison of data from the highest unrescaled triggers a fit to the QCD dijet rate for inner dijets. The prediction fits best when scaled up by 1.577 ± 0.026 .

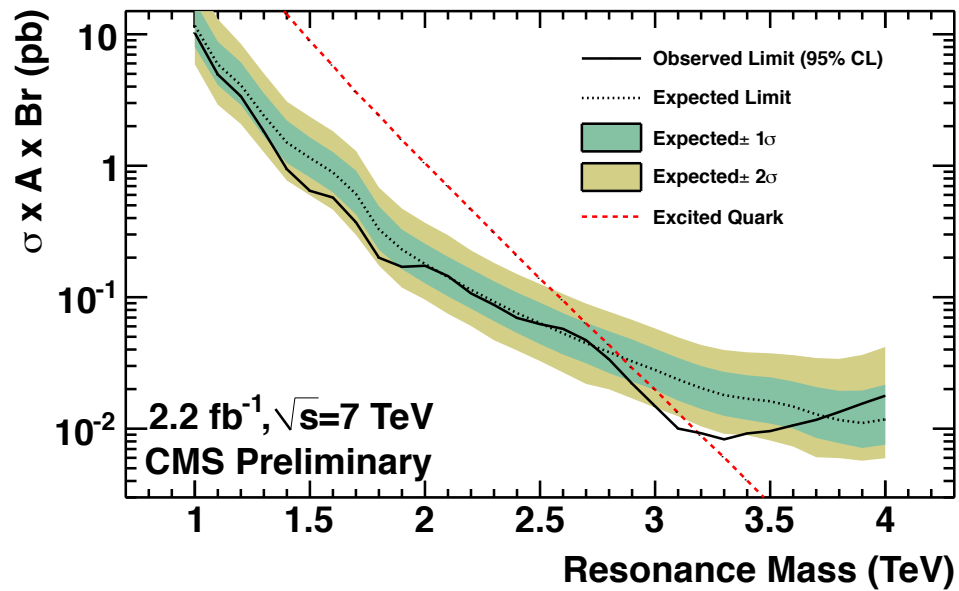


Figure 4.36: Limit summary plots for qg dijet resonances in the presence of systematic uncertainties. Scaled down by the acceptance, we show the lowest cross section excluded at 95% for the data (*solid black*), the SM expectation (*dashed black*) with 1 and 2 σ bands (*green/yellow*) versus M_{res} . For reference we show the theoretically predicted cross section for a massive excited quark decaying into this final state *dashed red*.

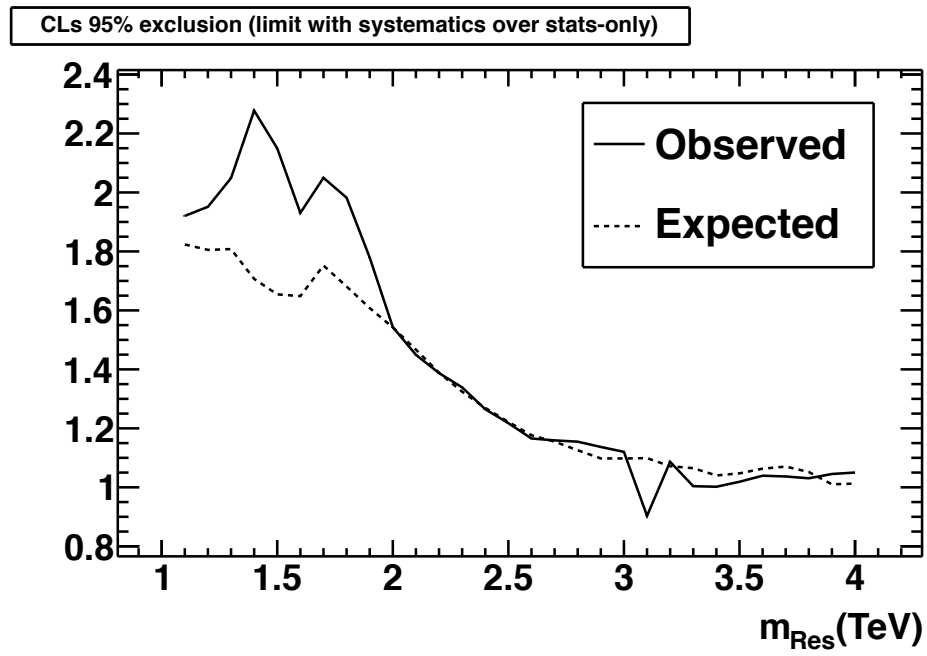


Figure 4.37: Ratio of lower exclusion limits with systematics to limit with statistical uncertainties only.

Chapter 5

Conclusion

We present here the results from the CMS experiment in measuring the dijet angular ratio $N(\textit{inner})/N(\textit{outer})$ in 7 TeV pp collisions at the LHC, using a data sample corresponding to 2.2 fb^{-1} . The dijet angular ratio observed in data is consistent with the corrected NLO prediction and the PYTHIA prediction. We exclude at 95% confidence the existence of an excited quark below 3.2 TeV. Our limits on cross sections between 10 and 10^{-2} pb extend to all narrow, spin- $\frac{1}{2}$ qg resonances between 1.0 and 4.0 TeV.

Appendices

Appendix A

Highest Mass Dijet Events

In this appendix we present event displays of the highest-mass dijets in our dataset. We do this separately for numerator (inner) dijets (Figs. ??-??) and denominator (outer) dijet events (Figs. ??-??). All look like healthy events with real, physical dijets.

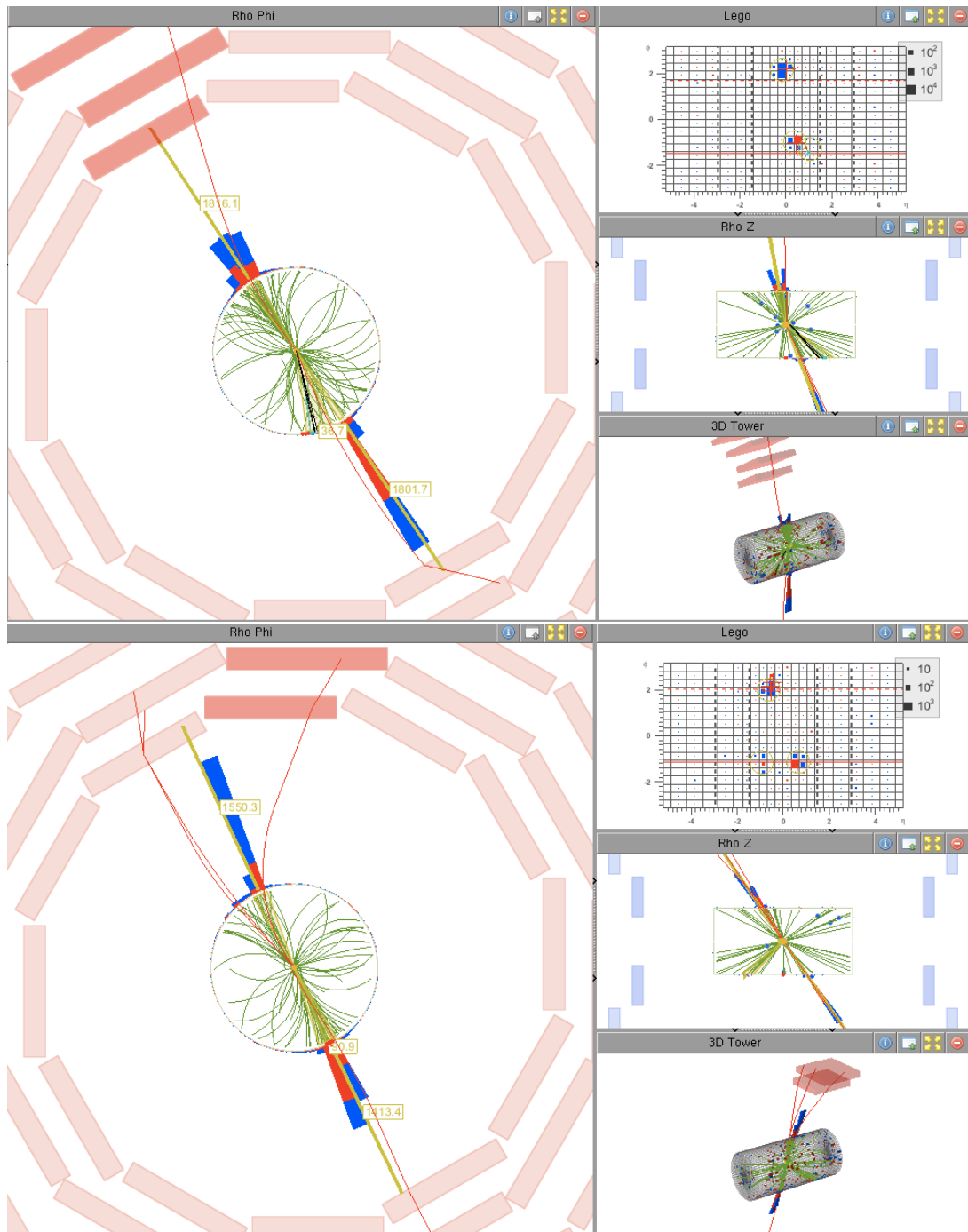


Figure A.1: The second pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??.

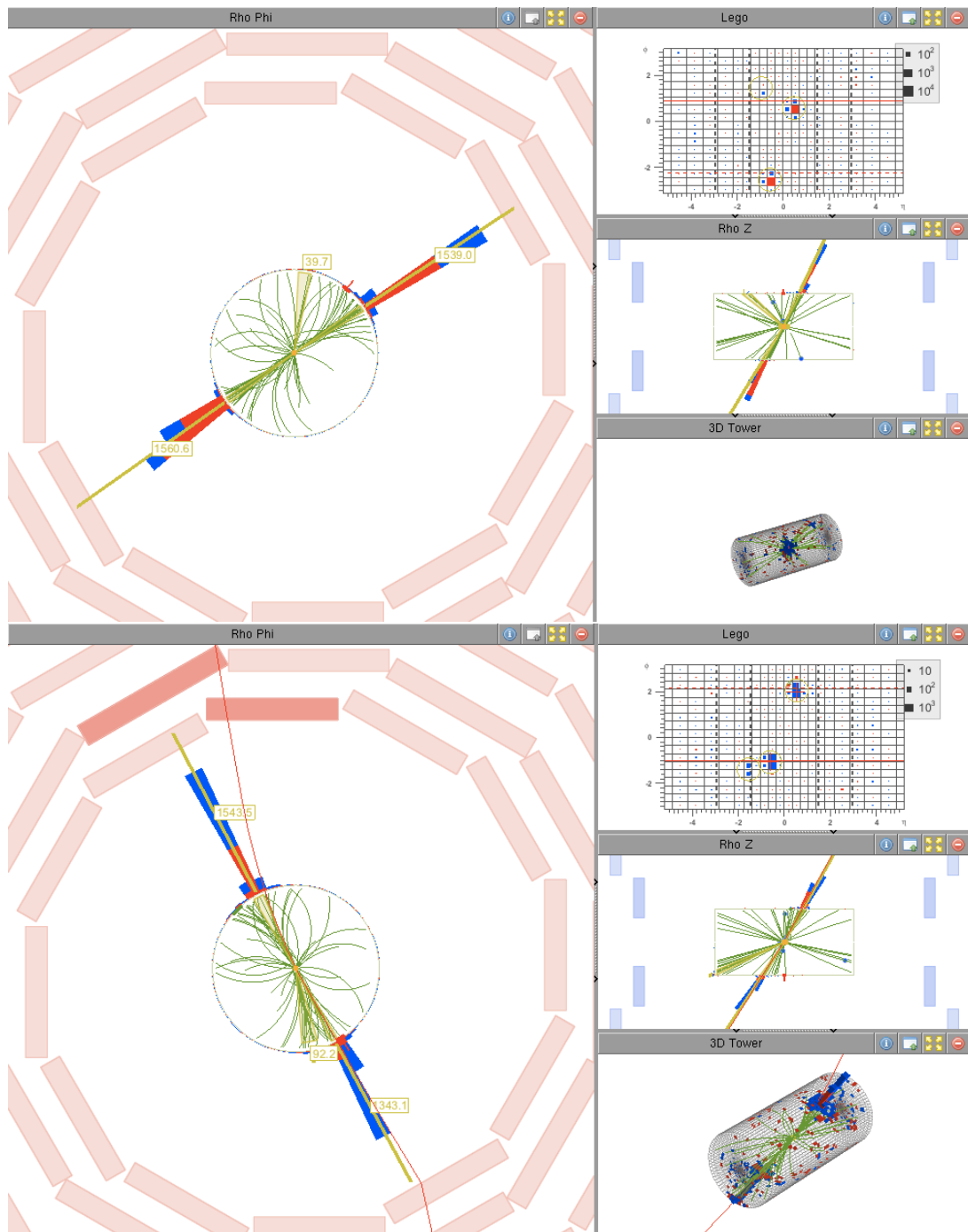


Figure A.2: The second pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??.

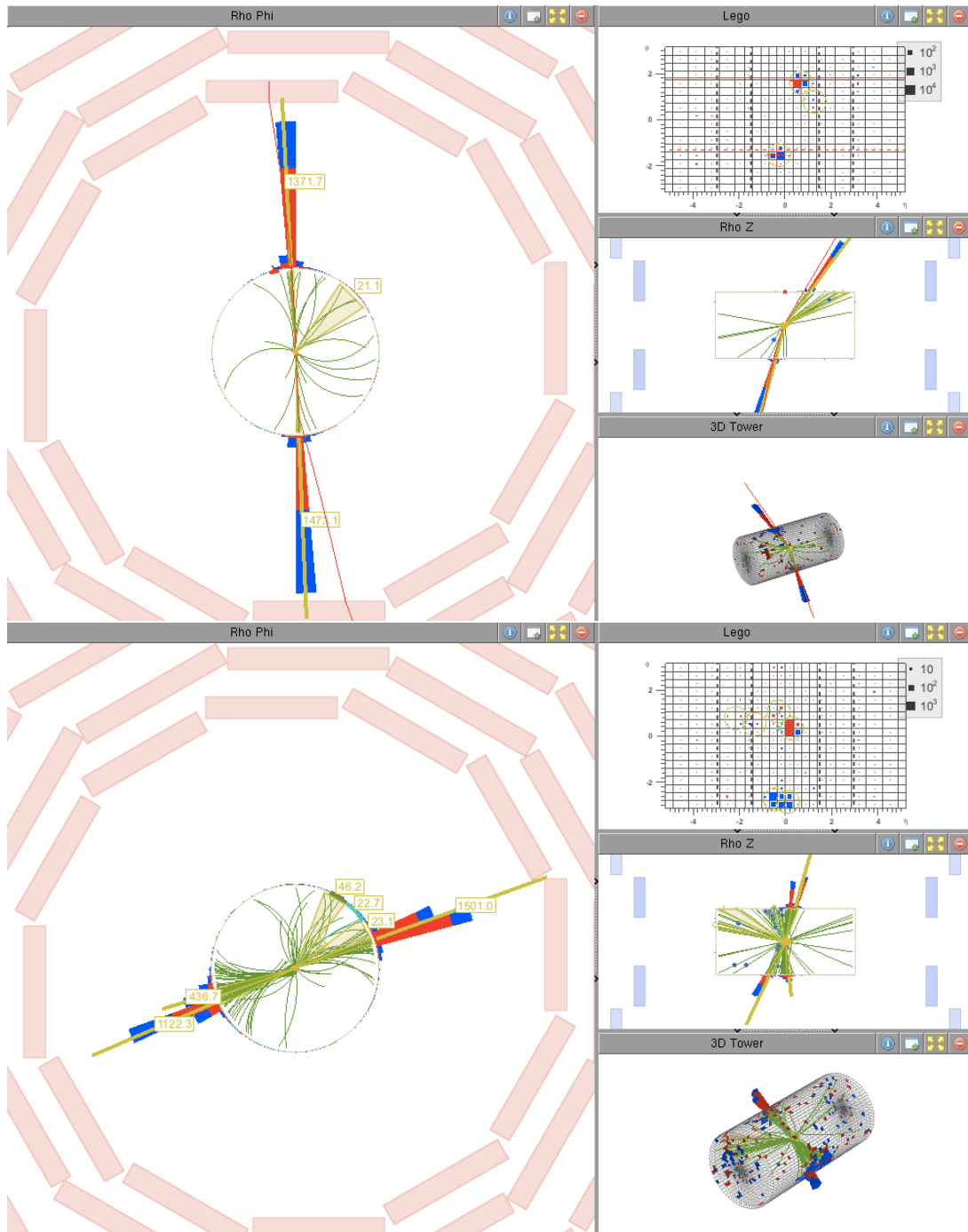


Figure A.3: The third pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??.

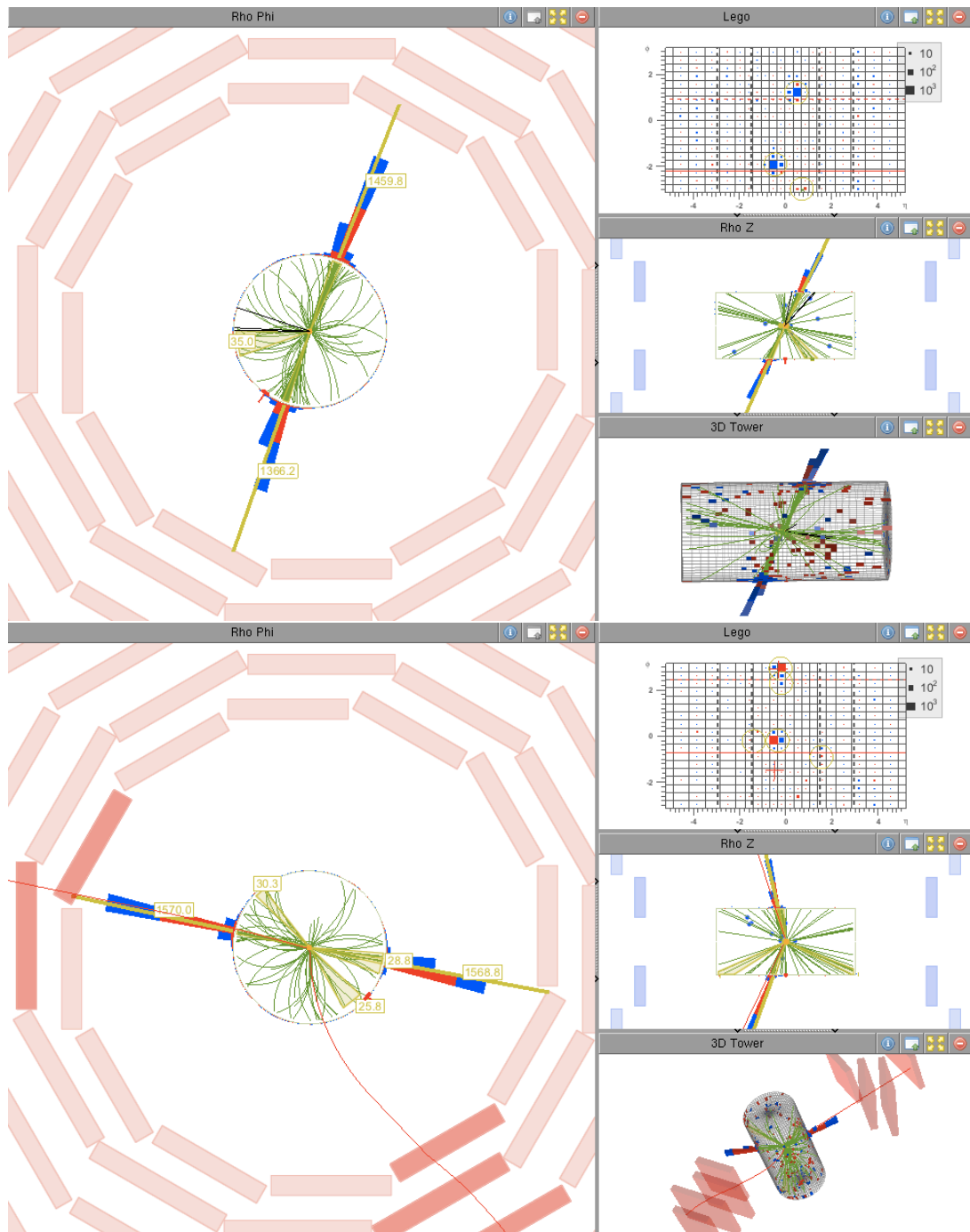


Figure A.4: The fourth pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??.

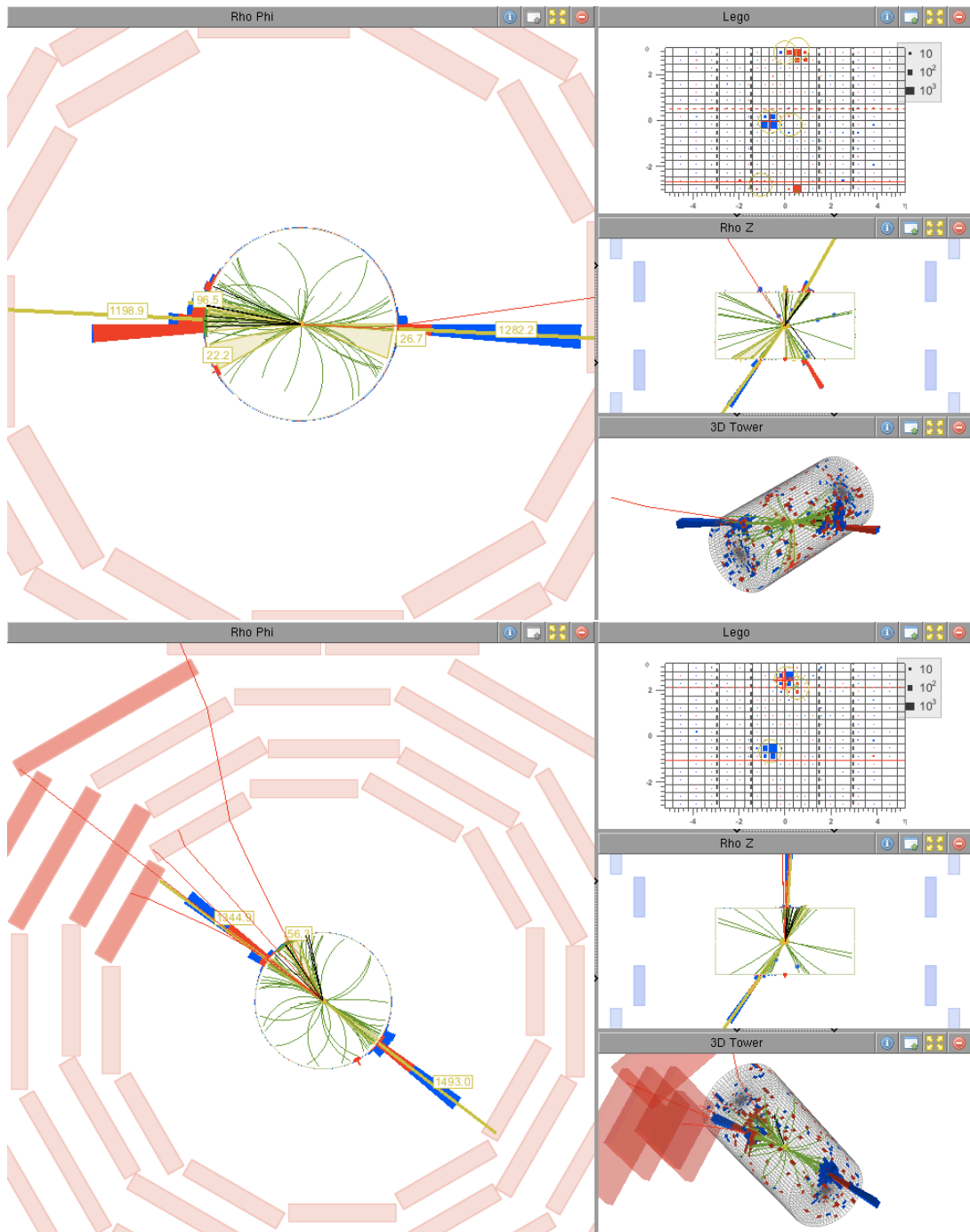


Figure A.5: The last pair of the ten highest-mass inner dijet events. Details of these events are given in Table ??.

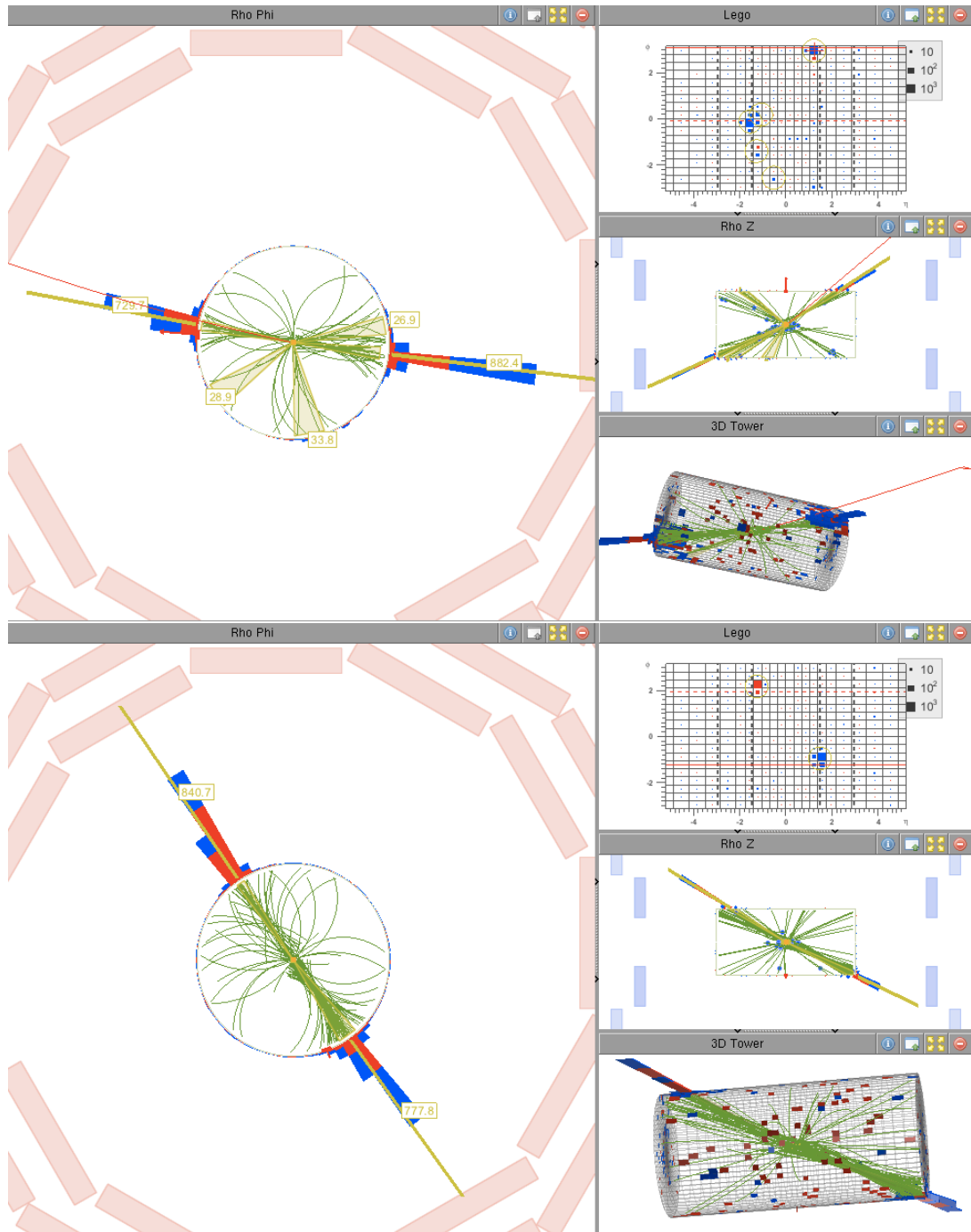


Figure A.6: The first pair of the eight highest-mass outer dijet events.

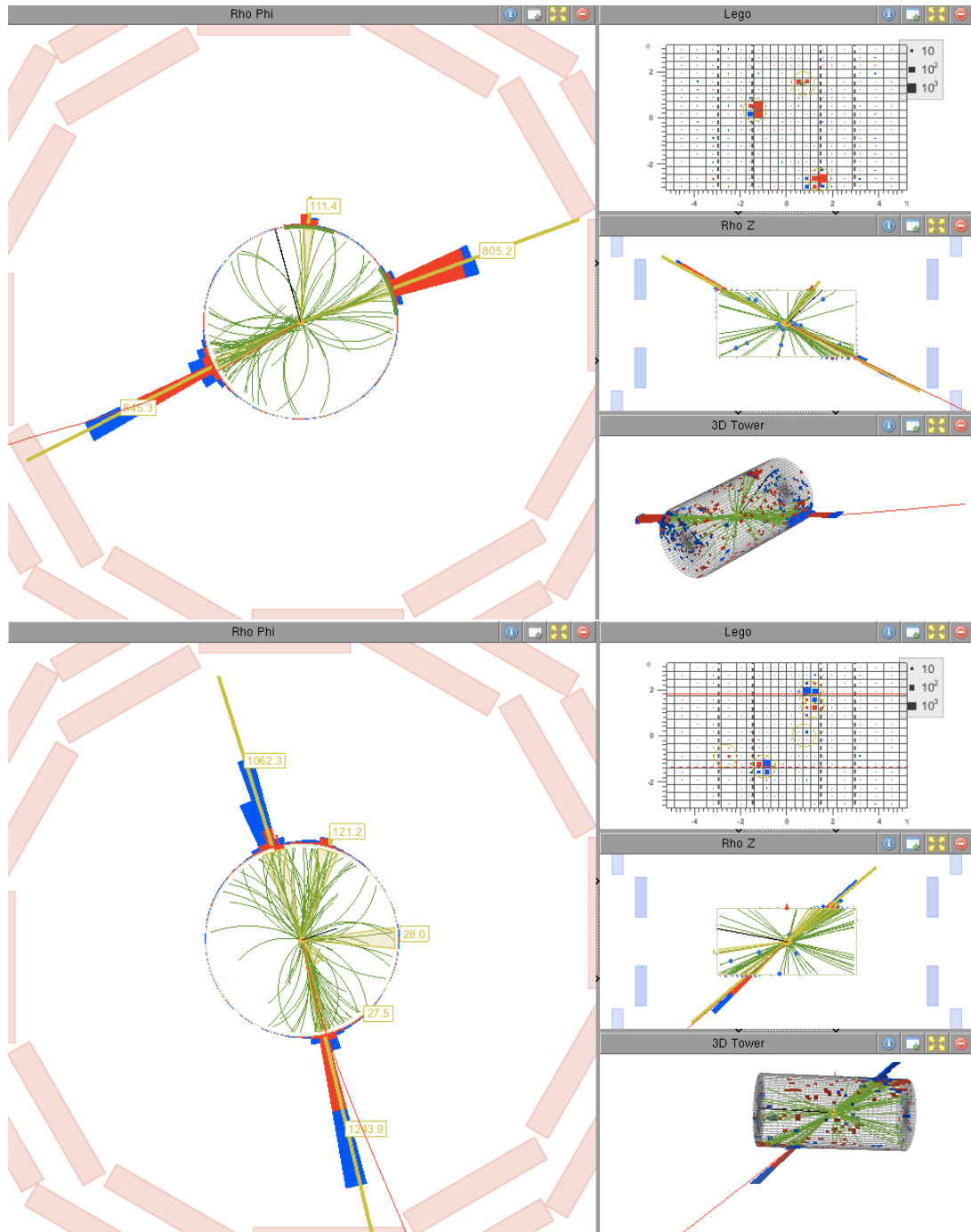


Figure A.7: The second pair of the eight highest-mass outer dijet events.

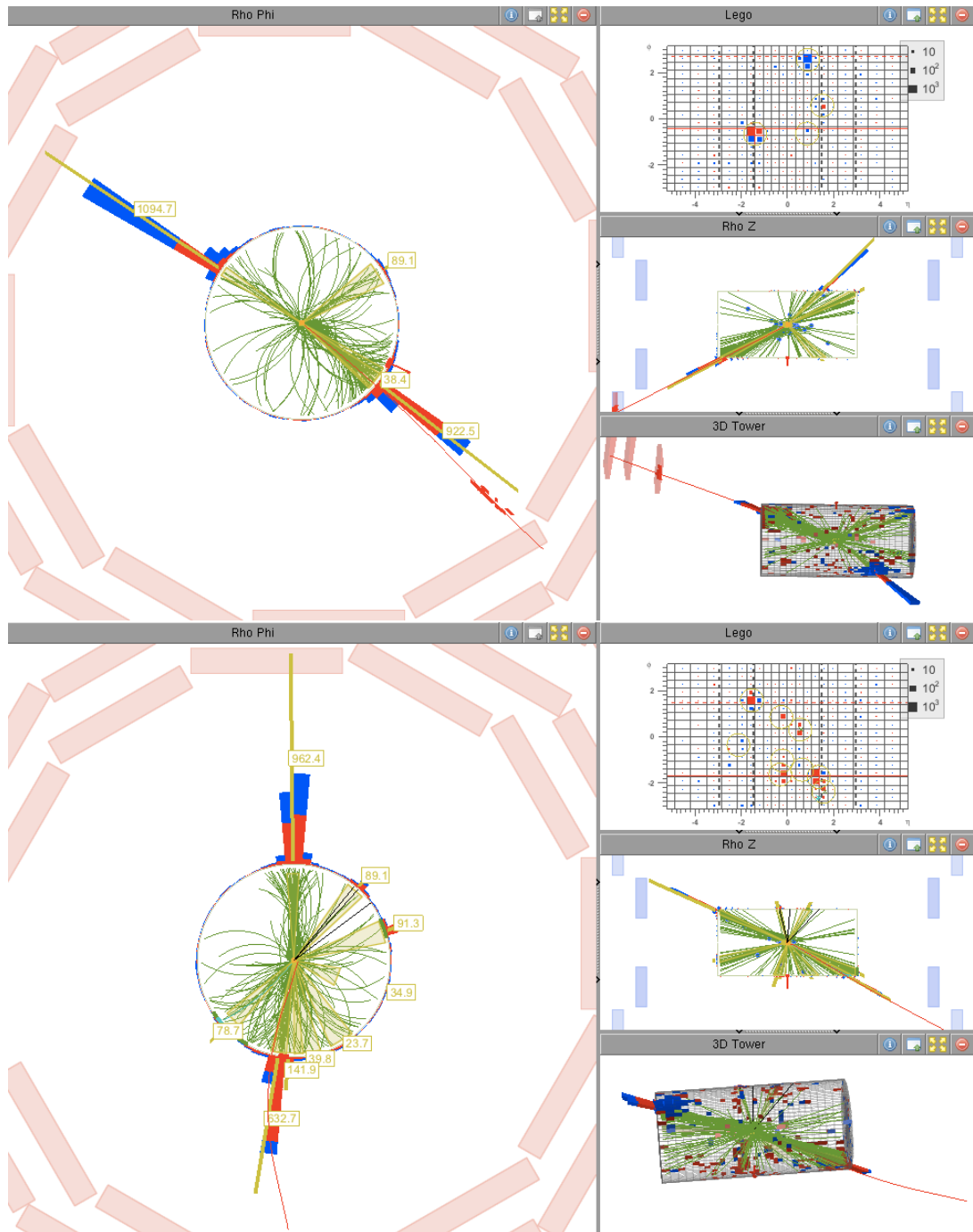


Figure A.8: The third pair of the eight highest-mass outer dijet events.

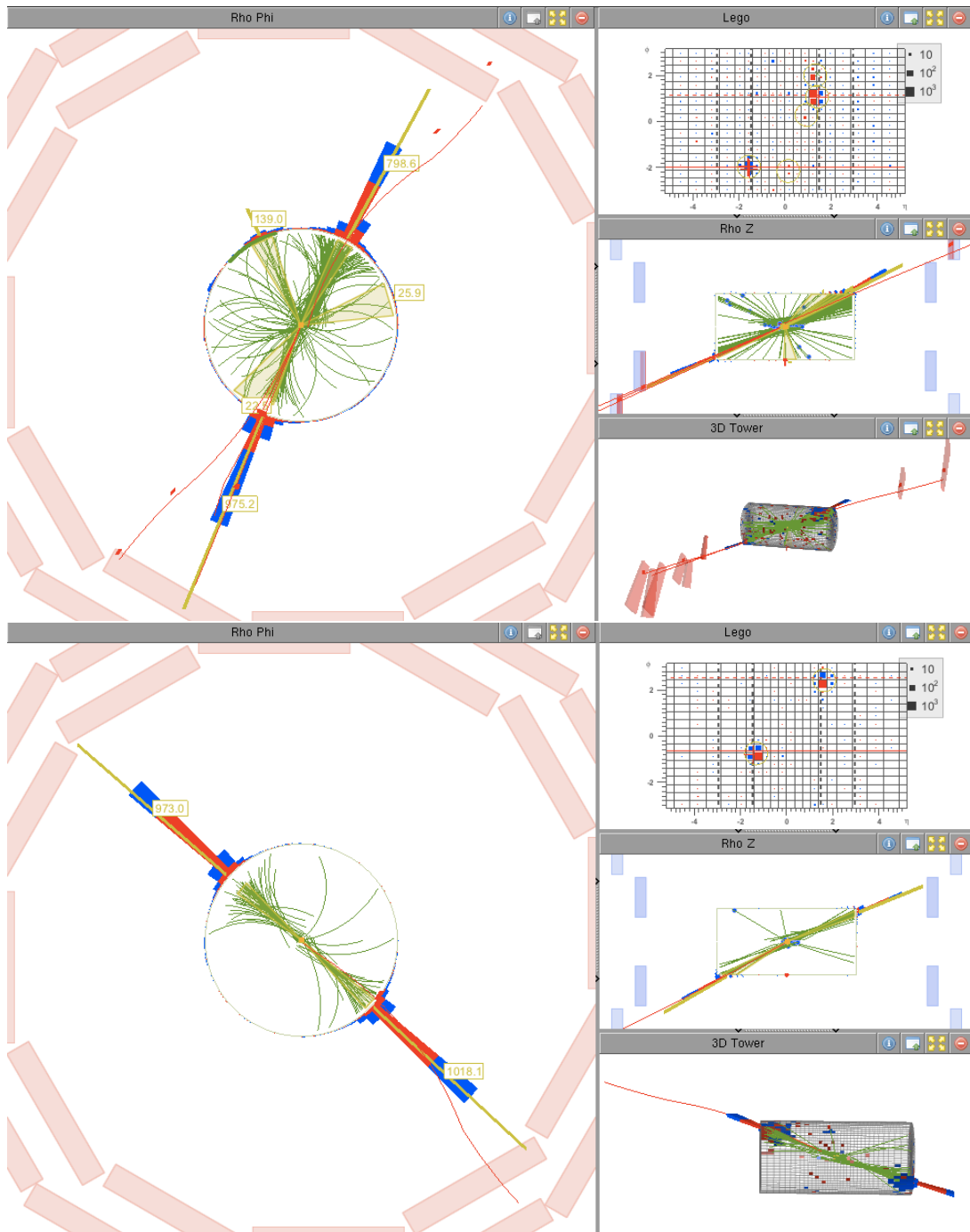


Figure A.9: The last pair of the eight highest-mass outer dijet events.

Appendix B

HCAL Readout Electronics System

This appendix is included to give the reader a working knowledge of the HCAL readout electronics chain. Such a knowledge will be of use in App. ??, which refers to these details and perhaps assumes the reader to be an avid fan of all HCAL electronic matters (as the author is).

Scintillation light from the tiles of HBHEHO is transmitted over colorless optical fibers to shine on the photocathode of a photodetector known as a Hybrid Photo-Diode (HPD). Light from the fibers of every scintillator tile in a calorimeter cell is optically added just before encountering the HPD. Nineteen hexagonally close-packed sites on the photocathode liberate photoelectrons which are accelerated as shown in Fig. ?? [?] to a depleted silicon surface. Incident electrons liberate electron-hole pairs [?], which allow momentary current between the incident-face contact and segmented contact pads at the back of the silicon. HPD pixels therefore act as electrical current sources in response to incident photoradiation.

In HF, Cherenkov light from quartz fibers is similarly carried to the photocathodes of photo-multiplier tubes (PMTs), where liberated photoelectrons initiate electron cascades along dynode chains. At the end of the cascade substantial voltage results from even single photoelectron liberation.

Here the readout chains of HCAL's subregions HB, HE, HF, and HO converge. Current from these two types of photoconverters is integrated in 25 ns windows by a module known as the QIE-8 [?]. The electrical current is asymmetrically distributed to four banks of integrating capacitors having different capacitances. The smallest capacitor which is found to be mid-range is selected for encoding; the integrated charge provides the mantissa while the capacitor used becomes an exponent.

Digitized, encoded signals are communicated via Gigabit Optical Link (GOL) [?] off the CMS detector body to the nearby electronics 'counting room', an artificial cave known as USC55 (Underground Service Cavern at Point 5). There the signals are received by the Hcal Trigger and Readout card (HTR, pronounced as "heater").

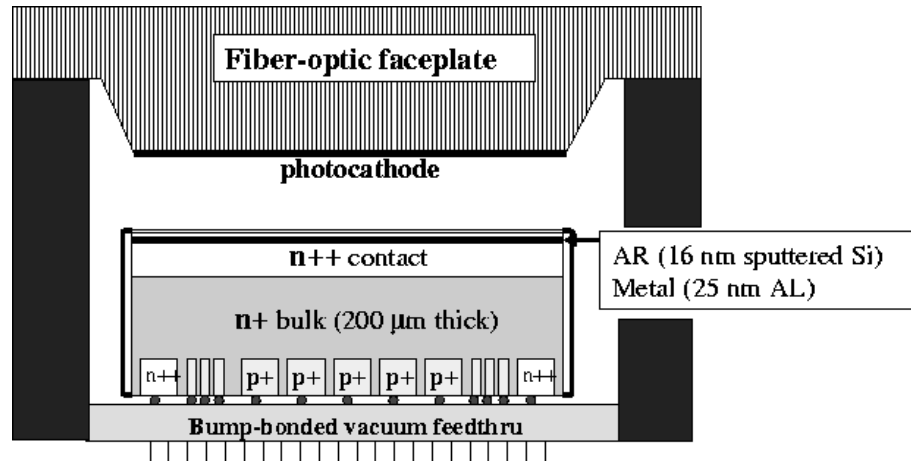


Figure B.1: Cutaway diagram of the HPD.

The HTR is responsible for parsing the bit stream it receives from the front end unit, supplying the hardware trigger with estimates of localized energy deposits in the calorimeter (more on that later), and supplying triggered events to the Data Concentrator Card (DCC).

All recorded CMS events have their assigned event number, which is simply a count of Level-1 Accept decisions from the start of the run. The DCC counts the lowest 24 bits of this number.

The Data Concentrator Card [?] lives in 9U VME racks with the HTR cards sending it input and the fanout card providing “Level-1 Accept” (L1A) signals from the hardware trigger electronics. It queues the event numbers of L1As in a first-in/first-out buffer (the “L1A FIFO”), while also receiving event data on all active input spigots. There are fifteen RJ-45 spigots, hosted in threes by five FPGA’s called Link Receiver Boards (LRBs) which implement FIFOs for the headers and data, independently. A block diagram showing the communication within the DCC is shown in Fig. ??.

At the heart of the DCC is the event builder FPGA (a Xilinx Spartan XC3SD1800A). If there is at least one L1A to take from the L1A FIFO, the event builder attempts to retrieve the header and data for the corresponding event from all active spigots. (The active spigot list is configurable during the start-of-run antics.) If it is successful, a Cyclical Redundancy Check (CRC) is performed on the assembled event, and the numerical value of the CRC is included in the event. The complete event is sent to the S-Link [? ?] transmitter board for relay to central DAQ.

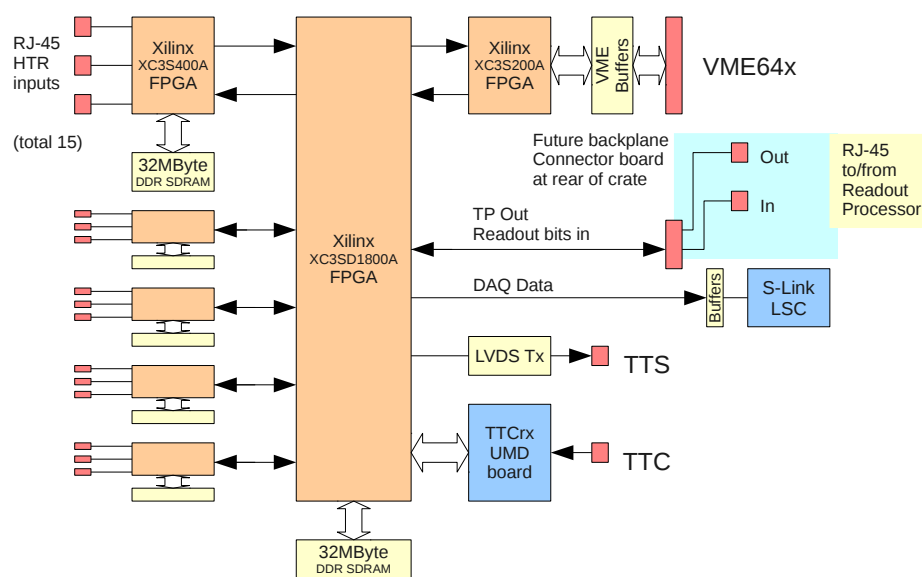


Figure B.2: Block diagram of the DCC.

Appendix C

Monitoring

We implemented code specific to the Data Concentrator Card (DCC) in two monitoring systems of the CMS HCAL. One system accesses registers on the DCC board over the backplane of the VME crates in which they live. This system is referred to as 'crate monitoring' or 'VME monitoring' and is treated in section ???. The other lives in a CMS-wide framework with a custom API examining fully-formed event data. It is known as DQM (Data Quality Monitoring), and normally runs live in a sampling mode on a dedicated data stream of the current CMS run, *online DQM*, and independently as a part of the CMS event reconstruction process, *offline DQM*. This discussion is found in Section ???.

C.1 DCC Crate Monitoring

The CMS data acquisition system is built with the XDAQ [?] software platform. Applications within this framework are responsible for running everything about the CMS subsystems, including the HCAL data-taking and electronics-monitoring functions. Several C++ classes in the HCAL XDAQ implementation configure, control, and monitor the DCC [?]. These classes all rely upon a Hardware Access Library (HAL) file [?], which describes every item and hardware address on the DCC over VMEbus and the PCI bridge. The HAL address table maps unique item names to locations in memory. Bit masks are provided in the HAL table for items which occupy less than the entire address.

C.1.1 Monitored Registers

The CMS HCAL DCC hardware is home to more than 900 monitorable items. With few exceptions, these may be considered in a hierarchy, classified as those which apply to the entire DCC, those for the Link Receiver Boards (LRBs) (five per DCC), and those for the input spigots (fifteen per DCC). The latter two groups of monitorable

items proliferate quickly because any status bit tallies and error flags which are kept for one LRB or spigot are also kept for all the others.

'Tables' of monitorables are used to gang items by their function, roughly reflecting the organization of the DCC HAL file. These tables are named for the DCC configuration, debugging variables, Trigger Throttling System (TTS) counters and timers, firmware versions, error counters, trigger counters, and error control configuration. For individual spigots and LRBs there are block and word counters, HTR status bit counters, and a large table of error counters generated by the DCC itself. They represent a digital smorgasbord.

C.1.2 The hcalDCCMonitoring Class

A single class implemented in the HCAL XDAQ packages bridges the gap from the HAL file describing the hardware to the 'flashlists' and monitoring tables described above. It relies upon the DCC and DCCManager classes for communication with a physical DCC over the VME crate backplane. After extensive setup functions at instantiation time, the workhorse of the hcalDCCMonitoring class is the update() method. This function is called by the XDAQ monitoring service at a regular, settable frequency, normally 10 Hz.

The Collector is a XDAQ service which interfaces with this monitoring class as well as others such as hcalHTRMonitoring. It passes along the regular results it finds to the Alarmer and Logger services, and may be called at the user's command by the MonVis service, a visualization package based on ROOT. An example of the MonVis output is given in Fig. ???. This plot is typical of the debugging work conducted during the commissioning of HCAL and also of so-called global runs. Although drab in color, the MonVis plots are an exciting and important tool of the HCAL operator. Their entries can be clicked to reveal information about which electronic components contribute to the clicked bin, and any information which is currently available about each of these from the log.

We configured the Logger and the Alarmer to react under certain rules to the input they receive. For example, the Alarmer raised an alarm upon the incrementation of the bad bunchcount error counter, and the logger updates the monitoring log when ten minutes have elapsed, as a 'heart-beat' check on system liveness.

In a global run many or all CMS subsystems are orchestrated by the central DAQ, and there were many incidental complications which were understood through the monitoring, logging, and visualization systems.

C.2 DQM: Data Quality Monitoring

The data transmitted to the CMS central DAQ over S-Link has a specific format, shown diagrammatically in Fig. ???. A software module we developed as part of the CMS Data Quality Monitoring (DQM) system [?] checks the raw data from

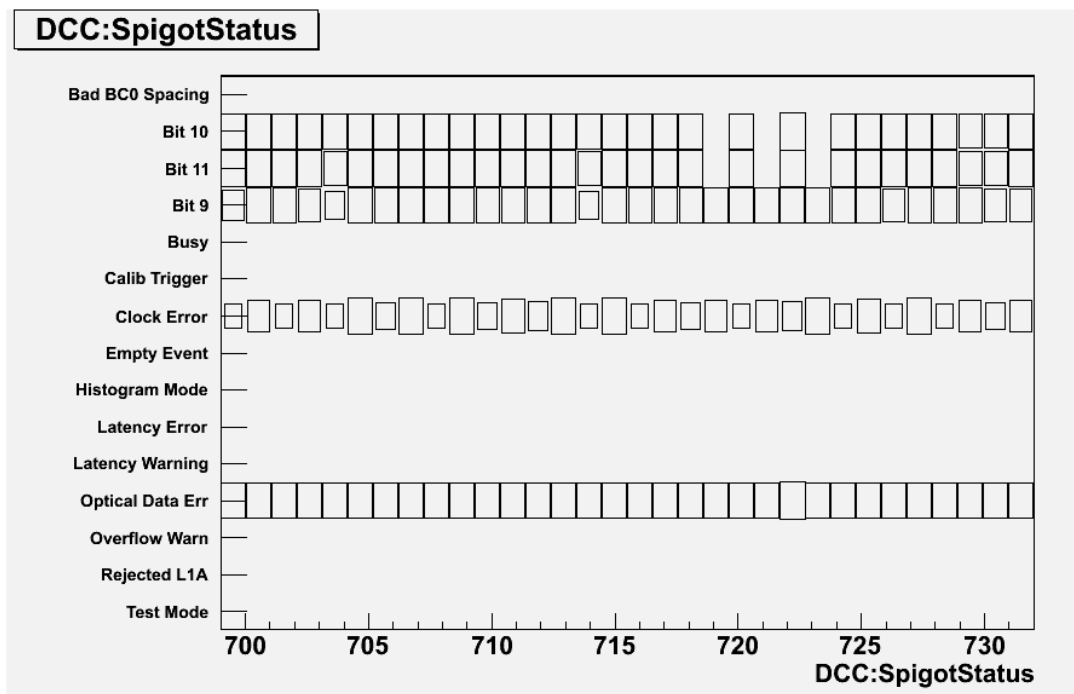


Figure C.1: An example plot from the electronics monitoring system MonVis displaying quantities from the internal registers of the DCC. Each column represents one DCC, and the status codes on the vertical axis are summed across each DCC’s active spigots.

HCAL for adherence to this specification, internal consistency, error flags, and a few quantities with diagnostic purposes. This module is the HCAL Raw Data Monitor.

Offs	Byte 7		Byte 6		Byte 5		Byte 4		Byte 3		Byte 2		Byte 1		Byte 0		
0x00	5	EvtTy	EvN				BcN				FED Number				FOV	1	--
0x08	0 - Reserved -																
0x10	DCC Rev*				DCC Status				0	HTR status				0	TTS	Format ver	
0x18	HTR Err		LRB Err		E P B V T C		HTR 1 nWords		HTR Err		LRB Err		E P B V T C		HTR 0 nWords		
0x20	HTR Err		LRB Err		E P B V T C		HTR 3 nWords		HTR Err		LRB Err		E P B V T C		HTR 2 nWords		
...	Additional headers for HTRs 4..13																
0x50	-0-				HTR Err				LRB Err		E P B V T C		HTR 14 nWords				
0x58	-0-																
0x60	HTR 0 Payload																
	HTR 1 Payload																
	Additional HTR payloads (one for each header where nWords .ne. 0)																
	A	--	Evt_Length (64 bit words)				CRC (16)				--	Stat	TTS	0	--		

Figure C.2: The format specification for DCC events at the level of raw bits.

CMS DQM processes sampled events live during data taking (about 10-20 Hz) from a dedicated instance of the Storage Manager [?]. The contributions from various triggers are specialized for this stream in an effort to make the online monitoring as useful as possible. A different configuration is used during the reconstruction of the CMS primary datasets. Primary datasets are defined by the triggers allowed to add events to them. They are processed in large numbers of parallel jobs whose results are merged. So-called offline DQM is executed as part of this processing and final merging. The DQM code we wrote, the HCAL Raw Data Monitor, is required to function in both scenarios.

C.2.1 Problem rates

For purposes of raising the alarm, the Hcal Raw Data Monitor produces 'Problem Plots' in the format standard to all HCAL DQM. These plots are arranged in the 'geographical' space of calorimeter cells: 72 azimuthal $i\phi$ angular intervals numbered 1 through 72, by 83 roughly polar $i\eta$ intervals numbered -41 through +41. At certain values of $i\eta$ the longitudinal segmentation of HCAL's subsections comprises more than one layer of calorimeter cells, known as a *depth*, such as the $i\eta = 29$ barrel-endcap transition region. Four plots suffice to show all of these, with some judicious labeling, as shown in Fig. ???. The color scale conveys the severity of any problems identified by the DQM software, ranging from tolerable (green) to quite severe. The beloved gray of ROOT defaults is used where there is nothing to note, and white conveys that no information is available.

C.2.2 Corruption in the Raw Data

The HCAL Raw Data Monitor produces plots pertaining to data corruption, data flow, and diagnostics. The corruption plots, Fig. ??, are meant to identify signs that the data are untrustworthy for use in physics. These signs of corruption, at the highest levels, are violations of the CMS-wide data format specification and that of the DCC. At an intermediate level they are the disagreement of event-counting numbers between the DCC and the HTRs on its spigots, and the error flagging of both the HTRs and the LRBs. Finally on the finest level, individual channels are monitored for flags identifying any intrustworthiness.

The last three plots in the corruption folder feature a device borrowed from Eric Hazen's EricDQM package, the Insidious Box. Quantities being displayed for each hardware unit on the plot are arranged in a regular way around a rectangular outline. Examples are shown in Fig. ???. These keys require significant abbreviation to describe the meaning of each quantity displayed in something like the same area in which it is actually displayed. The abbreviations are therefore given in Tables ??, ??, and ??. Fortunately the technical specifications for the HCAL electronics hardware suggests nearly all of these abbreviations, and few if any had to be concocted for use in DQM

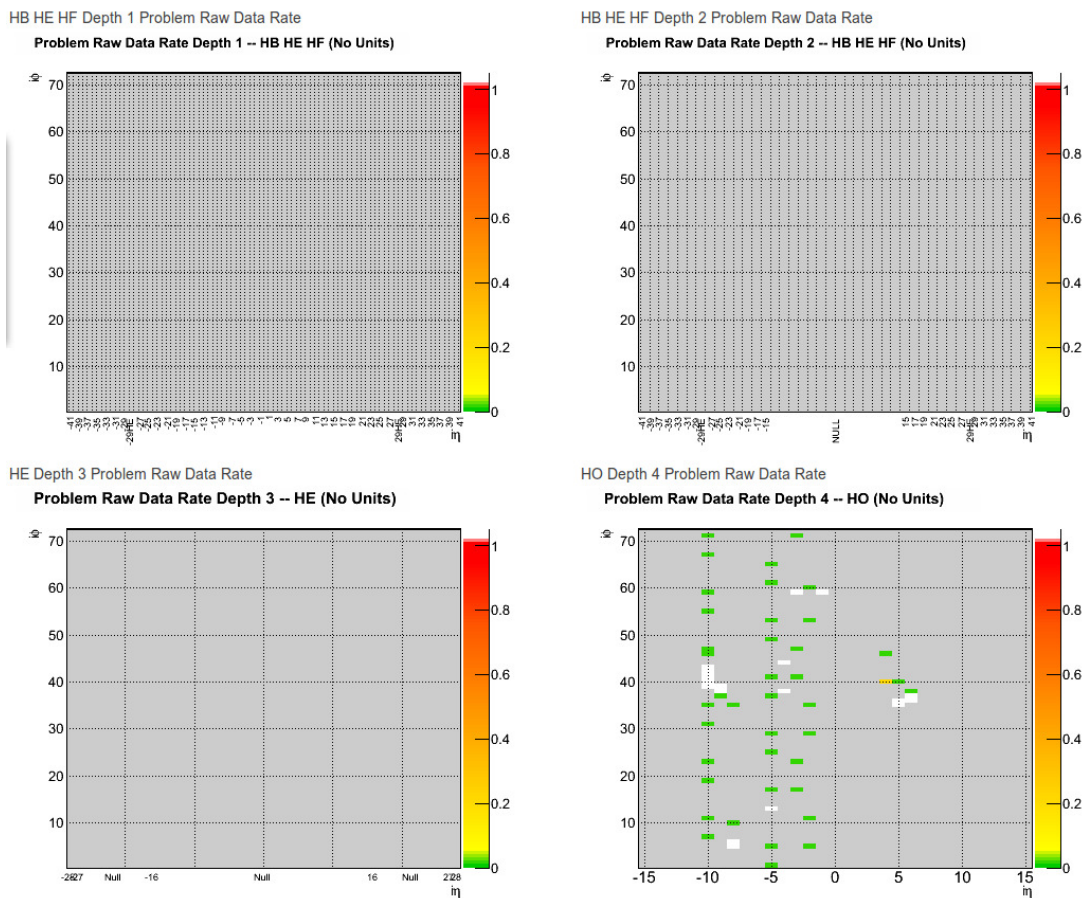


Figure C.3: The Problem Plots. Every cell of HCAL can be displayed in this standardized set of plots. Here there are scattered, low-level problems noted in the HO, as well as a few channels missing.

plots.

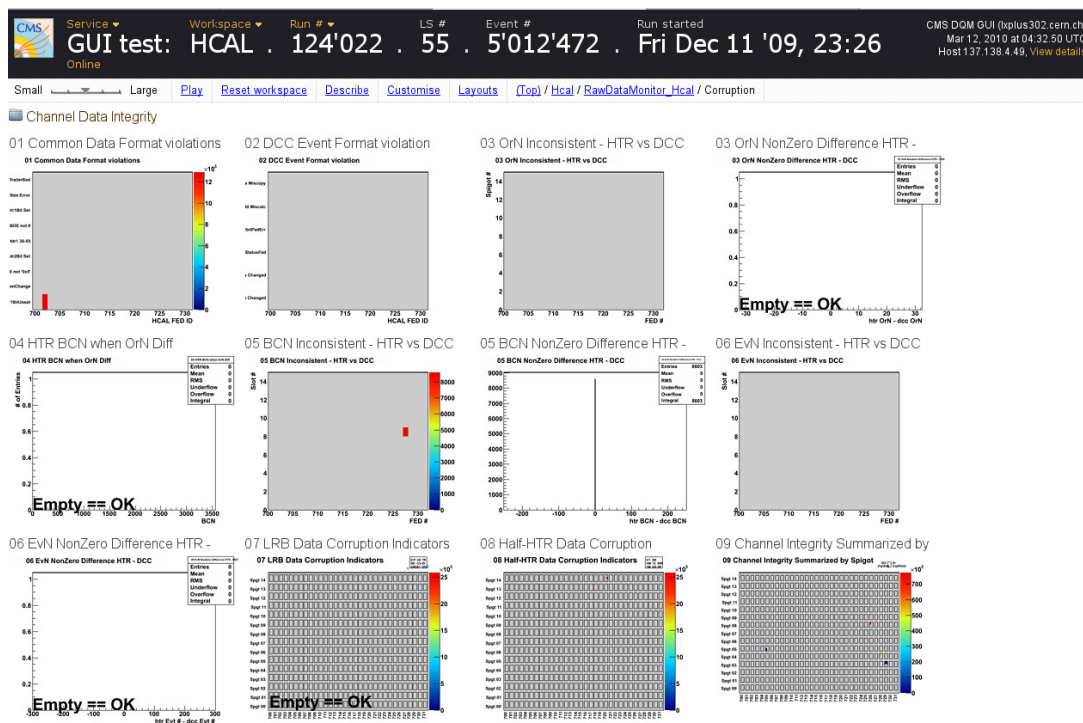


Figure C.4: The corruption folder.



Figure C.5: Left to Right: The keys for "07 LRB Data Corruption Indicators", "08 Half-HTR Data Corruption", and "09 Channel Integrity" demonstrating the Insidious Box. The abbreviated codes are explained in Tables ??, ??, and ??.

Table C.1: Abbreviations used in the plot "07 LRB Data Corruption Indicators".

E!P	The DCC spigot was enabled, but no data was received. This condition should never occur unless the DCC is set to "Auto-Resync" mode (not normally used) and a HTR for some reason did not respond to an L1A.
UE	An uncorrected error was seen in the HTR/DCC link. This could be caused by a bad or disconnected cable between the HTR and DCC.
TR	An event fragment longer than the maximum size of 2^{10} words was received from a HTR, and truncated by the DCC. This indicates a HTR or DCC firmware problem.
ND	An event fragment shorter than the minimum size of (4) 32-bit words was received from a HTR. This indicates a HTR (or DCC) firmware problem.
OV	The LRB input buffer on the DCC is in an overflow state. Data has been discarded by the DCC. This condition should not occur unless the TTS backpressure to the trigger is being ignored by the trigger.
ID	The low 8 bits of the EvN mismatch between the header and trailer in the HTR event fragment.
CRC	The DCC has detected a CRC error in an event fragment from the HTR. This implies that the DCC did not receive correctly what the HTR sent.
ST	The DCC received a header or trailer word from the HTR out of the correct sequence. If this is seen in combination with UE or CE conditions it could indicate a HTR/DCC cable problem.
ODD	The DCC received an odd number of 16-bit words from the HTR in an event fragment. This is not allowed, and indicates either a link problem or a HTR/DCC firmware problem.
T	The DCC event builder truncated the data received from the LRB portion of the DCC. Indicates a DCC firmware problem.

Table C.2: Abbreviations used in the plot "08 Half-HTR Data Corruption".

CT	This trigger was marked as a calibration trigger. This bit is likely no longer used, as it refers to a now-unused scheme for generating triggers using the TTC system. In particular, it does not indicate a legitimate orbit gap trigger. Thus, if this bit is set, it probably means that the data for the event was corrupted.
BE	The HTR did not see the expected number of BX in an orbit between BC0.
LW	Latency Warning. One LW per run per spigot may be ok.
HM	Histogramming Mode trigger. Should not occur in normal running.
15	Bit 15 in HTR status word was seen as '0'; should always be '1'.
WW	HTR word count is not consistent with "Empty Event" flag or "Compact Data" flag.
TM	HTR is set to Test Mode. Should not occur in normal running.
CK	HTR experienced a clock problem (delay lock loop on Xilinx not locked, or TTCrx not ready). Data should be considered suspect. Check for TTC (Trigger) system problems.
IW	Illegal Word Count Three lengths are considered valid: Length = 8 words (Empty Event size) Length = 12 + nTP + nDD (Compact event size) Length = 20 + nTP + nDD (Normal event size) Any other size will cause this bin to be filled.

Table C.3: Abbreviations used in the plot "Channel Integrity" plots, both the summary and the breakouts by DCC.

!DV	A HTR front-end link receiver reported non-valid data received from the front-ends. This is normal in calibration events, but high rates in only part of the hardware could mean that the link is experiencing errors (due perhaps to a bad optical fiber).
Er	A HTR front-end link receiver reported an error in the data received from the front-ends. This usually means that the link is experiencing errors.
DigiSize	The number of time-samples for a front-end channel does not match the number specified in the event header.
CapRotat	The CapID did not rotate correctly in the QIE data.

C.2.3 Monitoring Data Flow

In the Raw Data Monitor there are several plots of interest for monitoring the flow of HCAL's raw data, both its quality and its quantity. These are shown in Fig. ???. The first of these plots is another Insidious Box plot, whose key is given in Fig. ??. Two-character codes inside the box refer to the half-HTR: OverFlow Warning and Busy, the TTC states, and Empty Event. (EE is marked in the subsequent event.) Three-character codes refer to DCC states OverFlow Warning and Busy, the TTC states, and an overfull L1A FIFO. Below and to the left of the box, Corrected Errors are marked, referring to the HTR-DCC link's Hamming code, for error detection and correction.

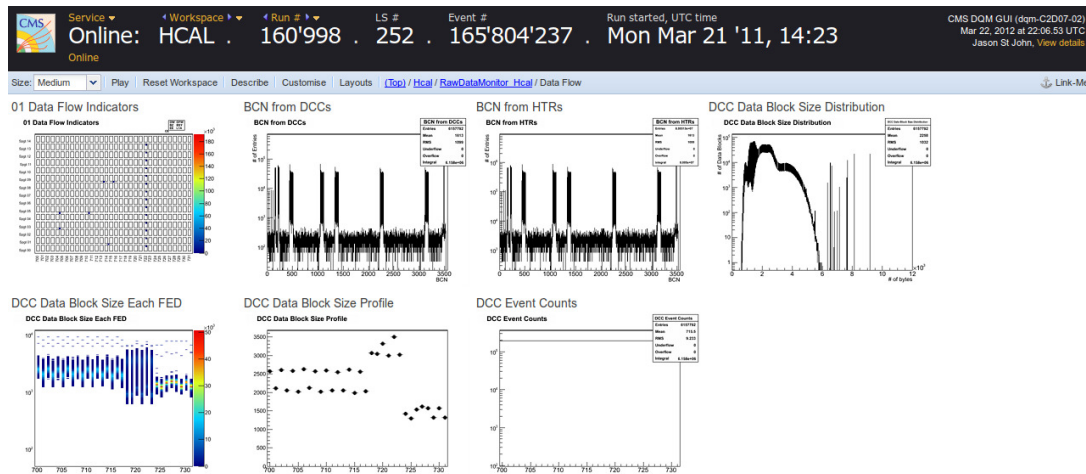


Figure C.6: The Data Flow folder.

OW	OFW
BZ	BSY
EE	L1A

CE

Figure C.7: The key for "01 Data Flow Indicators."

Bunch count number distributions are shown in semi-log scales for the half-HTRs and DCCs both. Peaks here demonstrate the bunch structure of LHC orbits as well as calibration triggers in the orbit's abort gap.

Three plots give the distributions of raw event fragment sizes for all DCCs together, individually, and individually on average. These plots make the effect of HCAL zero suppression quite evident. When channels are suppressed the size of the event fragment is lower, but rarely are all channels suppressed. A periodic, systematic 'unsuppression' gives events of maximal size. The two plots showing individual

DCCs also show clear structure from HBHE (FEDs 700-717), HF (FEDs 718-723), and HO (FEDs 724-732). Special calibration channels are included in every other DCC, giving a picket fence alternation to the fragment sizes of sequential FEDs. The different channel count and number of calorimeter time samples included per event gives this subsystem dependency of the data fragment size.

A final plot simply checks that every DCC reported the same number of events of any size, as required of all CMS FEDs.

C.2.4 Expert Diagnostics

The plots in the Diagnostics folder reflect quantities of interest only for a handful of experts on the HCAL electronics. They include histograms of the firmware and data format version numbers, status bits from the metadata whose meaning changes with the firmware version numbers, and the distribution of these status bits in the HCAL electronics crates. For non experts, these are perhaps not terribly illuminating. The plots are maintained for occasional post-mortem diagnostic use in close cooperation with one of these experts.

Bibliography

- [1] [CMS collaboration], G. L. Bayatian *et al.*, Journal of Physics **G34**:995–1579 (2007).
- [2] E. Rutherford, Philosophical Magazine **6**:21 (1911).
- [3] R. W. McAllister and R. Hofstadter, Physical Review **102** (1956).
- [4] P. A. M. Dirac, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **117(778)** (1928).
- [5] C. D. Anderson, Physical Review **43**:491–494 (1933).
- [6] [Http://hep.pa.msu.edu/cteq/public/cteq6.html](http://hep.pa.msu.edu/cteq/public/cteq6.html).
- [7] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, jet Pair Production (1995).
- [8] S. Mandelstam, Physical Review **112**:1344–1360, <http://link.aps.org/doi/10.1103/PhysRev.112.1344> (1958).
- [9] T. Sjöstrand, S. Mrenna, and P.Z. Skands, Journal of High Energy Physics **05**:026 (2006).
- [10] CTEQ, European Physical Journal **4(3)**:463, <http://dx.doi.org/10.1007/s100520050220> (1998).
- [11] V. N. Gribov and L. N. Lipatov, Soviet Journal of Nuclear Physics **15**:438–450 (1972).
- [12] G. Altarelli and G. Parisi, Nuclear Physics B **126** (1977).
- [13] Y. L. Dokshitzer, Soviet Physics JETP - Journal of experimental and theoretical physics of the Academy of Sciences of the USSR **46**:641–653 (1977).
- [14] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Physics Reports **97** (1983).

- [15] B. Andersson, *The Lund Model*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, <http://books.google.com/books?id=PrTFh7XvJfEC> (1998).
- [16] [JADE Collaboration collaboration], *Zeitschrift fur Physik C* **21** (1983).
- [17] [L3 Collaboration collaboration], *Physics Letters B* **561** (2003).
- [18] [DELPHI Collaboration collaboration], *European Physical Journal C* **51** (2007).
- [19] [D0 Collaboration collaboration], *Physics Letters B* **464** (1999).
- [20] J. W. Rohlf, *Nuclear Physics B* **171**, FERMILAB-THESIS-1979-20 (1980).
- [21] J. W. Rohlf, *Acta Physica Polonica B* **36** (2005).
- [22] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Journal of High Energy Physics* **9708**:001 (1997). [arXiv:hep-ph/9707323]
- [23] M. Cacciari, G.P. Salam, and G. Soyez, *Journal of High Energy Physics* **04**:63 (2008).
- [24] U. Baur, M. Spira, and P. M. Zerwas, *Physical Review D* **42**:815–824, <http://link.aps.org/doi/10.1103/PhysRevD.42.815> (1990).
- [25] U. Baur, I. Hinchliffe, and D. Zeppenfeld, *International Journal of Modern Physics* **A2**:1285 (1987).
- [26] CERN, *CERN faq: LHC the Guide*, CERN Communication Group (2006).
- [27] CMS, *Journal of Instrumentation* (**2008 JINST 3 S08004**) (2008).
- [28] K. Nakamura and P. D. Group, *Journal of Physics G: Nuclear and Particle Physics* **37(7A)**:075021, <http://stacks.iop.org/0954-3899/37/i=7A/a=075021> (2010).
- [29] [CMS collaboration], S. Chatrchyan *et al.*, *Physics Letters B* **704(3)**:123 – 142 (2011).
- [30] ATLAS Collaboration, *Physics Letters B* **708(12)**:37 – 54 (2011).
- [31] R. Brun and F. Rademakers, *Nuclear Instruments and Methods in Physics Research A* **389**, see also <http://root.cern.ch/> (1997).
- [32] D. Crockford, *The application/json Media Type for JavaScript Object Notation (JSON)*, <http://tools.ietf.org/html/rfc4627> (2006).

- [33] V. Rapsevicius and the CMS DQM Group, Journal of Physics: Conference Series **331(4)**:042038, <http://stacks.iop.org/1742-6596/331/i=4/a=042038> (2011).
- [34] CMS, Journal of Instrumentation **6(11)**:P11002, <http://stacks.iop.org/1748-0221/6/i=11/a=P11002> (2011).
- [35] CMS, CMS Physics Analysis Summary (**CMS-PAS-JME-09-008**), <http://cdsweb.cern.ch/record/1259924>, cERN, Geneva Switzerland (2010).
- [36] R. Harris et al., CMS AN-2010-108 **2010/108** (2010).
- [37] CMS, CERN Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET (2009).
- [38] R. M. Harris and K. Kousouris, International Journal of Modern Physics **A26**:5005–5055 (2011). [arXiv:1110.5302]
- [39] R. Harris, *Discussion of the dijet mass spectrum fit form.*, private communication (2012).
- [40] [UA2 collaboration], J. Alitti *et al.*, Zeitschrift fur Physik **C49**:17–28 (1991).
- [41] [UA2 collaboration], J. Alitti *et al.*, Nuclear Physics B **400**:3–24 (1993).
- [42] [CDF Collaboration collaboration], Physical Review Letters **74**:3538–3543, <http://link.aps.org/doi/10.1103/PhysRevLett.74.3538> (1995).
- [43] [(CDF Collaboration) collaboration], Physical Review D **55**:R5263–R5268, <http://link.aps.org/doi/10.1103/PhysRevD.55.R5263> (1997).
- [44] C.J. Clopper and E.S. Pearson, Biometrika **26**:404 (1934).
- [45] A. L. Read, Prepared for Workshop on Confidence Limits, Geneva, Switzerland, 17-18 Jan 2000 (2000).
- [46] R.D. Cousins and V.L. Highland, Nuclear Instruments and Methods **A320**:331 (1992).
- [47] R.D. Cousins, J.T. Linneman, and J. Tucker, Nuclear Instruments and Methods **A595**:480–501 (2008).
- [48] CMS, Physics Letters B **704(3)**:123 – 142, <http://www.sciencedirect.com/science/article/pii/S037026931101077X> (2011).
- [49] P. B. Cushman and A. H. Heering, SLAC www.slac.stanford.edu/pubs/icfa/fall02/paper1/paper1.pdf (2002).

- [50] C. Joram, Nuclear Physics B - Proceedings Supplements **78(1-3)**:407 – 415, <http://www.sciencedirect.com/science/article/B6TVD-40YSHKS-KX/2/033cdc5ffe46b2cd30705536adcf9f8>, advanced Technology and Particle Physics (1999).
- [51] T. Zimmerman and J. R. Hoff, IEEE Journal of Solid State Circuits **39**:895–905 (2004).
- [52] P. Moreira, T. Toiff, A. Kluge, G. Cervelli, A. Marchioro, and J. Christiansen, CERN Microelectronics group (2005).
- [53] E. Hazen, A. Heister, P. D. Lawson, J. Rohlf, A. Rubenstein, J. St. John, and S. X. Wu, Nuclear Instruments and Methods (2012).
- [54] O. Boyle, R. McLaren, and E. van der Bij, *The S-LINK Interface Specification*, <http://cmsdoc.cern.ch/cms/TRIDAS/horizontal/docs/s-link.pdf> (1997).
- [55] A. Racz, R. McLaren, and E. van der Bij, *The S-LINK 64 bit extention specification*, <http://cmsdoc.cern.ch/cms/TRIDAS/horizontal/docs/slink64.pdf> (2003).
- [56] <http://xdaq.web.cern.ch>, xDAQ generic data acquisition software (2009).
- [57] *Hardware Access Library*, <http://cmsdoc.cern.ch/~cschwick/software/documentation/HAL/index.html> (2007).
- [58] CMS Electromagnetic Calorimeter Group and E. di Marco, Nuclear Physics B Proceedings Supplements **197**:267–270 (2009).
- [59] G. B. et al., Journal of Physics: Conference Series **219(2)**:022038, <http://stacks.iop.org/1742-6596/219/i=2/a=022038> (2010).