

Florida State University Libraries

Electronic Theses, Treatises and Dissertations

The Graduate School

2011

Search for the Dark Matter Signature in the Lepton Jet Final State $\#(S) = 7$ TeV

Sergei V. Gleyzer



THE FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

SEARCH FOR THE DARK MATTER SIGNATURE IN THE LEPTON JET FINAL
STATE AT $\sqrt{S} = 7$ TEV

By

SERGEI V. GLEYZER

A Thesis submitted to the
Department of Physics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2011

The members of the committee approve the thesis of Sergei V. Gleyzer defended on May 10, 2011.

Vasken Hagopian
Professor Directing Thesis

Sanford A. Safron
University Representative

Harrison B. Prosper
Committee Member

Bernd Berg
Committee Member

Grigory Rogachev
Committee Member

Approved:

Mark Riley, Chair, Department of Physics

Joseph Travis, Dean, College of Arts and Sciences

The Graduate School has verified and approved the above-named committee members.

Dedicated to my family: Vladimir, Emma, Pearl, Michael, Vitaliy, Yana and Sonia

ACKNOWLEDGMENTS

I would like to thank and acknowledge everyone who helped make this dissertation come to fruition:

First and foremost, my wonderful advisor, Vasken Hagopian, to whom I would like to offer endless gratitude for his invaluable mentorship, guidance, care and intuition, that I often admire during our lengthy discussions of particle physics. These conversations went far beyond the scope of my dissertation research and have taught me a great deal about the field. I also would like to thank him for always being there whenever I needed anything, or wanted to know anything. For helping me always stay on track, focus and not forget the final goal. I also would like to thank Vasken for making me feel like a part of the family. Thank you, Vasken, for all of that. I will never forget how much you have done for me.

Sharon Hagopian, who has shown tremendous care to me in both professional and personal life. Thank you, Sharon, for your wonderful care and encouragement.

Harrison Prosper, for being my co-advisor, teacher and mentor in many ways. I will never forget how during my first summer of high energy research, Harrison asked me to study a book on decision trees and teach it back to him. I suspect that he already knew a great deal about decision trees, but it was an absolutely wonderful way to introduce me to the field and encourage me to pursue knowledge in this area. This early decision tree and rule ensemble project, together with Harrison's excellent explanations, encouraged my imagination and inventiveness and evolved into PARADIGM, a statistical and analytical framework that figures prominently in this dissertation. These events encourage and inspire me to this day. By direct example, hands-on application and often by typing on the keyboard next to me, Harrison introduced me to the various modern statistical methods and tools, an interest that I feel will take me beyond this dissertation work. I would like to thank Harrison for always caring, listening, encouraging me to explore my ideas to the fullest, for making me think outside of the box while maintaining sound grounding.

The members of my Committee: Grigory Rogachev, Bernd Berg, Sanford Safron. Thank you for following my progress over the years and offering constant encouragement in the process. Thank you, Grigory, for warm and lighthearted attitude as well as enlightening discussions over an occasional and welcome tennis match. Thank you, Bernd, for being a wonderful teacher as well. The theoretical knowledge you have passes on to me was of great help during the writing of this dissertation, far beyond the aspects of the Standard Model. Thank you, Sanford, for your constant and cheerful encouragement and excellent questions that made me question and test my understanding at every opportunity.

The entire CMS collaboration. It has been and I hope will continue to be a great adventure. It is very hard for me to contain all the excitement from looking at data at the

new energy frontier with a marvelous machine like the Large Hadron Collider. Especially, I would like to thank and acknowledge my collaborators from Princeton and Rice on the lepton jet analysis: Valerie Halyo, Nadia Adam, Paul Lujan, Chaouki Boulahouache, they have been a continuous pleasure to work with. For first bringing the topic of lepton jets to my attention, for being an early fan of PARADIGM and being a wonderful collaborator in all aspects of the lepton jet analysis, thank you, Valerie. Paul and Nadia, for wonderful discussions, for always being available for questions and discussions. Chaouki, for excellent discussions on all aspects of muon reconstruction and triggering.

The wonderful faculty members of the high energy physics group who I have not mentioned yet: Todd Adams, Susan Blessing, Horst Wahl, Howard Baer, Jeff Owens, Laura Reina, Andrew Askew and others. I would like to especially thank Horst, Todd and Andrew for their advice and encouragement with the various aspects of my analysis. I also would like to thank Andrew for his caring attitude over the years, it really helped along the way.

Professors at the FSU Physics department with whom I have developed many productive relationships that endured my tenure at CERN. In particular, I would like to thank Simon Capstick, Vladimir Dobroslavljevic, Per-Arne Rikvold, Nick Bonesteel, Jorge Piecarewitz, Stephan Von Molnar, Winston Roberts, Donald Robson, Samuel Tabor, Ingo Wiedenhover, Kirby Kemper, David Lind, Efstratios Manousakis, Mark Riley, Pedro Schlottman (for wonderful sense of humor especially) and Fred Huffer from the department of Statistics.

I would like to additionally mention Professors James Russ, Sara Majetich, Barry Luokkala, Hugh Young, Colin Morningstar, Manfred Paulini, Reinhard Schumacher, Indira Nair and especially Robert Kraemer and Fred Gilman for stimulating my interest in particle physics during my time at Carnegie Mellon. This happened prior to my beginning of graduate studies, however the spirit of academic and research rigour they conveyed, stayed with me throughout the years. Kunal Ghosh deserves a special mention, for providing valuable guidance at the best possible time. Also, I would like to thank Allan Blaer, my first college physics professor at Columbia University, who exemplified the value of clarity in lectures as well as Robert Quail, my first physics teacher, for stimulating my interest in both theoretical and experimental physics.

My friends and classmates who have provided endless support and encouragement. In particular, I would like to thank Lonnie Mott, Tiglet Besara, Eric Johnson, Bernard Fendler, Lei Chen, Khaled Adelaat, Shibi Rajagopalan, Stefano Profumo, Relja Vasic, Fernando Febres-Cordero, Paulo Rottman, Daniel Duggan and Ibrahim Abou-Hammad, for all the wonderful discussions we have had over the years. My classmates in Columbia: James Ulrich and Cyrus Golcaz for wonderful and fun discussions on physics topics. Max Negodaylov, Anthony Guselnikov, Dennis Mark, Val Trofimov and Nick Lazenko, with whom I always enjoyed a healthy academic competition while growing up.

I would like to additionally thank my friends and colleagues whom I have met while at CERN, in particular Ivan Furic, Itzhak Bars, Rene Brun, Thomas Speer, Lawrence Sulak, Edwin Norbeck, Duane Ingram, Darin Acosta, Giovanni Franzoni, Andreas Meyer, Ilaria Segoni, Lorenzo Agostino, Luca Malgeri, Jonatan Piedra, Ianna Osborne, Julia Yarba, Stephen Mrenna, Salavat Abdoulline, Martin Mulders, Pete Markowitz, Stephan Linn, German Martinez, Dragoslav Lazic, Serguei Vlassov, Alexander Singovsky, Luis Lebolo, Alexander Ledovskoy, Alexey Ferapontov, Chad Jarvis, Andy Kubik, Vladlen Timciuc,

Nick Kepreos, Ilknur Hos, Patrick Tsang, Yousi Ma, Mindaugas Janulis, Andres Abaudo Rodriguez, Sezen Sekmen, Hovhannes Khandanyan, Andrey Tsyganov, Edgar Sargsyan, Alec Avedisyan, Marie-Luce Falipou and many others. I would like to especially thank Aram Avetisyan for being a great sounding board of my various ideas and interesting discussions relevant to the CMSSW implementation of these ideas. Thank you for being a great chess partner as well.

I would like to thank Lorenzo Moneta, Sergey Bityukov and Kyle Kranmer for very insightful discussions about various statistical topics; Patrick Peplowski and Bryan field for carefully proofreading my dissertation and extremely useful comments, Igor Drozdovsky and Leo Georgiev for useful discussions on cosmology and astrophysics; Michael Peskin, Jesse Thaler, Lian-Tao Wang and Itay Yavin, for enlightening lepton jet discussions on theoretical aspects of this work; Jim Pivarski, for fruitful discussions about some of the early aspects of the analysis and other fun and entertaining discussions; Jeff MacDonald, for help with various computing issues over the years, especially in the beginning of my studies; Mike Polyakov for a wonderful piece of C++ on bit manipulation which made Paradigm at least a factor of 10 faster and more elegant; Denis Rubin, for interesting discussions at all hours of the day at night; Sherry Tointigh, Eva Crowdis, Sara Yount, Sara Stanley, Blake Sharin, Melissa Adams, Laurel Carlson, Kathy Mork, Maurizio Bertoldi, Kurtis Johnson and everyone without whom my time in Florida would not have been as pleasurable and memorable. Post-humously, I would like to thank and acknowledge Peter Sharp, Engin Isiskal, Jack Crow, Svetlana Zhutovskaya for the care they have taken in my academic and personal development.

Finally, I would like to thank my parents, Vladimir and Emma, who have supported me unconditionally from the very beginning; my brother Vitaliy, who provides constant and cheerful encouragement; my grandparents, Pearl and Michael, who are a lifelong inspiration; my dog Taz, for being an excellent messenger, not unlike the dark photons that appear in this dissertation. And finally, I would like to thank Yana, whose love, care and patience and have provided countless and limitless inspiration in ways which have made this work possible. The very last thank you is for Sophia, who appeared miraculously after the second chapter, and is an inspiration in of herself. You are a little ray of sunshine.

TABLE OF CONTENTS

List of Tables	xi
List of Figures	xii
List of Abbreviations	xvii
Abstract	xix
1 Introduction	1
1.1 Big Bang Theory	2
1.2 The Universe since the Big Bang	4
1.3 Dark and Visible Matter in the Universe	4
1.4 Dissertation Outline	5
2 Standard Model of Particle Physics and the Theory of Dark Matter	6
2.1 Standard Model of Particle Physics	6
2.1.1 Formulation	6
2.1.2 Probing Nature with Colliders	10
2.1.3 Challenges	10
2.2 Beyond the Standard Model	10
2.2.1 SuperSymmetry	11
2.3 Evidence for Dark Matter	11
2.3.1 Galactic rotation	12
2.3.2 Gravitational Lensing	12
2.3.3 CMB Anisotropy	14
2.3.4 Colliding Galaxies	15
2.4 Alternatives	16
2.4.1 Modification of Gravity and General Relativity	16
2.4.2 Superstring/M-Theory	16
2.5 Dark Matter Properties	16
2.5.1 Structure	17
2.5.2 Composition	17
2.6 Dark Matter Experimental Searches	18
2.6.1 Direct Detection	18
2.6.2 Indirect Detection	18
2.6.3 Collider Searches	20
2.6.4 Discussion	21
2.7 Supersymmetric Models of Dark Matter	21

2.7.1	Main production mechanisms	21
2.7.2	Other lepton jet production mechanisms	22
2.7.3	Lepton Jet Model	23
2.7.4	Analysis Strategy and Caveats	23
2.7.5	Monte Carlo Simulations	24
2.8	Lepton Jets	26
2.8.1	Algorithm	26
2.8.2	R-Parameter Optimization	26
2.8.3	Kinematical Distributions in Monte Carlo Signal	26
3	Description of the experimental apparatus	31
3.1	The Large Hadron Collider	31
3.2	The Compact Muon Solenoid Detector	34
3.2.1	Coordinate System	34
3.2.2	Magnet System	34
3.2.3	Inner Tracking System	34
3.2.4	Electromagnetic Calorimeter	36
3.2.5	Hadronic Calorimeter	37
3.2.6	The Muon System	39
3.2.7	Other Detectors	43
3.2.8	Trigger and Data Acquisition System	43
3.2.9	Software and Computing Model	44
3.2.10	Calibration and Alignment of the Detector	45
4	Event Reconstruction and Data Analysis	49
4.1	Event Model Data Format	49
4.1.1	RAW	49
4.1.2	RECO	49
4.1.3	AOD	49
4.2	Tracking of charged particles	50
4.3	Muon Reconstruction	51
4.3.1	Local Reconstruction	52
4.3.2	Segment Reconstruction	53
4.3.3	Seed Generation	54
4.3.4	Propagators	54
4.3.5	Standalone Muons	55
4.3.6	Global Muons	56
4.3.7	Tracker Muons	57
4.3.8	Arbitration	57
4.3.9	Calorimeter Based Muons	57
4.4	Jet Reconstruction	57
4.4.1	Jet Algorithms	58
4.4.2	Standard Jet Reconstruction in CMS	60
4.4.3	Jet Types	60
4.4.4	Jet ID	62

4.4.5	Jet Energy Corrections	63
4.5	Missing Transverse Energy Reconstruction	65
4.6	Event selection	65
4.6.1	Luminosity	65
4.6.2	Trigger Requirements	66
4.6.3	Vertex Requirements	68
4.6.4	Cosmic Rays	68
4.7	Muon Pre-Selection	68
4.7.1	Global Muons	68
4.7.2	Tracker Muons	69
4.7.3	Muon Acceptance Requirements	69
4.8	Lepton Jet Tagging	71
4.8.1	Kinematical Acceptance and Jet Tagging Efficiency	71
4.8.2	Muon Selection	72
4.9	Jet and Missing Transverse Energy Selection	72
5	Search for Dark Matter in Lepton Jets	74
5.1	Systematic Uncertainties and Efficiencies	74
5.1.1	Systematic Uncertainties	74
5.1.2	Efficiencies	75
5.2	Lepton Jet Kinematical Distributions in Data	75
5.2.1	Lepton Jet and Muon Multiplicity	75
5.2.2	Transverse Momentum, Pseudorapidity, Missing \cancel{E}_T and Angular Distributions	78
5.2.3	Lepton Jet Isolation	78
5.3	Muon Resonances with Lepton Jets	79
5.4	Final Lepton Jet Selection	79
5.5	Background Estimation	83
5.5.1	Testing the Background-only hypothesis	83
5.6	Paradigm	84
5.6.1	Variable Selection	84
5.6.2	Paradigm Predictions	85
5.6.3	Global Loss Function	85
5.6.4	Variable Boosting	87
5.7	Limit Setting on Cross Section times Branching Fraction	88
5.7.1	CLs Method	88
5.8	Results and Interpretation	89
6	Observation of Double J/ψ production	92
6.1	Theoretical Motivation	92
6.2	Double J/ψ Production	94
6.3	Event Selection	94
6.3.1	Dataset	94
6.3.2	Trigger Selection	95
6.3.3	Muon Selection	95

6.3.4	Vertex Requirements	95
6.3.5	Isolation	95
6.3.6	Final Selection	95
6.4	Event Displays	95
6.5	Invariant Mass Distributions	95
6.6	Kinematical Distributions of Double J/ψ events	97
7	Summary and Conclusions	99
A	Efficiency	100
A.0.1	Trigger	100
A.0.2	Muon efficiency	100
B	Paradigm	103
B.1	Introduction	103
B.2	Concept	103
B.3	Outline	104
B.4	Motivation	105
B.5	Definitions	105
B.6	Example 1	106
B.7	Example 2	108
B.7.1	Data Description	108
B.7.2	Results and Analysis	109
B.8	Decision Making Framework	114
B.8.1	Performance Optimization	114
B.9	Discussion and Summary	115
	Bibliography	118
	Biographical Sketch	125

LIST OF TABLES

2.1	Anti-Kt Lepton Jet Algorithm Parameters	26
4.1	Calorimeter Cell Thresholds	60
4.2	HLT Triggers Used in the Analysis	67
5.1	Systematic Uncertainties	74
5.2	Final Selection Cuts	80
5.3	Monte-Carlo Signal Efficiency after Final Selection	82
5.4	Variables selected for PARADIGM analysis	85
5.5	Signal-rich parameter hyperspaces	85
6.1	Selected Triggers	94
B.1	Hypothetical Example I Values	106
B.2	Hypothetical Example I Calculation	106
B.3	Hypothetical Example II Values	106
B.4	Hypothetical Example II Calculation	107
B.5	Single top quark variable set $\{V\}$	107
B.6	Degrees of Freedom in the Analysis	108
B.7	Optimal Set $\{S'\}$ for Reduction Using Cardinality Window $14 \leq S \leq 16$	113
B.8	Optimal Set $\{S'\}$ for Reduction Using Cardinality Window $23 \leq S \leq 24$	114

LIST OF FIGURES

1.1	Evolution of the Universe from Big Bang until Present Day [106]	2
1.2	Cosmic Microwave Background measured by WMAP [17]	3
1.3	Energy Budget of the Universe [106]	3
2.1	Three Generations of Matter	7
2.2	Typical Spiral Galaxy Rotation Curves: A. Predicted B. Observed	12
2.3	Images of the same object produced with the strong gravitational lensing of a foreground galaxy (Einstein Cross)	13
2.4	Strong Gravitational Lensing in Abel 1689 Indicating the Presence of Dark Matter.	13
2.5	Mass Distribution with Weak Lensing of the COSMOS Survey by the Hubble Space Telescope	14
2.6	Multipole Moments of the Cosmic Microwave Background [67]	15
2.7	Large scale dark matter distribution from a large N-body simulation, the Millennium Simulation, show hierarchical filamentary structure [110]	16
2.8	Measurements of energy dependent positron fractions from Pamela observations. Solid line shows the theoretical prediction using secondary positron production	19
2.9	Energy intensity as a function of gamma-ray energy measured by Fermi(red), ATIC (green) and PPB-BETS (purple). Dashed line shows the theoretical prediction using secondary positron production	19
2.10	Energy intensity as a function of gamma-ray energy as measured by INTEGRAL. Left: the central galactic region Right the periphery. Also shown is the fit to the power law plus positronium plus line [112].	20
2.11	The Feynman Diagram for squark production resulting in lepton jets and missing energy. \tilde{n}_2 and \tilde{n}_1 are dark fermions where $m_{\tilde{n}_2} > m_{\tilde{n}_1}$	22

2.12	Branching ratio of a light dark photon (γ_d). For dark photon masses $< 0.5 \text{ GeV}/c^2$ decays are predominantly into leptons [55].	23
2.13	Squark production cross section at the LHC	24
2.14	Event displays for typical simulated Monte-Carlo lepton jet events	25
2.15	Maximum ΔR between muons in a lepton jet Red : squark mass $650 \text{ GeV}/c^2$ Black : squark mass $250 \text{ GeV}/c^2$ [45]	27
2.16	Lepton Jet Multiplicity in MC signal	27
2.17	Number of Muons in Lepton Jets in Signal MC	28
2.18	Lepton Jet P_T in MC Signal	28
2.19	Lepton Jet η in MC Signal	29
2.20	Lepton Jet ϕ in MC Signal. Distribution is uniform, as expected.	29
2.21	Lepton Jet Missing Transverse Energy for MC Signal	30
3.1	The Large Hadron Collider Experiments a. Aerial view where two major detectors ATLAS and CMS shown b. Underground locations of all LHC experiments	32
3.2	The Compact Muon Solenoid Detector	33
3.3	CMS Inner Tracking System	35
3.4	Pixel detector elements	35
3.5	CMS ECAL Detector [16]	36
3.6	CMS HCAL Detector	37
3.7	η segmentation and size increase in HE (left) and HF (right)	38
3.8	Quarter View of the CMS Muon System	40
3.9	Transverse View of the CMS Muon System	41
3.10	CMS Drift Tubes	41
3.11	Honeycomb Structure of the DT SuperLayers	42
3.12	CMS Cathode Strip Chamber	42
3.13	Resistive Plate Chamber Layers	43
3.14	CMS Tier Computing Centers across the world.	44

3.15	CMS Software Framework Modular Structure and Event Setup	46
4.1	Tangential to the particle trajectory makes an angle λ with the magnetic field	50
4.2	CMS Magnetic Field Map	51
4.3	CMS Muon Reconstruction	52
4.4	Muon Reconstruction Workflow	53
4.5	Two Trajectory States on a Common Surface	56
4.6	Cone-type Jet Reconstruction: a. list of initial 4-vectors b. and c. stable cones found d. splitting/merging algorithm applied to assign 4-vectors in overlap regions to a single jet e. final jets	58
4.7	Left: collinear unsafety Right: infrared instability in jet algorithms	59
4.8	Jet Reconstruction. Left: Calorimeter jets Center: Jets-Plus-Tracks jets, Right: Particle-Flow jets	61
4.9	Factorized Jet Energy Corrections in CMS. Minimum requirements are shown in solid boxes, optional in dashed boxes [39].	63
4.10	Global Muon Trigger Selection of up to 4 Muon Candidates	66
4.11	Barrel Track Finders combine input vectors from DT stations and form tracks with well-defined p_T	67
4.12	Jet Evolution: from partons to reconstructed jets.	70
4.13	Lepton jet kinematical acceptance as a function of the boost and pseudorapidity	71
4.14	Lepton Jet Tagging efficiency as a function of pseudorapidity and energy. Top left: tracker muons. Top right: trackseg muons. Bottom: global muons. The empty bins at high pseudorapidity and low boost are due to the kinematic acceptance.	73
5.1	Lepton Jet Multiplicity in Data	75
5.2	Number of Muons in Lepton Jets in Data	76
5.3	Lepton Jet P_T in Data	76
5.4	Lepton Jet η in Data. Most of the lepton jets are produced centrally	77
5.5	Lepton Jet ϕ in Data. The distribution is uniform, as expected	77
5.6	Lepton Jet Missing Transverse Energy in Data	78

5.7	Lepton jet isolation cones. Black: “inner signal” cone (not applied) Grey: “outer hollow” cone	79
5.8	Invariant Mass Spectrum of Opposite-Sign Muon Pairs [40]	80
5.9	Lepton Jet Mass Spectrum: top left: ω , ρ , η and ϕ resonances top right: $J/\psi, \psi'$ resonances bottom left: Υ (1S, 2S, 3S) resonances bottom right: Z^0 peak.	81
5.10	Event displays for a typical event passing the final selection. Red lines Muons Green Lines Tracks Green Cones QCD Jets Blue Rectangles HCAL Energy Red Rectangles ECAL Energy Red Arrow \cancel{E}_T	82
5.11	Predicted data-driven background in the signal-enriched region $R < 0.1$ using a missing transverse energy \cancel{E}_T template derived from the background-enriched $(0.3 < R < 0.7)$ region.	83
5.12	Relative Variable Importance in the Lepton Jet Analysis using PARADIGM. Lepton Jet η appears to be irrelevant for this analysis.	86
5.13	Upper limit on squark production cross-section times branching ratio. Blue: Observed Red: Theoretically predicted. Experimentally constrained cross-section limits represent a significant improvement over theoretical predictions	89
5.14	Upper limit on γ_D cross-section times branching ratio. Blue: $M_{\tilde{q}} = 300 \text{ GeV}/c^2$ Magenta: $M_{\tilde{q}} = 550 \text{ GeV}/c^2$	90
6.1	Typical Feynman diagrams for $gg \rightarrow 2J/\psi$ reaction [18]	92
6.2	Event displays for a typical double J/ψ events in data using Fireworks.	93
6.3	Left: Two-dimensional plot of the invariant mass of the first lepton jet LJ_1 as a function of the second lepton jet LJ_2 . Right: Close-up in three-dimensions. By definition $P_T(LJ_1) > P_T(LJ_2)$	96
6.4	Left: The invariant mass of the first lepton jet LJ_1 Right: The invariant mass of the second lepton jet. By definition $P_T(LJ_1) > P_T(LJ_2)$. J/ψ peaks are clearly visible in both cases. Additionally the ψ' peak is visible in the left figure.	96
6.5	Left: The invariant mass of the first lepton jet LJ_1 when the mass of the second lepton jet LJ_2 is in the J/ψ mass window $3.0 < M_{LJ_2} < 3.2$. Right: The invariant mass of the second lepton jet LJ_2 when the mass of the first lepton jet LJ_1 is in the J/ψ mass window $3.0 < M_{LJ_1} < 3.2$. J/ψ peaks are clearly visible in both cases.	97

6.6	Left: The P_T of the first lepton jet LJ_1 when the mass of the second lepton jet LJ_2 is in the J/ψ mass window $3.0 < M_{LJ_2} < 3.2$. Right: The P_T of the second lepton jet LJ_2 when the mass of the first lepton jet LJ_1 is in the J/ψ mass window $3.0 < M_{LJ_1} < 3.2$.	98
6.7	The invariant mass of the two lepton jets (left) and pseudorapidity of the lepton jets when both are in the mass window $3.0 < M_{LJ} < 3.2$ (right).	98
A.1	Trigger efficiency for High Level Muon Triggers: Left: HLT_MU11. Right: HLT_MU15.	101
A.2	Muon identification efficiency for single muons as a function of η and P_T	101
A.3	Correction function for close-by muons as a function of ΔR between the two muons. Black: Tracker Muons with at least two arbitrated segments Red: Global muons.	102
B.1	A typical plot of relative variable importance. Variable ID is assigned in the decreasing order of RVI	104
B.2	Feynman diagrams for electroweak production of single top quarks for the s - and t - channels respectively.	108
B.3	ROC spectrum for sequential increase in cardinality of $\{S\}$. a) $2 \leq S \leq 12$ b) $10 \leq S \leq 18$	109
B.4	Comparison between variable importance measures provided by PARADIGM and RULEFIT on an absolute scale	110
B.5	a. relative variable importance for $\{V\}$ b. relative variable importance results with $IM(alljets - bestjet)$ removed	110
B.6	<i>Gloss function</i> results for a) $1 \leq S \leq 26$ b) $14 \leq S \leq 16$	111
B.7	Comparison of classifier performance using RVBoost with PARADIGM and RULEFIT to that without RVBoost, for a fixed number of classifiers.	112
B.8	Relative variable importances in the random-seed approach using 1% of the classifiers.	115
B.9	The non-linear relationship between the global loss function and the classifier performance measure $F(S)$	116

LIST OF ABBREVIATIONS

LHC Large Hadron Collider
CMS Compact Muon Solenoid
QGP Quark Gluon Plazma
CMB Cosmic Microwave Background
COBE Cosmic Background Explorer
WMAP Wilkinson Microwave Anisotropy Probe
HBB Hot Big Bang
SUSY Supersymmetry
LSP Lightest Supersymmetric Partner
SM Standard Model of Particle Physics
QED Quantum Electrodynamics
QCD Quantum Chromodynamics
MSSM Minimal Supersymmetric Standard Model
HDM Hot Dark Matter
WDM Warm Dark Matter
CDM Cold Dark Matter
WIMP Weakly Interacting Massive Particle
MACHO Massive Compase Halo Object
CDMS Cryogenic Dark Matter Search
Integral International Gamma Ray Laboratory
MC Monte-Carlo
LEP Large Electron Positron Collider
LINAC Linear Accelerator
SPS Super Proton Synchrotron
ATLAS Toroidal LHC Apparatus
ALICE Large Ion Collider Experiment
LHCb Large Hadron Collider Beauty Experiment
CP Charge-Parity
ECAL Electromagnetic Calorimeter
EB Electromagnetic Barrel Calorimeter
EE Electromagnetic Endcap Calorimeter
ES Electromagnetic Preshower
HF Hadronic Forward Calorimeter
HE Hadronic Endcap Calorimeter
HO Hadronic Outer Calorimeter

HB Hadronic Barrel Calorimeter
CSC Cathode Strip Chamber
DT Drift Tube
RPC Resistive Plate Chamber
CASTOR Centauro and Strange Object Search
ZDC Zero Degree Calorimeter
ASIC Application Specific Integrated Circuit
FPGA Field Programmable Gate Array
RAM Random Access Memory
LUT Look up table
HLT High Level Trigger
EDM Event Data Model
CAF CERN Analysis Facility
RecHit Reconstructed Hit
SisCone Seedless Infrared Safe Cone Algorithm
ADC Analog-To-Digital Count
JPT Jets Plus Tracks Algorithm
PF Particle Flow
HPD Hybrid Photo Diode
NHF Neutral Hadron Fraction
CHF Charged Hadron Fraction
CEF Charged Electromagnetic Fraction
NEF Neutral Electromagnetic Fraction
HLX HF Luminosity Transmitter
HTR HCAL Trigger and Readout Board
GMT Global Muon Trigger
RECO Reconstruction
GEN Generation
RVI Relative Variable Importance
ROC Reciever Operating Characteristic
GF Global Loss Function
RVBoost Relative Variable Importance Boosting
SLAC Stanford Linear Acceleration Center

ABSTRACT

The Large Hadron Collider is pushing high energy physics in to a brand new territory. This extraordinary era may bring discoveries of unprecedented magnitude, delivering validation or extreme disappointment to the physics theories of the previous decades. By colliding particles at more than 3.5 times the center of mass energy of the TEVATRON accelerator at Fermilab National Accelerator Laboratory, the CERN Large Hadron Collider aims to produce particles in the mass range above those that are already known. At the same time, there are exciting possibilities for new physics in the low-mass range that may have gone unnoticed until now. An example of this is a GeV-scale dark sector with a colorful spectrum of new particles. This physics model produces unique signatures of collimated leptons at the Large Hadron Collider energies. In the first part of this work, we describe the interesting astrophysical evidence that motivates a search for lepton jets and focus our attention on a minimal supersymmetric standard model with a GeV-scale dark sector that produces this exciting signature. In the next part of the thesis, we describe a search using the Compact Muon Solenoid (CMS) detector for evidence of dark matter in events containing muonic lepton-jets produced in 7 TeV proton-proton collisions at the Large Hadron Collider. We employ a novel lepton jet algorithm and find no evidence of an excess of such events with respect to the rate predicted by the Standard Model and interpret the null result in terms of a recently developed supersymmetric theory of dark matter. In doing so, we severely constrain the theoretical model and its parameters with the actual data from the Large Hadron Collider. In addition, we report the first observation of double J/ψ production, a new physical process discovery at the next energy frontier.

CHAPTER 1

INTRODUCTION

Dissertations in particle physics typically start with the story of Democritus, who first proposed the existence of atoms to explain the world around him. This work is at the intersection of cosmology and particle physics with deeper origins in observational astronomy. Therefore, it seems more appropriate to start the story much earlier, with the beginning of astronomy.

I would like to begin the journey in the ancient Stonehenge located in modern England [49]. The inhabitants of Stonehenge built the first known observatory. They also made some of the earliest astronomical observations. Others had looked up at the skies before them; however to notice astronomical events and track them with a purpose was a step forward for humanity. This way, astronomy, the oldest of all sciences, was born.

Astronomy is universal. Ancient cultures in places as far away from each other as Asia, Mesopotamia and Mesoamerica, have developed astronomical tools, solar and lunar calendars, star catalogs and various ideas about events in the sky. Ancient Chinese observers kept track of notable astronomical occurrences, such as arrivals of comets or new stars. This meticulous methodology helped preserve the first known observation of a supernova in 185 A.D., a feat repeated by Tycho Brahe in the Western world only in the 16th century. The ancient Greeks developed astronomy based on precise mathematics, an approach followed by the Arabic scientists in later centuries. Following the Renaissance, Western astronomy was resurrected by Galilei, Copernicus, Kepler, Brahe and Newton, among others. Galilei's observations in favor of the heliocentric model changed the modern view of the world and ushered in the modern Scientific Age. The real story of the Universe, starts much earlier.

“Our whole Universe was in a hot dense state,
Then nearly fourteen billion years ago expansion started. Wait...
The Earth began to cool,
The autotrophs began to drool,
Neanderthals developed tools,
We built a wall (we built the pyramids),
Math, science, history, unraveling the mysteries,
That all started with the big bang!
It’s expanding ever outward but one day
It will cause the stars to go the other way,
Collapsing ever inward, we won’t be here, it won’t be heard

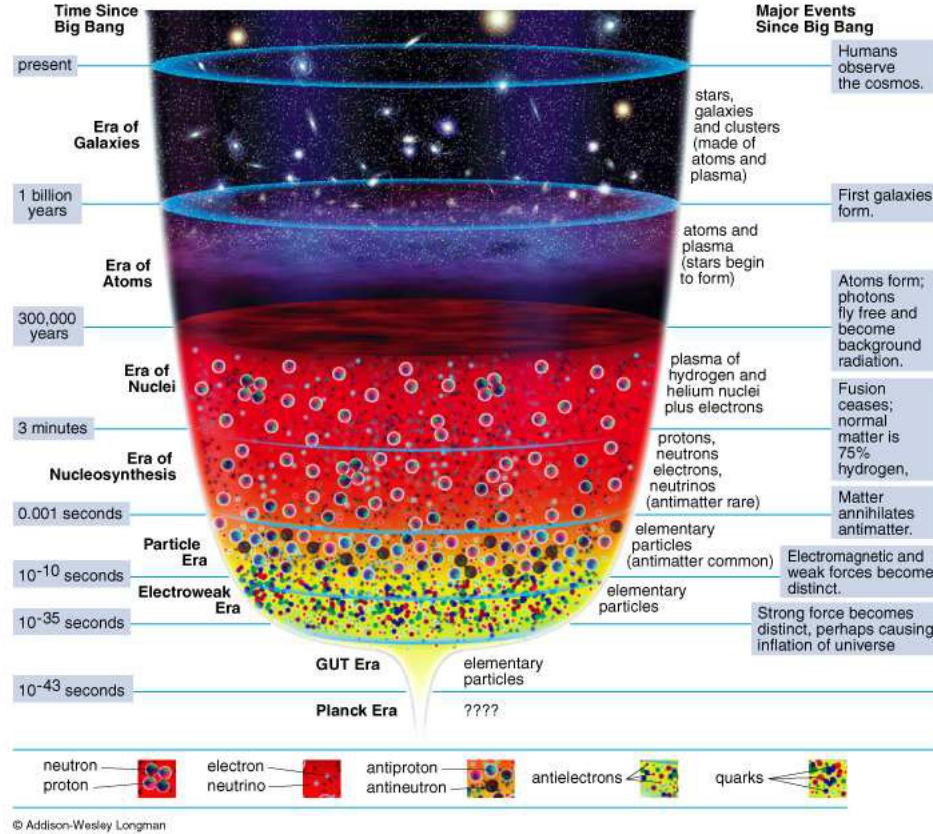


Figure 1.1: Evolution of the Universe from Big Bang until Present Day [106]

Our best and brightest figure that it'll make an even bigger bang!"

- Barenaked Ladies, theme song from The Big Bang Theory [78]

1.1 Big Bang Theory

As the song suggests, our Universe began as a hot dense explosion, appropriately called the Big Bang [87]. Initially everything was completely dominated by radiation, and the Universe underwent an extreme inflationary process known as the Hot Big Bang (HBB) [64]. During this stage, the Universe was extremely dense and hot, with mostly photons and neutrinos around, together with some matter in the form of particle-antiparticle pairs which spontaneously popped out of vacuum. At this early age of the Universe, there was continuous particle annihilation. Photons collided, creating particle pairs in a state of thermal equilibrium between matter and radiation. As the Big Bang evolved, the Universe expanded and the temperature dropped rapidly, bringing the average velocity of particles down [82].

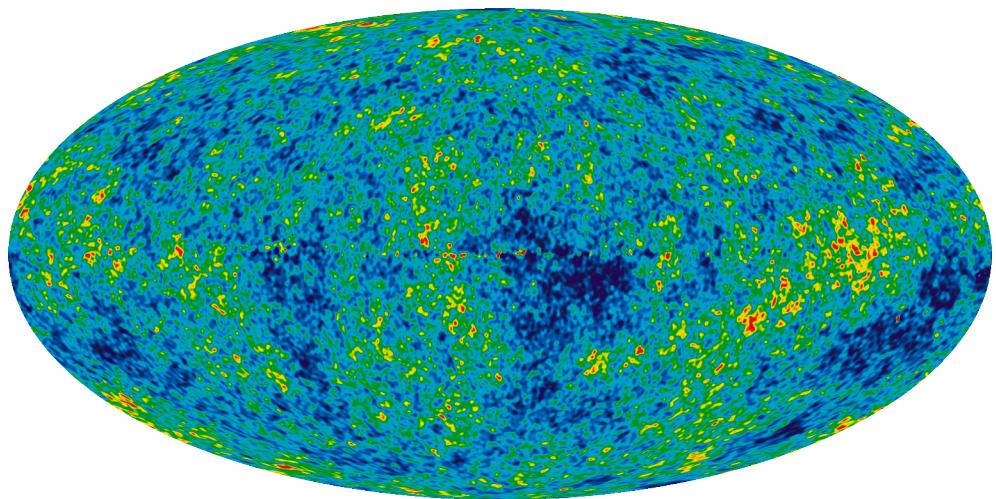


Figure 1.2: Cosmic Microwave Background measured by WMAP [17]

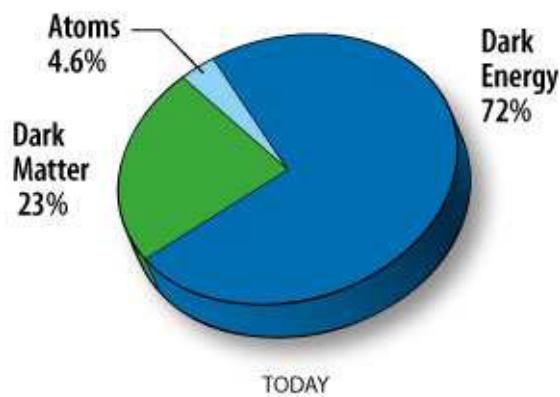


Figure 1.3: Energy Budget of the Universe [106]

1.2 The Universe since the Big Bang

The very young universe was very hot, with temperatures soaring above a billion Kelvin. Protons and neutrons cannot produce deuterium in such temperatures and instead formed into a quark-gluon plasma (QGP) as shown in Figure 1.1. Once the universe cooled and expanded, the earliest of the atoms formed.

Today the remnants of the Big Bang can be seen in the Cosmic Microwave Background (CMB), first discovered by Arno Penzias and Robert Wilson [92], and precisely mapped by first the Cosmic Background Explorer (COBE) and later, by the Wilkinson Microwave Anisotropy Probe (WMAP) [109, 17]. Figure 1.2 shows the results of WMAP observations.

Photons decoupled from baryons and escaped their initial confinement at about 380,000 years after the Big Bang [106]. This is the earliest we know about our Universe from direct observations. They cooled over time to about 2.7 K following the evolution of a typical black body spectrum [82]. Eventually, star formation started from nucleosynthesis, and larger structures, such as galaxies, began to form as the Universe evolved to its present state, ever expanding.

Recent observations show the Universe is accelerating as it expands [100], implying the existence of a previously unknown force, called the dark energy. This mysterious force expands the fabric of space, counteracting the gravitational forces that tend to bring the Universe together. Estimates suggest that dark energy accounts for $72.1 \pm 1.5\%$ of the energy budget of the Universe, as shown in Figure 1.3 [106]. In addition to the dark energy, there is another unsolved mystery in the Universe, called dark matter.

1.3 Dark and Visible Matter in the Universe

Our first observation of dark matter occurred in the 1930s. Fritz Zwicky noted that very large amounts of invisible matter were necessary to properly explain the galactic velocities in the Coma cluster [120]. Ever since, this non-luminous matter has been known as dark matter. It is invisible because dark matter does not produce or emit light, and doesn't interact with normal matter in any other way except for gravity. In the decades that followed, more evidence emerged of its existence and composition [101, 117, 12, 33, 32, 86].

Dark matter has never been created in the lab or directly observed. Instead, its existence has been inferred from its gravitational effects on galactic clusters, galaxies and individual stars. As Figure 1.3 illustrates, the current estimates for the Universe's energy budget reveal that the non-baryonic dark matter makes up $23.1 \pm 1.3\%$ of the total energy budget, while baryonic matter makes up only $4.3 \pm 0.1\%$ [106], a factor of 5 less.

Various direct and indirect experiments have been conducted to search for dark matter. Section 2.6 describes the results of these experiments and their implications. Particle accelerators have also been used in the search for dark matter, under the assumption that dark matter is made up of particles, as some of the leading theories and observational evidence suggest [10]. This work is an example of a dark matter search with a particle collider, namely the Large Hadron Collider, also known as the LHC.

1.4 Dissertation Outline

The remainder of the Dissertation is organized as follows: Chapter 2 describes the Standard Model of Particle Physics as well as the observational and experimental evidence for dark matter from cosmology. Section 2.2.1 describes a promising theory for the physics beyond the Standard Model, called Supersymmetry (SUSY), which may manifest itself at the higher energy frontier accessible to the LHC. In Supersymmetry, the lowest supersymmetric partner (LSP) represents a bonafide dark matter candidate, accompanied by unique and observable lepton-rich signatures [13, 70]. A hidden-sector model that unifies the experimental observations and theoretical considerations for dark matter and predicts a unique lepton jet collider signature is presented in Section 2.7. Chapter 3 details the Large Hadron Collider experiment and the general purpose Compact Muon Solenoid detector. Data, collected with the CMS detector, are presented in Chapter 4, while the results of the search for dark matter in the lepton jets channel are shown in Chapter 5. Previously unseen double- J/ψ events observed during data analysis are described in Chapter 6. Final remarks and conclusions are made in Chapter 7.

CHAPTER 2

STANDARD MODEL OF PARTICLE PHYSICS AND THE THEORY OF DARK MATTER

“By convention there is color,
By convention sweetness,
By convention bitterness,
But in reality there are atoms and space”
-Democritus [50](circa 400 BC)

2.1 Standard Model of Particle Physics

The current knowledge about particle physics came together during the 1970s into an encompassing field theory of particles and interactions, called the Standard Model (SM) of Particle Physics ¹. Despite its incompleteness, the Standard Model is the most precise and extraordinary well-tested theory of all time, with predictions that have been experimentally tested up to the 9th decimal point, or one part in a billion. This constitutes a very impressive feat for any modern theory in any scientific field [90].

The Standard Model combines quantum mechanics and special relativity into one field-theoretic framework that describes the interactions between the most fundamental constituents of matter, such as quarks and leptons, as well as the mediator particles, responsible for various interactions between them. In Section 2.1 the Standard Model is explained, together with the challenges it currently faces. Section 2.2 focuses on new physics theories beyond the Standard Model, while Sections 2.3 - 2.6 discuss the properties and evidence for the existence dark matter. Section 2.7 describes the theoretical modeling of dark matter for collider searches. We conclude the chapter with the description of the lepton jet algorithm in Section 2.8.

2.1.1 Formulation

The Standard Model is a combination of an Abelian and a Non-Abelian spontaneously-broken gauge theory based on a Lie Algebra group structure². It merges the Electroweak Theory, described by the $SU(2) \otimes U(1)$ gauge group, with the theory of strong interactions,

¹For a thorough review of the Standard Model please see Ref. [93]

²For a review of Lie Algebra and symmetry groups please see Ref. [65]

Three Generations of Matter (Fermions)			
	I	II	III
mass→	2.4 MeV	1.27 GeV	171.2 GeV
charge→	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
name→	up	charm	top
Quarks			
mass→	4.8 MeV	104 MeV	4.2 GeV
charge→	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
name→	down	strange	bottom
Leptons			
mass→	<2.2 eV	<0.17 MeV	<15.5 MeV
charge→	0	0	0
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
name→	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino
Bosons (Forces)			
mass→	0.511 MeV	105.7 MeV	1.777 GeV
charge→	-1	-1	-1
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
name→	electron	muon	tau
mass→	80.4 GeV		
charge→	± 1		
spin→	1		
name→	W^\pm weak force		
mass→			91.2 GeV
charge→			0
spin→			1
name→			Z^0 weak force

Figure 2.1: Three Generations of Matter

described by the $SU(3)$ symmetry group, in order to unify all of the known fields and forces except gravity.

At currently accessible energies all of the elementary particles appear to be pointlike and indivisible. Quarks and leptons are fermions: particles with spin angular momentum of $\frac{1}{2}\hbar$. In contrast, particles with integer spin angular momentum are known as bosons.

Fermions are divided into three families or generations, as shown in Figure 2.1. There are 6 known quarks in nature (fermions with a fractional charge): up(u), down(d), strange(s), charm(c), bottom(b) and top(t). The top quark, also known as the truth quark, was the last to be discovered by the CDF and D0 collaborations in 1995 and is by far the heaviest of the family. An electron(e) is the most recognizable lepton, but it also has two heavier cousins, called the muon(μ) and the tau(τ). Other leptons include the three generations on neutrinos(ν), which are treated as massless particles in the Standard Model. It has been experimentally shown in the previous decade that neutrinos have (albeit very small) mass, making this assumption incorrect.

Gluons(g), photons(γ), the W^\pm and the Z^0 particles are the mediators of the fundamental interactions, called gauge bosons. In other words, they are the force carriers of a particular field. For instance, the strong (electromagnetic) force is carried by a gluon (photon). Force carries can be both massive and massless depending on the gauge symmetry involved. The Electroweak $SU(2) \otimes U(1)$ gauge symmetry is spontaneously broken, therefore its gauge bosons, W^\pm and Z^0 , are massive. On the other hand, the strong and electromagnetic interactions are unbroken symmetries, resulting in massless gauge bosons. Leptons, as well as gauge bosons, are either neutral or have an integer charge. Various charges, masses, and spins of quarks, leptons and gauge bosons are additionally summarized in Figure 2.1.

Yang-Mills Theories. There are two types of field theories used in the Standard Model: Abelian and non-Abelian. In Abelian theories, such as the theory of quantum electrodynamics (QED), main elements of a symmetry group associated with local and gauge transformations commute with each other. In the 1950s, Yang and Mills used a symmetry group with non-commuting generators to construct a new type of field theory. This non-Abelian construction has been used extremely successfully in particle physics, in particular in the theory of strong interactions called quantum chromodynamics (QCD).

Electromagnetic Interactions. The quantum theory of electrodynamics can be constructed with the following elements: an electromagnetic potential vector field, A_μ , that transforms as:

$$A_\mu \rightarrow A_\mu' - \frac{1}{e} \partial_\mu \theta \quad (2.1)$$

where $\theta = \theta(x)$, the local space-time coordinate. Because θ is a function of x , the symmetry under consideration is a local symmetry.

A covariant derivative, D_μ , defined as:

$$D_\mu \equiv \partial_\mu + ieA_\mu \quad (2.2)$$

and a free Lagrangian of the form:

$$L = \bar{\psi}(i\partial - m)\psi \quad (2.3)$$

are also used. Using the field strength tensor definition:

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu \quad (2.4)$$

the complete QED Lagrangian can be written as:

$$L_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\partial - m)\psi - e\bar{\psi}\not{A}\psi \quad (2.5)$$

where

$$J_\mu \equiv \bar{\psi}\gamma_\mu\psi \quad (2.6)$$

is the electromagnetic current density and γ_μ are the Dirac matrices.

The general family of transformations of the kind $U(\alpha) = e^{i\alpha(x)}$ is called $U(1)$, the unitary Abelian group. That means that the quantum electromagnetism theory we just constructed and its associated Lagrangian are invariant under such transformations. In order to preserve local gauge invariance, the force carrier of the electromagnetic field (the photon) has to be massless at all orders of perturbations, while the familiar electric charge e appears as the gauge coupling constant in the theory.

Quantum Chromodynamics. The theory of quantum chromodynamics describes strong interactions and has roots in the quark model, proposed by Murray Gell-Mann in the 1960s. The model predicts that strongly-interacting particles are made up of constituents, called quarks, which are fractionally-charged fermions with an additional quantum number called 'color'.

The strong sector is invariant under the local $SU(3)$ gauge transformations and the strength of the interaction grows with distance and weakens at short distances in an effect known as asymptotic freedom. This behavior prohibits the existence of isolated quarks in nature. Instead they occur in bound states called hadrons which come in two varieties: baryons and mesons. The former are of the form (qqq) , while the latter are of the form $(q\bar{q})$. Leptons are not affected by strong interactions because they do not carry color.

The strong sector part of the Standard Model Lagrangian can be summarized as:

$$L_{QCD} = -\frac{1}{4}G_{\mu\nu}^\alpha G^{\mu\nu,\alpha} + \sum_f \bar{\psi}_{f,i}(i\not{\partial}_{ij} - m_f\delta_{ij})\psi_{f,i} \quad (2.7)$$

where $G^{\mu\nu}$ is the field strength tensor, the slash notation signifies a contraction of a vector and a γ matrix, used in a special kind of algebra called Clifford Algebra³, and the covariant derivative in QCD is defined as:

$$D_{ij}^\mu = \partial^\mu\delta_{ij} - igA^{\mu,\alpha}T_{ij}^\alpha \quad (2.8)$$

where $A^{\mu,\alpha}$ is the gluon field with the color index α and the T_{ij}^α generators of the $SU(3)$ that obey the following relation:

$$[T^a, T^b] = if_{abc}T^c \quad (2.9)$$

where f_{abc} are the $SU(3)$ structure constants.

³ For a review of γ matrices and Clifford Algebra see [81]

Electroweak Unification and the Standard Model. The theory of weak interactions, that describes nuclear β -decay and the interactions between quarks and leptons via the intermediate W^\pm and Z^0 gauge bosons, was unified by Glashow, Salam and Weinberg in the 1960s with quantum electrodynamics to form the complete $SU(2)_L \times U(1)_Y$ theory of electroweak interactions. Together with quantum chromodynamics they constitute the $SU(3)_c \times SU(2)_L \times U(1)_Y$ Standard Model theory of particle physics.

2.1.2 Probing Nature with Colliders

Particle colliders are designed to explore the laws of forces and interactions between elementary particles. Large Hadron Collider (LHC) is the world's largest and highest energy particle accelerator. The LHC is a discovery machine, expected to address the many unanswered questions of the Standard Model, as well as verify some of the Standard Model predictions, such as the Higgs Electroweak Symmetry breaking mechanism. The LHC will reach further into the TeV scale territory and hopefully provide answers to some of the outstanding questions.

2.1.3 Challenges

Although extraordinary successful, the Standard Model is fundamentally incomplete. Currently, the Standard Model cannot explain the existence of neutrino masses or their oscillations from one flavor to another. It can not predict any of the elementary particle masses. Today's searches with the state of the art particle colliders, such as the Tevatron and the Large Hadron Collider, focus on the elusive Higgs boson, which may complete the Standard Model by explaining the origin of particle masses if it exists [79] or force a new theory to emerge. Additionally, dark energy and dark matter do not fit into the Standard Model equation despite their prevalence in the Universe.

Another omission is gravity, which is not a part of the Standard Model. The unification of quantum mechanics and gravity requires a theory operating on the fundamental Planck scale of $1.22 \times 10^{28} eV$, a factor of 10^{16} higher than what is accessible to today's colliders. This suggests a possibility of a more encompassing theory that is unobservable at current energies. A large energy scale gap of this kind may prevent our full understanding of the Universe for a long time. We hope that Nature is kind to us and presents evidence of new physics at the TeV scale.

2.2 Beyond the Standard Model

The Standard Model (SM) of particle physics has been widely successful in explaining experimental data from accelerator experiments. During the last decade there has been increasing evidence for physics beyond the Standard Model. Most notable is the cosmological evidence for existence of dark matter and dark energy, which together take up about 95% of the energy budget of the Universe. To explain these new phenomena, extensions to the Standard Model have been proposed [29, 51, 95]. One promising and theoretically elegant extension to the Standard Model is the Theory of Supersymmetry [13, 51], another is String Theory [51, 63, 95].

2.2.1 SuperSymmetry

Supersymmetry (SUSY) postulates the existence of new partners for the known particles of the Standard Model with similar properties but different spin. In this theory, scalar partners for all fermions are proposed, and fermionic partners exist for all the bosons. This helps resolve various field theory divergences with additional corrections coming from the hidden symmetry. The superpartners are usually labeled by an “s” in the beginning or an “ino” at the end of the particle name. For example, selectrons and squarks are scalar partners of electrons and quarks, and gluinos and photinos are fermionic partners of gluons and photons.

Supersymmetry was first proposed by Miyazawa in 1986 and later independently rediscovered by others [103]. Because scalar partners of fermions have not been observed experimentally, supersymmetric particles must be much heavier than their partners, meaning that Supersymmetry is a broken symmetry. What makes Supersymmetry especially attractive from theoretical standpoint is the fact that the lightest SUSY particle (LSP) does not interact strongly and escapes detection, making it an excellent dark matter candidate [62, 54].

Minimal Supersymmetric Standard Model. In the literature, the simplest supersymmetrization of the Standard Model yields a theory called the Minimal Supersymmetric Standard Model (MSSM). Supersymmetry is a broken symmetry at currently accessible energies. To accommodate this fact, the MSSM contains explicit soft SUSY breaking terms, which presents a problem for the theory because they are put in by hand in the Lagrangian. In all, the theory is defined by 124 parameters, which makes it flexible but not particularly predictive. That is to be compared to the 20 free parameters of the Standard Model. MSSM parameters include seven 3×3 supersymmetric partner mass matrices, 3 gaugino masses, 3 mixing matrices and additional parameters in the Higgs sector. More constrained models based on various theoretical motivations have been proposed [31].

Recent SUSY models [15, 10] define a $U(1)$ gauge symmetry that is broken at low energy. This leads to dark sector phenomenology that exhibits a lepton-rich signature[31].

2.3 Evidence for Dark Matter

A body of evidence exists for the presence of dark matter in the Universe. Dark matter was initially proposed by Fritz Zwicky in 1934 to explain the large missing mass needed to properly describe galactic rotation velocities in galactic clusters [120]. Visible mass from stars, cosmic dust, brown dwarfs or other known sources is insufficient to describe the galactic rotation behavior correctly and therefore, a previously unknown mass had to be postulated. Further evidence, such as weak gravitational lensing, behaviour of colliding galaxies and Cosmic Microwave Background anisotropy, solidified the theory behind the existence of dark matter.

After some skepticism that followed Zwicky’s observation, with more evidence coming along from other directions, astronomers embraced the picture in which dark matter dominates the matter landscape. Current estimates of the energy density budget of the universe show that the dark matter composes $23.3 \pm 1.5\%$ of the total energy budget of the Universe [106]. The prevalent view is that most of the visible stars and cosmic dust occupy the inner

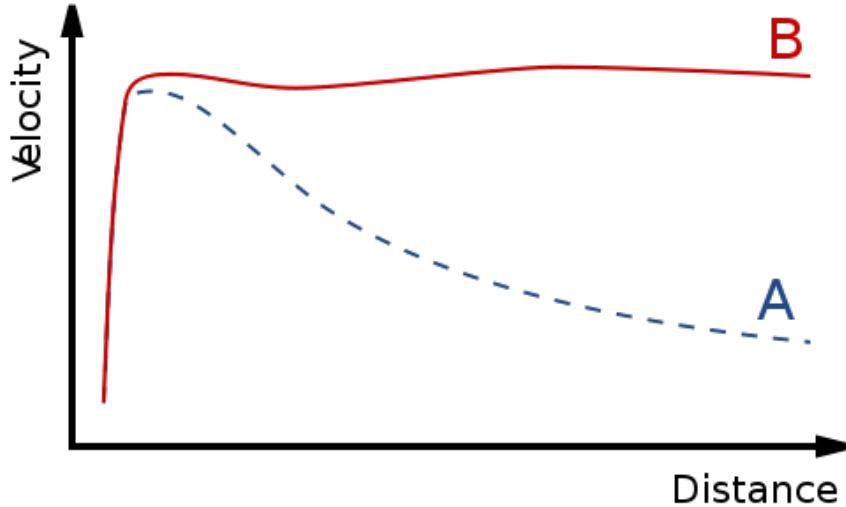


Figure 2.2: Typical Spiral Galaxy Rotation Curves: A. Predicted B. Observed

portion of galaxies, while a large spherical halo made up of dark matter permeates the galaxy in all directions [14].

2.3.1 Galactic rotation

Coma Cluster. Galactic rotation velocities were first studied by Zwicky in the context of the Coma cluster in 1934. He observed that the velocity dispersions of the galaxies in the cluster can not be adequately explained with visible matter and require a very large invisible mass to be present [120].

Stellar Doppler Shifts. In the 1970s, Vera Rubin studied the orbital velocities of individual stars in various galaxies by measuring their Doppler shifts. She found that stars further away from the galactic center do not travel slower, as expected [101]. Since most of the visible galactic mass is located near the center of the galaxy, the farther the star is from the center, the lower its orbital velocity should be. Rubin’s calculations showed that a correct explanation of observed Doppler shifts required a factor of two more galactic mass than what was visible.

Another noteworthy observation is the galaxy VIRGOHI21 which appears to be made up entirely of dark matter and contain almost no stars [86]. Such “dark” galaxies are theoretically predicted in the current theory of the Big Bang and structure formation; however, only a few have been detected so far.

2.3.2 Gravitational Lensing

Additional support for the dark matter picture comes from weak gravitational lensing of large structures such as the Abel 1969 cluster [52]. Because gravity warps space-time, light travelling in regions of gravitationally distorted space deviates from its normal direction of

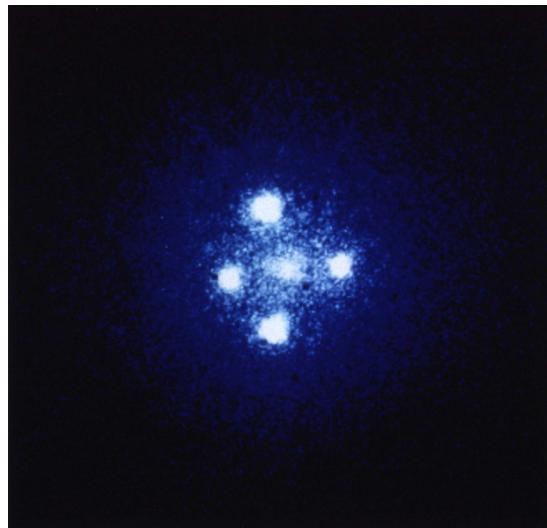


Figure 2.3: Images of the same object produced with the strong gravitational lensing of a foreground galaxy (Einstein Cross)

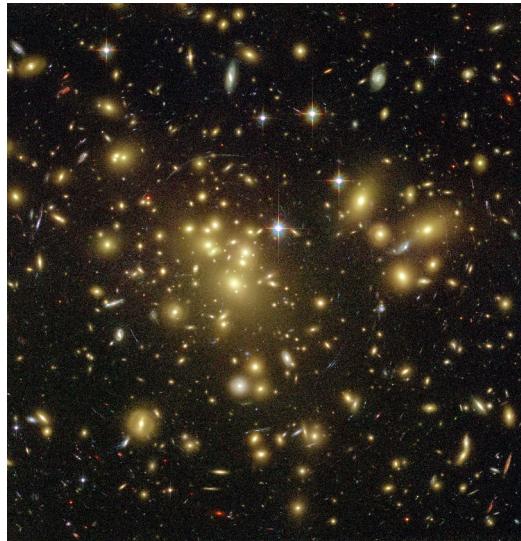


Figure 2.4: Strong Gravitational Lensing in Abel 1689 Indicating the Presence of Dark Matter.

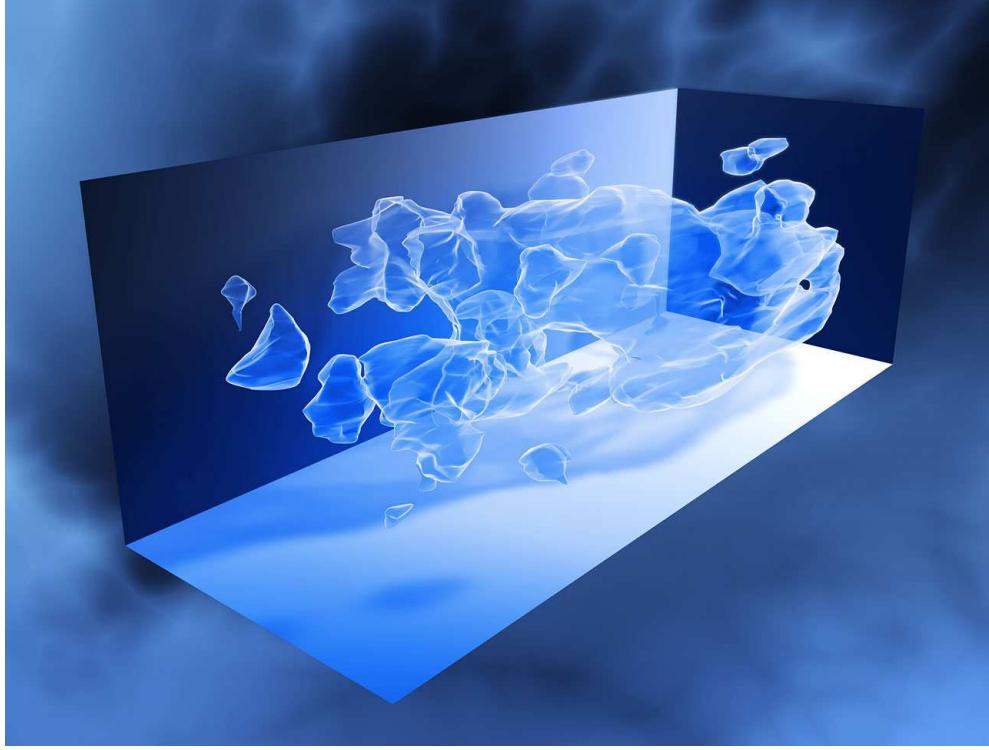


Figure 2.5: Mass Distribution with Weak Lensing of the COSMOS Survey by the Hubble Space Telescope

propagation, a phenomenon called gravitational lensing [89]. This effect makes it possible to detect gravitational effects of dark matter with a large structure like a cluster of galaxies. The idea of gravitational lensing was first theoreticized by Albert Einstein in his general theory of relativity [53]. Zwicky further suggested that one can use galactic clusters as gravitational lenses [121], an effect observed by Walsh et. al. in 1979 [89].

There are two types of gravitational lensing: strong and weak. In strong lensing the effects are visually pronounced with rings, arcs or sometimes multiple images of the same object, as shown in Figures 2.3 and 2.4. In weak lensing, the visual effects are less pronounced; however, a large statistical survey of objects shows a methodical distortion made by a gravitational lens on distant galaxies. This is used to make a map of the mass in an area, showing a background distribution of dark matter, as shown in Figure 2.5. Weak lensing observations of several galactic clusters, such as Abel 1969 and the Bullet Cluster, show that the total mass is much greater than the visible mass, and that the visible and dark matter do not coincide spatially [52, 32].

2.3.3 CMB Anisotropy

Cosmic Microwave Background was discovered in 1964 [92]. In 2009, Wilkinson Microwave Anisotropy Probe precisely measured the anisotropies in the Cosmic Microwave

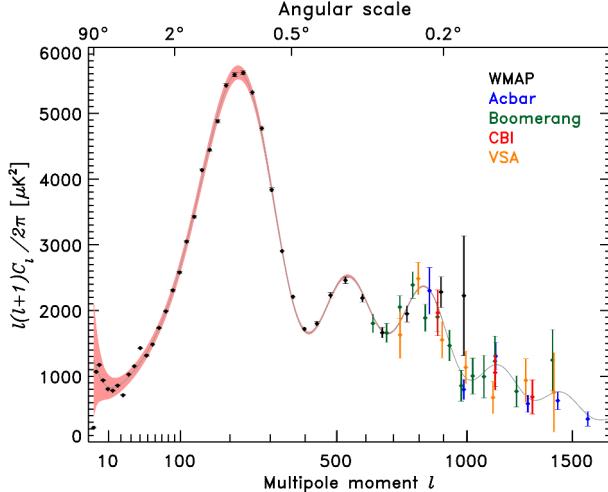


Figure 2.6: Multipole Moments of the Cosmic Microwave Background [67]

Background, ruling out several existing theories of cosmic structure formation and supporting theories that include dark matter [67]. The anisotropy of the CMB is explained by oscillations between photons and baryons at the time of decoupling when the Universe was about 380,000 years old. While normal matter interacts with the photons, dark matter does not, resulting in different and observable effects. To properly account for the power spectrum of the CMB shown in Figure 2.6, a large amount of dark matter is necessary, as in previous examples [67]. Calculations of the amount of dark matter needed to explain the CMB observations are consistent with contemporary measurements of dark matter content in the Universe [67, 82].

2.3.4 Colliding Galaxies

Observations of colliding galaxies, such as those in the Bullet cluster, are probably the best piece of evidence for dark matter [32]. When several galaxies collide, dark and visible matter behaves differently. Dark matter does not interact and passes through, while intergalactic gas and dust, which represents a large portion of the visible mass of the galaxies, undergoes electromagnetic interaction, slowing down significantly in the process. Because of the relatively large interstellar distances, stars are mostly unaffected by the collision. The study of the gravitational lensing shows a spatial discrepancy between the concentration of dark and visible matter. Moreover, changes of gravitational laws can not account for this discrepancy [32]. In the latter case, the lensing is predicted to follow the visible matter; however, the concentration of mass was demonstrated to be in different locations of the galaxies. Other colliding clusters support this picture as well [25].

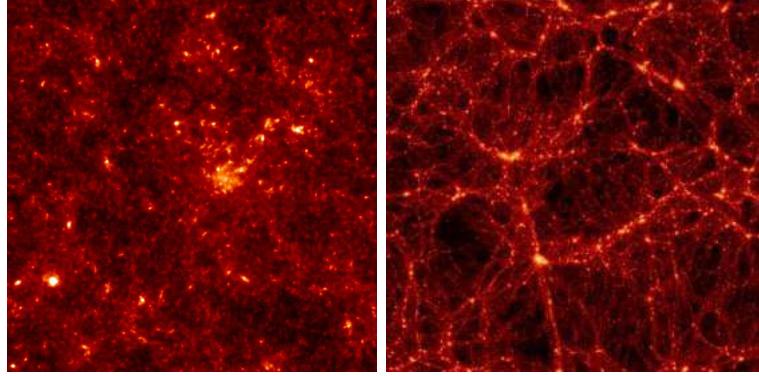


Figure 2.7: Large scale dark matter distribution from a large N-body simulation, the Millenium Simulation, show hierarchical filamentary structure [110]

2.4 Alternatives

2.4.1 Modification of Gravity and General Relativity

There have been proposals to modify the laws of general relativity and gravity that attempt to explain the observational evidence without dark matter [85]. However, recent observations of the merging of two galactic clusters described in Section 2.3.4 show that the visible matter and gravitational potential do not spatially coincide, meaning that dark matter's existence is a much more likely scenario than a modification of the laws of general relativity and Newtonian gravity.

2.4.2 Superstring/M-Theory

String theory is a promising theory that attempts to unite quantum mechanics and general relativity by proposing oscillating strings as the underlying elementary objects. Superstring theory is an extention of string theory that includes supersymmetry, providing a connection between bosons and fermions. String theory requires extra dimensions which may be too small to observe. Hypothetical particles that exist in those hidden dimensions may account for some of the effects currently attributed to dark matter. The main criticism of String Theory is that it is difficult to verify at the energy scale available to colliders.

2.5 Dark Matter Properties

Despite a large body of indirect evidence of its existence, dark matter has not been definitely observed. A fundamental property of dark matter is that its constituents do not interact with one another with the exception of gravity. Therefore, due to its lack of electroweak interactions, dark matter is not visible in the electromagnetic spectrum.

2.5.1 Structure

Dark matter particles do not bump or collide into one another, and tend to clump together. Recent models show that dark matter clusters in filaments, attracting normal matter to the intersections, as shown in Figure 2.7 [110]. These intersections appear to coincide spatially with large-scale structures in the Universe, such as large galactic clusters, adding to the importance of dark matter in galactic formation.

2.5.2 Composition

Baryonic vs. Non-baryonic. Weak gravitational lensing, galactic X-ray emission and velocity dispersions give much higher matter density [70] than the expected baryonic matter density from Big Bang nucleosynthesis [91]. Additionally, observations of the Cosmic Microwave Background estimate the total matter density to be 7 times larger than the estimated baryonic density [70], which implies that most of the dark matter is non-baryonic.

Several non-baryonic dark matter candidates have been proposed. They can be generally classified into three types: “hot”, “warm” and “cold”, depending on the speed with which a particle is moving during galactic formation.

Hot Dark Matter. A Hot Dark Matter (HDM) candidate moves at an ultra-relativistic speed when the galaxy is formed. An example of this candidate is a neutrino of any flavor. They are weakly interacting like WIMPs, described in Section 2.5.2, but have much smaller mass. Using CMB observations it is believed that the average neutrino mass does not exceed $0.3 \text{ eV}/c^2$ and therefore, the total matter density from non-relativistic neutrinos is $< 2 \times 10^{-2}$ [82]

Warm Dark Matter. A Warm Dark Matter (WDM) candidate was traveling at a slower relativistic speed $v_{dm} < .95 \text{ c}$ during galactic formation than the Hot Dark Matter candidate but faster than the Cold Dark Matter candidate ($v_{dm} > .1 \text{ c}$).

Neither hot nor warm dark matter can explain the galactic velocity dispersions because they move too fast. Similarly, they can not be used to explain the large scale galactic cluster formations because they cannot clump together due to their high velocity.

Cold Dark Matter. A Cold Dark Matter candidate moves at non-relativistic speed during galactic formation ($v_{dm} < .1 \text{ c}$), which makes it the most likely dark matter candidate. Principal CDM candidates are axions, Massive Compact Halo Objects (MACHOs) and Weakly Interacting Massive Particles (WIMPs).

Axions are hypothetical particles which have been postulated to address the strong charge-parity problem, which is a lack of strong processes that violate charge parity [107].

Massive Compact Halo Objects (MACHOs) are dense compact baryonic objects such as brown dwarfs. Studies of Big Bang nucleosynthesis suggest that baryonic MACHOs can not account for the large fraction of non-baryonic dark matter in the Universe [91]. Large black holes are ruled out as dark matter candidates due to gravitational lensing data. An extremely large number of tiny black holes remains an unlikely possibility for dark matter.

Weakly Interacting Massive Particles (WIMPs) are the leading dark matter candidates. They can be used to explain the large ratio of baryonic to non-baryonic matter in the Universe due to the relic abundance of dark matter WIMPs after a dark matter “freeze out”. Initially, there is thermal equilibrium between baryonic and non-baryonic matter

because of annihilation reactions of the type $X\bar{X} \rightarrow L\bar{L}$, where X is a dark matter particle and L is a standard model particle. The “freezout” occurs when dark matter annihilation stops when the initially hot Universe cools down, bringing the WIMP number density down [70]. Calculations with WIMPs that have TeV-scale mass show a relic abundance that is very close to the expectations from cosmological considerations [91]. This coincidence makes stable WIMPs an appealing dark matter candidate and experimental WIMP searches have been made, described below.

2.6 Dark Matter Experimental Searches

Experimental searches for Dark Matter can be classified into three categories: direct detection, indirect detection and collider searches. Direct detection experiments search for the scattering of dark matter particles off nuclei. Indirect detection experiments search for dark matter annihilation products. Collider searches attempt to find WIMPs produced in high energy collisions such as those in the LHC.

2.6.1 Direct Detection

Direct detection experiments are usually located deep underground to combat the large cosmic ray background on the surface. Examples of direct detection experiments located underground include Gran Sasso, Soudan Mine, SNOLAB, and others.

Direct detection experiments also include DAMA/LIBRA and Cryogenic Dark Matter Search (CDMS). The two differ in the types of detectors used in the scattering: CDMS uses germanium while DAMA uses sodium iodide crystals. Additionally DAMA looks for annual signal modulations due to a possible WIMP wind as the Solar System passes through the WIMPs, similar to the search for ether a century ago [20]. DAMA has reported a WIMP signal which is in the region excluded by CDMS [105]. In order to explain this conflicting result a new theory called Inelastic Dark Matter has been proposed, whereby the dark matter interaction with normal matter requires an energy level splitting and transitions into a higher energy state [56]. For a particular range of the splitting of about 100 keV it is possible to see a signal with DAMA but not with CDMS, due to the different masses of the crystals. Despite such explanations many questions remain about the DAMA result [115].

2.6.2 Indirect Detection

Indirect detection experiments include Pamela, ATIC, Integral, IceCube, and others. They generally look for the products of WIMP annihilation, such as gamma rays, anti-protons and positrons.

Pamela. Pamela is a space-based experiment with a silicon calorimeter that can distinguish the charge of incoming particles and measure their momenta. Pamela has observed an excess in the positron fraction up to the energies of 100 GeV, as shown in Figure 2.8, while simultaneously not observing any excess of anti-protons. This is significant because excess positrons could be produced in dark matter annihilations.

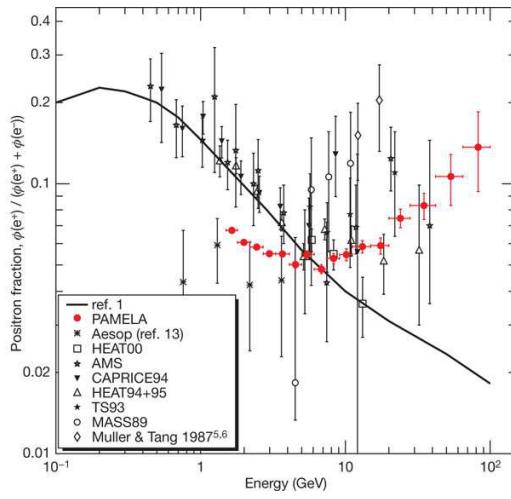


Figure 2.8: Measurements of energy dependent positron fractions from Pamela observations. Solid line shows the theoretical prediction using secondary positron production

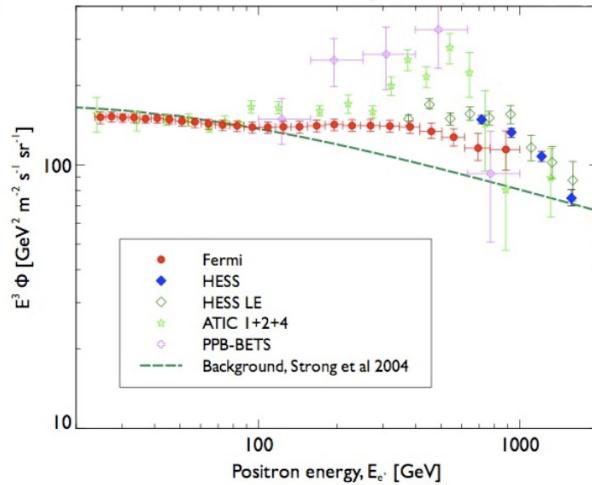


Figure 2.9: Energy intensity as a function of gamma-ray energy measured by Fermi (red), ATIC (green) and PPB-BETS (purple). Dashed line shows the theoretical prediction using secondary positron production

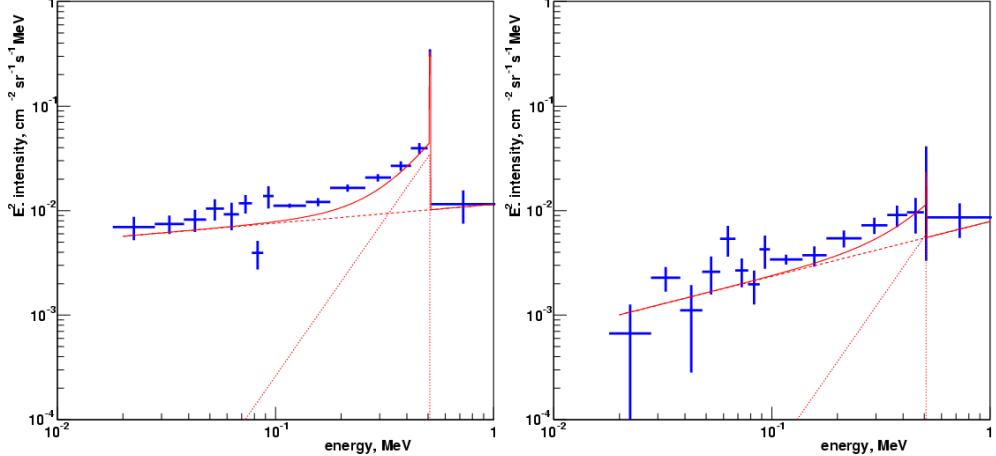


Figure 2.10: Energy intensity as a function of gamma-ray energy as measured by INTEGRAL. **Left:** the central galactic region **Right** the periphery. Also shown is the fit to the power law plus positronium plus line [112].

ATIC. ATIC is a balloon experiment that measures cosmic rays. It likewise found an excess of positrons or electrons (ATIC can not distinguish sign of the charge) in the 300 - 800 GeV range, shown in Figure 2.9. Other balloon experiments, such as Fermi and PPB-BETS, also observe electron or positron excess in the same energy range [6, 114].

INTEGRAL. The International Gamma Ray Laboratory (INTEGRAL), an orbiting gamma-ray observatory, looks for γ -ray emissions in the 20 KeV - 8 MeV energy range. INTEGRAL has reported a 511 KeV ridge in the center of our galaxy (the Milky Way) that is more pronounced than in the galactic periphery, shown in Figure 2.10 [112]. This suggests a large unknown positron emission source in the center of our galaxy that can possibly come from dark matter annihilation.

IceCube. IceCube is a high-energy neutrino telescope that is looking for distinct high-energy neutrinos coming from a WIMP annihilation from the center of the Sun where a large amount of WIMPs may have accumulated through atomic scattering and related energy loss due to high matter density in the core of the Sun.

2.6.3 Collider Searches

Dark matter can be directly produced in the lab in collider experiments. Because of their heavy mass(es), WIMPs are not expected to be directly accessible to experiments with a lower center of mass. The high center-of-mass energy of the Large Hadron Collider allows the possibility of direct WIMP production and their inferred detection via missing energy measurements [15].

2.6.4 Discussion

Observational evidence shows that dark matter is likely non-baryonic. Data from Ice-Cube, PAMELA, ATIC and INTEGRAL experiments suggest that it is composed of WIMPs [5]. Other conventional explanations for the observed lepton excesses that do not involve dark matter annihilation are possible but they appear unlikely given the amount of evidence that points to excess lepton production at various energy levels [57].

2.7 Supersymmetric Models of Dark Matter

Many recent theories suggest that dark matter is made up of previously unknown particle(s) on the scale of weak interactions [10]. These proposals address the recent astrophysical observations, outlined in Section 2.6, that show an anomalous excess of cosmic-ray leptons, by proposing a TeV-scale dark matter annihilation process [10] with a “lighter” GeV-scale hidden dark sector that couples to the Standard Model and may be accessible to experimental probes with the LHC. These dark matter states can provide distinct collider signatures with cascade-decays of highly collimated leptons or possibly pions as well. Addition of Supersymmetry (SUSY)[11, 15] makes the signatures even more interesting with various lepton-rich topologies. This theoretical framework is similar to hidden-valley theories [111] and others [102] that discuss a GeV-scale hidden or dark sector. A generic feature of all such supersymmetric dark matter frameworks is a cascade decay down to the lightest stable dark sector particle, such as a neutralino or a dark fermion, which escapes detection as missing energy and represents a cold dark matter candidate.

2.7.1 Main production mechanisms

There are several possible production mechanisms that involve SUSY in the dark sector. One such mechanism involves electroweak-inos [31] that tend to cascade down to the lightest neutralino χ_0^1 , producing (color) particles along the way, resulting in QCD jets in addition to lepton jets in the model’s collider signature. Another production mechanism involves squark pair production and subsequent decay [15]. In this model, which we will consider in more detail, squarks represent the lightest supersymmetric partner (LSP) of the MSSM that decays into a quark and lighter dark sector fermion \tilde{n}_2 , which then cascade-decays into the lightest dark sector fermion \tilde{n}_1 and either a dark photon γ_d or a dark higgs h_d that decays to two dark photons. Dark photons then in turn decay back into the standard model states, producing collimated leptons or pions. The lightest dark sector fermion becomes the true LSP which escapes detection as a cold dark matter candidate. The Feynman diagram of this process is shown in Figure 2.11.

Previous dark sector searches at the Tevatron focus on the electroweak production of charginos and neutralinos [4], while the higher center-of-mass energy of the LHC allows to focus the search on colored SUSY production as we will set out to do, because of higher color production cross sections, where squarks become the lightest superpartners.

Dark Showering. Additional off-shell emission of dark photons increases the multiplicity of the final state leptons via dark sector showering. Dark sector showering, or dark showering, occurs when a highly boosted dark particle radiates a dark photon, which is the

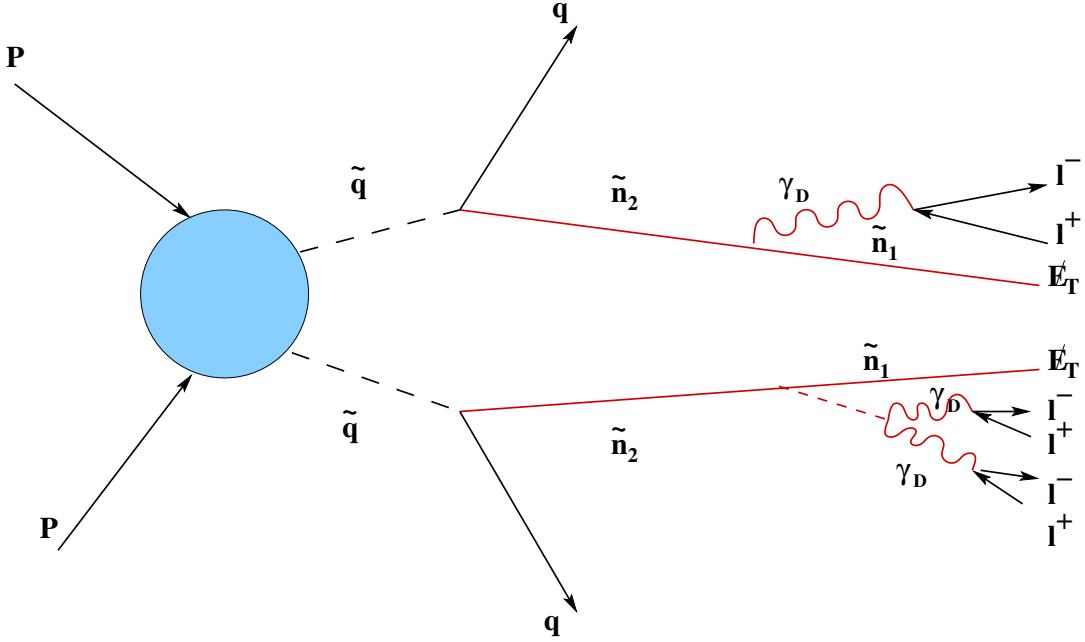


Figure 2.11: The Feynman Diagram for squark production resulting in lepton jets and missing energy. \tilde{n}_2 and \tilde{n}_1 are dark fermions where $m_{\tilde{n}_2} > m_{\tilde{n}_1}$

highest dark gauge boson available in the dark sector. This process can be described by the following equation:

$$N_{b_\mu} \approx \frac{\alpha_d}{2\pi} \log\left(\frac{M_{EW}^2}{M_{Dark}^2}\right)^2 , \quad (2.10)$$

Where N_{b_μ} is the expected number of dark photons emitted within an energy window defined with M_{EW} and M_{Dark} , M_{EW} is the initial dark state invariant mass, M_{Dark} is the mass scale of the dark sector, assumed to be 1/100th of the M_{EW} , and $\alpha_d \approx \alpha_{EM}$ in calculations.

Pion Contamination. If the dark vector boson mass is close to the ρ resonance then, as shown on Figure 2.12, it will preferentially decay to pions rather than leptons. Such scenario is not explicitly a part of this search, but should be studied by further analyses. It is important to note that potential pion contamination can reduce the efficiency of a lepton jet search with hadronic isolation. In light of this, the inner “signal” cone hadronic isolation described later is not enforced.

2.7.2 Other lepton jet production mechanisms

Another possible supersymmetric topology that produces a detector signature with lepton jets involves anomaly mediated supersymmetry breaking. Here, electromagnetic corrections split the degeneracy between the charged and neutral W-inos resulting in lepton jets

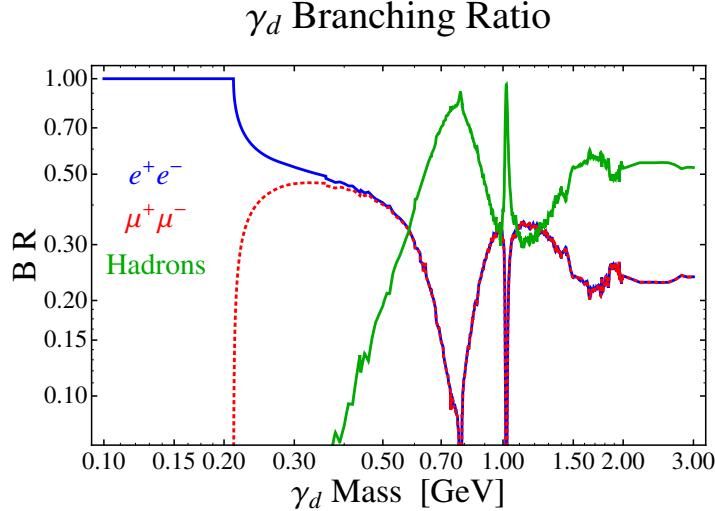


Figure 2.12: Branching ratio of a light dark photon (γ_d). For dark photon masses < 0.5 GeV/ c^2 decays are predominantly into leptons [55].

with a displaced vertex in the final state from the \tilde{W}^0 decay [31]. Such displaced-vertex scenarios are not considered in this work because they require a separate analysis strategy and trigger requirements for displaced-vertex leptons. In another low-scale SUSY-breaking scenario a gravitino can be sufficiently light for neutralino to decay in to it, producing several lepton jets, a real hard photon and isolated leptons [15]. This scenario was previously considered by the D0 collaboration [3].

2.7.3 Lepton Jet Model

Lepton Jet Signature. A dark particle can do one of the following things: decay into an even number of leptons, become missing energy or a mixture of the two. As we have seen in Figure 2.11, a dark sector cascade results in multiple jets of collimated leptons, QCD jets plus missing energy, which leads to the following unique detector signature:

- ≥ 1 Boosted Isolated Lepton Jet + 2 QCD jets + \cancel{E}_T

2.7.4 Analysis Strategy and Caveats

The lepton jet analysis needs to be independent of the features of the dark sector, resonance production and the charges of the particles in the lepton jet. It should be additionally insensitive to variations between two- or three-body decays. Because lepton jets are boosted and collimated, a charge constraint that leptons have to be opposite-sign inside the lepton jet can not be enforced, because lepton charges can be easily mis-identified when they are spatially close. Because of possible presence of electron jets and potential pion contamination, a single boosted muon jet scenario is considered further.

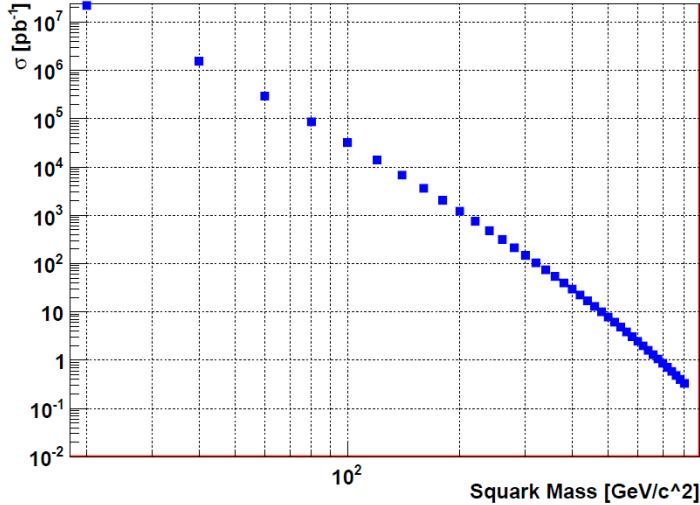


Figure 2.13: Squark production cross section at the LHC

2.7.5 Monte Carlo Simulations

Lepton Jet Signal. Monte-Carlo methods generally employ repeated random sampling to simulate various physical phenomena and are used often in high energy physics simulations⁴. Madgraph Event Generator [9] is used together with Bridge [84] to simulate the lepton jet signal, while Pythia Event Generator is used for the showering and hadronization part [108]. Events are then passed through full CMS Detector reconstruction based on Geant4 [8].

Three relevant processes are considered: squark pair-production, squark-gluino production and gluino pair-production. Squark masses are chosen to be degenerate in the 200 - 800 GeV/c^2 range at 50 GeV/c^2 intervals, together with a fixed dark photon mass of 0.5 GeV/c^2 . Gluino mass is 20% greater than the squark mass. The dark photon masses were varied from 0.5 - 20 GeV/c^2 for a fixed squark mass. Squark production cross section at the LHC is shown in Figure 2.13. Sample Monte-Carlo simulated lepton jet events are shown in Figure 2.14. Event displays are made using the CMS Fireworks software.

Background. Main backgrounds to lepton jet production come from bottom quark double semi-leptonic decay and gluon splitting. Track isolation in an annulus cone around the lepton jet is used to reduce the standard model QCD background. A data-driven way is chosen for background estimation, presented in Section 5.5.

⁴Please see reference [19] for a thorough review of modern Monte-Carlo methods

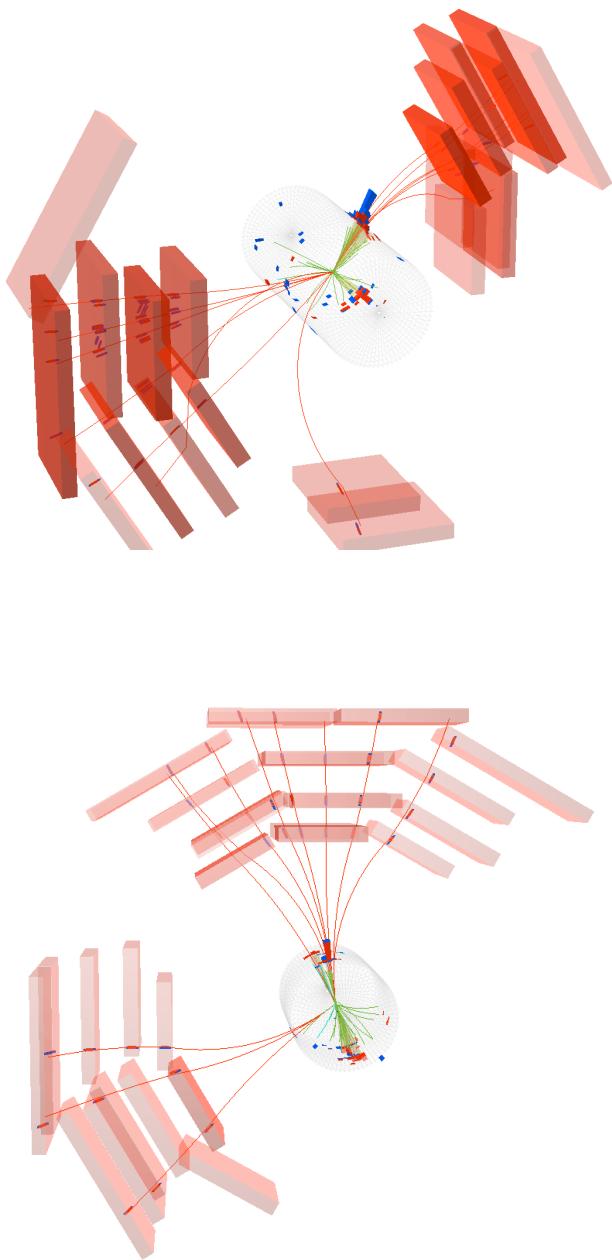


Figure 2.14: Event displays for typical simulated Monte-Carlo lepton jet events

Table 2.1: Anti-Kt Lepton Jet Algorithm Parameters

Parameter	Value
Jet P_t^{min}	.5
CaloTower input E_t^{min}	0
CaloTower input E^{min}	0
Primary Vertex Correction	No
Pileup with Offset Correction	No
NSigmaPU	1.0
RadiusPU	0.5
DoAreaFastJet	No
DoRhoFastJet	No
ActiveAreaRepeats	1
GhostArea	0.01
GhostEtaMax	5.0

2.8 Lepton Jets

2.8.1 Algorithm

Because the idea of grouping hadrons into a jet is similar in nature to reconstruction of dark sector particles that produce collimated groups of leptons as well, the jet reconstruction idea is extended to the leptons. The advantage of this idea is that jet clustering algorithms have been extensively studied and many useful properties of jet reconstruction can be readily utilized. Variables relevant to jets, such as jet mass, size etc, are directly applicable to the lepton jets as well.

The algorithm used to group leptons in a jet is of the anti- k_t variety, which is described in detail in Section 4.4.1. It has the desired property of producing hard jets by first merging the soft particles (or subjets) together with a hard particle, before attempting to group them together.

The input to the lepton jet algorithm for each event passing the trigger is a clean collection of tracker muons with muon and track quality cuts discussed in 4.7.1 applied to them. Also, muons are required to have a minimum transverse momentum p_T of 5 GeV. The details of other parameters of the lepton jet algorithm are given in Table 2.1.

2.8.2 R-Parameter Optimization

As Figure 2.15 shows, essentially all of the signal is contained within the 0.1 radius cone. Therefore, $R = 0.1$ is chosen as the operating lepton jet size.

2.8.3 Kinematical Distributions in Monte Carlo Signal

Jet and Muon Multiplicity. Figure 2.16 shows the multiplicity of lepton jets passing the selection criteria. All of the events have one associated lepton jet in them. Figure 2.17 shows the multiplicity of muons inside lepton jets that pass the selection criteria. Most of them have 2 muons, however there are some also with 3 and 4 muons.

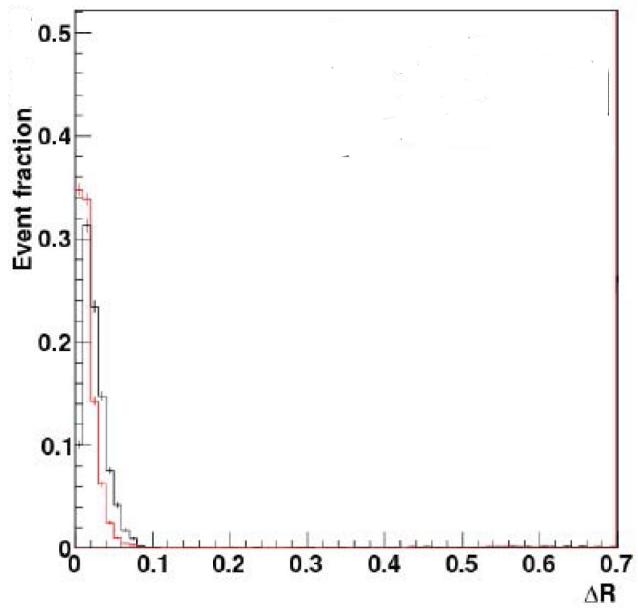


Figure 2.15: Maximum ΔR between muons in a lepton jet **Red:** squark mass 650 GeV/c^2 **Black:** squark mass 250 GeV/c^2 [45]

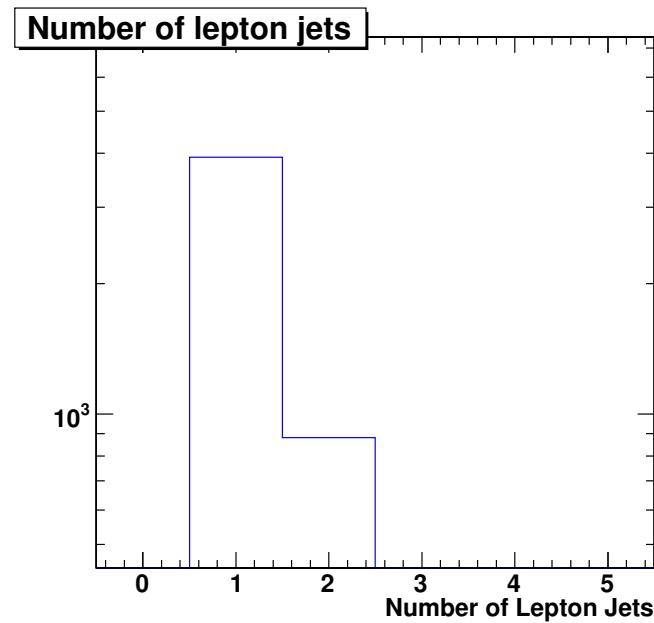


Figure 2.16: Lepton Jet Multiplicity in MC signal

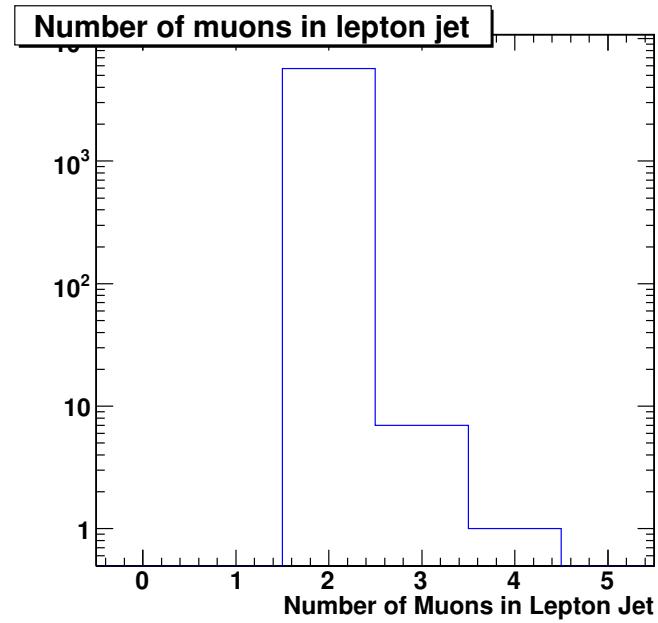


Figure 2.17: Number of Muons in Lepton Jets in Signal MC

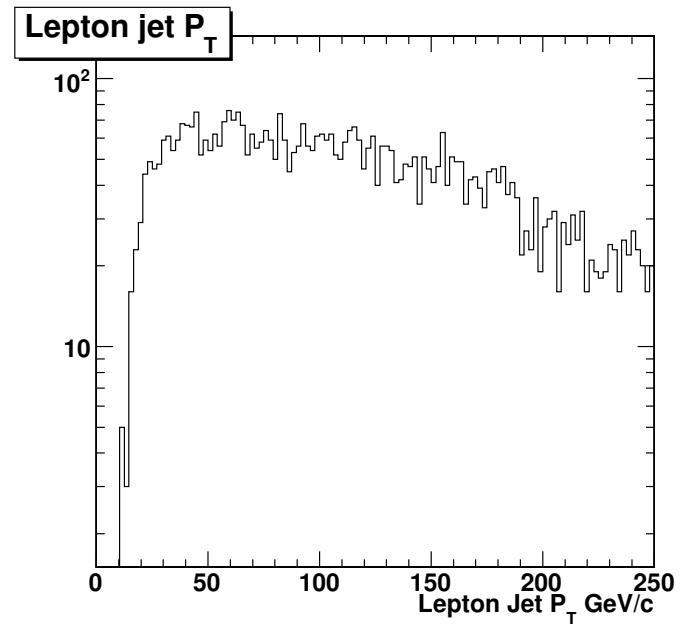


Figure 2.18: Lepton Jet P_T in MC Signal

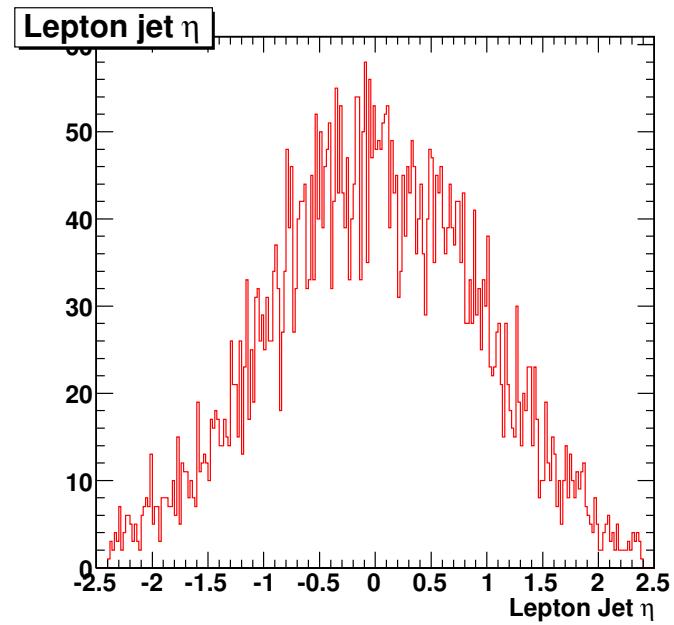


Figure 2.19: Lepton Jet η in MC Signal

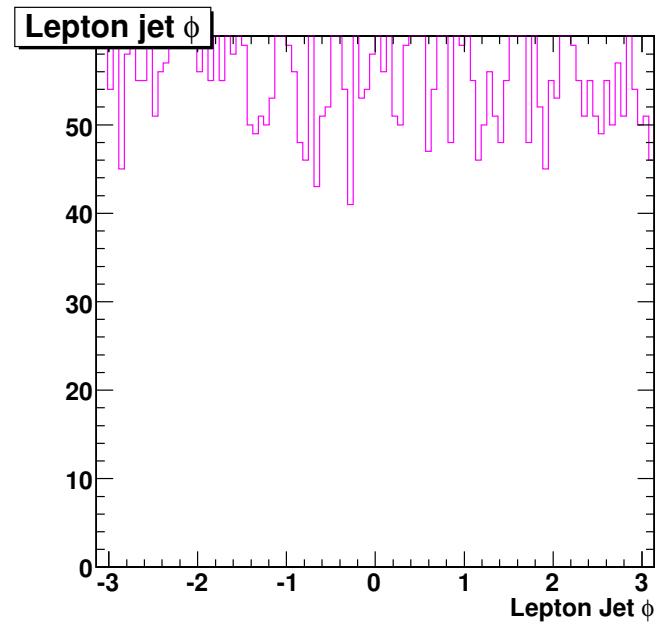


Figure 2.20: Lepton Jet ϕ in MC Signal. Distribution is uniform, as expected.

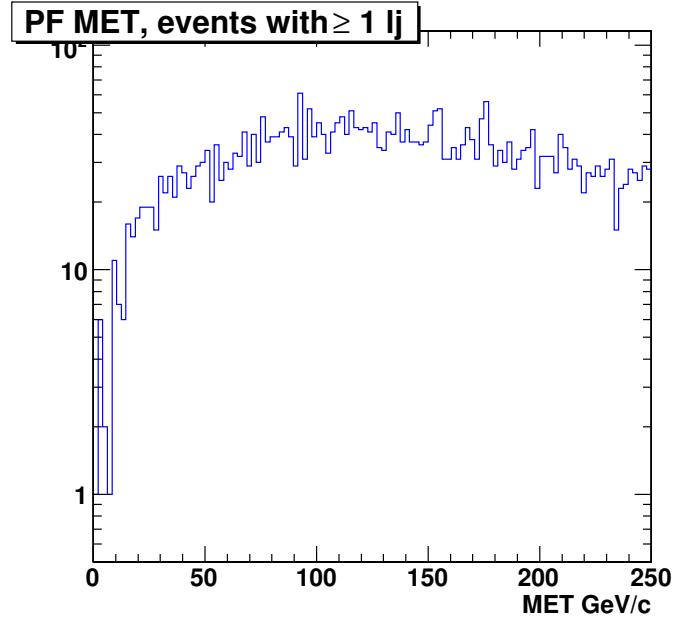


Figure 2.21: Lepton Jet Missing Transverse Energy for MC Signal

Transverse Momentum, Pseudorapidity, Missing \cancel{E}_T and Angular Distributions. Transverse momentum, pseudorapidity and angular ϕ distributions of lepton jets for the MC signal are shown in Figures 2.18, 2.19 and 2.20. Figure 2.21 shows the \cancel{E}_T distribution for the selected events in the MC signal. As Figure 2.21 suggests, signal events tend to have very large missing transverse energy.

CHAPTER 3

DESCRIPTION OF THE EXPERIMENTAL APPARATUS

3.1 The Large Hadron Collider

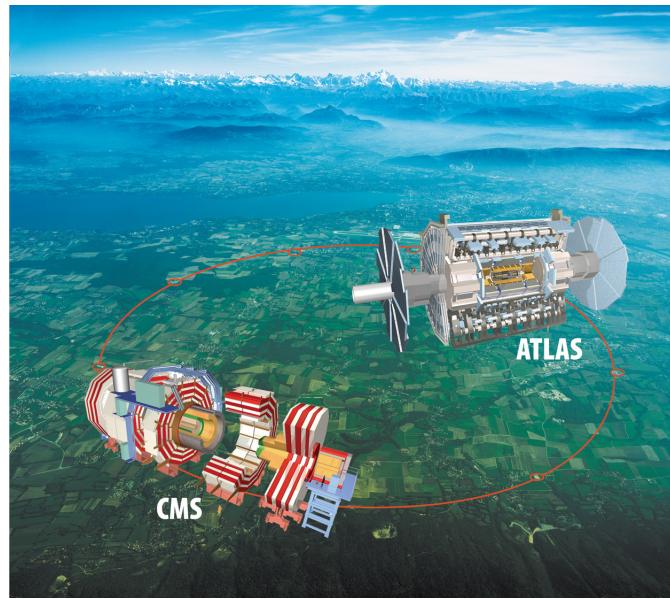
The Large Hadron Collider (LHC) is located about 100 meters underground on the border of France and Switzerland, near Geneva. It is currently the world's highest energy particle accelerator with 7 TeV center-of-mass collision energy and more to come in the upcoming years. It uses a pre-existing Large Electron Positron (LEP) Collider tunnel that is 27 km in circumference. The LHC is primarily designed for proton-proton collisions, although heavy-ion collisions are recorded about one month per calendar year.

The LHC layout is shown in Figure 3.1. Decision to collide protons over electrons was made in order to achieve higher energies than those possible with LEP. Prior to entering the ring, protons are accelerated in the Super Proton Synchrotron (SPS) ring, which along with a series of preliminary linear accelerators (LINACs) and boosters, provides initial acceleration of 450 GeV.

Beam is conditioned by a series of magnets in the main LHC ring, the most critical ones being the dipoles and the quadrupoles. The former are used to guide the particle beams, while the latter are used to focus the beams for collisions in specified locations in the ring that have detectors in place to measure the aftermath of the collisions.

There are four main experiments housed in the LHC tunnel: ALICE, ATLAS, CMS and LHCb. ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) are general purpose detectors designed to search for new physics, while ALICE (Large Ion Collider Experiment) is primarily used to study the quark-gluon plasma produced in heavy-ion collisions, reminiscent of the state of matter in the early Big Bang era. LHCb (Large Hadron Collider Beauty) experiment's goal is to study the Charge-Parity (CP) violation by focusing on b quarks.

A total of 43 pb^{-1} (inverse picobarn) of data were collected by the LHC experiments at 7 TeV center-of-mass energy in 2010. The LHC resumed its operation in March 2011 at 7 TeV center-of-mass energy and will continue running for about 18 months collecting data. Then it will be shut down to allow the final work to be done to increase the center-of-mass energy to the design value of 14 TeV.



Overall view of the LHC experiments.

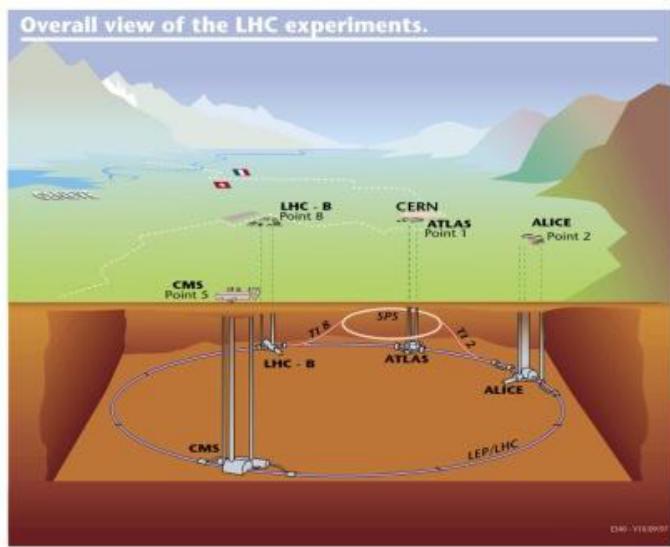


Figure 3.1: The Large Hadron Collider Experiments a. Aerial view where two major detectors ATLAS and CMS shown b. Underground locations of all LHC experiments

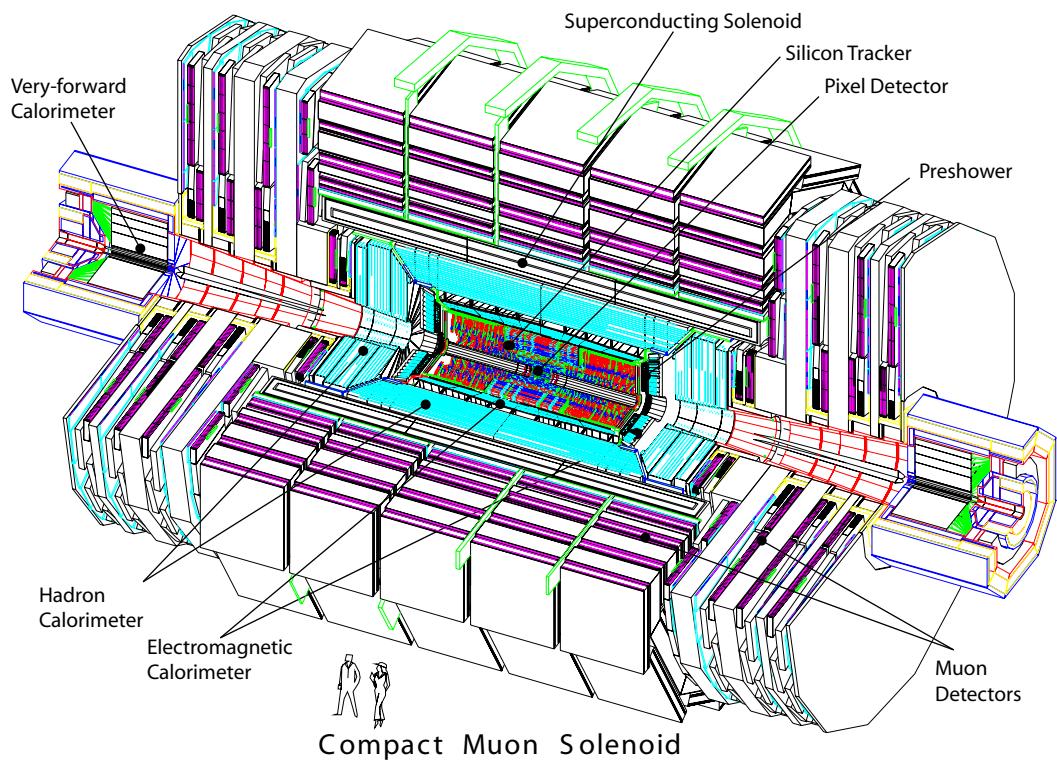


Figure 3.2: The Compact Muon Solenoid Detector

3.2 The Compact Muon Solenoid Detector

CMS is a general purpose experiment, shown in Figure 3.2, designed to explore physics at the TeV energy scale. It weighs about 12,600 tonnes, has a diameter of 20 m and is about 25 m long. CMS consists of various particle detectors that surround the primary Interaction Point (IP) where high-energy collisions occur. In the order of proximity to the interaction point, the first detector is the Inner Tracking System, which consists of an Inner Pixel Silicon and the Outer Silicon Microstrip Tracker detectors. It is used to measure the position and momentum of individual charged particles that leave tracks as they traverse the tracking system. Behind the tracker is the Electromagnetic Calorimeter (ECAL), used to measure the energy of the electromagnetic particles (electrons and photons) that deposit their energy in the volume of the calorimeter. Next is the Hadronic Calorimeter (HCAL), which is used to measure the energy deposited by the hadrons and consists of four separate detector components: Forward (HF), Barrel (HB), Endcap (HE) and Outer (HO). Behind the HB/HE calorimeters is a 3.8 T solenoid magnet coil. The coil curves the trajectories of charged particles, crucial for their precise momentum measurement. And last, but not least, is the Muon System, used to measure the momentum and position of muons, the only Standard Model particles, other than neutrinos, that make it outside of all the detectors. Muon detector system is the only full detector system outside the solenoid magnetic field. Muons are crucial probes of new physical phenomena, as multiple models of new physics predict energetic muons in their signatures. Therefore the CMS Muon System was designed with excellent efficiency and strong background rejection in mind.

3.2.1 Coordinate System

The coordinate system used by CMS is shown on the bottom right of Figure 3.9. The origin is the interaction point. The x-axis is horizontal pointing to the center of the LHC, y-axis is vertical (up), while the z-axis is horizontal along the beam axis pointing west. The azimuthal ϕ direction is measured in the x-y plane and the angle θ is measured with respect to the z-axis. Angle θ is rarely used and instead the use of pseudorapidity $\eta = -\ln(\tan(\theta/2))$, a relativistic approximation to the true rapidity, is prevalent. Any components in the transverse direction, such as the transverse energy or momentum components are denoted with a subscript T, for instance P_T , in the case of transverse momentum.

3.2.2 Magnet System

CMS has a superconducting solenoid magnet that provides a strong magnetic field of 3.8 T. The magnet system consists of a superconducting coil in a vacuum tank, as well as cryogenics and other supplying materials. It weighs 12000 tonnes and at 6 m in diameter and 13 m in length, is the largest superconducting magnet ever built. The purpose of the magnet is to bend charged particles under the magnetic field such that their momenta can be measured from their trajectories.

3.2.3 Inner Tracking System

The Inner Tracking System is pictured in Figure 3.3. It is cylindrically shaped and measures 5.4 m in length and 2.4 m in diameter. It consists of the largest array of segmented

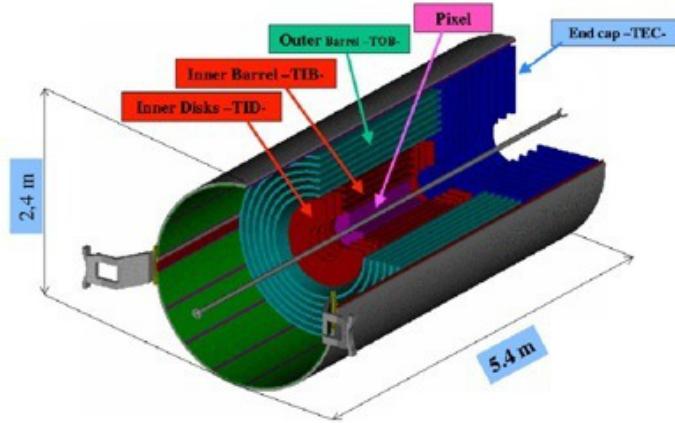


Figure 3.3: CMS Inner Tracking System

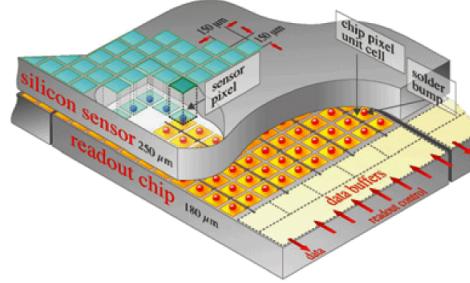


Figure 3.4: Pixel detector elements

silicon detectors ever built. Silicon detectors are semiconductors which measure ionization of silicon atoms by charged particles, allowing for a measurement of energy loss within the detectors, which can be used to determine the energy and momentum of incoming particles. The main purpose of the inner tracking system is to measure the momenta, trajectories and impact parameters (perpendicular distances of closest approach) of charged particles that are produced in a collision event. Its surface area is 200m^2 and it consists of 16588 modules. The tracking system consists of two distinct detector systems: the Inner Pixel Silicon Detector and the Silicon Microstrip Tracker.

Inner Pixel Silicon Detector. The Inner Pixel Silicon Detector consists of 1400 silicon pixel sensor modules and covers the pseudorapidity range of $-2.5 < \eta < 2.5$. It is the closest detector to the interaction point and is highly segmented to allow for precise measurement of particle's impact parameter and position. Each sensor consists of 52×53 arrays of $100 \times 150 \text{ micron}^2$ pixels, shown in Figure 3.4, which amounts to 18M pixels in the forward region alone and 66M pixels in total.

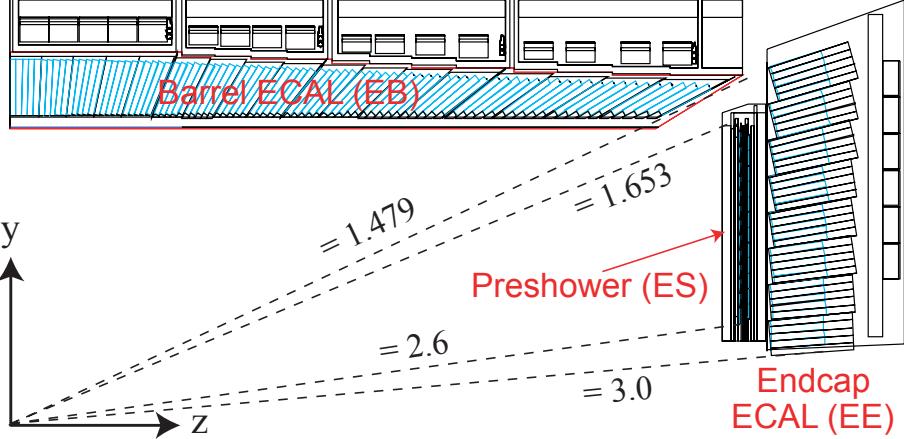


Figure 3.5: CMS ECAL Detector [16]

Silicon Microstrip Tracker Detector. The Silicon Microstrip Tracker has four basic subdetectors. Tracker Inner Barrel (TIB) consists of 4 cylindrical layers where strips of different geometries are used. Tracker Inner Disk (TID) contains six disc structures. In each disk, the modules are arranged in 3 rings, placed alternately in the forward and back parts of the structure. The Tracker Outer Barrel (TOB) consists of 6 layers. Each layer is made of rods, with 6 modules inside each rod. The 2 innermost layers have double-sided modules. The Tracker End Cap (TEC) has 9 disks, each disk made of 8 petals. Inside the petals the modules are arranged in rings.

Basic track reconstruction consists of 5 parts: clustering of strips or pixels into hit reconstruction, seed generation, trajectory building, ambiguity resolution and the final track fit. More details on track reconstruction can be found in [16].

3.2.4 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) measures the energy deposited by electrons and photons in the active volume of the detector. It consists of three main parts: Barrel (EB), Endcap (EE) and Preshower (ES) as shown in Figure 3.5. Its basic component is a clear dense and fairly radiation-hard lead tungstate PbW_04 scintillator. The crystals are placed in a quasi-projective geometry such that they subtend a small angle (3°) with respect to the interaction point. The front face area of each crystal is $22 \times 22\text{mm}^2$ and its length is 23 cm, which corresponds to 25.8 radiation lengths.

ECAL Barrel. The ECAL Barrel subdetector consists of 36 identical supermodules, each subtending 20° in ϕ , and covers the pseudorapidity range $|\eta| < 1.479$, as shown in Figure 3.5.

ECAL Endcap. The ECAL Endcap subdetector covers the pseudorapidity range $1.479 < \eta < 3.0$. Its basic unit is a supercrystal which contains a 5×5 array of crystals. The endcap is divided into two halves, or Dees. Each Dee contains 3662 crystals.

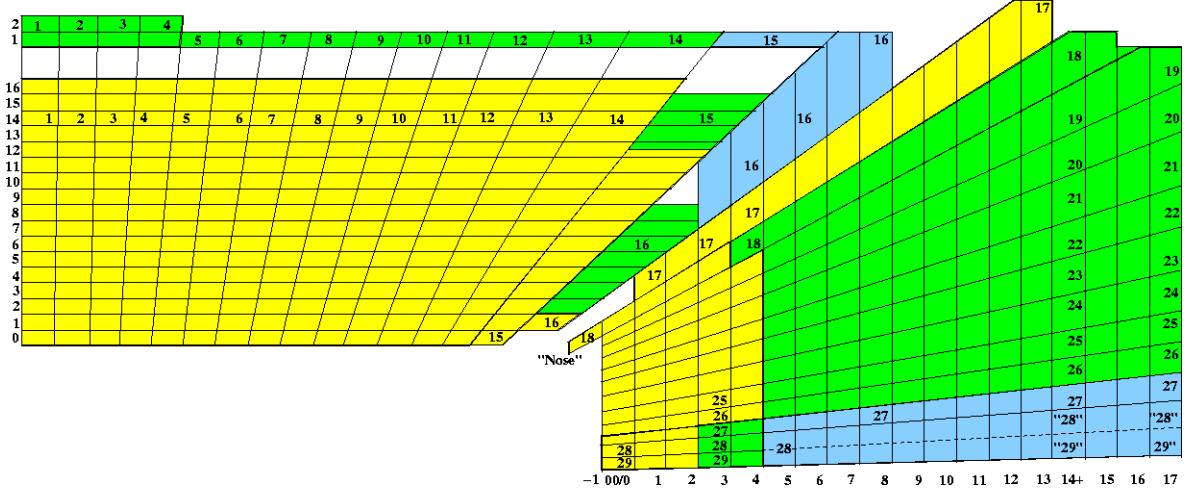


Figure 3.6: CMS HCAL Detector

ECAL Preshower. In front of the ECAL Endcaps is the ECAL Preshower, covering the fiducial region of $1.653 < \eta < 2.6$ as shown in Figure 3.5. The preshower is a sampling calorimeter, made up of two layers: lead radiators and strip sensors. Incoming particles shower as they enter the radiators and silicon sensors measure their energy deposition and shower profiles. Its main aim is to identify π^0 s as well as improve the position determination of electrons and photons.

3.2.5 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is depicted in Figure 3.6. Its goal is to measure the energy deposited by hadrons and together with ECAL measure the energy of jets and determine missing transverse energy. It consists of four main parts: barrel (HB), endcap (HE), outer (HO) and forward (HF). HB covers the pseudorapidity range of $|\eta| < 1.4$ while HE extends that range to 3.0, providing overlap with HB in the η range of $1.3 < \eta < 1.4$. HF is located 11.2 m from the interaction point and provides coverage up to $|\eta| < 5.0$, while HO provides additional coverage outside of the magnet coil.

HCAL Barrel. The Hadronic Barrel Calorimeter consists of two halves, each composed of 18 identical wedges that cover 20° in ϕ . The first and last absorber layers are made of stainless steel. Each wedge contains brass alloy absorber plates and 16 layers of active plastic scintillator tiles located in between the stainless steel and brass absorber plates. The total absorber thickness at 90° is 5.82 interaction lengths. The brass layer thickness is 5.0 - 5.5 cm and scintillators are 3.7 mm thick, except for layers 1 and 16, which are 1.0 cm thick. Individual scintillator tiles are equipped with wavelength shifting fibers, optically added from each active layer into an ‘‘HCAL tower’’ in η and ϕ , with the exception of tower 15 and 16 at the edge of the barrel, where multiple optical readouts are present. The tower size in $\Delta\eta \times \Delta\phi$ is $.087 \times .087$.

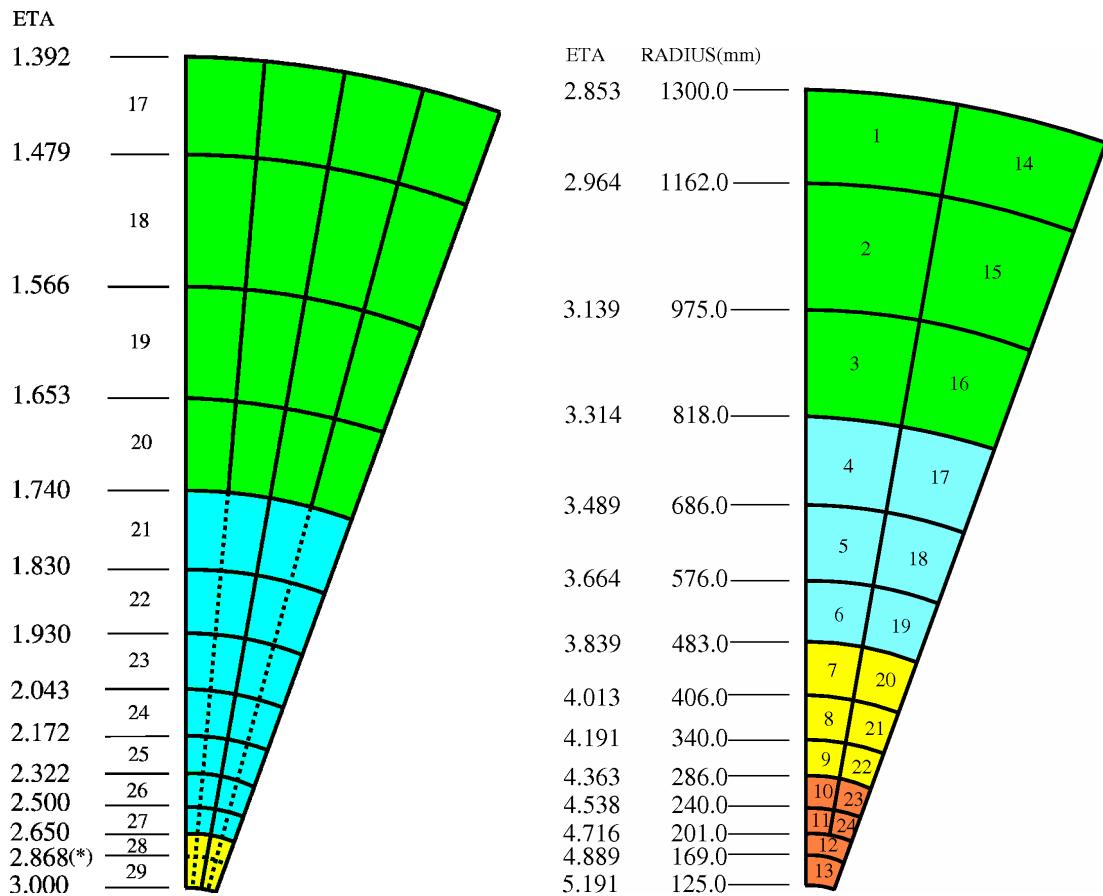


Figure 3.7: η segmentation and size increase in HE (left) and HF (right)

HCAL Endcap. The Hadronic Endcap Calorimeter is composed entirely of brass absorber plates and overlaps with the Hadronic Barrel Calorimeter in tower 16. It has 19 plastic scintillation layers and its 18-fold geometry in ϕ matches the geometry of HB. The thickness of the plates is 78 mm and scintillator thickness is 3.7 mm. The $\Delta\eta \times \Delta\phi$ tower segmentation for pseudorapidity range $|\eta| < 1.74$ matches that of the barrel, while for higher η the size doubles, as shown in Figure 3.7.

HCAL Forward. The Hadronic Forward Calorimeter is located 11.2 m from the interaction point and is made of steel absorbers and radiation hard quartz fibers. Long (1.65m) and short (1.43m) quartz fibers are alternately placed and readout separately with phototubes. Each HF module has 18 wedges in a non-projective geometry, with quartz fibers located in parallel to the beam axis along the length of the iron absorbers.

HCAL Outer. The Hadronic Outer Calorimeter consists of arrays of scintillator located outside of the magnet coil. It consists of five rings 2.5 m wide along the z-axis, with 12 sectors in each ring. The HO covers the pseudorapidity range $|\eta| < 1.26$ and its primary function is to measure the energy leakage from the HB.

3.2.6 The Muon System

The Muon System consists of three separate gaseous detector systems: Aluminum Drift Tubes (DT), Resistive Plate Chambers (RPC) and Cathode Strip Chambers (CSC). It is the largest and one of the most important systems of the CMS detector. In the Barrel region $|\eta| < 1.2$, Drift Tubes are used. In the Endcaps, where these rates are higher and the magnetic field is also high, Cathode Strip Chambers are used. The CSCs cover the pseudorapidity range up to $|\eta| < 2.4$. In addition to those, Resistive Plate Chambers are deployed in both the barrel and the endcap regions. The DTs and CSCs are used to obtain a precise position measurement and the bending angle of a muon, while the RPCs are useful for triggering because of their rapid response with precise timing measurement.

The one quarter view of the CMS Muon System is shown in Figure 3.8 and the transverse view is shown in Figure 3.9. There are 4 cylindrical Muon Stations integrated with the magnet return yoke in the barrel region. In the endcaps, the CSCs and RPCs are arranged in four disks perpendicular to the beam direction, and in concentric rings (3 in the innermost station and 2 in others). The fact that the muon chambers are located within the yoke of the magnet allows for an independent momentum measurement without the use of the inner tracker.

Muon Barrel. The Muon Barrel Detector consists of 250 chambers in 4 layers labeled MB1, MB2, MB3 and MB4 inside the return yoke of the magnet. The segmentation along the beam axis reflects the 5 wheels of the yoke (YB-2 for the farthest in -z, and YB+2 for the farthest in +z). Each wheel is divided into 12 sectors, corresponding to 30° each, labeled in the order of increasing ϕ . The two innermost stations, MB1 and MB2, have DT chambers sandwiched in between two RPCs. In the two outermost stations, MB3 and MB4, the DT chambers are coupled to a layer of 1, 2 or 4 RPCs depending on the sector and station.

Muon Endcaps. The Muon Endcap Detector consists of 468 Cathode Strip Chambers. Each Endcap consists of 4 stations with chambers, labeled ME1, ME2, ME3 and ME4,

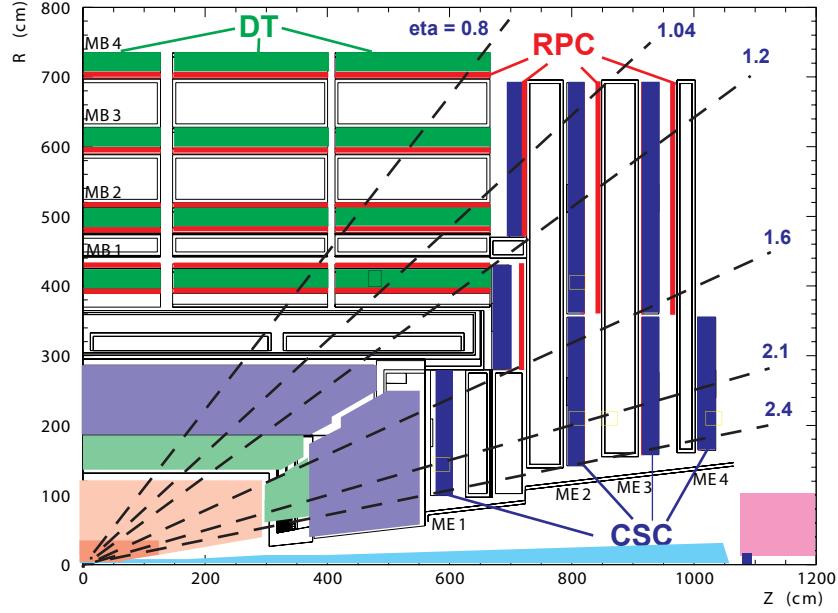


Figure 3.8: Quarter View of the CMS Muon System

in the increasing order of distance from the interaction point. The stations are mounted perpendicular to the direction of the beam on the disks enclosing the magnet. In each disk, the chambers are divided into rings around the beam axis (two for ME2-4 and three for ME1). Each of the rings has 36 chambers, except for the innermost rings of ME2-ME4, that have 18 chambers. Similarly to the Muon Barrel, there are layers of the double-gap RPCs in the Endcaps.

Drift Tube Chambers. Each DT chamber consists of 12 layers of drift tubes subdivided into 3 groups of 4 super-layers as shown in Figure 3.11. The tubes inside a super-layer are staggered by half-tube. Two super-layers measure the $r\phi$ coordinate of the trajectory while the 3rd super-layer measures the z -coordinate parallel to the beam axis. The last station, MB4, has only 2 super-layers that measure the $r\phi$ coordinate. The DTs are coupled to the RPCs and a high-pt muon crosses up to 6 RPC modules and 4 DT chambers on its way through the DT-system.

Cathode Strip Chambers. Cathode Strip Chambers are gas proportional counters located in the Endcaps of the Muon System. Each CSC has 6 gas gaps, while each gap has a plane of cathode strips and anode wires perpendicular to the strips. It is trapezoidal in shape as shown in Figure 3.12. As a charged particle traverses the plane of each chamber, ionization of the gas and consequent electron avalanche produces a charge in the anode wire and an image charge in the cathode strips. This allows an accurate measurement of the $(r, \phi$ and z) coordinates in each of the 6 layers. All CSCs overlap in ϕ (except for those in ME1/3) to avoid any gaps in muon acceptance.

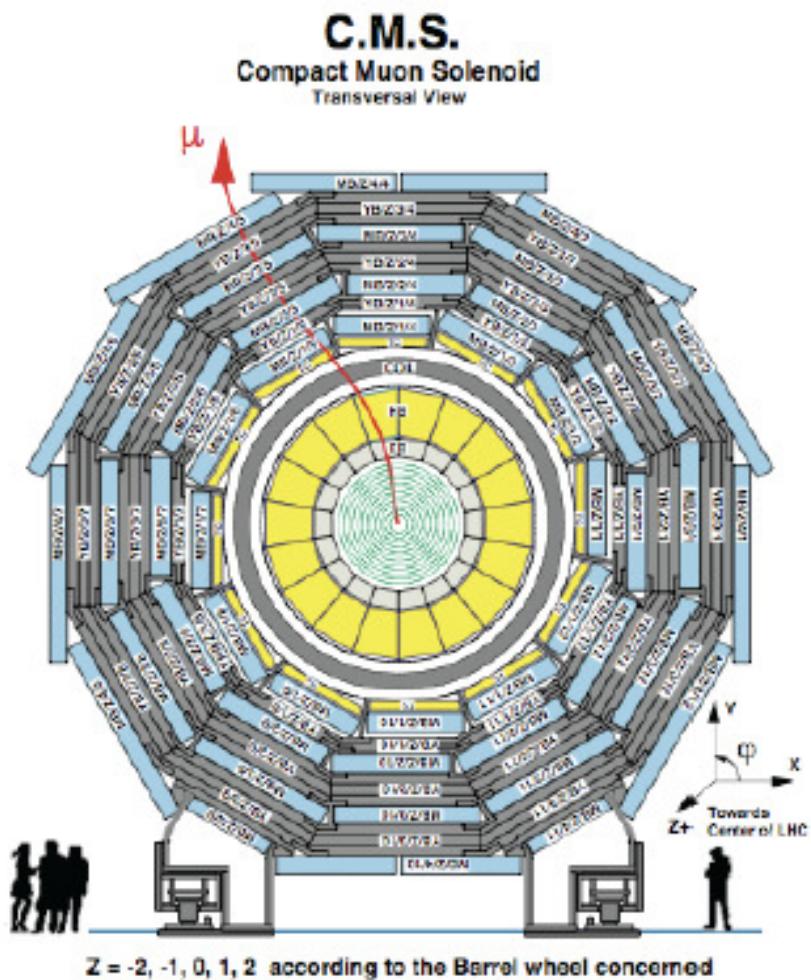


Figure 3.9: Transverse View of the CMS Muon System

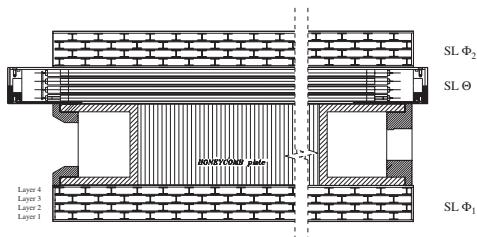


Figure 3.10: CMS Drift Tubes

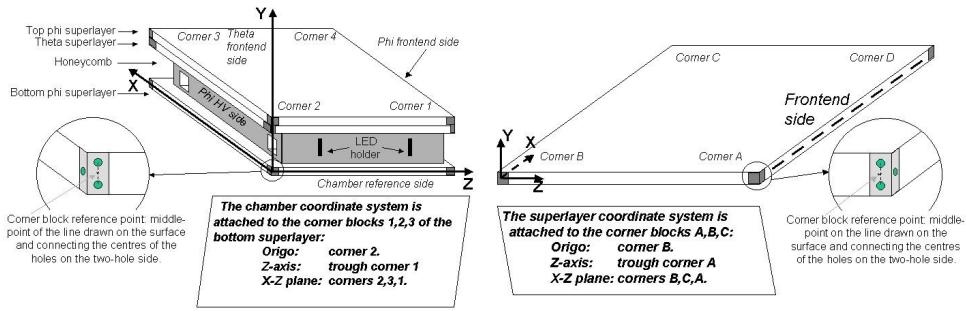


Figure 3.11: Honeycomb Structure of the DT SuperLayers

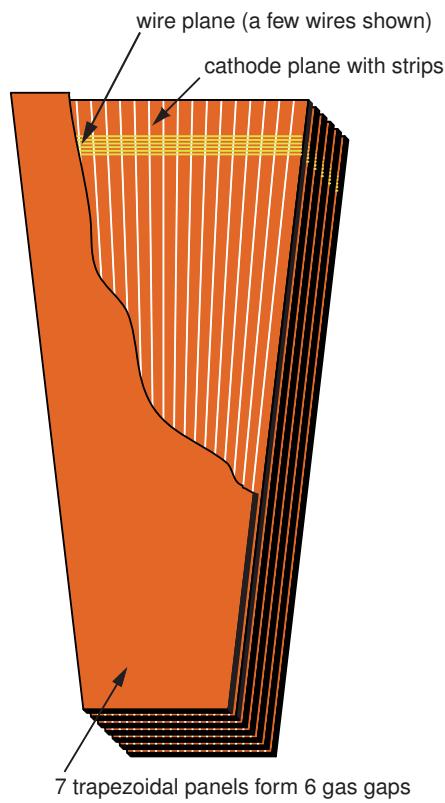


Figure 3.12: CMS Cathode Strip Chamber

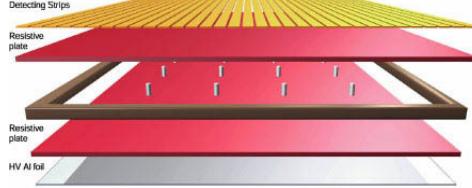


Figure 3.13: Resistive Plate Chamber Layers

Resistive Plate Chambers. The Resistive Plate Chamber detectors consist of a double-gap bakelite chamber (Figure 3.13) operating in avalanche mode. The width of the gap is 2 mm. The length of the strip is 130 cm, except in MB2, where it is either 85 cm or 130 cm. There are 480 RPCs in the barrel and 1020 total. RPCs provide fast response and an accurate timing measurement which is used in the muon trigger.

3.2.7 Other Detectors

In addition to the main detectors, Centauro and Strange Object Research (CASTOR) and Zero Degree Calorimeters (ZDC) are located further away from the interaction point. CASTOR is a Cherenkov sampling calorimeter located 14.3 m away from the interaction point that provides additional pseudorapidity coverage for $-6.6 < \eta < -5.2$, while the ZDCs are located 140 m away from the interaction point on both sides of CMS. They are designed to measure the very forward neutral particles such as neutrons, photons and π^0 s coming from the collisions.

3.2.8 Trigger and Data Acquisition System

The CMS experiment only records 300 Hz out of a much larger fraction of events/collisions that occur (factor of 10^6 at design luminosity) due to the limitation of archival media available. This reduction in data rate is achieved with a robust Trigger and Data Acquisition system which is designed to select only interesting events for further analysis.

This system consists of four parts: detector electronics, the Level-1 Trigger, readout and online High Level Trigger Filter.

Level 1 Trigger. The total allocated time to make a decision whether to keep or reject an event is $3.2 \mu\text{s}$. During this time the trigger data is collected from front-end electronics and reduced by a factor of 1000 with the Level-1 Trigger, which takes about $1 \mu\text{s}$. Special hardware processors which involve the calorimetry and muon systems as well as some correlation between these systems, are used in making this decision. It is based on the presence of a “trigger primitive” objects, such as photons, electrons, muons and jets with pt above certain thresholds, as well as sums of transverse energy E_t and missing transverse energy E_t^{miss} . The logic of the Level-1 trigger system is contained in Application Specific Integrated Circuits (ASICs), Field Programmable Gate Arrays (FPGAs), Programmable Logic Devices (PLDs) and Random Access Memories (RAMs) used in the memory Look-Up-Tables (LUTs).



Figure 3.14: CMS Tier Computing Centers across the world.

Readout. After the data are received from the Level-1 Trigger, they are passed into the front-end readout buffers. Upon further signal processing, zero-suppression and compression, data are placed in dual port memories for DAQ system access. The information for each event is contained in several hundred readout buffers and is transferred to an HLT processor.

High Level Trigger. The final reduction in event data rate from 100 kHz to 250 Hz occurs at the High-Level Trigger (HLT). An HLT processor farm is used for this purpose. Instead of full event reconstruction, local (calorimeter towers and muon chambers) and regional (combination of local) event reconstruction is performed and used whenever possible. This approach leads to several levels of triggering: initially using calorimetry and muon information, then incorporating partial tracking information, and in the end, full event reconstruction with full tracking. Depending on the luminosity, some of the trigger paths are pre-scaled. This term refers to the probability of acceptance of events that pass the trigger and are recorded. Due to the limited available bandwidth event rates for some trigger paths are reduced by pre-scale factors.

3.2.9 Software and Computing Model

The CMS Computing covers many important tasks such as storage, access, reconstruction and analysis of data. Given the huge amount of data which is collected by the detector every year, CMS has opted for a distributed computing model, utilizing global resources located in many countries, in a tiered architecture. Primary “Tier-0” is located at CERN, while “Tier-1”, “Tier-2” and “Tier-3” computing centers are located worldwide at universities and laboratories. Figure 3.14 shows the global distribution of the CMS tier centers.

Tier 0. The Tier-0 (T0) is located on site in CERN and its purpose is to store the RAW data straight from the detector. There are no analysis capabilities available at this level.

A CERN Analysis Facility located at CERN provides the resources for critical calibration and alignment tasks as well as prompt analysis of data and data quality monitoring.

Tier 1. Tier-1 centers are distributed globally, at major national laboratories such as Fermilab or DESY. A copy of the RAW data is stored at Tier-1s as well as a fraction of simulated and reconstructed data. Tier-1 sites also provide a focal transfer point for data going to Tier-2 and Tier-3 centers as well as some storage for analysis results coming from Tier-2 and Tier-3 centers.

Tier 2 and 3. Tier-2 and Tier-3 centers are places where final analyses on data are performed. The results are either stored on-site or transferred back to Tier-1 for storage, depending on their size.

This architecture relies heavily on the recent development of grid technology [94] designed to optimally take advantage of a distributed computing power of the different computing centers.

CMS Software Framework. CMS has developed a modular software framework called CMSSW centered around the event data model (EDM). It performs data processing, filtering, selection and all of the necessary analysis tasks both online and offline. EDM modules can access any existing event information, perform complex operations and put additional information back into the event. The main module types are: input-output, producers, filters, selectors and analyzers. Input-output modules provide I/O tasks to the data, while producers produce new data products to be put back into the event. Filters and selectors filter and select events based on some criteria and analyzers perform analysis level tasks.

Various modules are independent of each others and can be performed separately or in conjunction. There is a configuration file specified in a Python programming language, which allows a user to configure the modules and the order of their execution, as shown in Figure 3.15.

3.2.10 Calibration and Alignment of the Detector

Calibration and alignment of the detector are monumental tasks which assure optimal performance of the different detectors. Calibration aims at equalization of the response of the different detector channels as well as achievement of correct scales for various physical quantities, for instance the jet energy scale. The goal of alignment is to properly align the tracker and the muon system for optimal detection and measurement.

Figure 3.15 shows the role of the calibration and alignment in the CMS offline workflow. During the data-taking phase dedicated calibration streams are automatically produced and promptly analyzed during the first several hours in the CERN Analysis Facility (CAF). The results of calibration are channel-by-channel constants (correction factors) which are loaded into a conditions database for further use.

Calibration of ECAL. Calibration of the electromagnetic calorimeter is crucial for a potential discovery of the Higgs boson in the $H^0 \rightarrow \gamma\gamma$ channel, one of the important physics goals of CMS. During the construction phase, each of the assembled supermodules have been pre-calibrated using laboratory measurements of the light yield and photodetector gain. 9 out of the 27 supermodules (25%) passed through electron test beam, which has determined

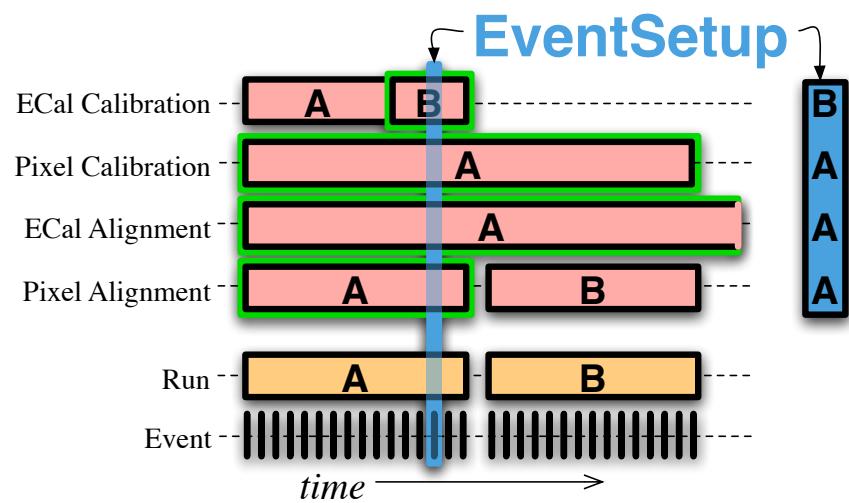
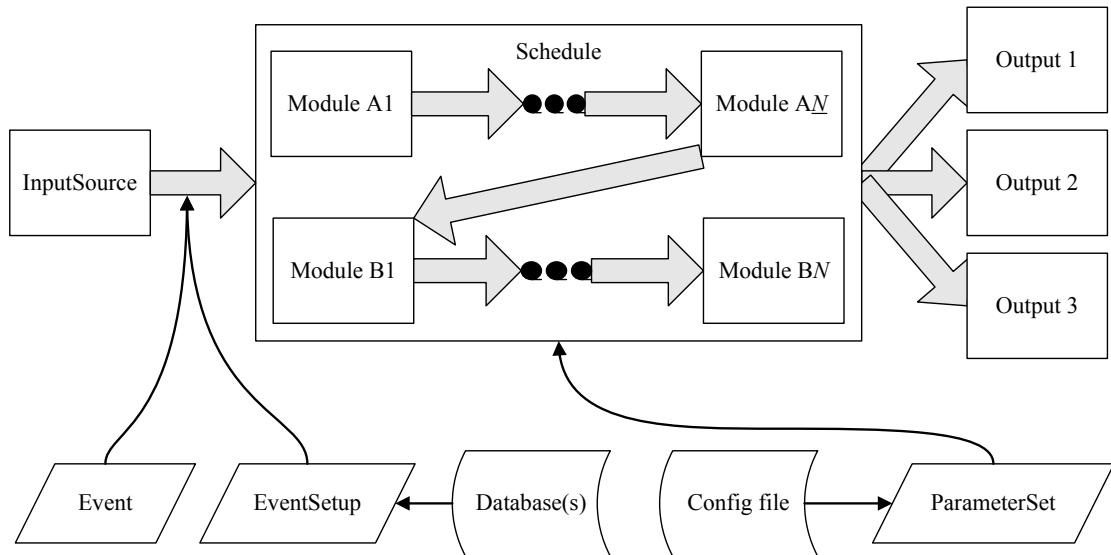


Figure 3.15: CMS Software Framework Modular Structure and Event Setup

their pre-calibration to 0.3%. Additional pre-calibration constants were determined using cosmic rays with 1.5% precision. Prior to collisions, events from stopped circulating beams (beam splashes) were used to independently verify the pre-calibration for the entire barrel at 1.6%

During data-taking in-situ calibration of the crystals is performed with $\pi^0 \rightarrow \gamma\gamma$ decays and the ϕ symmetry method. In the first method, the mother particle mass constraint is used to calibrate the crystals in photon reconstruction. The second method relies on the homogeneity of the average energy deposition for a fixed pseudorapidity as a function of ϕ . Total transverse energy (ΣE_t) deposited in one crystal is compared to the average ΣE_t of other crystals with the same pseudorapidity. Precision of 1.2% is obtained with the π^0 method, while 1.4% precision is obtained with the ϕ -symmetry method. Combining the methods and the pre-calibration from the beam dumps leads to a 0.6% precision for $|\eta| < 0.8$. The absolute energy scale calibration is determined with $Z^0 \rightarrow e^+e^-$ events.

ECAL energy resolution has been measured in test beam to be:

$$\frac{\sigma(E)}{E} = \frac{2.8\%}{\sqrt{E}(GeV)} \oplus \frac{0.12GeV}{E} \oplus 0.3\% , \quad (3.1)$$

The three parts correspond to the stochastic, noise and constant terms. ECAL laser calibration system installed in situ provides real-time corrections and monitoring of calibration constants.

Calibration of HCAL. Before installation in situ, pre-calibration of HCAL was performed with a radioactive 5 mCi Co^{60} source. The source was placed at the tip of a stainless steel wire and inserted into a tube mounted on the surface of the tower tiles. The relative response of each tower was then weighted according to a shower profile. Cosmic-ray events, test-beam measurements, and beam “splashes”, where single beams are impinged on a closed collimator near CMS, were used for pre-calibration leading to 5% – 12% inter-channel pre-calibration precision for HB, 10% for HE, 12% for the HF and 5% for the HO. During startup a ϕ symmetry method is used for calibration of HCAL rings. The absolute energy scale calibration is done with isolated charged hadron tracks in di-jet events by comparing calorimetry and tracker information. Initial absolute scale calibration is from the test beam pions from 3 to 300 GeV/c. Calibration of the forward region is performed with di-jet events, where jet energy balancing is used [35]. The calibration of jets is achieved with $\gamma + jet$ and $Z+jet$ events. With initial data the precision for energy scale calibration in jet reconstruction was shown to be $10\% + 2\% \times |\eta|$ using only the calorimeters, and $5\% + 2\% \times |\eta|$ for algorithms that combine calorimeter and tracker information.

Combined HCAL and ECAL energy resolution was measured in test beam to be [35]:

$$\frac{\sigma(E)}{E} = \frac{84.7\%}{\sqrt{E}(GeV)} \oplus 7.4\% , \quad (3.2)$$

The two parts correspond to the stochastic and constant terms.

Calibration and Alignment of the Tracker/Muon System. Alignment of both the inner tracker and the muon system are crucial for physics analyses in CMS. In order to attain optimal track-parameter resolution, the position of the tracker modules needs to be known with μm precision. Alignment of pixel modules is key to precise event vertex

reconstruction. Proper alignment of the Muon System is crucial for the measurement of the muon transverse momentum.

During the assembly, surveys were conducted to measure the global position of the inner tracker modules with a precision of about $500\ \mu\text{m}$. Remaining misalignment was corrected with cosmic rays and collisions using track-based procedures. Additionally, a laser-based monitoring system continuously monitors tracker components and help correct their positions.

CMS uses two methods to align the inner tracking system using charged particle tracks. The first method, called HIP, fits alignment parameters of each individual tracker module and then applies the corrections and repeats the fitting until convergence is reached. A global method, called Millipede-II, performs a simultaneous fit to all track and alignment parameters at once. Tracker alignment has been performed both during cosmic-ray data-taking in early 2010 and with early 7-TeV collisions. Using cosmic rays a 3 to 4 micron precision was achieved for the barrel and a 3 to 14 micron for the endcaps. Further details on tracker alignment are discussed in [44, 58].

Using the tracks as position references the Muon System can be aligned by individually adjusting the muon chambers. Similar to the inner tracker alignment a fitting procedure is performed to compute these corrections. Initial alignment of the DTs and CSCs was conducted with the magnet off and later with the magnet on. Prior to collisions, cosmic-ray muons were used to improve alignment the Barrel part of the Muon system, while muons coming from the beam halo are used to align the CSCs in the Endcaps. Inter-channel synchronization of the DTs and the calibration of CSC front end electronics are achieved with test-pulses during the inter-collision machine abort gap.

The various alignment algorithms run on the CAF, taking into account the inter-dependent nature of inner tracker and muon alignment and calibration. First the tracker is aligned, then the Muon system is calibrated and later aligned.

CHAPTER 4

EVENT RECONSTRUCTION AND DATA ANALYSIS

In the following chapter, we describe the event reconstruction in the CMS detector and the analysis of the data in the search for lepton jets. We focus on the reconstruction of muons and jets, objects that are directly relevant to the lepton jet analysis. Event data model is described in Section 4.1. Section 4.2 discusses charged particle tracking, while Sections 4.3 - 4.5 explain muon, jet and missing transverse energy reconstruction. Luminosity, trigger and vertex requirements are discussed in Section 4.6. Basic muon, jet and missing transverse energy selections for lepton jet reconstruction are described in Sections 4.7 - 4.9. Lepton jet tagging and muon selection are discussed in Section 4.8.

4.1 Event Model Data Format

A basic unit of the CMS computing model is an event, which contains raw digitized data, reconstructed products of higher level objects, as well as the provenance (origin) for all created objects. There are three main data formats: RAW, RECO and AOD, which are explained below.

4.1.1 RAW

The RAW data format contains all the original detector information in the event including the Level-1 and High Level Trigger results. This format is not suitable to be used directly for data analysis, and therefore, higher-order data formats are described below.

4.1.2 RECO

The RECO data format contains information on low and high level reconstructed objects. An example of a low-level reconstructed object is a muon chamber hit, while jets and photons are examples of high-level reconstructed objects.

4.1.3 AOD

The AOD data format is a subset of the RECO collection that contains only high level physics objects and quantities relevant to a broad physics analysis.

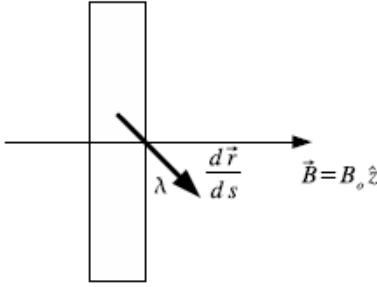


Figure 4.1: Tangential to the particle trajectory makes an angle λ with the magnetic field

4.2 Tracking of charged particles

In order to measure the relativistic momentum ($\vec{p} = \gamma m \vec{v}$) and charge q of a particle in the magnetic field, its trajectory needs to be measured. The Lorenz Force $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$, which relates the particle momentum to its motion in the magnetic field, is used to determine the equation of motion of the charged particle. In parametrized form, using parametrized distance $s(t)$ along the trajectory, the particle trajectory is given by:

$$\frac{d^2\vec{r}}{ds^2} = \frac{q}{p} \frac{d\vec{r}}{ds} B(r) , \quad (4.1)$$

where

- $\frac{d\vec{r}}{ds}$ is the unit length tangential to the trajectory,
- $\frac{d^2\vec{r}}{ds^2}$ is the curvature of the trajectory.

Therefore, for a given magnetic field \vec{B} , the momentum at any point (x, y, z) is determined by measuring the tangent to the trajectory and the curvature of the trajectory. The tangential to the trajectory makes an angle λ with the magnetic field as shown in Figure 4.1, and the solution to equation 4.1 yields three equations for $x(s)$, $y(s)$, $z(s)$ that describe a helix in space, parametrized by x , y , z , λ and p .

The above parametrization is in ideal conditions and does not take into account the effects of non-homogeneous magnetic field and particle's energy loss as it traverses the material and multiple scattering.

The inhomogeneous magnetic field $\vec{B}(x, y, z)$ is shown in Figure 4.2 and incorporated into the equations of motion. Additionally, the resulting parametrization scales with multiple scattering effects. Therefore, the trajectory is solved for recursively with the Runge-Kutta method described in Section 4.3.4.

The average energy loss of a charged particle heavier than the electron is given by the *Bethe-Bloch* formula [21]:

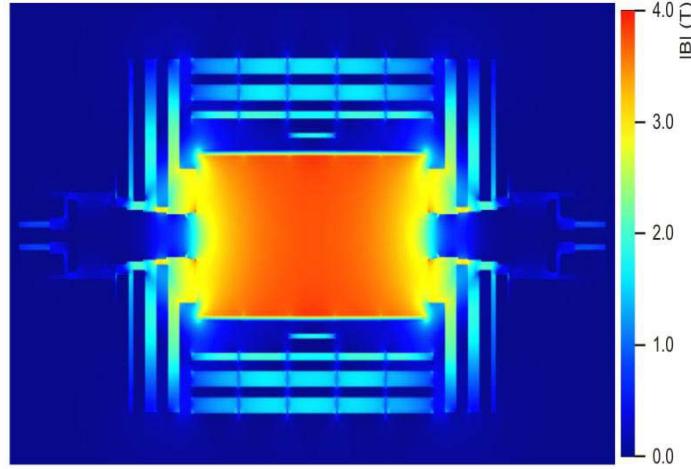


Figure 4.2: CMS Magnetic Field Map

$$-\frac{dE}{dx} = \frac{const}{\beta^2} \times \left[\ln\left(\frac{2m_e c^2 \beta^2 \gamma}{I}\right) - \beta^2 - \delta(\beta) \right], \quad (4.2)$$

where

- the constant term is roughly independent of material,
- the term inside of the logarithm corresponds to the mean excitation E
- $\delta(\beta)$ is the density effect

This formula provides the statistical energy loss per unit x (*density* \times *length*) and has to be incorporated into the equations of motion. For muons, a minimum value of $2 \text{ MeV cm}^2/g$ is reached at $E_\mu = 0.35 \text{ GeV}$. At higher muon momenta the ionization loss increases, such that at 100 GeV , the loss is $2.4 \text{ MeV cm}^2/g$.

4.3 Muon Reconstruction

In CMS, tracking, muon spectrometer and calorimetry information are used for muon reconstruction. This produces a final collection of high-level muon objects that contains three types of muon classification: standalone, tracker and global.

The reconstruction of muons starts with location and definition of a muon “seed”, as shown in Figure 4.3. Muon hit positions are locally reconstructed in individual muon detectors (DT, CSC, RPC). Then, reconstructed hits within each DT and CSC chamber are matched to form “segments”. Segments are then collected and matched to generated

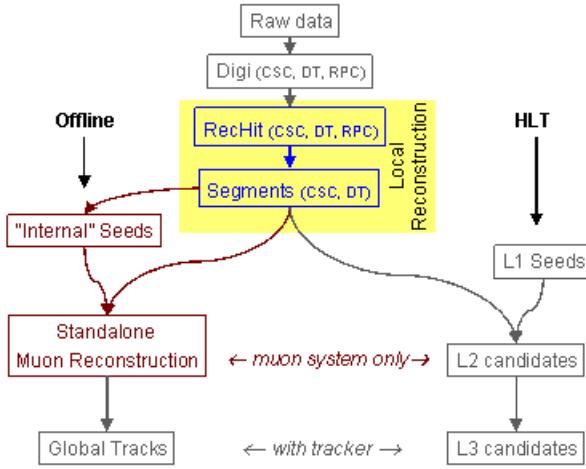


Figure 4.3: CMS Muon Reconstruction

seeds used to create a muon track. The result of this fit to the DT, CSC and RPC hits is a standalone muon. Standalone muons are matched with the tracks from the Tracker to form global muons. Tracker muons use a special algorithm which starts from a silicon track in the tracker and searches for compatible segments in the muon system. Muon reconstruction workflow is shown in Figure 4.4.

Seeding is done differently in the online mode, where the high-level trigger uses muon candidates found by the Level-1 muon trigger as seeds for muon reconstruction. This defines a region of interest where local reconstruction is performed for those chambers compatible with a seed.

4.3.1 Local Reconstruction

Local reconstruction is performed in individual muon detectors.

DT. The objects which result from DT local reconstruction are points within the DT cell. The distances from the wires are computed by converting drift times to drift distances using a linear time-to-distance parametrization with a constant drift velocity $v_{drift} = 54.3 \mu\text{m/ns}$ [96]. Errors in the reconstructed hit positions are calculated with a Gaussian fit to the distribution of residuals (differences between simulated and reconstructed distances from a wire).

CSC. The input to the local reconstruction in the CSCs are the signals in the cathode strips and anode wires. As a muon passes through the chamber the charge is typically distributed among three to five strips. The charge distribution is reconstructed and the position of the muon hit is determined from the peak position of the distribution. A

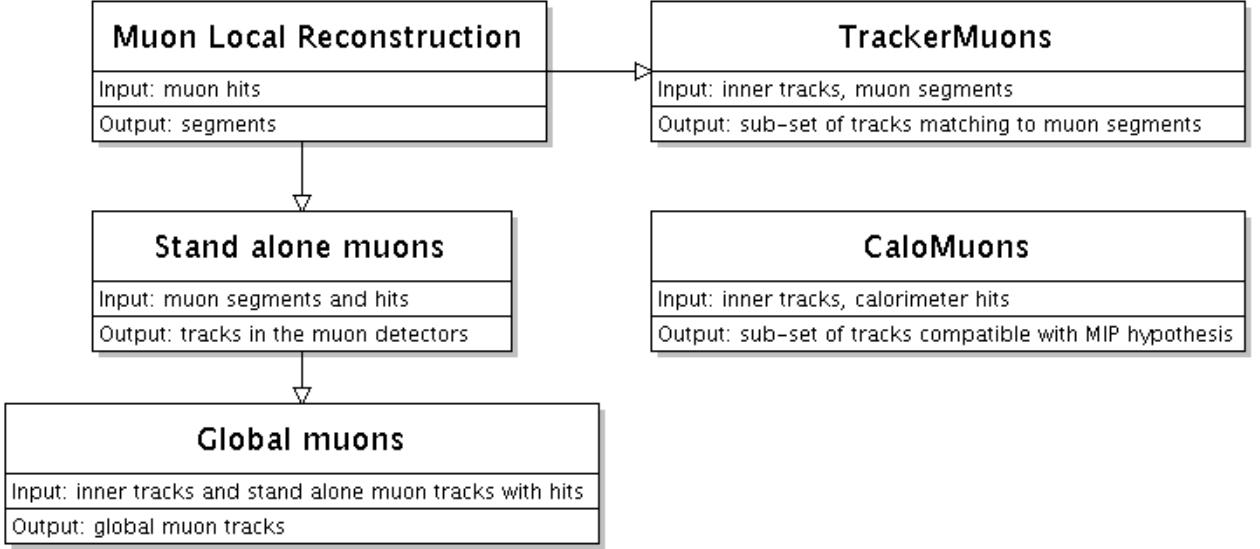


Figure 4.4: Muon Reconstruction Workflow

two-dimensional reconstructed hit (RecHit) is created for each intersection of a three-strip cluster and an anode wire group individually for each layer of the CSC chamber.

RPC. In the RPCs, local reconstruction produces points in the detector plane. A clustering procedure is applied which collects all the adjacent strips that have a signal. After the clusters are formed, a “center of gravity” is computed for each cluster. For the rectangular barrel strips this point is in the center of a rectangle, however for the endcap the computation is more complex because of the trapezoidal shapes of varying sizes.

4.3.2 Segment Reconstruction

DT. Segment reconstruction in DTs is performed in three steps. First, segment candidates are built from individual reconstructed hits in different layers, starting with the pair of hits with the greatest separation. If the angle of a segment candidate is compatible with a track that points to the interaction point, then the pair of hits is kept. For each pair, additional compatible hits are searched for in other layers, and a linear fit is performed to all the hit collections using their errors. Only one such segment candidate with the maximum number of hits and minimum χ^2 is kept for each collection.

This procedure produces a set of segment candidates. Then a test is performed to see if any of the segments share hits. From such segments, the best candidates are selected based on a number of hits and smallest χ^2 . Finally, hit reconstruction is updated with the new segment information and segments are refitted again.

CSC. Segments in the CSCs are built from the RecHits in each of the six layers in the chamber. The algorithm starts with the first and last hits in the chamber (using a local

layer x-coordinate) and constructs a straight line between them, while the $r\phi$ separation is required to be at least 1 cm. Then for each of the intermediate layers, the algorithm attempts to add a hit and update the linear fit accordingly. Only linear fits with reasonable χ^2 are considered and the hit most compatible to the segment is kept in a layer. Segments with at least four hits are considered further.

4.3.3 Seed Generation

From the segments obtained in the DTs and CSCs, muon seeds are generated. The algorithm searches for a pattern of segments using geometrical criteria (if only one segment passes the criteria it is kept as well). Then the transverse momentum of the seed is parametrized with the following equation:

$$p_T = A - \frac{B}{\Delta\phi} , \quad (4.3)$$

where

- $\Delta\phi$ is the angle of the DT segment with respect to the vertex position if the seed is in MB1 or MB2.
- If segments from both MB1 and MB2 exist then the weighted average of two estimated transverse momenta is used.
- If the seed candidate has segments only in MB3 and MB4, then the angular difference between the segments in the two stations is used to compute the p_T .

In the CSCs:

- the seeds are built from segments in the first two stations or the first and third station
- $\Delta\phi$ is the angular difference in the ϕ position of the two segments

4.3.4 Propagators

Several propagators are used during muon track reconstruction to predict the muon state vector (muon position, momentum and direction) given its initial state vector. The propagators properly account for the muon's energy loss as it traverses the material as well as for the presence of the magnetic field. They are used to predict the expected path of the muon through the detector and to propagate the initial state errors, expressed as a covariance matrix, to the final state point, including energy loss fluctuations and effects of multiple scattering.

CMS uses two kinds of propagators at different stages of muon reconstruction: the *Runge-Kutta* propagator and the *stepping-helix* propagator. The first is used inside the silicon tracker volume, while the latter is used for track propagation outside the tracker to the calorimeters and the muon system.

Runge-Kutta. The Runge-Kutta propagator is a numerical method that accounts for the non-uniformity of the magnetic field. The state of the trajectory is updated by choosing a boundary surface with subsequent propagation of the current trajectory state to this surface and introduction of material effects. The non-uniformity of the magnetic field is accounted for by solving the equation of motion with the fourth order Runge-Kutta method [23].

Stepping helix. The *stepping-helix* propagator provides a solution to the muon's equations of motion by using steps of finite helix length with all the parameters updated after each step. The magnetic field and material effects are updated at each step at the middle point, which is equivalent to a second order Runge-Kutta method. The propagator stops at all material and magnetic volume boundaries. Detector material navigation is based on the model of the CMS detector volume. The silicon tracker is modeled as several volumes of constant density. Similarly, the calorimeters, solenoid volume, yoke and endcap muon chambers are represented by volumes of constant material density. The barrel muon detector volume utilizes the magnetic field model: each magnetic and non-magnetic volume boundary is taken from the magnetic field description. If the field is greater than 0.6 T at each point then the volume is considered to be a solid iron. If the volume is non-magnetic, then it is treated as a chamber with constant density corresponding to the average density in the volume.

Material effects described in section 4.2 are accounted for during transport. The muon energy loss and its dependence on muon momentum correspond to the mean energy loss for iron. The values are fit by a single function:

$$\frac{dE}{dx} = -(11.4 + 0.96 \cdot |\ln(2.8p)| + 0.033p(1 - p^{-\frac{1}{3}})) \text{ MeV/cm} , \quad (4.4)$$

Energy loss in all non-iron volumes is represented by a fraction of the energy loss in iron. The fluctuations in energy loss is incremented at each step by $\Delta E^2/x(1 + p \cdot 10^{-3})$, where ΔE is the energy loss within that step in GeV.

Multiple scattering effects are accounted for based on material radiation length and the value of muon momentum using the following formula [118]:

$$\theta_0 = \frac{13.6 \text{ MeV}}{p} \sqrt{x/X_0} (1 + 0.38 \ln(x/X_0)) , \quad (4.5)$$

where X_0 is the radiation length of the material

To keep this dependence over multiple steps the propagator keeps track of the integrated path length as a function of material radiation lengths. Similar to energy loss, material radiation length is based on values relative to that of iron.

4.3.5 Standalone Muons

In the standalone muon reconstruction only the information from muon system is used - tracker is not used at all. The DTs, CSCs and RPCs participate in this procedure. After local reconstruction, muon track segments are made in the muon chambers. Afterwards, muon trajectory is built inside-out, using a Kalman filter procedure [71], where the extrapolated value of the muon's state vector is compared to the measured value at the next station's

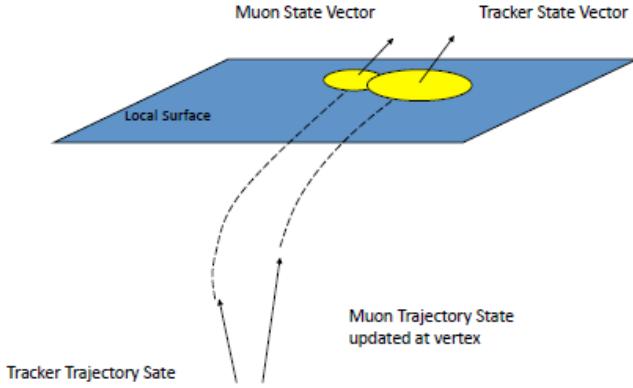


Figure 4.5: Two Trajectory States on a Common Surface

surface and updated accordingly. In the barrel DTs, the reconstructed track segments are used in the Kalman filter procedure, while in the endcap CSCs the individual 3D hits of the segments are used instead, due to the inhomogeneity of the magnetic field. A χ^2 -cut is applied to reject bad hits. If no hits are found at a particular station, the search continues at the next. The state vector of a muon is propagated from one station to another, taking into account muon energy loss in the material, effects of scattering as well as the nonuniformity of the magnetic field, as shown in Figure 4.5. All track parameters and errors are updated iteratively at each step, until the outermost station's surface is reached. Then a backward Kalman filter is applied, from outside-in, until the innermost surface is reached. Finally, the muon track is extrapolated to the interaction point and a final vertex-constraining fit is performed to muon's track parameters.

4.3.6 Global Muons

In the global muon reconstruction, silicon tracker hits are also used together with the information from the muon system. Beginning with a standalone muon, its trajectory is extrapolated from the innermost station to the outermost tracker surface, taking into account muon energy loss and effects of multiple scattering. Once the silicon tracker layers compatible with the muon's trajectory are found, regional track reconstruction proceeds in these layers, using an additional constraint that a muon originates from the interaction point. Initial candidates are built from pairs of reconstructed hits. Starting from the regional seeds, a track-reconstruction algorithm based on the Kalman-filter technique, is applied to reconstruct the tracks in the region of interest. The track trajectory is built, cleaned (all ambiguities resolved) and smoothed with a final fit. The trajectory builder creates possible trajectories propagating from one layer to the next, while the cleaner resolves ambiguities

on the basis of number of hits and the χ^2 of the fit. The final fit includes the reconstructed tracker hits and the original standalone muon reconstruction. A second cleaning step resolves possible ambiguities and selects final muon candidates with a χ^2 cut. Therefore, for each standalone muon track there is a maximum of one reconstructed global muon.

4.3.7 Tracker Muons

In some cases, the information in the muon system used for the standalone reconstruction is minimal and the reconstruction fails. This is particularly the case for low p_T muons (below 5 - 7 GeV/c) that either leave very few hits in the muon system or do not have enough momentum to reach the muon system. In this case, an alternative approach that begins the muon reconstruction with the tracks in the silicon tracker and looks for compatible hits in the muon system is used. These muons are called “Tracker Muons”.

The algorithm starts with reconstructed silicon tracker tracks above a minimum p_T threshold. Each track is propagated to the calorimeter and then to the muon system using the stepping helix propagator, such as that used in the standalone muon reconstruction, again taking into account the magnetic field, multiple scattering and muon energy loss as the muon passes through the material. During this propagation stage, the algorithm keeps track of the chambers being crossed by the track as well as the local x and y coordinates of the segments, their slopes and uncertainties on those quantities.

A tracker muon is stored if there is a minimum number of associated segments, default being one segment. In subsequent muon selection, tracker muons with at least two matched segments are used.

4.3.8 Arbitration

If there are two or more tracks close to each other, which may happen to muons inside jets, it is possible for several segments to be associated with more than one track at a time. Muon arbitration procedure resolves this problem. This algorithm uses geometrical criteria (ΔR between the track and a segment) to pick the optimal match between a segment and an extrapolated track.

4.3.9 Calorimeter Based Muons

Calorimeter based muons, or “calo muons” are a subset of muons that have a correspondence of a silicon tracker track to energy deposition in the calorimeters characteristic of a minimum ionizing particle. Because of a high fake rate, calorimeter-based muons are not used in the following selection.

4.4 Jet Reconstruction

As described in Section 2.7, there are QCD jets, in addition to lepton jets, produced in the final state of dark sector cascade decay. A jet is a cone of hadrons and other particles produced during the hadronization process of quarks and gluons as shown in Figure 4.12. Because of confinement, quarks that carry a color charge cannot exist freely and are forced to hadronize. The resulting collection of hadrons traveling in a cone deposits energy in

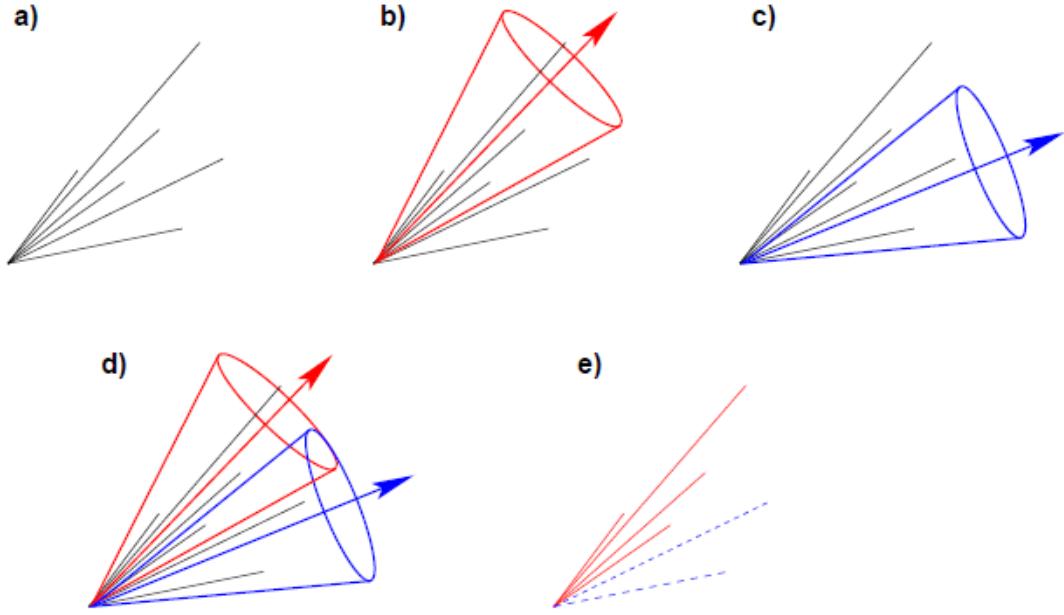


Figure 4.6: Cone-type Jet Reconstruction: **a.** list of initial 4-vectors **b.** and **c.** stable cones found **d.** splitting/merging algorithm applied to assign 4-vectors in overlap regions to a single jet **e.** final jets

the calorimeters and is reconstructed as a particle jet, allowing the measurement of the underlying parton level process of the original quark or gluon.

4.4.1 Jet Algorithms

During early analyses of collider experiments, jet algorithms assumed that particle jets were simple cones. Over time, more sophisticated jet algorithms emerged [30, 104]. Currently, there are two main classes of jet algorithms: cone and sequential recombination algorithms. In the cone algorithm the input objects (particles or calorimeter towers) are grouped into intermediate states, while the calculation of jet properties, such as jet energy and direction, is done in one final step after all of the jets are found. In contrast, during sequential recombination the input objects are merged into potential jets at each iteration step and jet quantities are computed continuously.

Cone algorithms typically use a seeding approach with the initial seed defined as the center of a cone or as midpoints between previously found stable cones. An exception to this class is the Seedless Infrared Safe Cone Algorithm (SisCone) [104], which does not rely on a seed.

SisCone. The cone algorithm, as the name suggests, uses a concept of a stable cone with a circle of radius ΔR in the $\eta-\phi$ space and the sum of all the momenta of the particles

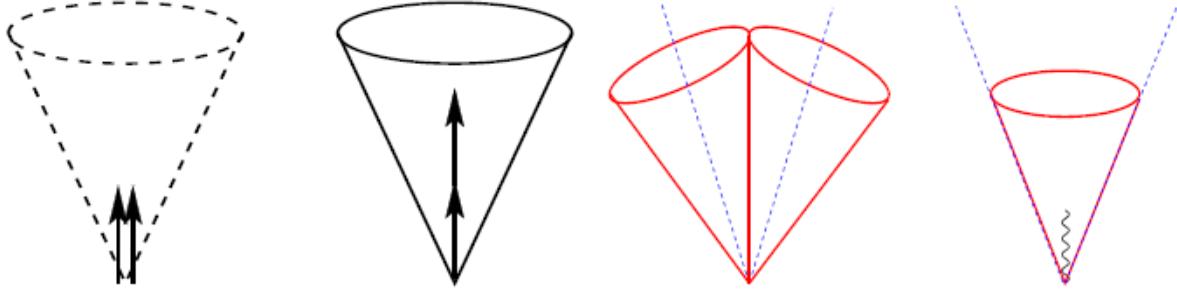


Figure 4.7: **Left:** collinear unsafety **Right:** infrared instability in jet algorithms

within the cone points in the direction of the center of the circle. The algorithm attempts to identify all the stable cones as shown in Figure 4.6. In contrast to seeded algorithms that miss some stable cones when soft particles are added or when energy of a single particle is distributed to two collinear particles, which leads either to infrared or collinear unsafety as shown in Figure 4.7 (divergence in perturbative expansions), the seedless algorithm finds all stable cones and is therefore infrared and collinear safe. This satisfies jet algorithm requirements set forth in the Snowmass accords [69].

Sequential Recombination. This algorithm class defines a distance between pairs of particles and makes successive recombinations of the pair of the closest particles, stopping when the resulting final objects are too far apart. There are three main sequential recombination algorithms: k_t , Cambridge-Aachen and anti- k_t . They differ in the way the distance d_{ij} between entities (particles, pseudojets) i and j is defined:

$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2) , \quad (4.6)$$

where

- $p = 1$ for the k_t algorithm
- $p = 0$ for the Cambridge-Aachen algorithm
- $p = -1$ for the anti- k_t algorithm.

A particular recombination algorithm searches for the smallest d_{ij} . Once the smallest d_{ij} is found, objects i and j are removed from a list of particles used for clustering and merged into a new jet object. This object is then re-inserted in the list of objects to cluster. This procedure is repeated until all of the objects are included in the jets.

The main feature of the anti- k_t algorithm is that it is infrared and collinear safe as required by the Snowmass accords. Another important part is that there is a connection to the cone algorithms: because of the distance definition in the anti- k_t algorithm, softer particles are first recombined with hard ones before being (re)combined together, resulting

Table 4.1: Calorimeter Cell Thresholds

Section	Threshold (GeV)
HB	0.7
HE	0.8
HO	1.1/3.5(Ring 0/Ring 1,2)
HF(long)	0.5
HF(short)	0.85
EB	0.07(per crystal, double-sided)
EE	0.3(per crystal, double-sided)
EB sum	0.2
EE sum	0.45

in regular, hard jets. Because of this feature as well as its speed and simplicity, the anti- k_t algorithm is chosen as the most appropriate algorithm for lepton jets.

4.4.2 Standard Jet Reconstruction in CMS

Jet reconstruction in CMS proceeds in the following steps. First, Analog-to-Digital (ADC) counts are converted to energy in individual calorimeter cells, which is done differently for HCAL and ECAL, as discussed in section 3.2.10. Then HCAL and ECAL cells are combined into projective towers, called CaloTowers, following the general granularity of HCAL. In the barrel region of the calorimeter ($|\eta| < 1.4$) an unweighted sum of a single HCAL cell and a 5×5 matrix of ECAL crystals form a projective calorimeter tower. In the endcap region ($1.4 < |\eta| < 3.0$) the association between HCAL cells and ECAL crystals is more complex and follows the respective subdetector geometry.

The final part of the reconstruction process is a clustering algorithm that produces the jets. These algorithms are generic and can run on any set of 4-momenta, which separates the clustering algorithm from the physical nature of the jet constituents. Towers are treated as massless particles, where the particle energy is given by the tower's energy and direction is defined by the center of a tower and the interaction point. “E-scheme” is used as the recombination scheme: the energy and momentum of a jet is defined as the sums of the energies and momenta of jet constituents.

There are three jet clustering algorithms chosen by CMS: K_t , anti- k_t and SisCone, described in the previous sections.

4.4.3 Jet Types

Due to the non-compensating nature of the HCAL calorimeter described in Section 3.2.5, the response to charged pions is different to that of electrons. The high e/h ratio leads to a non-linear response in calorimeter jets. Therefore additional algorithms that use other detectors, in particular the silicon tracker, have been introduced to improve the jet energy resolution. For each of the jet clustering algorithms four types of jets are produced in the standard reconstruction: calorimeter jets, Jet-Plus-Track (JPT) jets, Particle-Flow jets and track jets.

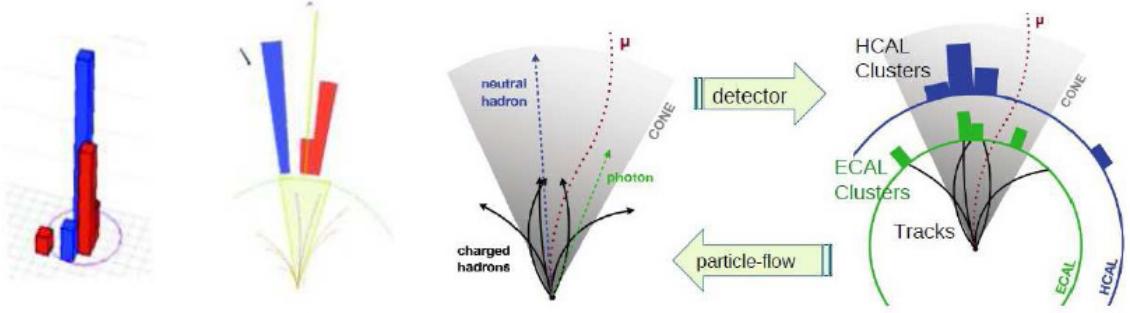


Figure 4.8: Jet Reconstruction. **Left:** Calorimeter jets **Center:** Jets-Plus-Tracks jets, **Right:** Particle-Flow jets

Calorimeter Jets. Calorimeter jets, or CaloJets, use calorimeter towers (CaloTowers) as input to the jet clustering as shown in Figure 4.8 left. They also contain information about the electromagnetic fraction of the jet. To reduce the noise in the HCAL towers and the amount of data transferred forward, zero-suppression is applied. This procedure means an application of a lower cut on the energy or the transverse energy of a calorimeter readout cell or calorimeter tower built from a calorimeter cell. Towers and cells with energy that is below the cut are not used in jet reconstruction. These thresholds are listed in Table 4.1. For calorimeter jets, to reduce the contribution from pileup (extra proton-proton interactions in the same bunch-crossing) calorimeter towers with the transverse energy of $E_T^{Towers} < 0.3$ GeV are not used in jet reconstruction.

Jet-Plus-Tracks. In the Jet-Plus-Tracks algorithm, calorimeter jets are reconstructed as above, and then particle tracks are associated to the jet based on the separation in $\eta - \phi$ space between the jet axis and the track momentum measured at the interaction vertex as shown in Figure 4.8 center. Tracks are then projected on the surface of the calorimeter and classified as “in-cone”, if they are within the cone of the jet, or “out-cone”, if they are outside the cone. Both the momenta of the in-cone and out-cone tracks are added to the energy of the jet. For the in-cone tracks, the average energy contribution in the calorimeters is subtracted depending on the jet momentum. Then the final direction of the jet axis is corrected as well.

The Jet-Plus-Tracks correction can be summarized in the following equation:

$$E = E_{jet}^{calo} \times f_{ZSP} + \sum_{in} (p_{tk} - \langle E_{tk}^{calo} \rangle) + \sum_{out} p_{tk} + \Delta E_{tk}^{ineff} + \sum_{\mu} (p_{\mu} - 2\text{GeV}) \quad (4.7)$$

where

- E_{jet}^{calo} is the raw calorimeter jet energy,
- f_{ZSP} is a correction function for zero-suppression,

- $\langle E_{tk}^{calo} \rangle$ is the average calorimeter energy as a function of the track momentum,
- ΔE_{tk}^{ineff} is a correction for track finding inefficiency.
- The algorithm also adds the momenta of muons identified inside the jet while subtracting their average calorimeter deposition, which amounts to 2 GeV.

More details about the Jets-Plus-Tracks algorithm can be found in [36, 41].

Particle Flow Jets. Particle Flow Jets use particle flow objects in the jet clustering. They additionally contain information about the number of various particle flow objects and their energy contribution. The Particle Flow algorithm [37, 42] uses the information of all CMS sub-detectors to identify and reconstruct all the particles in the event: muons, electrons, photons, charged and neutral hadrons. Particle Flow Jets are then reconstructed from the list of reconstructed particles as shown in Figure 4.8 right, using the anti- k_t algorithm with a jet cone size of $R = 0.5$. Particle flow algorithm extrapolates tracks reconstructed in the silicon tracker to the surface of the calorimeters, which are then associated to calorimeter energy depositions. Neutral hadrons and photons are reconstructed from calorimetric clusters. The strong point of the particle flow algorithm is that it allows the identification of charged hadrons and photons inside jets.

Track Jets. Track jets are constructed from the reconstructed charged particle tracks in the tracker. The algorithm uses well-measured tracks associated to the primary vertex and is independent of calorimetric measurements. Track jets have very good angular resolution and because of their association to a single primary vertex, are transparent to pileup effects from additional proton collisions producing tracks in other vertices.

Track jets use the anti- k_t algorithm with a jet cone size of $R = 0.5$ and are similar to the charged component of particle flow jets. The main difference is, while particle flow jets aim for precise energy measurement and therefore include very low momentum tracks with large impact parameters, the track selection is stricter in track jets, leading to relatively lower track multiplicity and energy response.

4.4.4 Jet ID

CMS has developed jet quality criteria (Jet ID) in order to retain the vast majority of real jets and reject most fake jets from calorimeter and detector electronics noise. To pass the Jet ID, calorimeter jets must have the electromagnetic fraction $EMF > 0.01$ if they are in the calorimeter fiducial region of $|\eta| < 2.6$. This reduces the contribution from “soft fakes”, which are low p_T jets consisting of a few calorimeter cells just above the reconstruction threshold without energy deposition in ECAL. “Hard fakes” that have less than 20 calorimeter towers with a large electromagnetic fraction are likely to originate from muons. The remaining Jet ID requirements are:

- 90% of the jet energy must be contained in more than one calorimeter cell
- Jet energy fraction attributed to the hottest hybrid photodiode (HPD), f_{HPD} , must be less than 0.9.

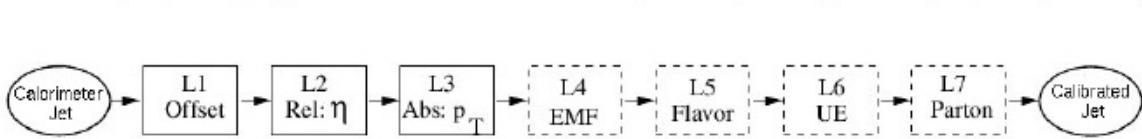


Figure 4.9: Factorized Jet Energy Corrections in CMS. Minimum requirements are shown in solid boxes, optional in dashed boxes [39].

- Particle Flow jets are required to have the charged fraction $CHF > 0.0$ if within the fiducial tracking region $|\eta| < 2.4$,
- Particle Flow neutral hadron fraction $NHF < 1.0$,
- Particle Flow charged electromagnetic fraction $CEF < 1.0$
- Particle Flow neutral electromagnetic fraction $NEF < 1.0$

Jet-Plus-Track jets are treated as calorimetric jets for the purpose of jet identification.

4.4.5 Jet Energy Corrections

Because of the non-linear response of the calorimeters as well as electronics noise and presence of pileup from other pp collisions in the same bunch-crossing, the translation between the true parton energy and the measured jet energy is not straightforward. Jet energy corrections allow an improvement in the accuracy of this procedure.

There are several levels of factorized corrections which are implemented in CMS [34, 39] as shown in Figure 4.9. Each of them addresses a different effect. Generally, each layer of correction represents a scaling of the jet four-momentum depending on various jet quantities such as p_T , η , flavor, e.t.c. These corrections are applied sequentially and in a fixed order:

$$E_{Corrected} = (E_{Uncorrected} - E_{Offset}) \times C_{rel}(\eta, p_T'') \times C_{abs}(p_T') , \quad (4.8)$$

- where p_T'' is the transverse momentum of the jet corrected for offset,
- p_T' is the transverse momentum of the jet corrected for offset and pseudorapidity dependence.
- C_{rel} is the relative correction obtained by removing variations in the jet response relative to the central region.
- C_{abs} is the absolute correction for variations in jet response versus p_T .

Level 1 Detector Noise and Pileup. The first jet energy correction accounts for the pileup and noise effects by removing the energy of the pileup events and noise. The noise contribution is estimated with random triggers using a veto on a real collision event, while pileup is estimated from the Minimum Bias trigger events. For both types of events, the average calorimeter p_T in the cone of size 0.5 are summed up and then subtracted.

Level 2 Relative Correction. The second correction attempts to make the jet energy response flat as a function of the pseudorapidity η . This is achieved by correcting a jet in arbitrary η relative to a jet in the central region of $|\eta| < 1.3$ with the Monte-Carlo truth and a data-driven dijet balancing technique. In the dijet balancing procedure, p_T balance is used between back-to-back dijet events with one jet in the central region of the calorimeter $|\eta| < 1.3$ and the other in an arbitrary η . The two leading jets are required to be azimuthally separated by $\Delta\phi > 2.7$ radians and no additional hard jets are allowed.

Level 3 Absolute Correction. The next correction makes the jet energy response flat in p_T . After the level two correction, the jet is corrected back to the particle level using the Monte-Carlo truth or the Z^0 +jet/photon+jet data driven technique, requiring that the corrected CaloJet (JPT or Particle Flow Jet) p_T is on average equal to the GenJet p_T . In the photon+jet procedure, p_T balance between a back-to-back photon and a jet is used to derive the correction. An isolated photon candidate with $p_T > 15$ GeV in the central calorimeter region $|\eta| < 1.3$ and a recoiling jet in the same barrel region are used. Similarly, in the Z^0 +jet technique, p_T balance between a jet and a Z^0 that decays to two muons is used to obtain the correction. At low transverse momentum, calorimeter jets require a large multiplicative factor (up to 2) because of the non-compensating nature of the calorimeters, while JPT, Particle Flow and Track Jets do not, because they rely heavily on tracking information.

Level 4 Electromagnetic Fraction Correction. The goal of the next correction is to improve the jet energy resolution by correcting dependence on the jet electromagnetic fraction (EMF) of the jet energy response that shows significant deviations for low and high values of EMF.

Level 5 Jet Flavor Dependence Correction. Level five jet energy correction addresses the dependence of response on the jet flavor since different corrections are needed for different jet flavors. For instance, jets from a W boson hadronic decay consist of up, down, strange and charm quarks, which on average have a higher response than the bottom quark and gluon jets, because light quarks fragment into higher momentum particles. These corrections are determined from the QCD and top quark pair events.

Level 6 Underlying Event Correction. Level six optional correction corrects for underlying event energy due to soft interactions involving spectator partons.

Level 7 Parton Correction. The final jet energy correction corrects back to the parton level, making the CaloJet (JPT or Particle Flow Jet) p_T equal on average to the original parton p_T . Gluons, which radiate more than the light quarks, have lower jet energy response because more of the final state radiation falls outside the jet. This correction uses matching between the original parton and a GenJet. The response of the GenJet depends on the size of the jet. For cone algorithms, the response increases with the size

parameter. Cone and sequential recombination algorithms with comparable size parameters have similar parton corrections.

4.5 Missing Transverse Energy Reconstruction

The missing transverse energy \cancel{E}_T is a quantity that is extremely important to the lepton jet dark matter analysis. This is because the dark particle can escape detection and become missing energy as described in Section 2.7.3. Presence of missing energy is an indicator of non-interacting particles which may be dark matter, and therefore its accurate measurement is paramount to identification of such particles. Understanding the missing transverse energy spectrum is critical to the separation of a possible new physics signal with real missing energy from other detector effects.

Due to the momentum conservation during the collision, the net transverse momentum should be zero. Because of detector resolution effects, noise, the non-compensating nature of the CMS calorimeters and other non-linearities, the missing transverse energy is usually non-zero. Neutrinos escape detection and provide an additional contribution to the \cancel{E}_T . Missing transverse energy can be defined as a sum of all the deposited calorimeter energy in the x- and y- directions multiplied by a factor of (-1) :

$$\cancel{E}_T \equiv \sqrt{\cancel{E}_x + \cancel{E}_y} \quad (4.9)$$

where

$$\cancel{E}_{x,y} = - \sum_i E_i^{x,y} \quad (4.10)$$

CMS employs various methods to measure the missing transverse energy \cancel{E}_T , such as Calo \cancel{E}_T , PF \cancel{E}_T and TC \cancel{E}_T [47], which follow closely the several types of jet reconstruction outlined in Section 4.4.3. We choose the particle-flow-based technique called PFMET because of its optimal resolution across the energy spectrum [47]. The PF \cancel{E}_T is computed with the scalar sum of transverse energies of all the particle-flow particles. More details about the PF \cancel{E}_T can be found in Reference [42].

4.6 Event selection

The data sample was collected by the Compact Muon Solenoid detector at the 7 TeV center-of-mass energy in proton-proton LHC collisions during the data-taking period of 2010. It corresponds to integrated luminosity of $36 \pm 4 \text{ pb}^{-1}$. Section 4.6.1 describes the luminosity measurement with the CMS detector, while Sections 4.6.2 and 4.6.3 discuss the trigger and vertexing requirements.

4.6.1 Luminosity

The integrated luminosity is measured with the luminosity system designed to monitor performance of the Large Hadron Collider at CMS. The system is located in the hadronic forward calorimeter 11.2m from the interaction point and utilizes an additional hardware

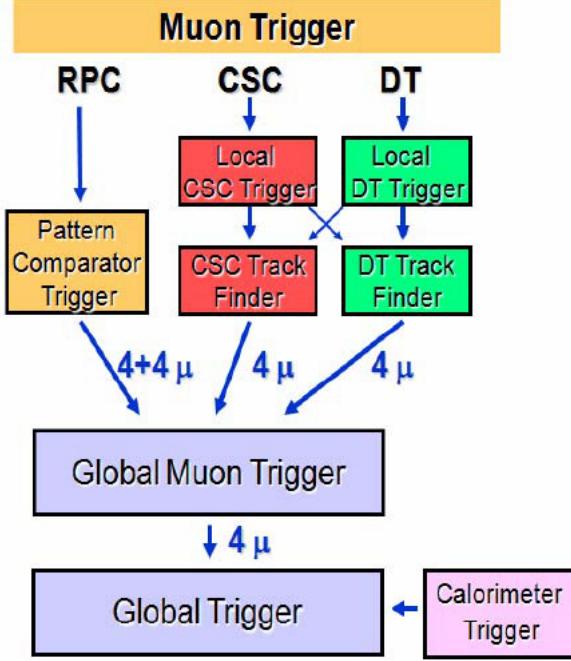


Figure 4.10: Global Muon Trigger Selection of up to 4 Muon Candidates

board called the HF Luminosity Transmitter (HLX), mounted on the HCAL Trigger and Readout (HTR) board.

To measure instantaneous luminosity the system uses the fact that the average number of interactions per bunch crossing μ is proportional to the luminosity L :

$$\mu = \frac{\sigma L}{f_{BX}} , \quad (4.11)$$

where f_{BX} is the bunch-crossing rate. Details of the luminosity (lumi) system hardware and measurement can be found in [43].

4.6.2 Trigger Requirements

Muon Trigger. All the muon systems discussed in Section 3.2.6 participate in the trigger. Barrel DT chambers provide track segments in the ϕ -projection and hit patterns in the η -projection for local triggers. The Cathode Strip Chambers deliver three-dimensional track segments in the Endcaps. The local trigger information is delivered to the Regional Muon Trigger, that consists of DT and CSC track finders as shown in Figure 4.10. These algorithms join track segments into track candidates and assign their physical attributes as shown in Figure 4.11. In addition, the RPC chambers provide track candidates based on regional hit patterns with excellent time resolution. The Global Muon Trigger combines the information from the three muon detectors.

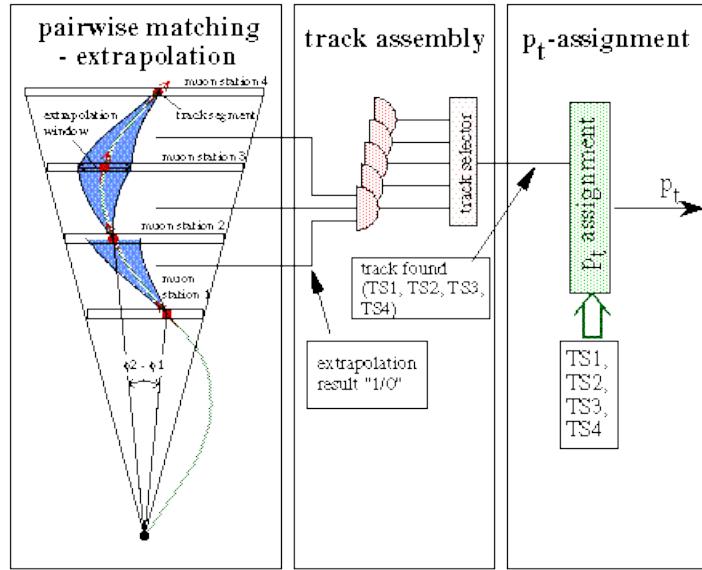


Figure 4.11: Barrel Track Finders combine input vectors from DT stations and form tracks with well-defined p_T

Table 4.2: HLT Triggers Used in the Analysis

HLT Trigger	Start: Run	End: Run	Recorded Luminosity pb^{-1}
HLT Mu9	135921	146116	4
HLT Mu11	146428	148058	11
HLT Mu15	148882	149442	21

Trigger Selection. For simplicity, only un-preserved lowest threshold high-level single muon triggers are used (see section 3.2.8 for high-level trigger pre-scaling discussion). The value of the minimum threshold increased several times during data taking from 9 GeV to 11 GeV to 15 GeV as shown in Table 4.2, reflecting the increase in the LHC luminosity.

Whenever available, the un-preserved 15 GeV single muon high-level trigger is used. In all other cases, this requirement is emulated by requiring a presence of a muon candidate of at least 15 GeV.

4.6.3 Vertex Requirements

Each event passing initial pre-selection is required to have at least one primary vertex (PV). The vertex is also required to have $|r| < 2$ cm and $|z| < 24$ cm to efficiently reject cosmic-ray muons which do not originate from the interaction point. In other words, projection of the muon is required to be within 2 cm of the proton beam axis and within 24 cm along the z-axis from the center.

4.6.4 Cosmic Rays

It is possible for a cosmic ray to pass through the CMS detector in coincidence with an LHC bunch crossing event. The main feature of such muon is that it passes through the whole detector as opposed to originating from the interaction point. Therefore, a presence of another track of almost equal p_T and opposite direction is an indication of a cosmic muon, and they can be adequately rejected by requiring that the muons are not back-to-back. Muon timing is used as well to reject cosmic rays that do not come in coincidence with the collision.

Another possibility for a cosmic-ray event to interfere with a collision event is if some of the segments from a collision muon are mismatched to those of a cosmic-ray muon. Basic muon identification cuts listed below are effective in reducing this possibility to near zero.

4.7 Muon Pre-Selection

Prior to grouping muons into jets, muon quality criteria are applied to remove various backgrounds such as decays in flight and punch-through. Punchthroughs are hadrons which penetrate the calorimeters to the muon system leaving hits reminiscent of a muon. Decays in flight are primarily pions, kaons and sometimes strange baryons that decay in flight into muons a significant distance away from the primary and secondary vertices.

4.7.1 Global Muons

Global Muon χ^2 . Normalized global muon χ^2 is a powerful variable to reject both decays in flight and punch-through. In what follows, normalized $\chi^2 < 10$ is applied.

Track Quality Cuts. To reject any remaining decays in flight, the following track quality cuts are applied:

- Track impact parameter $d_0 < 3\text{cm}$

- Track impact parameter $z_0 < 15\text{cm}$
- Normalized $\chi^2 < 10$
- Number of silicon hits in the *track* > 10
- Number of pixel *hits* ≥ 1
- Number of valid muon *hits* ≥ 1

A cut on track impact parameter d_0 is very efficient for the rejection of decays in flight, as well as muons from bottom (b) and charm (c) decay.

4.7.2 Tracker Muons

Arbitration. As discussed in section 4.3.8 tracker muons used in the following selection are arbitrated to ensure a unique segment for each muon track which eliminates muons from accidental overlaps and punchthrough.

Segment Match. As discussed in section 4.3.7 tracker muons that have at least two segments matching the track are used.

TMLastStation Algorithm. Since muons pass through the whole muon system while hadrons do not, one of the requirements is that one of the matching segments has to be in the last station of the muon system.

The following summarize track quality requirements for tracker muons:

- Track impact parameter $d_0 < 3\text{cm}$
- Track impact parameter $z_0 < 15\text{cm}$
- Track $\chi^2 < 4$
- Number of silicon hits in the *track* > 10
- Number of pixel *hits* ≥ 1
- Number of matched *segments* ≥ 2

4.7.3 Muon Acceptance Requirements

The minimum acceptance requirements for an event is to contain at least one reconstructed muon of $p_T > 15\text{ GeV}$ and $|\eta| < 2.4$. The former ensures a trigger efficiency close to 100%, and the latter ensures that the muon candidate is in the fiducial region of the CMS detector.

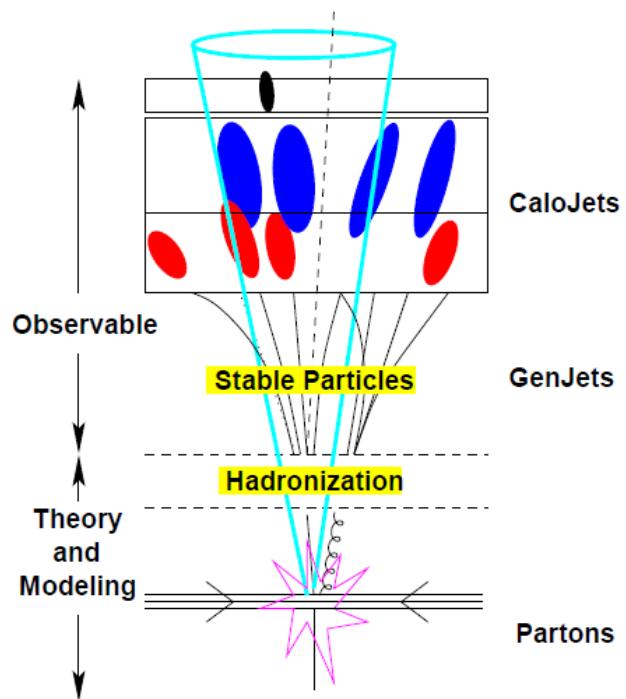


Figure 4.12: Jet Evolution: from partons to reconstructed jets.

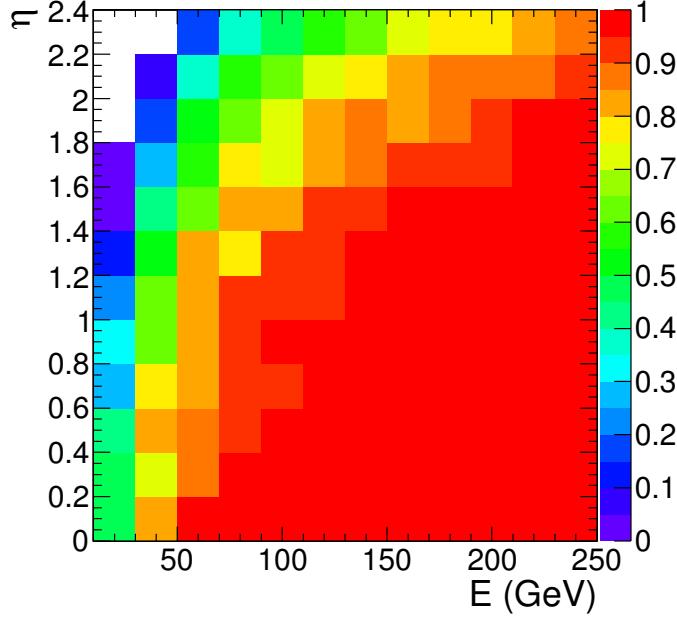


Figure 4.13: Lepton jet kinematical acceptance as a function of the boost and pseudorapidity

4.8 Lepton Jet Tagging

To evaluate the acceptance and reconstruction efficiency of lepton jets, a model independent lepton jet tagging study is conducted using the expected boost and production angle of a light resonance in the CMS detector. A modified J/ψ -like resonance with a mass of 0.5 GeV is produced and decays to two opposite-sign muons. Lepton jets are reconstructed using the anti- k_t algorithm, described in Section 4.4.1, for various values of the boost and pseudorapidity of the original resonance.

Lepton jets are reconstructed on two levels: the reconstruction (RECO) level and generator (GEN) level. In the latter case, the lepton jet algorithm with the same parameters as in Table 2.1 is applied directly to a collection of Monte-Carlo generator level muons.

4.8.1 Kinematical Acceptance and Jet Tagging Efficiency

The likelihood of finding both generator level muons within the size of the lepton jet with cone size $R = 0.1$ is shown as a function of the boost and pseudorapidity in Figure 4.13. The signal model described in Section 2.7 produces boosted muon jets with $\gamma > 100$, and therefore lies in the region of high acceptance. For each bin in Figure 4.13 at least 10,000 events were simulated.

We define lepton jet tagging efficiency as:

$$\epsilon_{TAG} = \frac{N_{RECO}^{LJ}(2\mu)}{N_{GEN}^{LJ}(2\mu)} \quad (4.12)$$

where

- $N_{RECO}^{LJ}(2\mu)$ is the number of reconstructed lepton jets that contain exactly two muons
- $N_{GEN}^{LJ}(2\mu)$ is the number of generator level lepton jets that contain exactly two muons

4.8.2 Muon Selection

Three choices of muon selection are considered in the context of lepton jets: tracker, global and “trackseg”. Tracker and global muons are described in Section 4.3. Trackseg is an optimized muon selection, described in Section 4.3.7, where at least two matched arbitrated segments are required for tracker muons. General track quality criteria listed in Section 4.7.1 are applied. The muons are furthermore required to be within the geometrical acceptance of the detector.

As Figure 4.14 shows, simple tracker muons have the best reconstruction efficiency; however, as discussed in Section 4.3.7, they also account for a greater fake rate and lower background rejection than global muons. Trackseg muons have a lower fake rate and higher background rejection power than simple tracker muons, while maintaining a high jet-tagging efficiency rate.

Due to the combined high lepton jet tagging efficiency and strong background rejection power, trackseg is chosen as the optimal muon selection. Global muons are not considered any further because of their low reconstruction efficiency. It is worth noting that global muon reconstruction efficiency has improved substantially for the 2011 data, but not retroactively applicable to the 2010 dataset. For future analyses on 2011 data, trackseg will still remain the optimal selection, despite the improvements in global muon reconstruction, although the difference should be much smaller.

4.9 Jet and Missing Transverse Energy Selection

Particle Flow (PF) jets, described in Section 4.4.3, with the Jet Energy Corrections, outlined in Section 4.4.5, are selected for further analysis due to their robust performance across the energy spectrum. A minimum transverse momentum of $P_T > 40$ GeV/ c is required for both QCD jets. As discussed in section 4.5 Particle Flow based Missing Transverse Energy algorithm called PFMET is used for the determination of the \cancel{E}_T [42].

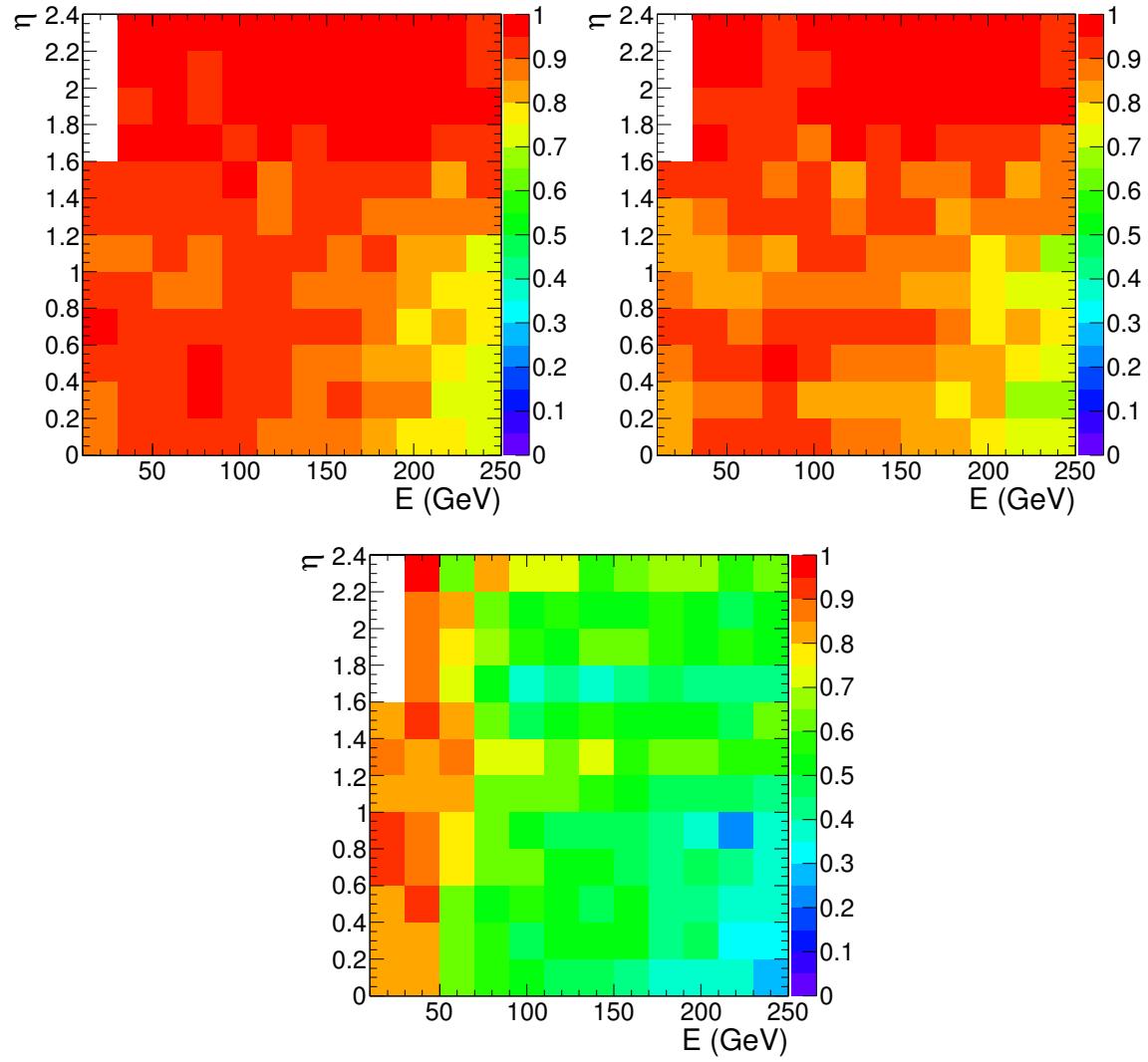


Figure 4.14: Lepton Jet Tagging efficiency as a function of pseudorapidity and energy. **Top left:** tracker muons. **Top right:** trackseg muons. **Bottom:** global muons. The empty bins at high pseudorapidity and low boost are due to the kinematic acceptance.

CHAPTER 5

SEARCH FOR DARK MATTER IN LEPTON JETS

In order to draw any reasonable conclusions about the presence of new physics in the lepton jet events that were found during the analysis, systematic uncertainties, efficiencies as well as estimates of the background in the signal region need to be taken into account. Systematic uncertainties are presented in Section 5.1. Efficiencies related to trigger and muon identification are discussed in Section 4.8. Lepton jet kinematical distributions and performance in the opposite-sign dimuon channel are shown in Sections 5.2 - 5.3. Final event selection is described in Section 5.4. Background estimate is described in Section 5.5. A decision-making framework is discussed in Section 5.6. Because predictions of the Standard Model agree with the data, a frequentist technique in Section 5.7 is used to set upper limits on the supersymmetric particle production in the context of the dark matter model. The results are presented in Section 5.8.

5.1 Systematic Uncertainties and Efficiencies

5.1.1 Systematic Uncertainties

The main sources of systematic uncertainties, summarized in Table 5.1, come from the measurements of luminosity, background estimation, signal acceptance and efficiency calculations. The uncertainty on the luminosity measurement is estimated to be 11% during the 2010 run. This is a measurement uncertainty independent of the analysis. The uncertainty in the background estimates is determined from the largest difference obtained by using background regions of various sizes from 0.1 to 0.7. Using this methodology a 30% error

Table 5.1: Systematic Uncertainties

Source	Uncertainty(%)
Luminosity	11
Background Estimation	30
Signal Acceptance	1
Trigger Efficiency	3

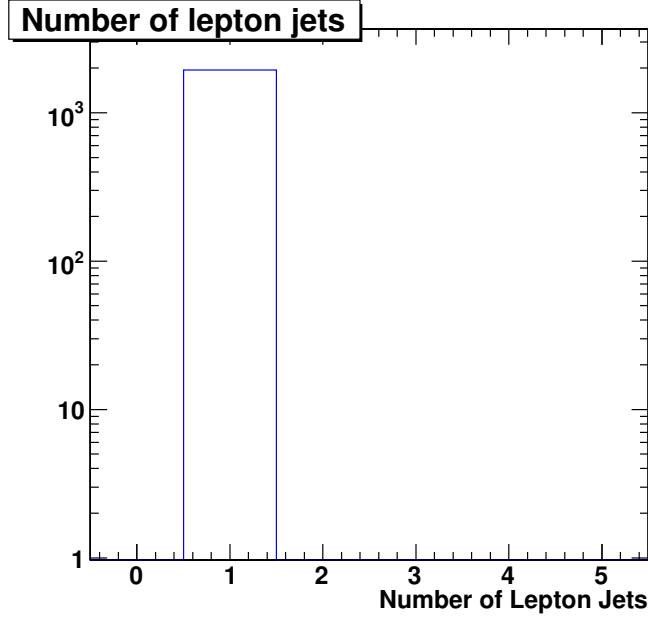


Figure 5.1: Lepton Jet Multiplicity in Data

is obtained. To measure the systematic uncertainty on the signal model acceptance, the signal is reproduced for various squark masses with two different probability distribution functions, CTEQ6 (default) and MRST. Then the difference in the errors is computed and added in quadrature. Following this procedure a 1% systematic error estimate is obtained on the signal acceptance. Trigger efficiency uncertainty is computed from the difference in the data and Monte-Carlo measurements and found to be 3%. These total uncertainty values are consistent with other estimates from comparable CMS analyses of the 2010 data [46].

5.1.2 Efficiencies

The various efficiencies relevant to the lepton jet analysis such as the trigger efficiency and muon identification efficiency are described in Appendix A.

5.2 Lepton Jet Kinematical Distributions in Data

5.2.1 Lepton Jet and Muon Multiplicity

Figure 5.1 shows the multiplicity of lepton jets passing the initial selection criteria. All of the 1938 events have one associated lepton jet in them. Figure 5.2 shows the multiplicity of muons inside lepton jets that pass the selection criteria. Most of them have 2 muons, however there are 11 events with 3-muon lepton jets.

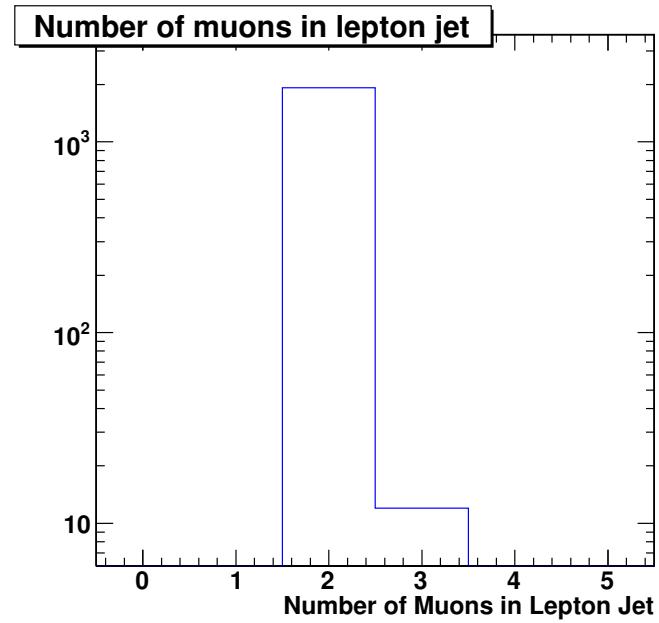


Figure 5.2: Number of Muons in Lepton Jets in Data

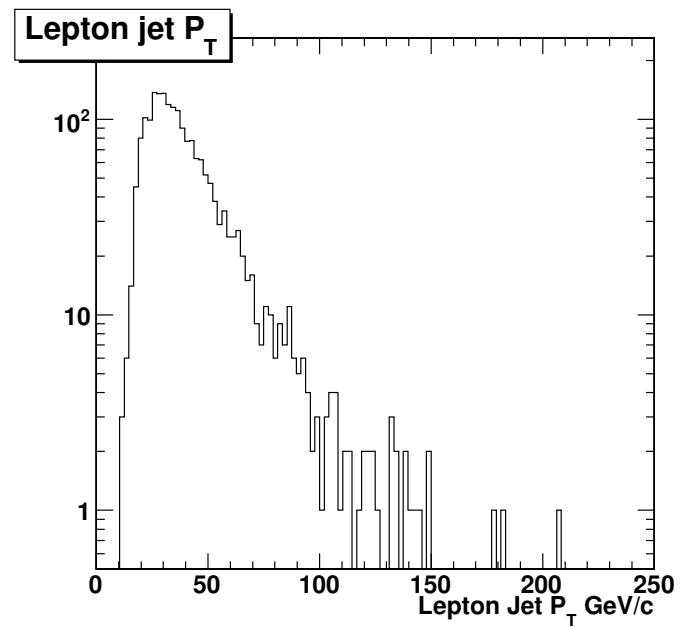


Figure 5.3: Lepton Jet P_T in Data

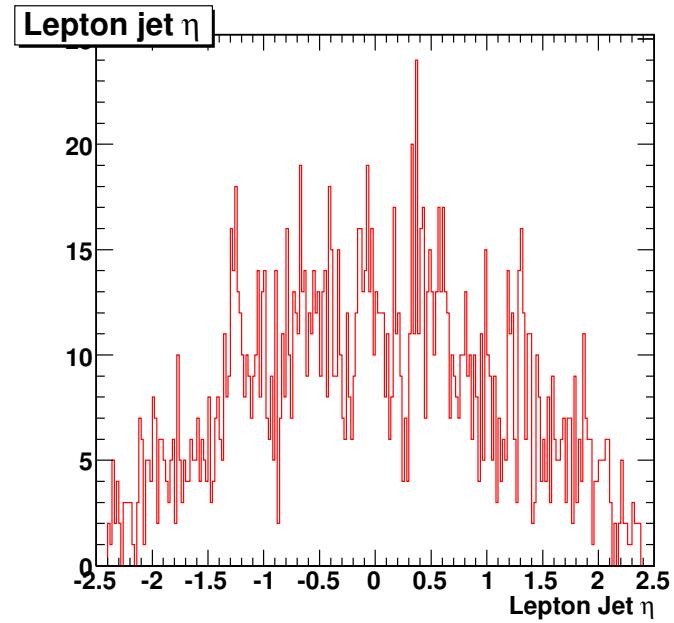


Figure 5.4: Lepton Jet η in Data. Most of the lepton jets are produced centrally

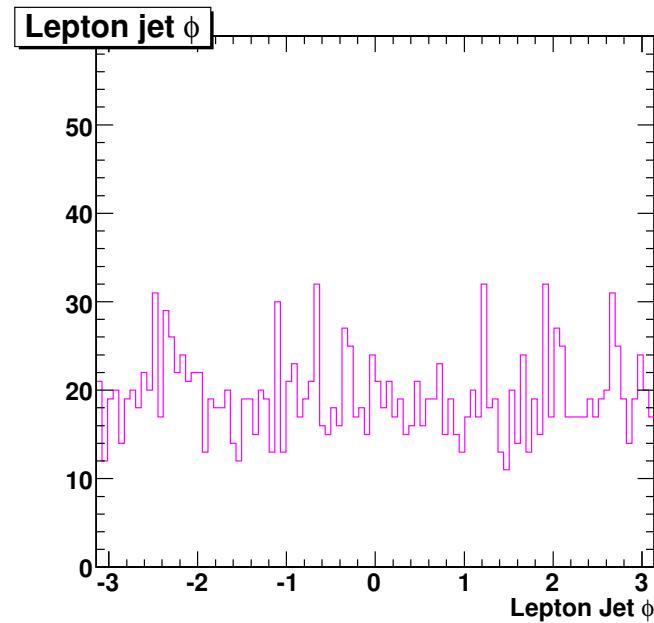


Figure 5.5: Lepton Jet ϕ in Data. The distribution is uniform, as expected

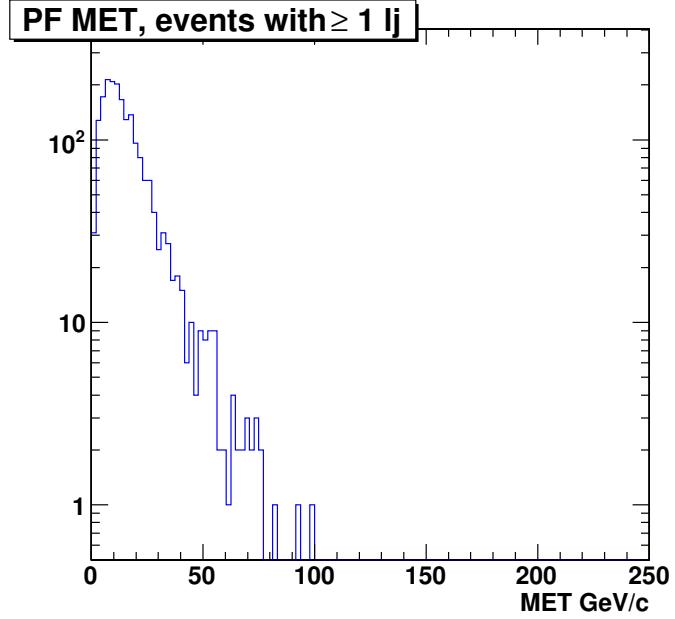


Figure 5.6: Lepton Jet Missing Transverse Energy in Data

5.2.2 Transverse Momentum, Pseudorapidity, Missing E_T and Angular Distributions

Transverse momentum, pseudorapidity and angular ϕ distributions of lepton jets are shown in Figures 5.3, 5.4 and 5.5. As Figure 5.4 shows most of the lepton jets are produced centrally. Figure 5.6 shows the E_T distribution for the 1938 selected events. These observed distributions significantly differ from the corresponding kinematical distributions of the signal model described in Section 2.8.3. For example, much higher lepton jet P_T and missing transverse energy E_T are expected from the signal model as shown in Figures 2.18 and 2.21 .

5.2.3 Lepton Jet Isolation

There are two jet isolation criteria that are applicable to lepton jets, as illustrated by Figure 5.7. First is the isolation from other objects in the event, called “outer cone” isolation, computed as the sum of hadronic or electromagnetic energy in the annulus cone of ΔR of 0.3 outside the lepton jet cone of $R = 0.1$. Because QCD jets are geometrically larger than the signal lepton jets and have a lot of hadronic and electromagnetic activity associated with them typically within the size of $R = 0.7$, applying outer-cone isolation significantly reduces this important background.

The other isolation criterion, called “inner cone” isolation, is the sum of the hadronic or electromagnetic energy inside the lepton jet. QCD events are again expected to produce large electromagnetic and hadronic activity within the $R = 0.1$ cone, while signal lepton

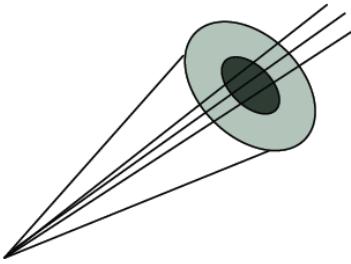


Figure 5.7: Lepton jet isolation cones. **Black:** “inner signal” cone (not applied)
Grey: “outer hollow” cone

jets are expected to be relatively clean - containing mostly leptons. However, due to the possibility of vector boson decays to pions described in Section 2.7 the “inner cone” isolation is not applied.

5.3 Muon Resonances with Lepton Jets

The performance of the lepton jet algorithm is verified by checking whether known muon resonances are reconstructed in lepton jets that contain two opposite-sign muons. Figure 5.8 shows the invariant mass spectrum of the opposite-sign muon pairs showing the various muon resonances observed by the CMS collaboration: Z^0 , Υ (1S, 2S, 3S), ψ' , J/ψ , ϕ , ω , ρ and η .

Figure 5.9 shows the same resonances reconstructed using the mass of the lepton jets that contain two opposite-sign muons. At this point only kinematical properties of the jet are used to infer about interesting physical quantities, rather than the individual properties of the constituent particles in the jet, in this case, muons. The algorithm finds all known muon resonances and shows excellent performance in muon resonance reconstruction.

5.4 Final Lepton Jet Selection

Final analysis selection cuts are listed in Table 5.2. Event signature discussed in Section 2.7.3 dictates the first five of the items, namely at least 1 outer-isolated lepton jet of size $R = .1$ and at least 2 QCD jets with transverse momentum ≥ 50 GeV/ c . Isolation size and \cancel{E}_T cuts were determined by maximizing the s/\sqrt{b} fraction before and after applying the outer-cone isolation criteria.

The Monte-Carlo signal efficiency after the final selection is summarized in Table 5.3.

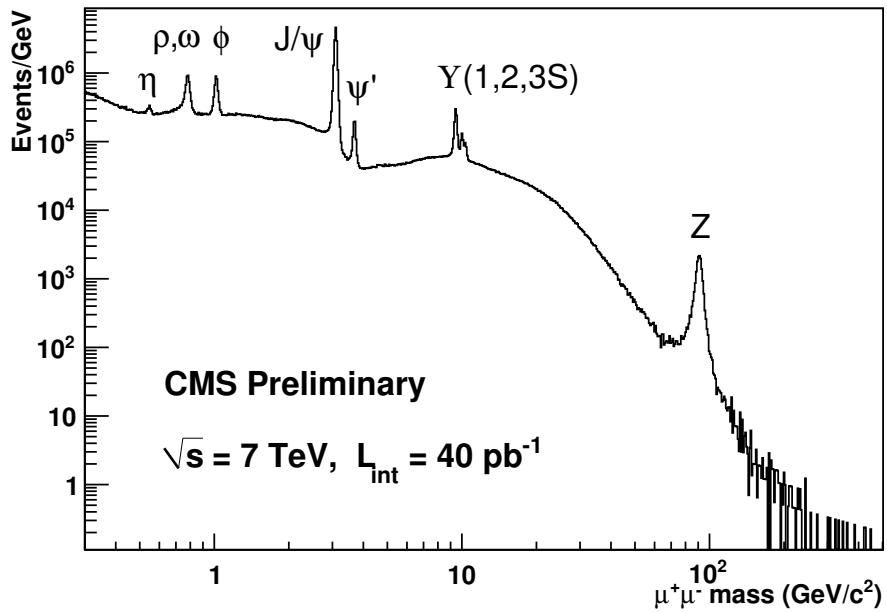


Figure 5.8: Invariant Mass Spectrum of Opposite-Sign Muon Pairs [40]

Table 5.2: Final Selection Cuts

Variable	Cut Value
Number of Lepton Jets	≥ 1
Lepton Jet Size (Rparam)	.1
Number of QCD Jets	≥ 2
QCD Jet P_T	≥ 50 GeV/c
P_T sum in hollow isolation cone	≤ 1 GeV/c
Lepton Jet Isolation size (hollow cone)	.3
\cancel{E}_T	≥ 60 GeV/c 2

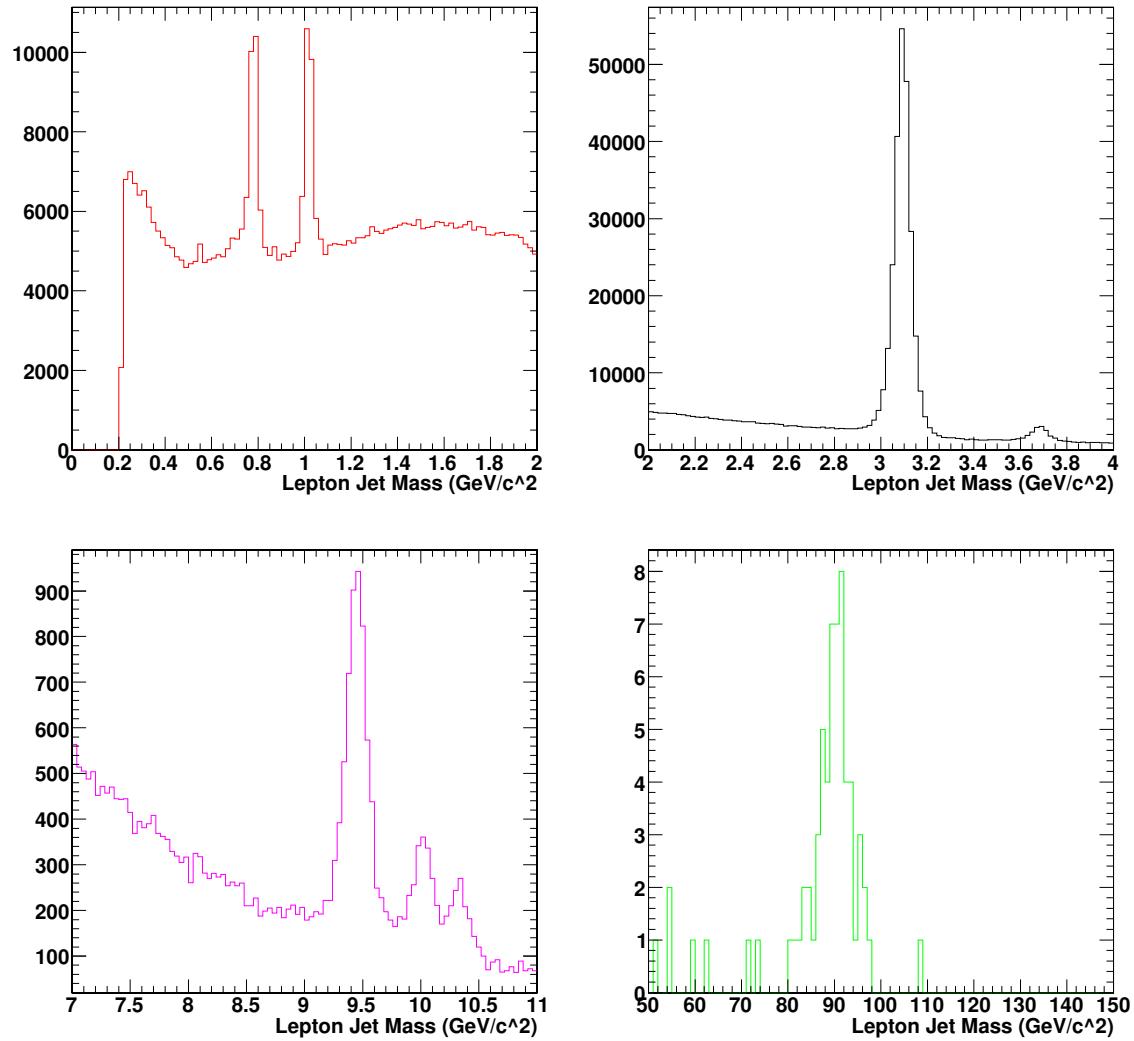


Figure 5.9: Lepton Jet Mass Spectrum: **top left:** ω , ρ , η and ϕ resonances **top right:** $J/\psi, \psi'$ resonances **bottom left:** Υ (1S, 2S, 3S) resonances **bottom right:** Z^0 peak.

Table 5.3: Monte-Carlo Signal Efficiency after Final Selection

$M_{\tilde{q}}$ (GeV/c 2)	Efficiency (%)
250	30.1
300	37.6
350	43.2
400	47.8
450	50.6
500	52.4
550	55.2
600	57.4
650	57.6
700	58.3
750	59.4

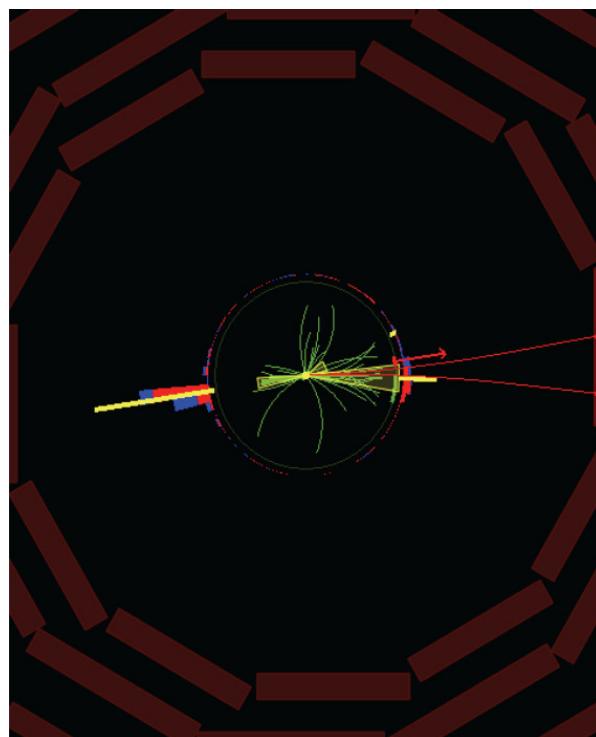


Figure 5.10: Event displays for a typical event passing the final selection. **Red** lines Muons **Green Lines** Tracks **Green Cones** QCD Jets **Blue Rectangles** HCAL Energy **Red Rectangles** ECAL Energy **Red Arrow** E_T

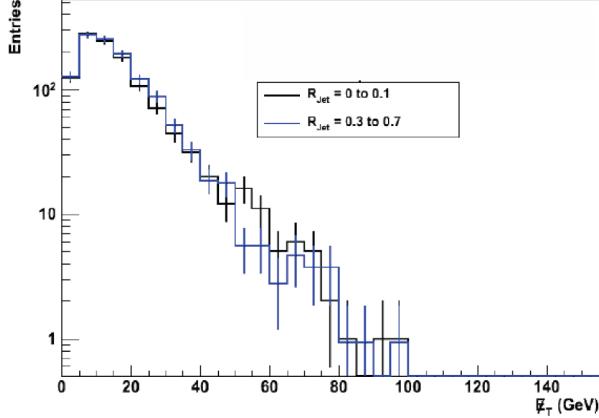


Figure 5.11: Predicted data-driven background in the signal-enriched region $R < 0.1$ using a missing transverse energy \cancel{E}_T template derived from the background-enriched ($0.3 < R < 0.7$) region.

5.5 Background Estimation

In order to properly estimate the background, we need to clearly define the signal region we are interested in. As Figure 2.15 suggests, all of the signal lepton jets are within $R < 0.1$ cone, regardless of the squark mass. This becomes our signal region of interest. Based on this assumption, we define a background-enriched region with $0.3 < R < 0.7$. There is an additional buffer region of $0.1 < R < 0.3$ that separates the signal and background-enriched regions to allow for possible signal events spilling outside of the signal-enriched region.

Despite all of our efforts in Section 5.2.3 to remove unwanted QCD background in the lepton jets, the final selection still shows clear signs of background events. In order to estimate the background contribution to the signal region, we construct a template of the missing transverse energy, \cancel{E}_T , for data events in the background region with $0.3 < R < 0.7$ and use this template to estimate the background contribution in the signal region. The value of the first bin of the template $\cancel{E}_T < 10$ GeV, where we do not expect to find the signal, is used to normalize the two distributions. As Figure 5.11 shows, a scaled \cancel{E}_T template can be successfully applied to the signal region to estimate the number of background events in it.

5.5.1 Testing the Background-only hypothesis

In order to test the background-only conclusion about the events passing the final selection, we consider all of the 21 events passing the final selection one-by-one and find a b-jet associated to all of the lepton jets. Figure 5.10 is an event display of a typical event that passes the final selection which shows a b-jet associated with the lepton jet. This reaffirms our initial finding in Section 5.5 that there are no signal events in the data.

5.6 Paradigm

PARADIGM is a multivariate decision-making framework that uses variables or features to construct classifiers of various sizes and dimensions. PARADIGM provides the researcher with easy to interpret criteria, relevant to different analysis tasks.

PARADIGM relies on several concepts that have their roots in information and decision theories. First is called *relative variable importance*, useful for tasks not associated with parameter space reduction and used in PARADIGM in the variable boosting algorithm. The other is the *global loss function*, relevant for parameter space reduction and classifier selection. PARADIGM is by design classifier-choice independent. A researcher can and should initially choose any or all of the classifiers available to her, such as neural networks, decision trees or rule ensembles, as long as a performance measure can be assigned to all or some of the classifiers. A common choice for this performance measure is the area under the receiver operating characteristic (ROC) curve [26].

5.6.1 Variable Selection

Relative Variable Importance. Relative variable importance reflects the relevance of a particular variable to a given task relative to all other variables. PARADIGM’s relative variable importance exhibits the main virtues of other relative importance algorithms [60, 76, 27], such as linear separability and order-independence, and provides additional sensitivity from the inclusion of individual variable effects in classification and the capability to identify noisy and adverse features.

PARADIGM computes input variable ranking based on how much classification power a variable contributes globally in various classifiers and in conjunction with other variables. PARADIGM achieves that by comparing classifier performance with the feature included and excluded. Stronger variables should be kept and weaker ones can be removed or have their role lessened without any considerable loss of classification power.

For the initially chosen variable set $\{V\} = \{X_1, \dots, X_N\}$, relative variable importance (RVI) is defined to be:

$$RVI(X_i) \equiv \sum_{S \subseteq V: X_i \in S} F(S) \cdot W_{X_i}(S), \quad (5.1)$$

where $F(S)$ is a general classifier performance measure¹, the sum encompasses subsets $\{S\}$ of $\{V\}$ that contain the variable X_i , and

$$W_{X_i} \equiv 1 - \frac{F(S - X_i)}{F(S)}, \quad (5.2)$$

is a weight that accounts for individual variable’s share of the classifier performance measure $F(S)$. This weight is defined as a fractional performance loss (or gain) in $F(S)$ if the variable X_i is removed from a classifier. The final RVI values are additionally normalized:

$$N \equiv \sum_{X_i} F(S) \times W_{X_i}(S) \quad (5.3)$$

¹The range of the performance measure may vary. For the area under the ROC curve the range of $F(X_i, \dots, X_j)$ is from 0.5 to 1

Table 5.4: Variables selected for PARADIGM analysis

Name	Symbol
Lepton Jet Transverse Momentum	$LJ P_T$
Lepton Jet Pseudorapidity	$LJ \eta$
Leading QCD Jet P_T	$J_1 P_T$
Sum of the leading two QCD Jet P_T	$J_1 P_T + J_2 P_T$
Invariant Mass of the Lepton Jet and Leading QCD Jet	$IM(LJ, J_1)$
Invariant Mass of the Lepton Jet and Second Leading QCD Jet	$IM(LJ, J_2)$
Missing Transverse Energy	\cancel{E}_T
Angle ϕ between Lepton Jet and \cancel{E}_T	$\Delta\phi(LJ, \cancel{E}_T)$

Table 5.5: Signal-rich parameter hyperspaces

Cut1	Cut2	Cut3	Cut4	SNR
$LJP_T < 42.9$	$J_1 P_T < 120.7$	$\cancel{E}_T > 53.7$	$IM(LJ, J_2) > 9.7$.98
$LJP_T > 74.6$	$J_1 P_T < 120.7$	$\cancel{E}_T < 33.2$	$J_1 P_T + J_2 P_T > 168.0$	1.0
$LJP_T > 42.9$	$J_1 P_T < 120.7$	$\cancel{E}_T > 33.2$	$J_1 P_T + J_2 P_T > 151.0$.97
$LJP_T > 46.0$	$J_1 P_T > 120.7$	$\cancel{E}_T > 27.0$.99
$LJP_T > 81.0$	$J_1 P_T > 120.7$	$\cancel{E}_T < 27.0$.95
$LJP_T > 175.0$	$J_1 P_T > 45.8$	$\cancel{E}_T < 23.5$	$IM(LJ, J_2) < 42$	1.0

so the RVIs sum to unity.

For the lepton jet analysis, the following variables, summarized in Table 5.4, were pre-selected for PARADIGM analysis. For this feature set, relative variable importance results are shown in Figure 5.12.

Variables involving missing transverse energy \cancel{E}_T are very strong, while others are relatively weaker. In particular, the angular difference between the lepton jet and \cancel{E}_T is shown to be the most powerful and relevant variable in building classifiers that separate signal from background. PARADIGM also shows something that should have been obvious beforehand but wasn't. The lepton jet η variable does not play a significant role in the separation of signal and background and can be easily removed from the analysis.

5.6.2 Paradigm Predictions

PARADIGM predicts the following signal-enriched parameter hyperspaces, shown in Table 5.5. The parameter hyperspaces are defined by cuts and also the signal-to-noise ratio SNR. SNR is defined to be the ratio of signal events to the sum of signal and background events. Some of the parameter hyperspaces do not have any events in the 2010 dataset but should be watched, nevertheless, with new data.

5.6.3 Global Loss Function

Another key piece of information that PARADIGM provides is the global loss function. The global loss or gloss function (GF) is an information measure specific to variable reduc-

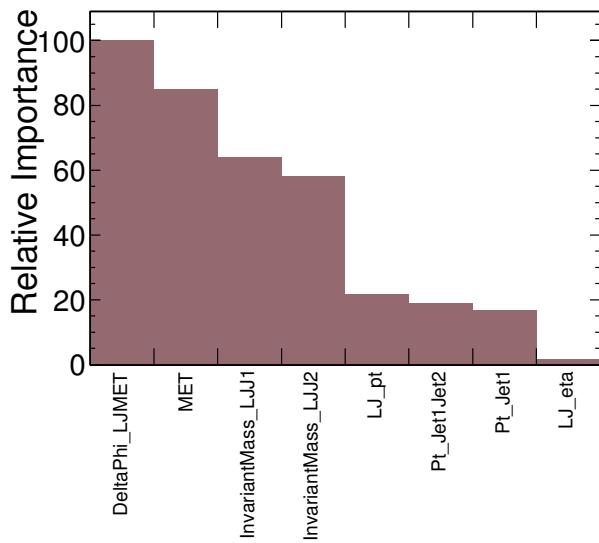


Figure 5.12: Relative Variable Importance in the Lepton Jet Analysis using PARADIGM. Lepton Jet η appears to be irrelevant for this analysis.

tion that allows a researcher to make sound decisions by incorporating variable interactions. Given a subset to be reduced, the global loss function measures the predictive power loss relative to an upper bound of achievable performance of classifiers that remain:

$$GF(S') \equiv 1 - \frac{\sum_{S \subseteq (V-S')} F(S)}{2^{|V-S'|}}, \quad (5.4)$$

where $S' \subset V$ is the subset considered for reduction and the absolute scale limit in the denominator is given by²:

$$\sum_{S \subseteq (V-S')} F(S)_{max} = 2^{|V-S'|} \quad (5.5)$$

In other words, given the initial variable set $\{V\}$ from which a variety of classifiers can be built, how much classification performance would be lost if one removes subsets $\{S'\}$ of $\{V\}$ of various sizes? The answer is precisely the gloss function. The lower its value, the lower the loss of classification power resulting from the removal of the subset $\{S'\}$.

A characteristic plot of the global loss function is shown in Fig. B.6. This information is very useful when an analysis involves many explanatory variables and some parameter reduction is desired. Lepton jet analysis is fairly low-dimensional, but analyses requiring a greater number of variables, such as the single top analysis described in Section B.7, benefits from such a reduction. An interesting point is that the straightforward exhaustive computation of both the relative variable importance and global loss functions is completely untractable for high-dimensional parameter spaces due to combinatorics which grows as 2^n where n is the number of input variables.

PARADIGM's innovative solution to this challenge comes from utilization of randomization theory. Instead of building the exhaustive set of possible combinations of classifiers, the algorithm produces a much smaller number of random seeds and builds only those classifiers that differ by one dimension (variable) from the seed. This saves a great deal of time and removes any issues of scalability. As Figure B.8 shows, there is no measurable loss from using the extremely fast random-seed approach, making it an extremely attractive solution for high-dimensional analyses. The scalability issue can be handled this way all the way up to hundreds of thousands of variables, more than sufficient for most, if not virtually all, analyses.

5.6.4 Variable Boosting

Another valuable idea explored in PARADIGM is called variable boosting. In this new classification technique relative variable importance information (from Figure 5.12 for example) is fed back into the original classification process, creating, on average, more powerful classifiers. This is achieved by placing a weight proportional to the relative importance of the variable at each juncture of the classifier-building process. For example in decision trees relative variable importance information is introduced at each decision-making junction, to influence the votes used to split the branches. As shown in Figure B.7, variable

²that follows from: $\sum_{k=0}^n \binom{n}{k} = 2$ and $F(S)_{max} = 1$. If $F(S)_{max} \neq 1$, the right hand side of Eq. 5.5 and the denominator in Eq. 5.4 instead become $2^{|V-S'|} \times F(S)_{max}$

boosting leads to stronger classifiers built from the same available parameter space. The variable boosting technique can, in principle, be applied standalone or together with other boosting and/or boot-strapping techniques, such as boosted decision-trees or random forest algorithms. This is a future line of research in this area.

5.7 Limit Setting on Cross Section times Branching Fraction

Because the Standard Model shows excellent agreement with the data, this null result can be used to set an upper limit on the signal cross-section times branching fraction in to muons. We proceed to do this using a hybrid *Bayesian* and *Frequentist CLs* technique, previously used in LEP and Tevatron experiments [99].

5.7.1 CLs Method

The CLs construction uses a negative log-likelihood ratio of Poisson probabilities as a test-statistic used to distinguish the *signal plus background* hypothesis from the *background-only*, also called null, hypothesis. The original test-statistic Q is by construction monotonically increasing for signal-like experiments and is defined to be:

$$Q \equiv \frac{L(s+b)}{L(b)} , \quad (5.6)$$

Instead of using this test-statistic directly, a logarithmic form $-2\ln Q$ is usually used for convenience. Confidence in the signal plus background hypothesis is given by the probability that the test statistic value is less than or equal to the experimentally observed value Q_{obs} :

$$CL_{s+b} = P_{s+b}(Q \leq Q_{obs}), \quad (5.7)$$

Small values of CL_{s+b} indicate unlikely compatibility of the *signal plus background* hypothesis, therefore favoring the *background-only* hypothesis. By the same token, the confidence level in the *background-only* null hypothesis is given by:

$$CL_b = P_b(Q \leq Q_{obs}), \quad (5.8)$$

Probability density functions for the test statistic in the presence and absence of signal are constructed with ensembles of toy experiments and integrated to obtain $1 - CL_b$, the criterion used to evaluate the confidence in the null hypothesis, while the approximate confidence in the *signal plus background* hypothesis is evaluated with:

$$CL_s = \frac{CL_{s+b}}{CL_b} , \quad (5.9)$$

the *modified frequentist measure* that gives origin to the name of the method. By taking the ratio of confidences in Equation 5.9, the confidence level in the signal plus background hypothesis is effectively normalized with the confidence level in the null hypothesis.

A signal hypothesis is excluded at a particular confidence level CL when:

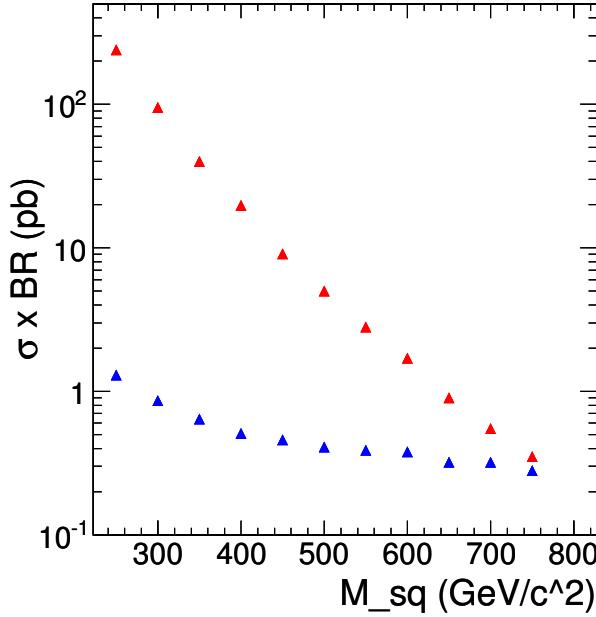


Figure 5.13: Upper limit on squark production cross-section times branching ratio. **Blue:** Observed **Red:** Theoretically predicted. Experimentally constrained cross-section limits represent a significant improvement over theoretical predictions

$$1 - CL_s \leq CL, \quad (5.10)$$

In other words, a signal hypothesis can be excluded at 95% confidence level when $CL_s \leq .05$. Each source of systematic error is approximated with a Gaussian distribution and properly taken into account during the toy experiments. The CL_s method is implemented in a hybrid calculator macro available within the statistical extension of the Root software and analysis framework, called RooStats[88], which is used to set the final upper limits in this analysis.

5.8 Results and Interpretation

In the end of the analysis, we find 21 events that contain lepton jets passing the final selection. By comparing the missing transverse energy \cancel{E}_T distribution of the events in the signal region to that predicted using a background estimate computed from a background region, we conclude that there is no visible excess of muonic lepton jet events above the Standard Model expectation. We use this fact to compute the upper limit of the cross-section times branching ratio into muons with 95% confidence level for squarks and dark photons, using the limit-setting technique described in Section 5.7.1. This technique incorporates the signal efficiency as well as systematic uncertainties outlined in Section 5.1.1.

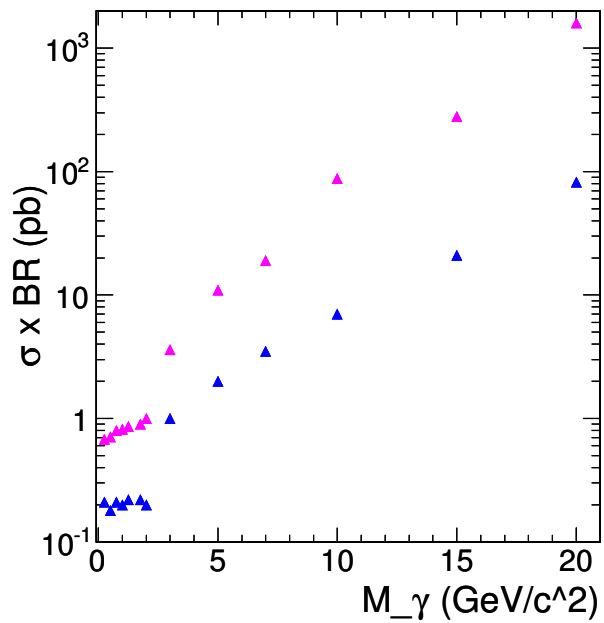


Figure 5.14: Upper limit on γ_D cross-section times branching ratio. **Blue:** $M_{\tilde{q}} = 300 \text{ GeV}/c^2$ **Magenta:** $M_{\tilde{q}} = 550 \text{ GeV}/c^2$

The upper limits set on the cross-sections are shown in Figures 5.13 and 5.14. As the Figure 5.13 indicates, it is possible to rule out the existence of squarks with the mass up to 750 GeV/c^2 in the context of the dark model we have considered in Chapter 2.

CHAPTER 6

OBSERVATION OF DOUBLE J/ψ PRODUCTION

6.1 Theoretical Motivation

As we described in Chapter 2, isolated quarks do not occur in nature. This is due to the effect known as the asymptotic freedom, discussed in detail in Section 2.1.1, when the strength of the strong interaction grows with distance. Instead they form bound states called mesons, of the form $(q\bar{q})$ or hadrons, in the form (qqq) . It is the existence of so many hadrons prior to the discovery of quarks in deep inelastic scattering experiments at SLAC that led Enrico Fermi to exclaim: "If I could remember the names of all these particles, I would be a botanist".

In reality, the complex and rich bound-state particle botany can be completely described with the 6 elementary quarks: up, down, strange, charm, bottom and top. When the electrically neutral J/ψ meson was discovered in 1974, its heavy mass of about 7000 times that of the electron, coupled with an unusually long mean lifetime of 7×10^{-21} seconds, indicated that there was another generation of elementary quarks in addition to the light up, down and strange quarks. This was an important discovery during the period of rapid changes in particle physics called the "November Revolution", that paved the way for future discoveries of the heavier bottom and top quarks, culminating in the top quark discovery in 1995.

The J/ψ meson is a charm-anticharm pair that constitutes the first excited state of char-

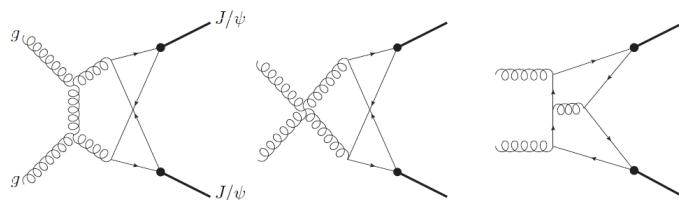


Figure 6.1: Typical Feynman diagrams for $gg \rightarrow 2J/\psi$ reaction [18]

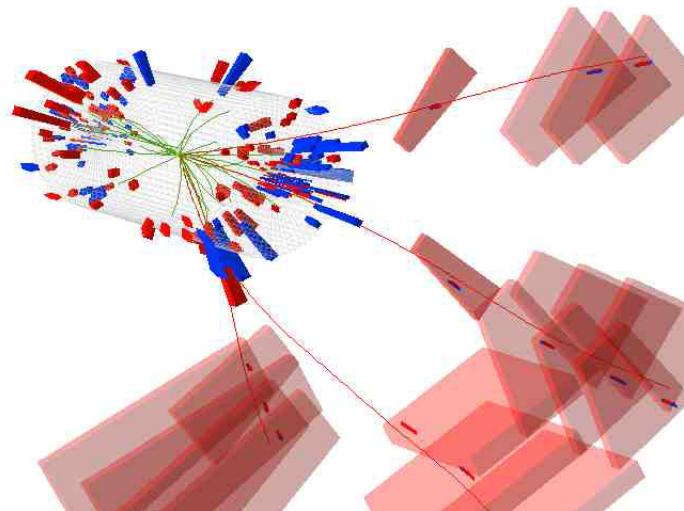
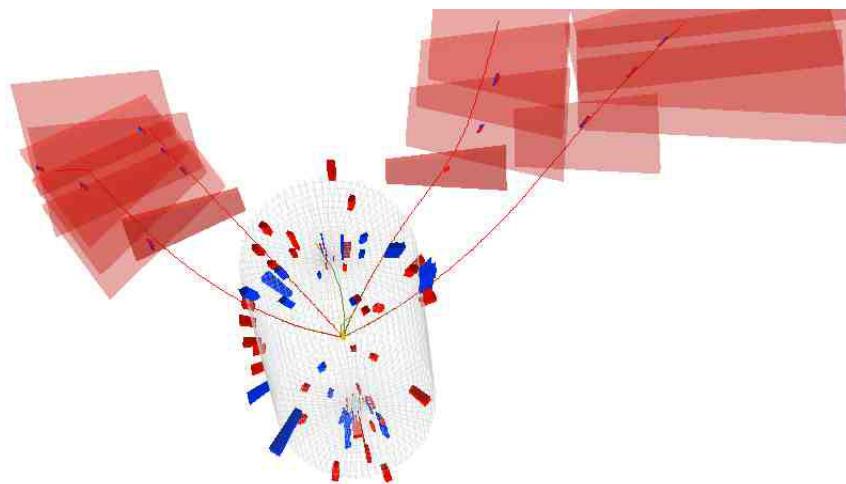


Figure 6.2: Event displays for a typical double J/ψ events in data using Fireworks.

Table 6.1: Selected Triggers

Trigger	Description
<i>Single_Mu3</i>	Single muon trigger with minimum P_T threshold of 3 GeV/c
<i>Double_Mu3</i>	Di-muon trigger with minimum P_T thresholds of 3 GeV/c

monium ($c\bar{c}$). The two names came about because it was independently discovered by two research groups, one in Stanford Linear Accelerator Laboratory and another at Brookhaven National Laboratory, who each named their newly discovered particle differently.

The hadronic decays of the J/ψ meson are heavily suppressed because of the OZI rule, proposed by Okubo, Zweig and Iizuka in the 1960s. This rule states that strong processes with final states that are reached via quark-antiquark annihilation are suppressed. This is the reason for the increased mean lifetime of the J/ψ and gives its very narrow decay width of 93.2 ± 2.1 KeV. Such scenario allows the J/ψ electromagnetic decays to compete with the hadronic decays. Due to this strong suppression the J/ψ has a significant branching ratio into leptons, in particular muons, the main concentration of our analysis.

6.2 Double J/ψ Production

In Figures 5.8 and 5.9 we showed the single J/ψ -meson production and decay in to muon pairs, as observed with the CMS detector using either simple di-muons or alternatively the muonic lepton jet algorithm. The J/ψ pair production processes obey similar selection rules to those in single J/ψ decays. That is why J/ψ and η -meson pair production is theoretically allowed, while the combination of J/ψ and η -meson or J/ψ and χ -meson is not allowed. Double J/ψ events are theoretically predicted in [68, 18, 97] but have never been previously observed in hadron collisions. They may also be a hint of other exotic states such as $\Upsilon - \Upsilon$ and others. Figure 6.1 shows the theoretically predicted $gg \rightarrow 2J/\psi$ channel, which is expected to dominate in the LHC because of enhanced color production.

The J/ψ -pair production is not directly related to the dark matter search with muonic lepton jets, presented in Chapter 5. However, it was observed during the main lepton jet analysis using the robust lepton jet algorithm, described in Chapter 5. The fact that the lepton jet reconstruction allows unambiguous di-muon pair identification of individual J/ψ 's, as opposed to having to do combinatoric selection with simple di-muon pairs, allows a very clean measurement of the double J/ψ production with little background contamination.

6.3 Event Selection

6.3.1 Dataset

We again use the full 2010 dataset described in Section 4.6 that corresponds to integrated luminosity of $36 \pm 4 \text{ pb}^{-1}$. We use trigger selection and vertexing requirements appropriate for the J/ψ -pair production analysis, described in Sections 6.3.2 and 6.3.4, while continuing to use the same muon selection as that presented in Section 4.8.2 due to its robustness.

6.3.2 Trigger Selection

The muons from the double J/ψ -pairs are not expected to have a very high P_T or boost, such as those in the dark matter lepton jet analysis, which justifies the loosening of the trigger selection. We use the triggers shown in Table 6.1 for the J/ψ -pair analysis. The first trigger is a single high-level (HLT) muon trigger with the minimum P_T threshold of 3 GeV/c, smallest P_T threshold available above 0 GeV/c. Similarly, the second trigger is a double muon trigger with the minimum P_T threshold of 3 GeV/c for each muon candidate.

6.3.3 Muon Selection

Due to its excellent performance we continue to employ the same muon selection as before, described in Section 4.8.2.

6.3.4 Vertex Requirements

Same vertex selection criteria are applied as in Section 4.6.3. Additionally, all four muons are required to come from the same vertex.

6.3.5 Isolation

“Outer”-cone isolation described in Section 5.2.3 is applied individually to each lepton jet.

6.3.6 Final Selection

The final selection consists of 2 outer-isolated lepton jets with invariant masses of the individual lepton jets in the J/ψ window of $3.0 - 3.2$ GeV/c.

6.4 Event Displays

As Figure 6.2 shows, the events passing the final selection have two well-isolated lepton jets, each containing two opposite-sign muons with masses consistent with a J/ψ ($3.0 < M_{LJ} < 3.2$). No QCD jets are observed near the J/ψ , as expected due to the “outer” cone isolation requirements described in Section 6.3.

6.5 Invariant Mass Distributions

Lepton jets are labeled 1 or 2 in the decreasing order of transverse momentum. Therefore, the lepton jet with higher transverse momentum is called LJ_1 and the other LJ_2 . Figure 6.3 shows the invariant mass of the first lepton jet LJ_1 as a function of the invariant mass of the second lepton jet LJ_2 . Double J/ψ events are clearly visible in Figure 6.3 left and Figure 6.3 right at the intersection of the horizontal and vertical J/ψ invariant mass bands.

Figure 6.4 left shows the invariant mass of the first lepton jet LJ_1 , while Figure 6.4 right shows the invariant mass of the second lepton jet LJ_2 . The J/ψ peak is clearly visible in both figures. Additionally, the ψ' peak is visible in the Figure 6.4 left. Figure 6.5 left shows

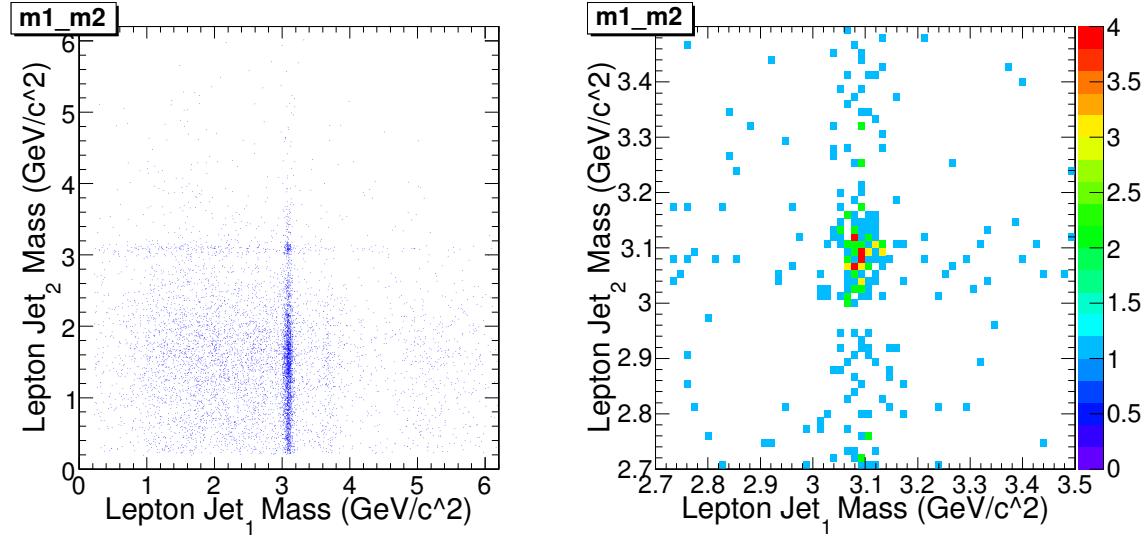


Figure 6.3: **Left:** Two-dimensional plot of the invariant mass of the first lepton jet LJ_1 as a function of the second lepton jet LJ_2 . **Right:** Close-up in three-dimensions. By definition $P_T(LJ_1) > P_T(LJ_2)$

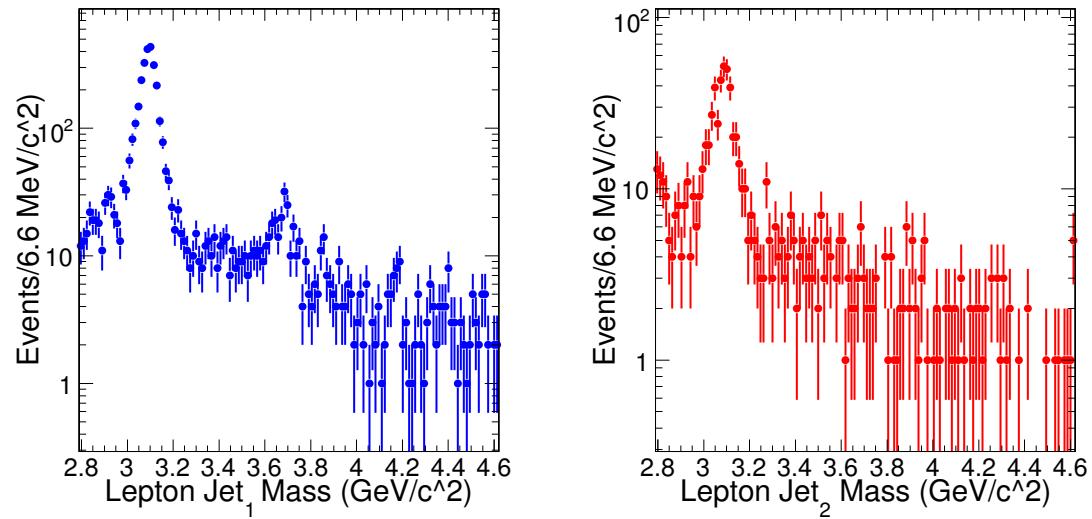


Figure 6.4: **Left:** The invariant mass of the first lepton jet LJ_1 **Right:** The invariant mass of the second lepton jet. By definition $P_T(LJ_1) > P_T(LJ_2)$. J/ψ peaks are clearly visible in both cases. Additionally the ψ' peak is visible in the left figure.

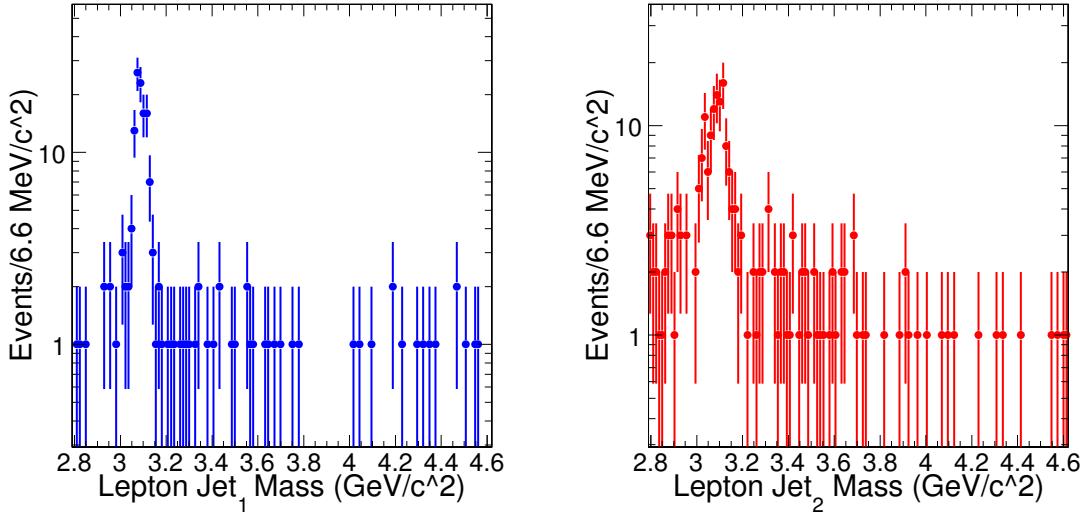


Figure 6.5: **Left:** The invariant mass of the first lepton jet LJ_1 when the mass of the second lepton jet LJ_2 is in the J/ψ mass window $3.0 < M_{LJ_2} < 3.2$. **Right:** The invariant mass of the second lepton jet LJ_2 when the mass of the first lepton jet LJ_1 is in the J/ψ mass window $3.0 < M_{LJ_1} < 3.2$. J/ψ peaks are clearly visible in both cases.

the invariant mass of the second lepton jet LJ_2 for $3.0 < M_{LJ_2} < 3.2$. Despite the limited statistics, the double J/ψ events are clearly visible in both Figures 6.5 left and 6.5 right, due to the narrow J/ψ resonance width, which additionally confirms the observation of the double J/ψ events in Figure 6.3.

6.6 Kinematical Distributions of Double J/ψ events

Figure 6.6 left shows the P_T distribution of the first lepton jet LJ_1 when the second lepton jet LJ_2 is in the J/ψ mass window $3.0 < M_{LJ_2} < 3.2$, while Figure 6.6 right shows the P_T distribution of the second lepton jet LJ_2 when the first lepton jet LJ_1 is in the J/ψ mass window $3.0 < M_{LJ_1} < 3.2$. These distributions show a falling transverse momentum spectrum. Figure 6.7 shows the pseudorapidity of the individual lepton jets in the J/ψ mass window. As it shows, the majority of the J/ψ are produced with $1 < |\eta| < 2$. Because the double J/ψ are more forward, other LHC experiments with higher pseudorapidity coverage, such as ALICE and LHCb, should observe more events of this kind.

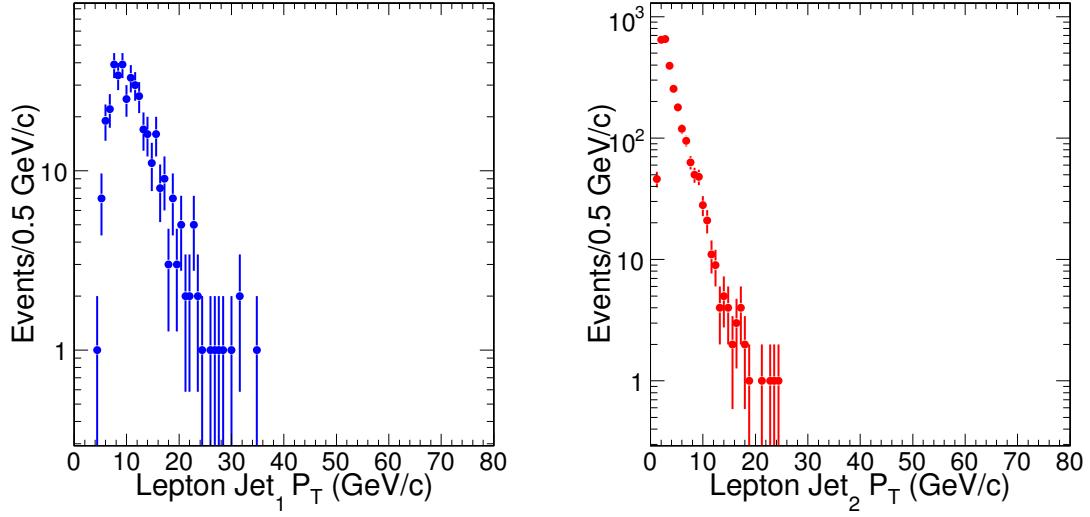


Figure 6.6: **Left:** The P_T of the first lepton jet LJ_1 when the mass of the second lepton jet LJ_2 is in the J/ψ mass window $3.0 < M_{LJ_2} < 3.2$. **Right:** The P_T of the second lepton jet LJ_2 when the mass of the first lepton jet LJ_1 is in the J/ψ mass window $3.0 < M_{LJ_1} < 3.2$.

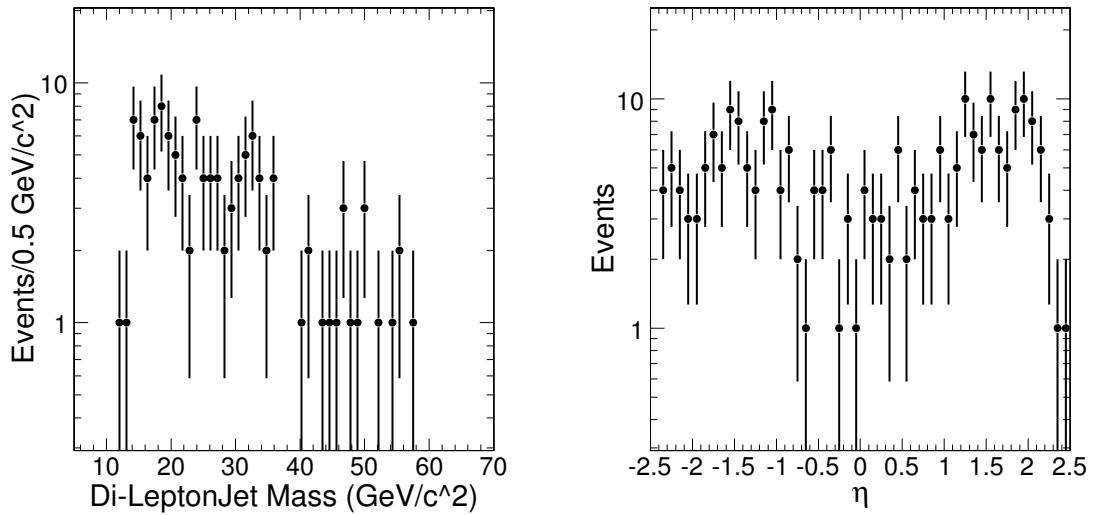


Figure 6.7: The invariant mass of the two lepton jets (**left**) and pseudorapidity of the lepton jets when both are in the mass window $3.0 < M_{LJ} < 3.2$ (**right**).

CHAPTER 7

SUMMARY AND CONCLUSIONS

We performed a dark matter lepton jet search with the Large Hadron Collider and the CMS detector. Lepton jets were observed in the analysis, however evidence for dark matter was not found. Various novel techniques were applied in the search, such as the first application of the anti- kt jet algorithm to leptons, and a high-level decision making framework called PARADIGM. Because of excellent agreement between the Standard Model background and the muonic leptons observed in the data, we established a 95% confidence level limit on the squark (\tilde{q}) and dark photon (γ_d) production cross-sections. These upper cross-section limits constitute a significant constraint on the supersymmetric dark matter model described in this dissertation and elsewhere. While performing the analysis, we observed the theoretically-predicted but previously unseen double J/ψ production decaying into muons in proton-proton collisions. This constitutes a discovery of a previously theoretized process at the next energy frontier.

APPENDIX A

EFFICIENCY

A.0.1 Trigger

Trigger efficiency is measured as a function of transverse momentum P_T with the Tag and Probe (TP) technique, as shown in Figure A.1, for data and Monte-Carlo. Tag and Probe technique is discussed in detail in reference [38].

A.0.2 Muon efficiency

Individual muon identification efficiency is shown as a function of pseudorapidity η and transverse momentum P_T in Figure A.2. The overall single muon efficiency in data is $98.9 \pm 0.1\%$. Because multiple muon identification is affected by muon proximity in the detector, it needs to be corrected by a function that depends on the ΔR between the muons, as shown in Figure A.3, resulting in the following muon reconstruction efficiency:

$$\epsilon_{reco} = \epsilon_{single_\mu}^2 \times f_{collimated_\mu}, \quad (\text{A.1})$$

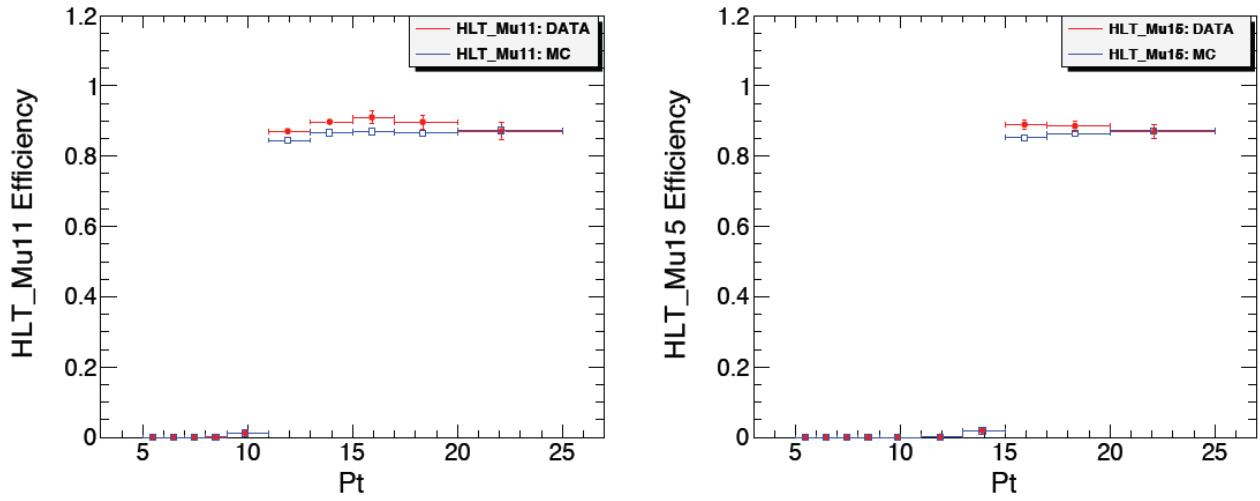


Figure A.1: Trigger efficiency for High Level Muon Triggers: **Left:** HLT_MU11. **Right:** HLT_MU15.

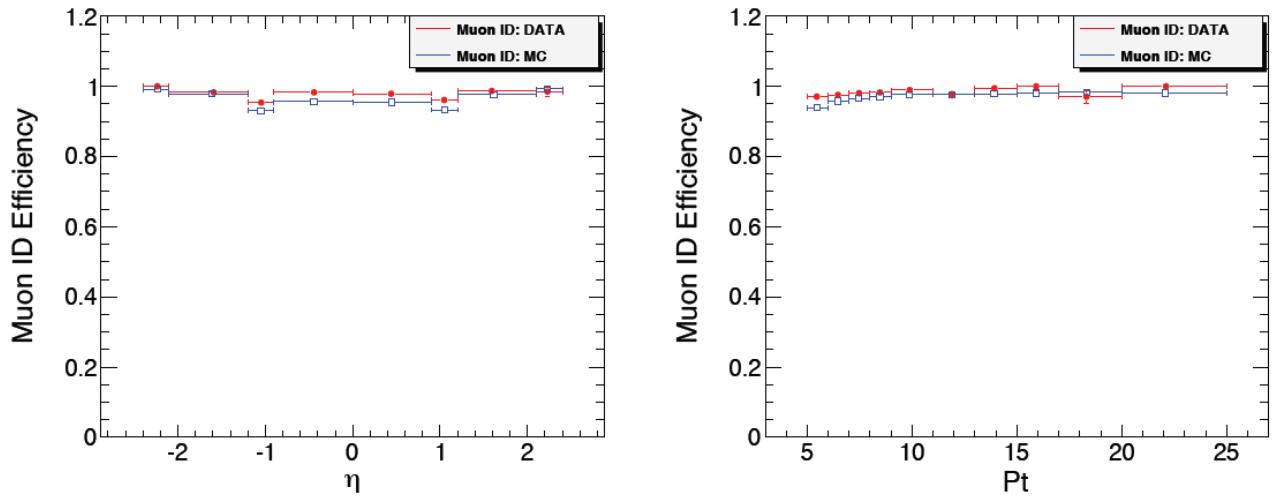


Figure A.2: Muon identification efficiency for single muons as a function of η and P_T

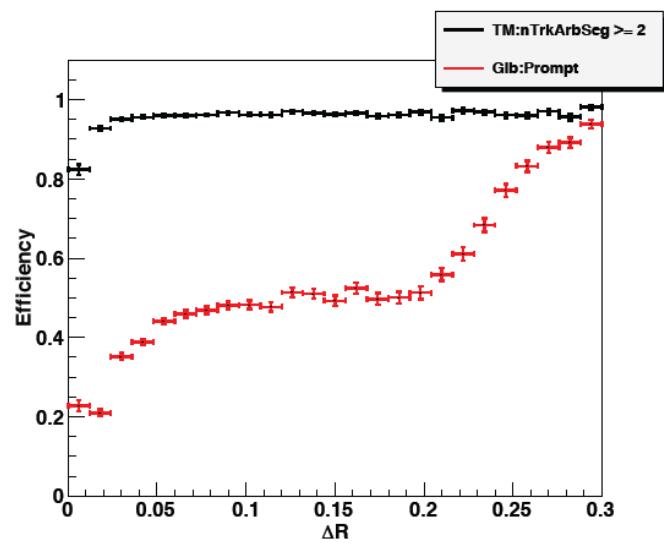


Figure A.3: Correction function for close-by muons as a function of ΔR between the two muons. **Black:** Tracker Muons with at least two arbitrated segments **Red:** Global muons.

APPENDIX B

PARADIGM

In what follows, we present the multivariate decision-making framework called PARADIGM that uses information theory measures to give the researcher easy to interpret criteria, relevant to most analyses.

B.1 Introduction

In high energy physics, variable selection and reduction are key to conducting robust multivariate analyses. Initial variable selection often results in variable sets with greater cardinality or size than the number of degrees of freedom of the underlying model. This motivates the need for variable reduction, and more fundamentally, for a consistent decision-making framework. We illustrate the common pitfalls of variable selection and reduction, such as variable interactions and variable shadowing, and show that PARADIGM gives consistent results in their presence. We discuss the application of PARADIGM to several searches for new phenomena in high energy physics and compare the performance of different measures of relative variable importance, in particular of those based on binary regression. Finally, we describe a technique called variable boosting and show how PARADIGM can be used to improve classification performance.

B.2 Concept

The *relative variable importance* concept appears in several texts [80, 113, 77]. Despite various attempts at generalization, the definitions of variable importance are for the most part model-specific [77]. Several questions are often encountered that require resolution. Are all the features necessary to achieve a particular analysis performance goal? If a variable set can be reduced, what is the optimal size of the final feature set? Is there any tolerance to noise in the variables? What is the optimal analysis strategy? Some questions are easier to answer than others, but, fundamentally, a consistent decision-making framework is highly beneficial for such circumstances.

Classification-based criteria are widely used for variable selection and related decision-making [22, 72, 60]. Other measures can be derived directly from data without the use of classification, as in [76]. The classification-independent approach is well known to be robust and less accurate than its classification-based counterpart [22]. Most traditional methods

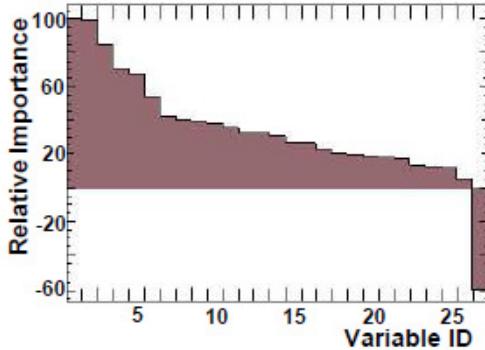


Figure B.1: A typical plot of relative variable importance. Variable ID is assigned in the decreasing order of RVI

determine relative variable importance by either averaging over all orderings of variable subsets, permuting class labels randomly, or via variable cross-correlations [60, 27, 76]. Effects of constituent variables or features on individual classifiers are often neglected for problems of high dimensionality whenever classification-based criteria are used. This can be explained by the amount of complexity growth that comes from the increase in variable ordering and the resulting difficulty in classifier micro-management. To make matters worse, some of the most popular classifiers, such as decision trees, have greedy algorithm structure [98] that causes significant inaccuracy for the majority of the variable importance measures developed for individual classifiers. It also limits the sensitivity needed to identify and remove problematic variables.

B.3 Outline

A decision-making framework is proposed that has the main virtues of other relative importance algorithms such as linear separability and order-independence, while at the same time maintains feature space reduction and identification of noisy and adverse features a high priority. We previously applied this framework to a search for single top quarks[1], vector boson fusion [7] and lepton jets in Chapter 5. In Sections B.4 and B.5 we explain and formulate the framework, followed by the description and analysis of various examples in Sections B.6 and B.7. We discuss the intrinsic limitations of relative variable importance criteria for parameter space reduction and advocate the use of the global loss functions for this purpose in Section B.7. A summary follows in Section B.9.

B.4 Motivation

PARADIGM is by design classifier-choice independent. A researcher can and should initially choose any or all of the classifiers available to her, such as neural networks, decision trees or rule ensembles, as long as a performance measure can be assigned to all or some of the classifiers.

Any of the popular machine-learning based classifiers can be used in the following arguments. One often faces the task of separating a desired function Ψ_S from another function Ψ_B , typically referred to as signal and background. Both Ψ_S and Ψ_B may depend on explanatory variables (X_1, \dots, X_N) , also known as features, that are subject to measurement. A variety of classifiers can be built out of (X_1, \dots, X_N) in order to maximize the separation between $\Psi_S(X_1, \dots, X_N)$ and $\Psi_B(X_1, \dots, X_N)$. Some techniques combine sets of classifiers to achieve as optimal classification performance on unseen data as possible [27, 28]. Others implement adaptive learning to create classifiers that are more powerful than their predecessors [59]. A classical measure of individual classifier performance is the “Receiver Operator Characteristic” (ROC) that shows the dependence of correctly classified positive cases to incorrectly classified negative cases [26]. The greater the separation of $\Psi_S(X_1, \dots, X_N)$ from $\Psi_B(X_1, \dots, X_N)$, the closer the ROC curve is to the 1 – 1 axis. Therefore, the *Area Under the ROC Curve* is a commonly chosen quantitative measure of the classifier’s predictive power [26, 74]. Since the choice of such performance measure varies across different fields, we proceed with the convention of a more general performance measure $F(S)$.

B.5 Definitions

Variable importance definitions are given in Section 5.6 and therefore, not repeated in detail here. Instead we focus on motivation behind the equations. We reproduce equation 5.1 for relative variable importance (RVI) here:

$$RVI(X_i) \equiv \sum_{S \subseteq V: X_i \in S} F(S) \cdot W_{X_i}(S), \quad (\text{B.1})$$

$F(S)$ is a general classifier performance measure, the sum runs over the subsets $\{S\}$ of $\{V\}$ that contain the variable X_i , and

$$W_{X_i} \equiv 1 - \frac{F(S - X_i)}{F(S)}, \quad (\text{B.2})$$

is a weight that accounts for individual variable’s share of the classifier performance measure $F(S)$.

A technique with a similar goal in mind is found in the context of rule-based regression, a framework that condenses classifiers into sets of (if, then, else) rules that can be used as ensemble predictors [60]. Notably, even if all the rules are poor but marginally better than random guessing, in a large ensemble they become excellent predictors [27, 28, 60]. A binary rule-based regression tool called RULEFIT [60] is selected for a comparative study. Analogous to how relative variable importance is defined in RULEFIT [60], Eqn. (B.1) without the weight, relates relative variable importance to the performance of classifiers in which a variable participates. A drawback is that the classifier’s measure of performance

Table B.1: Hypothetical Example I Values

S	$F(S)$
A, B	.88
A	.57
B	.63
\emptyset	.50

Table B.2: Hypothetical Example I Calculation

X_i	$RV\bar{I}(X_i)$	$RV\bar{I}(X_i)$
A	$\frac{F(A) - F(0) + F(A, B) - F(B)}{F(A) - F(0) + F(A, B) - F(A) + F(B) - F(0) + F(A, B) - F(B)}$	$\frac{.57 - .5 + .88 - .63}{2 * .88 - 2 * .5} = .42$
B	$\frac{F(B) - F(0) + F(A, B) - F(A)}{F(A) - F(0) + F(A, B) - F(A) + F(B) - F(0) + F(A, B) - F(B)}$	$\frac{.63 - .5 + .88 - .57}{2 * .88 - 2 * .5} = .58$

is evenly distributed among all the variables in set $\{S\}$. To address this, a weight W_{X_i} that accounts for individual variable's share of $F(S)$ is added. The weight is defined as the fractional loss (or gain) in $F(S)$, if a variable is removed from a classifier. The addition of the weight makes *relative variable importance* more sensitive to effects of individual variables within classifiers and allows for identification of variables that may have a negative effect on classification.

B.6 Example 1

Consider the following hypothetical example with just 2 explanatory variables, A and B, to illustrate the RVI criterion. Table B.1 shows the $F(S)$ for the four possible subsets, and Table B.2 shows the calculation of RVI for A and B. Another hypothetical example shows how the *relative variable importance* criterion can be used to identify variables that may have an adverse effect on classification. Table B.3 shows the $F(S)$ for the eight possible subsets of 3 explanatory variables A, B and C. Table B.4 shows the calculation of RVI for A, B and C. In this example variable C shows adversity and should be considered for removal.

Table B.3: Hypothetical Example II Values

S	$F(S)$
A, B, C	.79
A, B	.84
A, C	.57
B, C	.64
A	.59
B	.68
C	.55
\emptyset	.50

Table B.4: Hypothetical Example II Calculation

X_i	$RI(X_i)$
A	$\frac{F(A)-F(0)+F(A,B)-F(B)+F(A,C)-F(C)+F(A,B,C)-F(B,C)}{3*F(A,B,C)+F(A,B)-F(A)+F(A,C)-F(B)+F(B,C)-F(C)+F(A,B)-F(B)-3*F(0)}$
B	$\frac{F(B)-F(0)+F(A,B)-F(A)+F(B,C)-F(C)+F(A,B,C)-F(A,C)}{3*F(A,B,C)+F(A,B)-F(A)+F(A,C)-F(B)+F(B,C)-F(C)+F(A,B)-F(B)-3*F(0)}$
C	$\frac{F(C)-F(0)+F(B,C)-F(B)+F(A,C)-F(A)+F(A,B,C)-F(A,B)}{3*F(A,B,C)+F(A,B)-F(A)+F(A,C)-F(B)+F(B,C)-F(C)+F(A,B)-F(B)-3*F(0)}$

 Table B.5: Single top quark variable set $\{V\}$

Variable	Motivation
$\cos(\text{jet1,lepton})_{\text{lab}}$	Angular
$\cos(\text{lepton}, Q(\text{lepton}) \times Z)$	Angular
$\Delta R(\text{jet1,jet2})$	Angular
$\cos(\text{jet1, alljets})_{\text{alljets}}$	Angular
$\cos(\text{jet2, lepton})_{\text{lab}}$	Angular
$\cos(\text{notbest_jet, alljets})_{\text{alljets}}$	Angular
BestTopMass	Event Kinematics
$H_t(\text{alljets} - \text{bestjet})$	Event Kinematics
$H_t(\text{jet1,jet2}, W)$	Event Kinematics
$H(\text{alljets}, W)$	Event Kinematics
$IM(\text{alljets})$	Event Kinematics
$IM(\text{alljets} - \text{bestjet})$	Event Kinematics
$IM(\text{bestjet}, W)$	Event Kinematics
$Q(\text{lepton}) \times \eta$	Event Kinematics
$H_t(\text{alljets})$	Event Kinematics
$H_t(\text{jet1, jet2})$	Event Kinematics
$H(\text{alljets})$	Event Kinematics
$H(\text{jet1,jet2})$	Event Kinematics
$IM(\text{alljets}, W)$	Event Kinematics
$IM(\text{jet1,jet2})$	Event Kinematics
$M_t(\text{jet1, jet2})$	Event Kinematics
$M_t(W)$	Object Kinematics
$P_t(\text{jet2})$	Object Kinematics
$P_t(\text{alljets} - \text{bestjet})$	Object Kinematics
$P_t(\text{jet1})_{\text{notbest}}$	Object Kinematics
$P_t(E_t)$	Object Kinematics
$P_t(\text{jet1} + \text{jet2})$	Object Kinematics

Table B.6: Degrees of Freedom in the Analysis

Object	Quantity	Type	Degrees
Jet	2,3	4-vector	8,12
Lepton	1	4-vector	3
Missing E_t	1	2-vector	2

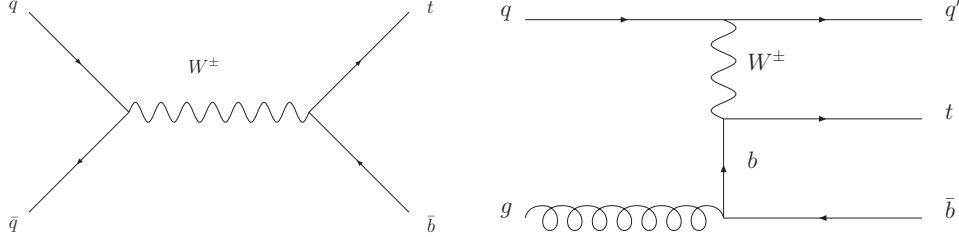


Figure B.2: Feynman diagrams for electroweak production of single top quarks for the s - and t - channels respectively.

B.7 Example 2

The framework is illustrated in a search for single top quarks[61]. Single top quarks have been theoretically predicted, and recently discovered [2].

In the Standard Model (SM) single top quarks can be produced in proton antiproton collisions via the following two processes shown in Fig. B.2: $p\bar{p} \rightarrow t\bar{b} + X$ (s-channel) and $p\bar{p} \rightarrow t q\bar{b} + X$ (t-channel) [116, 119]. The single top quark can decay into a bottom quark (b) and a W boson, where the W subsequently decays into a lepton and a neutrino. Bottom quarks in both diagrams can be experimentally detected as jets, and the neutrino can be identified by high missing transverse energy. Additionally in the t -channel a light quark is present[1]. Important backgrounds for these final states are W +jets, the top quark pair ($t\bar{t}$), and QCD multijet production.

The germane issue for single top quark discovery is successful separation of single top quark signal $\Psi_S(X_1, \dots, X_N)$ from the corresponding backgrounds $\Psi_B(X_1, \dots, X_N)$. From Quantum Chromodynamics (QCD) considerations, branching ratios and kinematics, an initial variable set $\{V\} = (X_1, \dots, X_N)$ can be pre-selected for analysis [1], summarized in table B.5. The varying motivation above leads to a feature set with a significantly greater cardinality than the underlying degrees of freedom exhibited by the problem, listed in Table B.6 [1]. Complexity of analysis grows rapidly with the size of $\{V\}$, and its reduction without a significant loss of classification accuracy is the cornerstone task for efficient analysis.

B.7.1 Data Description

Monte Carlo simulated data sets for both production and decay of single top quarks and the corresponding backgrounds $\Psi_B(X_1, \dots, X_N)$ are used [24, 83, 108]. The signal $\Psi_S(X_1, \dots, X_N)$ events with two distinct degrees of freedom, 13 and 17, determined by the

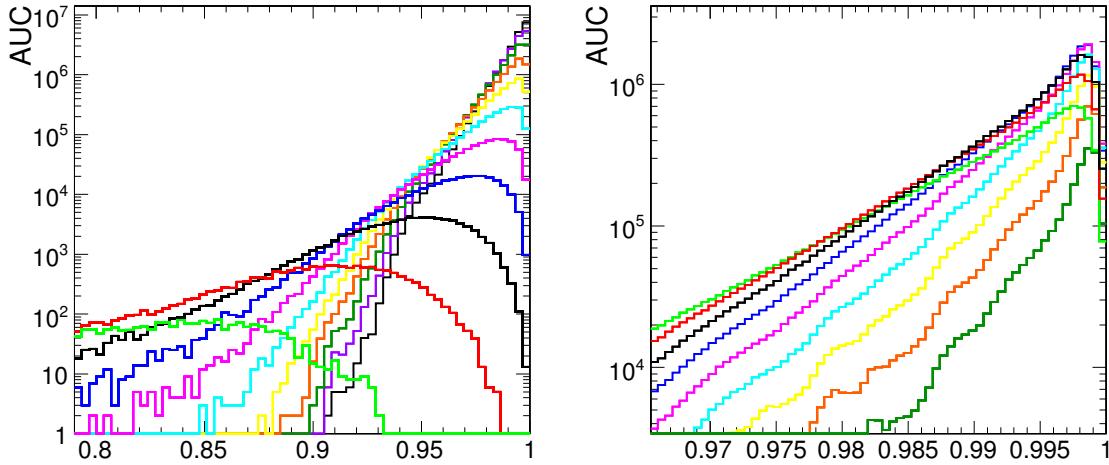


Figure B.3: ROC spectrum for sequential increase in cardinality of $\{S\}$. a) $2 \leq |S| \leq 12$ b) $10 \leq |S| \leq 18$

number of jets in the event, are intermixed. For the following example we choose the popular and intuitive classifier - a decision tree and use a well-known statistical package called C4.5 to build a full combinatoric set of classifiers from the data [98]. The choice of the classifier is entirely arbitrary, as long as a classifier performance measure $F(S)$ can be established. The framework allows for easy comparison of various popular classifier choices such as decision trees, neural networks, rule-based regression, and others, using the *gloss function* criteria.

B.7.2 Results and Analysis

$F(S)$ Incremental Addition and Cardinality. For highly dimensional classification problems the following procedure, based on incremental addition of variables, can be performed to help identify the optimal variable set size for reduction. Classifiers built from subsets of $\{V\}$ can be grouped by cardinality and plotted sequentially. This construction conveys the average improvement in the $F(S)$ spectrum from the increase in feature size. There is a sharp edge at 1, which is the maximum $F(S)$ value for an individual classifier. The amount of positive slope reflects the proportion of discriminants better than the average, to those worse than the average. When there is no additional improvement from the addition of variables, the slope saturates, and if such critical point exists, it can be used to identify the optimal feature set size.

We apply the procedure based on incremental addition of variables to help identify the optimal variable set size for reduction. Classifiers built from subsets of $\{V\}$ are grouped by cardinality and plotted sequentially as in Fig. B.3. The amount of positive slope reflects the average improvement of the classifiers. As Fig. B.3 b shows, there is an increasing trend until 15, the average between the two known degrees of freedom of $\Psi_S(X_1, \dots, X_N)$, and thereafter, no increase is evident. This implies a logical size for the optimal variable set for

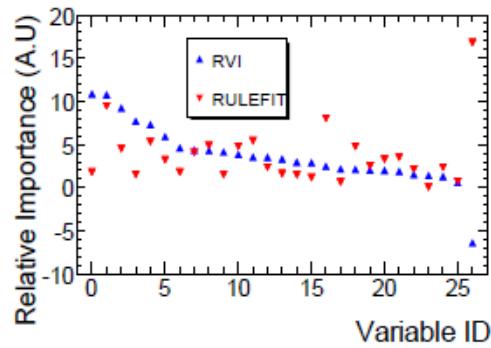


Figure B.4: Comparison between variable importance measures provided by PARADIGM and RULEFIT on an absolute scale

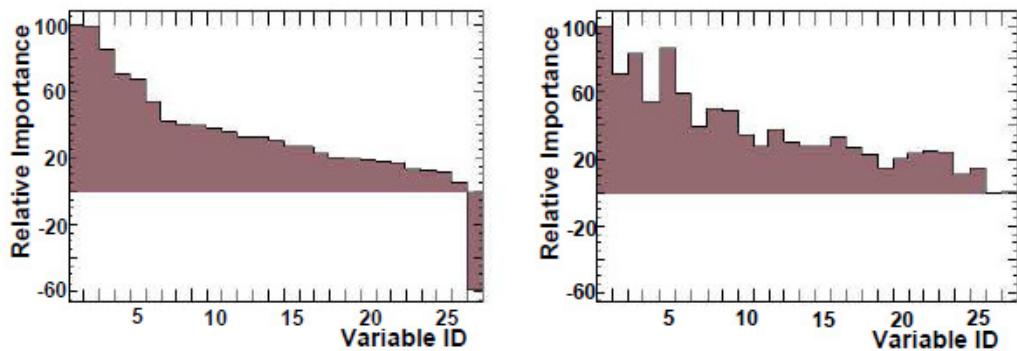


Figure B.5: a. relative variable importance for $\{V\}$ b. relative variable importance results with $IM(alljets - bestjet)$ removed

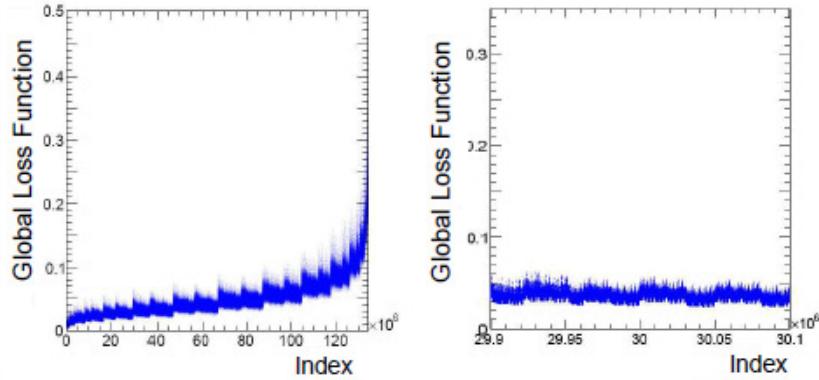


Figure B.6: *Gloss function* results for a) $1 \leq |S| \leq 26$ b) $14 \leq |S| \leq 16$

this example, where variable addition does not improve the overall predictive power of the set. From the above considerations a window of cardinality $14 \leq |S| \leq 16$ can be preselected for step 3 of the gloss function analysis.

Comparison between RVI and RULEFIT. Analogous to how relative variable importance is defined in RULEFIT, the RVI is directly tied to the performance of classifiers containing the variable in question. However, in contrast to RULEFIT, the weight W_{X_i} allows the RVI to be more sensitive to the effects of individual variables during classification and permits the identification of features that have a negative effect on classification. On an absolute scale, PARADIGMS RVI exhibits both similarities and differences to RULEFIT's variable importance measure (Fig. B.4). Overall, the two criteria appear consistent with one another with significant outliers. The notable exception is a variable on the extreme right of Fig. B.4, unambiguously identified by PARADIGM to be adverse to classification. As is true for all absolute value criteria, the sign of variable importance is unattainable with RULEFIT. In order to gain insight into relative variable importance, it is worthwhile to consider what makes one relative variable importance measure preferable to another. Comparisons, such as that in Figure B.4, can be used to infer the differences and similarities between the two candidate measures, but nothing more.

Variable Boosting Technique. The optimal way to address this quantitatively rather than qualitatively is to consider the amount of useful information provided by the two criteria and show how that information can be used to achieve analysis goals, for instance, to maximize the classification power of a given set of variables. We proceed to feed back the relative variable importance information into the classification process, a procedure called RVBoost, that stands for relative variable importance boost or amplification. This approach requires creation of new classifiers that use relative variable importance in direct decision making during the classification process. For example in decision trees relative variable importance information is introduced at each decision-making junction, to

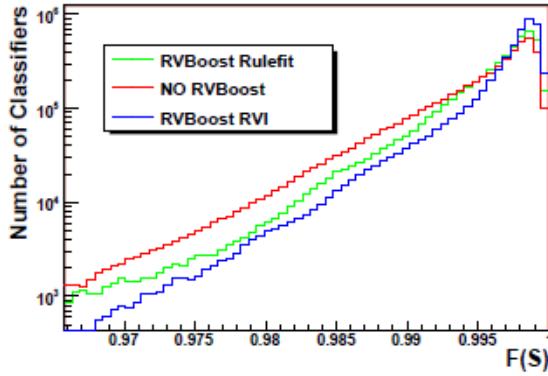


Figure B.7: Comparison of classifier performance using RVBoost with PARADIGM and RULEFIT to that without RVBoost, for a fixed number of classifiers.

influence the votes used to split the branches. The same boosting procedure can be easily applied to all known classifiers.

If the relative variable importance measure in question contains information that may be used to further the classification goals of the analysis, it is clearly beneficial. That is the case for both relative variable importance and RULEFIT but in differing amounts (Fig. B.7). A fixed number of classifiers in this example is RVBoosted with the RVI and the RULEFIT measures. The performance of the new classifiers is compared to that of the original classifiers. As Figure B.7 shows, PARADIGM's RVI criteria outperforms RULEFIT's variable importance and both outperform the original classifiers when relative variable importance boosting is considered.

Subtlety in Variable Reduction. To illustrate a common caveat in multivariate analysis involving classification, the last variable in Figure B.5 a. is removed and the figure itself redrawn in B.5 b. A variable that was previously marginally useful became adverse to classification and the order and magnitudes of relative variable importance have changed. This behavior can be explained by the presence of multiple interactions among variables, a common behavior during classification. For instance, in decision trees this can lead to a phenomenon known as variable shadowing, when a presence of one strongly interacting variable partially or entirely shadows its interacting partner, making it appear irrelevant.

The fact that interacting variables influence the performance of their partners in both directions, can be used to explain the common occurrence illustrated in Fig. B.5. A classical formulation of variable interactions on the basis of risk analysis is found in Ref. [48]. There are several methods to quantify the strength of variable interactions. For example, RULEFIT uses the concept of partial dependencies [60]. As Figure B.5 shows, the variable importance landscape becomes distorted by the removal of interacting variables. Presence of variable

Table B.7: Optimal Set $\{S'\}$ for Reduction Using Cardinality Window $14 \leq |S| \leq 16$

Variable
$\cos(\text{jet1, alljets})_{\text{alljets}}$
$\cos(\text{notbest_jet, alljets})_{\text{alljets}}$
$H_t(\text{alljets} - \text{bestjet})$
$\text{IM}(\text{alljets} - \text{bestjet})$
$Q(\text{lepton}) \times \eta$
$H_t(\text{alljets})$
$H(\text{alljets})$
$\text{IM}(\text{jet1, jet2})$
$M_t(\text{jet1, jet2})$
$P_t(\text{jet1})_{\text{notbest}}$

interactions significantly reduces the effectiveness of criteria that do not directly take them into account, such as RULEFITS variable importance or the RVI, when it comes to parameter space reduction. Ignoring this subtlety is a common mistake researchers make. One instead should choose measures for parameter space reduction that implicitly incorporate variable interactions, such as the *global loss function*, described in the next section.

Gloss Function Analysis. In addition to the *relative variable importance* criteria, in Section 5.6.3 we introduced a global measure of the global loss (*gloss*) function (GF) that measures the predictive power loss relative to an upper bound of achievable performance of classifiers that remain in the chosen set. Both the gloss function and the RVI are *wrapper* type algorithms that rely entirely on classification to produce criteria for parameter space reduction [22, 73]. The other type, known as *filter* algorithms, rely only on the raw data for this task [75, 22]. A review and comparison of the two methods is given in [66].

The *gloss function* for the full spectrum and the critical cardinality region, $14 < |S| < 16$, is shown in Fig. B.6. The $\{S'\}$ subsets are arranged in the order of increasing cardinality, and within regions of equinumerality, they are arranged by their binary index, as defined in Section B.8.1. Therefore, on the horizontal axis of Fig. B.6 a. the leftmost set is the empty set $\emptyset \equiv \{000\dots000\}$. There, the gloss function is close to 0 yet finite. The rightmost set is the full set $\{V\} \equiv \{111\dots111\}$, with the minimum gloss function value of 0.5. By convention, $\{S'\}$ is the subset to be reduced.

If a gloss function curve is considered, it resembles an inverted ROC curve. The closer the area under the gloss function is to 0, or conversely the upper area to 1, the more effective the overall feature selection. Therefore, the area under the gloss function curve is an additional powerful criterion to select and compare initial feature selections.

For the critical region of cardinality selected above, the local gloss function minimum is found to be 0.0167, which corresponds to the optimal subset for reduction $\{S'\}$, summarized in Table B.7. The reduction of this subset maximizes the discriminatory performance of classifiers that remain.

As a test case, we selected different cardinality windows for the *gloss function* analysis. The $\{S'\}$ s found were again subsets of the $\{S'\}$ from the critical region above. See, for

Table B.8: Optimal Set $\{S'\}$ for Reduction Using Cardinality Window $23 \leq |S| \leq 24$

Variable
IM(alljets - bestjet)
IM(jet1,jet2)
$P_t(\text{jet1})_{\text{notbest}}$

instance, Table B.8, which shows the $\{S'\}$ for the cardinality window $23 \leq |S| \leq 24$. The fact that the optimal reduction subsets for increasing $|S'|$ are supersets of each other gives additional merit to the gloss function method's consistency and implies that the gloss function is monotonic. Also noteworthy is the fact that the feature identified as adverse by the RVI method³ is likewise identified for reduction by the *gloss function* approach. That is the case not only in the window of cardinality around the critical region, but for very small $\{S'\}$ as well, which shows that although motivated differently, the RVI and the gloss function methods show consistency in identification of the noisy and adverse variables for reduction.

B.8 Decision Making Framework

By combining the gloss function and relative variable importance, a powerful decision making framework can be made. *Relative variable importance* allows optimization of performance of individual subsets, while the *global loss function* provides the capability to identify noisy and adverse variables during parameter space reduction. It is important to note that if the analysis requires or desires a reduction in parameter space, as we discuss in section B.7, only the gloss function criteria is fully advocated for this task, while the relative variable importance criteria is a locally informative measure used as a compliment to the gloss function approach. The structure of the framework is as follows:

- As described in Section B.4, suitable classifiers are selected
- An optimal classifier is chosen with the global loss function
- If parameter reduction is desired, the S' subset with the minimum gloss function value is chosen for reduction and its compliment is kept for further analysis
- Relative variable importance is used to cross-check that all adverse variables are included in the S' subset to be reduced
- Once the final variable set is selected, the RVBOOST procedure described in Section 5.6.4 applied to maximize the performance of the classifiers built from this set.

B.8.1 Performance Optimization

To process efficiently the $F(S)$ of Eqs. 5.1 and 5.4, PARADIGM employs the following data structure. A binary identifier is assigned to each subset $\{S'\}$, such that each digit

³also ranked as the most important feature by RULEFIT

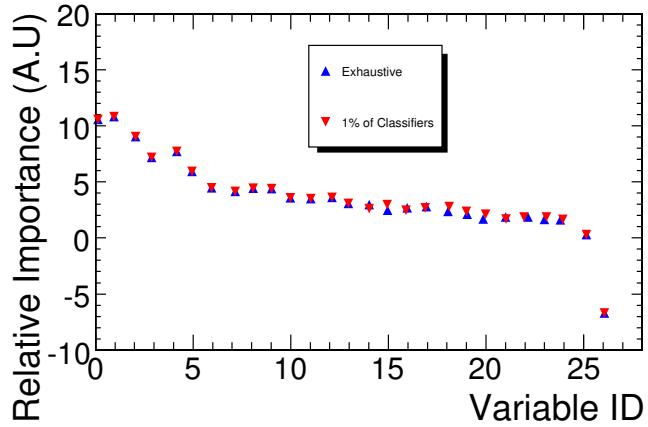


Figure B.8: Relative variable importances in the random-seed approach using 1% of the classifiers.

signifies the presence or absence of a corresponding variable in $\{S'\}$. All the $F(S)$ values are stored in an array, indexed with this identifier. Bit manipulation provides fast access to all the $F(S)$ derived from a particular identifier that differ in one or more bits, as required by the RVI analysis, and the $F(S)$ which lack a particular sequence of bits, as required by the gloss function analysis.

As described in Section 5.6, PARADIGM employs randomization theory to drastically reduce the computing power needed to compute the relevant criteria such as the RVI and the global loss function. The idea is to choose a fixed number of random seeds represented by bit sequences in a signal binary number, and only build and compute the difference in classifier performance of classifier that differ by one variable (bit) from the seed. This is then used to compute the RVI and the global loss function. Such elegant solution allows a multifold and extremely competitive reduction over an exhaustive approach, which albeit correct, may be intractable for a large number of variables, usually above 30. We used the single-top example with 27 features and as shown in Figure B.8, the results obtained with the random-seed approach provide the same information as the fully exhaustive approach.

B.9 Discussion and Summary

It is worthwhile to note that minimization of the global loss function is not equivalent to maximization of the classifier performance measure $F(S)$, i.e. finding the highest performing classifier and its constituent variables (Fig. B.9). Some researchers attempt a quick search

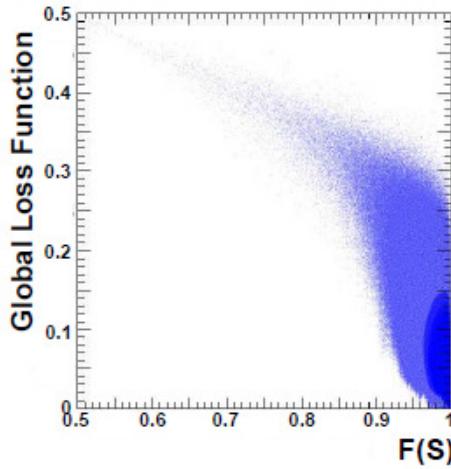


Figure B.9: The non-linear relationship between the global loss function and the classifier performance measure $F(S)$

for high performing classifiers, typically by adding or subtracting variables with forward selection/backward elimination methods [72]. Once such a classifier is found its constituent parameter space is declared optimal for further analysis. This approach, besides neglecting the variable interactions, is inflexible. In realistic searches for new phenomena that occur in nature classifiers are typically trained on simulated (usually Monte-Carlo) data for at least one of the major classes of events, usually the one related to the previously unseen object or model. If the researcher limits herself to only one classifier, or alternatively to only one of the many possible combinations of the reduced parameter space, without considering the associated loss of information, she becomes limited in options if the search does not yield the desired result. Making a choice of the parameter space based on the global loss function criterion, that consistently produces a strong family of classifiers out of its constituents, allows one to step back and modify the parameter space slightly and maintain a required high performance level, without having to repeat the classifier search. This becomes crucial when the models that are being probed come in significant variety (such as Supersymmetry) and contain free parameters with unknown values. In this case, flexibility, tied with high performance, becomes a crucial aspect of a successful search.

We introduced the *relative variable importance* criterion that has sensitivity to individual variable effects within machine-learning based classifiers and showed its usefulness in identification of adverse variables if such are present. This quality makes it an essential part of the decision-making framework.

If non-negligible interactions among some of the variables are present, relative variable importance criteria are shown to be unsuitable for the decision making pertinent to the reduction in parameter space. The presence of interactions complicates the task above due

to the ripple effect on remaining interacting partners caused by the removal of interacting features. One possible solution is a combination of relative importance and interaction criteria into one, which constitutes a balanced approach, but does not lead to an optimal measure, because the nature of the two criteria are distinct. Only for the case of weak interactions the combined measure has a significant decision making utility.

We propose the use of the *gloss function* for parameter space reduction due to its advantageous design for problems with and without significant variable interactions. The gloss function implicitly incorporates interaction effects by selecting entire variable sets for global removal and compares the amount of predictive loss on the scale of upper limit performance of the remaining classifiers. This approach achieves global optimization of predictive performance of the reduced parameter space.

The combination of the gloss function and relative variable importance criteria constitutes the heart of PARADIGM's decision-making framework. Within the framework, various types of popular classifiers such as decision trees, neural networks, binary regression, can be compared on the same performance scale with the gloss function. Completely unrelated variable sets as well as different phenomena can be compared on the absolute scale of classification performance. For instance, given parameter space (X_1, \dots, X_N) , consider the problem of comparing the extraction of various phenomena from a common background, or conversely, given the phenomenon Ψ_X , a comparison among different parameter spaces (X_{i1}, \dots, X_{iN}) for its extraction from the background. The integrated gloss function can be useful in both applications.

In summary, PARADIGM is a robust parallelized framework that provides decision-making information to assist and improve modern day multivariate analyses.

BIBLIOGRAPHY

- [1] V. M. Abazov et al. Evidence for production of single top quarks. *Phys. Rev.*, D78:012005, 2008.
- [2] V. M. Abazov et al. Observation of single top-quark production. *Phys. Rev. Lett.*, 103:092001, 2009.
- [3] V. M. Abazov et al. Search for dark photons from supersymmetric hidden valleys. *Phys. Rev. Lett.*, 103:081802, 2009.
- [4] V. M. Abazov et al. Search for events with leptonic jets and missing transverse energy in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ tev. *Phys. Rev. Lett.*, 105:211802, 2010.
- [5] R. Abbasi et al. Search for dark matter from the galactic halo with the icecube neutrino observatory. 2011.
- [6] Aous A. Abdo et al. Measurement of the cosmic ray e+ plus e- spectrum from 20 gev to 1 tev with the fermi large area telescope. *Phys. Rev. Lett.*, 102:181101, 2009.
- [7] N. E. Adam et al. Higgs working group summary report. 2008.
- [8] S. Agostinelli et al. G4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.
- [9] Johan Alwall et al. New developments in madgraph/madevent. *AIP Conf. Proc.*, 1078:84–89, 2009.
- [10] Nima Arkani-Hamed, Douglas P. Finkbeiner, Tracy R. Slatyer, and Neal Weiner. A theory of dark matter. *Phys. Rev.*, D79:015014, 2009.
- [11] Nima Arkani-Hamed and Neal Weiner. Lhc signals for a superunified theory of dark matter. *JHEP*, 12:104, 2008.
- [12] David J. Bacon, Alexandre R. Refregier, and Richard S. Ellis. Detection of weak gravitational lensing by large-scale structure. *Mon. Not. Roy. Astron. Soc.*, 318:625, 2000.
- [13] H. Baer and X. Tata. *Weak Scale Supersymmetry*. Cambridge University Press, 2006.

- [14] G. Battaglia, A. Helmi, H. Morrison, P. Harding, E. W. Olszewski, M. Mateo, K. C. Freeman, J. Norris, and S. A. Shectman. The radial velocity dispersion profile of the galactic halo: constraining the density profile of the dark halo of the milky way. *Mon. Not. Roy. Astron. Soc.*, 364:433–442, 2005.
- [15] Matthew Baumgart, Clifford Cheung, Joshua T. Ruderman, Lian-Tao Wang, and Itay Yavin. Non-abelian dark sectors and their collider signatures. *JHEP*, 04:014, 2009.
- [16] G L Bayatian et al. *CMS Physics Technical Design Report Volume I: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.
- [17] C. L. Bennett et al. First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Astrophys. J. Suppl.*, 148:1, 2003.
- [18] A. V. Berezhnoy, A. K. Likhoded, A. V. Luchinsky, and A. A. Novoselov. Double j/psi-meson production at lhc and 4c-tetraquark state. 2011.
- [19] Bernd A. Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis (With Web-Based Fortran Code)*. World Scientific, 2004.
- [20] R. Bernabei et al. *Phys. Let. B*, 1, 2000.
- [21] H. Bethe and J. Ashkin. *Experimental Nuclear Physics*. J. Wiley, 1953.
- [22] A.L. Blum and P. Langley. *Artificial Intelligence*, 97:245, 1997.
- [23] R. K. Bock et al. Data analysis techniques for high-energy physics experiments. *Camb.Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 11:1, 2000.
- [24] E.E. Boos et al. *Phys. Atom. Nucl.*, 69:1317, 2006.
- [25] Marusa Bradac et al. Revealing the properties of dark matter in the merging cluster macsj0025.4-1222. 2008.
- [26] J. Bradley. *Pattern Recognition*, 30:1145, 1997.
- [27] L. Breiman. *Machine Learning*, 24:123, 1996.
- [28] L. Breiman. *Machine Learning*, 45:5, 1996.
- [29] W. Buchmuller. Neutrinos, grand unification and leptogenesis. 2002.
- [30] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti kt jet clustering algorithm. *JHEP*, 04:063, 2008.
- [31] Clifford Cheung, Joshua T. Ruderman, Lian-Tao Wang, and Itay Yavin. Lepton Jets in (Supersymmetric) Electroweak Processes. *JHEP*, 04:116, 2010.
- [32] Douglas Clowe et al. A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006.

- [33] Douglas Clowe, S. W. Randall, and M. Markevitch. Catching a bullet: direct evidence for the existence of dark matter. *Nucl. Phys. Proc. Suppl.*, 173:28–31, 2007.
- [34] CMS Collaboration. *CMS-PAS-JME-07-002*, <http://cms-physics.web.cern.ch/cms-physics/public/JME-07-002-pas.pdf>, 2008.
- [35] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [36] CMS Collaboration. *CMS-PAS-JME-09-002*, <http://cdsweb.cern.ch/record/1190234>, 2009.
- [37] CMS Collaboration. *CMS-PAS-PFT-09-001*, <http://cdsweb.cern.ch/record/1194487>, 2009.
- [38] CMS Collaboration. *CMS-AN-2009-111*, 2009.
- [39] CMS Collaboration. *CMS-PAS-JME-10-010*, <http://cdsweb.cern.ch/record/1308178>, 2010.
- [40] CMS Collaboration. *CMS-PAS-MUO-10-002*, <http://cdsweb.cern.ch/record/1279140>, 2010.
- [41] CMS Collaboration. *CMS-PAS-JME-10-006*, <http://cdsweb.cern.ch/record/1275133>, 2010.
- [42] CMS Collaboration. *CMS-PAS-PFT-10-002*, <http://cdsweb.cern.ch/record/1279341>, 2010.
- [43] CMS Collaboration. *CMS-PAS-EWK-10-004*, 2010.
- [44] CMS Collaboration. Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays. *JINST*, 5:T03009, 2010.
- [45] CMS Collaboration. *CMS-PAS-EXO-11-011*, 2011.
- [46] CMS Collaboration. *CMS-PAS-EXO-11-013*, 2011.
- [47] CMS Collaboration. Missing transverse energy performance of the cms detector. *arXiv1106.5048*, 2011.
- [48] L.A. Cox. *Management Science*, 31:800–813, 1985.
- [49] B. Cunliffe and C. Renfrew. *Science and Stonehenge*. Oxford University Press, The British Academy 92, 1997.
- [50] Democritus. *from Sextus Empiricus*.
- [51] M. Dine. *Supersymmetry and String Theory: Beyond the Standard Model*. Cambridge University Press, 2007.

- [52] S. Dye et al. Gravitational lens magnification by abell 1689: Distortion of the background galaxy luminosity function. *Mon. Not. Roy. Astron. Soc.*, 321:685, 2001.
- [53] Albert Einstein. Lens-like action of a star by the deviation of light in the gravitational field. *Science*, 84, 1936.
- [54] J. Ellis, J. Hagelin, D.V. Nanopoulos, K. Olive, and B. Srednicki. *Nucl. Phys. B*, 238:453, 1984.
- [55] Adam Falkowski, Joshua T. Ruderman, Tomer Volansky, and Jure Zupan. Hidden higgs decaying to lepton jets. *JHEP*, 05:077, 2010.
- [56] Douglas P. Finkbeiner, Tongyan Lin, and Neal Weiner. Inelastic dark matter and dama/libra: An experimentum crucis. *Phys. Rev.*, D80:115008, 2009.
- [57] Douglas P. Finkbeiner and Neal Weiner. Exciting Dark Matter and the INTEGRAL/SPI 511 keV signal. *Phys. Rev.*, D76:083519, 2007.
- [58] G. Flucke et al. CMS silicon tracker alignment strategy with the Millepede II algorithm. *JINST*, 3:P09002, 2008.
- [59] Y. Freund and R. E. Schapire. *Proceedings of the 13th International Conference on Machine Learning (Morgan Kaufmann)*, page 148, 1996.
- [60] J.H. Friedman and B.E. Popescu. <http://www.stat.stanford.edu/people/friedman/index.html>. 2005.
- [61] S. V. Gleyzer and H. B. Prosper, editors. *PARADIGM, a Decision-Making Framework for Variable Selection and Reduction in High Energy Physics, Proceedings of the XII International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Erice, Italy, November 3-7, 2008*.
- [62] H. Goldberg. *Phys. Rev. Lett.*, 50:1419, 1983.
- [63] M. B. Green, J. H. Schwartz, and E. Witten. *String Theory*. Cambridge Monographs on Mathematical Physics, 1987.
- [64] Alan H. Guth. The inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev.*, D23:347–356, 1981.
- [65] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer Press, 2003.
- [66] M. Hall and L. Smith. Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper. *In Proceedings of the Florida Artificial Intelligence Symposium.*, 1999.
- [67] G. Hinshaw et al. Five-year wilkinson microwave anisotropy probe (wmap) observations: data processing, sky maps and basic results. *Astrophys. J. Suppl.*, 180:225–245, 2009.

[68] B. Humpert and P. Mery. $\psi\psi$ production at collider energies. *Z. Physics. C*, 20:83, 1983.

[69] J.E. Huth et al. *Fermilab-Conf-90-249-E*, 1990.

[70] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Phys. Rept.*, 267:195–373, 1996.

[71] R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Eng.*, 82, 1980.

[72] R Kohavi and G. H. John. *The Wrapper Approach in Feature Selection for Knowledge Discovery and Data Mining*, edited by H. Liu and H. Motoda. Kluwer Academic Publishers, Norwell MA, 1998.

[73] Ron Kohavi and George John. Wrappers for subset feature selection. *Artificial Intelligence*, 1997.

[74] S. Kok. *Proceedings of the 22nd International Conference on Machine Learning*, page 441, 2005.

[75] D. Koller and M. Sahami. *Proceedings of the 13th International Conference on Machine Learning*. Morgan Kaufmann, 1996.

[76] W. Kruskal. *The American Statistician*, 41:6, 1987.

[77] W. Kruskal and R. Majors. *The American Statistician*, 43:2, 1989.

[78] Barenaked Ladies. Big bang theory theme. *Amazon Digital Services*, <http://www.amazon.com/dp/B000XKMQZY>, 2007.

[79] L. Lederman. *The God Particle: If the Universe Is the Answer, What Is the Question*. Delta Publishing, 2006.

[80] R. H. Lindenman et al. *Introduction to bivariate and multivariate analysis*. Scott, Foresman and Company, Glenview IL, 1980.

[81] Pertti Lounesto. *Clifford algebras and spinors*. Cambridge University Press, 2001.

[82] D. Lyth and A. Liddle. *The Primordial Density Perturbation*. Cambridge University Press, 2009.

[83] M. L. Mangano et al. *J. High Energy Phys.*, 0307:001, 2003.

[84] Patrick Meade and Matthew Reece. Bridge: Branching ratio inquiry / decay generated events. 2007.

[85] M Milgrom. *Astrophys. J.*, 270, 1983.

[86] Robert Minchin et al. A dark hydrogen cloud in the virgo cluster. *Astrophys. J.*, 622:L21–L24, 2005.

- [87] S. Mitton. *Fred Hoyle: A Life in Science*. Aurum Press, 2005.
- [88] L. Moneta et al., editors. *The RooStats Project*, 2009.
- [89] Ramesh Narayan and Matthias Bartelmann. Lectures on gravitational lensing. 1996.
- [90] NIST. Fundamental physical constants. *CODATA*, 2006.
- [91] K. Olive et al. *Astrophys. J.*, 376:51, 1991.
- [92] Arno A. Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080- mc/s. *Astrophys. J.*, 142:419–421, 1965.
- [93] Donald H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 2000.
- [94] P. Plaszczak and R. Wellner. *Grid Computing*. Elsevier/Morgan Kaufmann, 2005.
- [95] A. M. Polyakov. *Gauge Fields and Strings*. Harwood Academic Press, 1987.
- [96] J. Puerta-Pelayo et al. Parametrization of the response of the muon barrel drift tubes. *CMS-NOTE-2005-018*, 2005.
- [97] C.F. Qiao, L.P. Sun, and P. Sun. Testing charmonium production mechanism via polarized j/ψ pair production at the lhc. *J. Phys. G*, 37:075019, 2010.
- [98] J.R. Quinlan. *c4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo CA, 1993.
- [99] A. L. Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28, 2002.
- [100] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.
- [101] V. C. Rubin, N. Thonnard, and Jr. Ford, W. K. Rotational properties of 21 sc galaxies with a large range of luminosities and radii, from ngc 4605 /r = 4kpc/ to ugc 2885 /r = 122 kpc/. *Astrophys. J.*, 238:471, 1980.
- [102] Joshua T. Ruderman and Tomer Volansky. Decaying into the hidden sector. *JHEP*, 02:024, 2010.
- [103] Weinberg S. *Quantum Theory of Fields*. Cambridge University Press, 1999.
- [104] Gavin P. Salam and Gregory Soyez. A practical seedless infrared-safe cone jet algorithm. *JHEP*, 05:086, 2007.
- [105] R. W. Schnee, R. Abusaidi, D. S. Akerib, P.D. Barnes, Jr., D. A. Bauer, A. Bolozdynya, P.L. Brink, R. Bunker, B. Cabrera, D.O. Caldwell, J.P. Castle, R.M. Clarke, and P. Colling. Results of the cryogenic dark matter search. 2000.
- [106] M. Seeds. *Foundations of Astronomy*. Thompson Brooks/Cole, 2008.

- [107] P. Sikivie. Dark matter axions. *Int. J. Mod. Phys.*, A25:554–563, 2010.
- [108] Torbjorn Sjostrand et al. High-energy physics event generation with pythia 6.1. *Comput. Phys. Commun.*, 135:238–259, 2001.
- [109] D. N. Spergel et al. Wilkinson microwave anisotropy probe (wmap) three year results: Implications for cosmology. *Astrophys. J. Suppl.*, 170:377, 2007.
- [110] Volker Springel et al. Simulating the joint evolution of quasars, galaxies and their large-scale distribution. *Nature*, 435:629–636, 2005.
- [111] Matthew J. Strassler and Kathryn M. Zurek. Echoes of a hidden valley at hadron colliders. *Phys. Lett.*, B651:374–379, 2007.
- [112] Andrew W. Strong et al. Gamma-ray continuum emission from the inner galactic region as observed with integral/spi. *Astron. Astrophys.*, 444:495, 2005.
- [113] H. Thiel and C. Chung. *The American Statistician*, 42:249, 1988.
- [114] S. Torii et al. High-energy electron observations by ppb-bets flight in antarctica. 2008.
- [115] David Tucker-Smith and Neal Weiner. Inelastic dark matter at dama, cdms and future experiments. *Nucl. Phys. Proc. Suppl.*, 124:197–200, 2003.
- [116] S. Willenbrock and D.A. Dicus. *Phys. Rev. D.*, 34:155, 1986.
- [117] X. Wu, T. Chiueh, L. Fang, and Y. Xue. A comparison of different cluster mass estimates: consistency or discrepancy? *Monthly Notices of the Royal Astronomical Society*, 301(3):861–871, December 1998.
- [118] W. M. Yao et al. Review of particle physics. *J. Phys. G*, 2006.
- [119] C.P. Yuan. *Phys. Rev. D.*, 41:42, 1990.
- [120] F. Zwicky. Spectral displacement of extra galactic nebulae. *Helv. Phys. Acta*, 6:110–127, 1933.
- [121] F. Zwicky. On the probability of detecting nebulae which act as gravitational lenses. *Phys. Rev.*, 51:679–679, 1937.

BIOGRAPHICAL SKETCH

EDUCATION

- Ph.D. Particle Physics (2011), Florida State University, Tallahassee
- M.S. Particle Physics (2006), Florida State University, Tallahassee
- B.S. Physics (2004), Carnegie Mellon University, Pittsburgh

RESEARCH EXPERIENCE

- Research Assistant, Florida State University, May 2005 - present
- Teaching Assistant, Florida State University, September 2004 - May 2005
- Research Assistant, National High Magnetic Field Laboratory, May - September 2004
- Teaching Assistant, Carnegie Mellon University, September 2003 - May 2004
- Intern, Hypres Incorporated, May - September 2003
- Intern, Transwitch Corporation, May 2000 - September 2001

PROGRAMMING AND SOFTWARE EXPERIENCE

- ROOT, R, MadGraph, Alpgen, Pythia, C++, Lisp, Python, Basic

FELLOWSHIPS, HONORS AND AWARDS

- Hagopian Endowment Award for Outstanding Student in High Energy Physics, April 2009
- American Physical Society Division of Particles and Fields Travel Award, April 2007
- Golden Key Honor Society, GPA 3.97
- Carnegie Mellon University Peer Tutor Excellence Award, April 2003
- Carnegie Mellon University Dean's List

TEACHING EXPERIENCE

- Course: Introduction to Astronomy Laboratory (section of 35 students) Spring 2005

- Course: Introduction to Astronomy Laboratory (section of 30 students) Fall 2004
- Course: Experimental Physics Laboratory (section of 25 students) Spring 2004
- Course: Experimental Physics Laboratory (section of 25 students) Fall 2003

Sergei was born on December 20, 1980 in a small town in South-Western Ukraine that borders Romania on one side and is surrounded by Poland, Belarus, Russia, Moldova, Slovenia and Hungary, on the others. While going through some old books and journals, at the age of 10, Sergei found a book called the “General Theory of Probability”. He confused the name with the “General Theory of Relativity”, but nevertheless read it. After questioning his father, Sergei found out about the existence of the real “Theory of Relativity”, which they read together. That early exposure to advanced physics topics inspired Sergei to pursue his academic interests in the subject later. In addition to providing continuous motivation for higher learning, Sergei’s parents taught Sergei at an early age that aside for various sports, chess and astronomy, there are important and wonderful things in life, such as books and music. Countless hours of practicing piano did not go wasted on him. There was a Nobel laureate in particle physics born in the same town that attended the same high school as Sergei in Cleveland Heights, Ohio. While there, Sergei was fortunate to have met Robert Quail, his first physics teacher, whose clear explanations furthered Sergei’s interests in advanced mathematics, physics and astronomy. Sergei likes looking at stars, with and without a telescope. His younger brother, Vitaliy, always joined him in his adventures, whenever science was involved. Other important sources of Sergei’s academic inspiration are Sergei’s grandmother Pearl and grandfather Michael, because of their hard work, persevering attitude and a passion for learning at any age.