

Anwendung von B-Erkennung auf die modellunspezifische Suche in CMS

von
Thomas Esch

Bachelorarbeit in Physik

vorgelegt der

Fakultät für Mathematik, Informatik und
Naturwissenschaften der RWTH Aachen

im Juli 2011

angefertigt am
III. Physikalischen Institut A

bei
Prof. Dr. Thomas Hebbeker

Ich versichere, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate kenntlich gemacht habe.

Aachen, den 18.07.2011

Thomas Esch

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	3
2.1	Das Standardmodell der Elementarteilchenphysik	3
3	Das CMS Experiment am CERN	7
3.1	Wichtige Größen des CMS-Detektors	8
3.2	Der CMS-Detektor	9
3.2.1	Silizium-Spurdetektor	10
3.2.2	Elektromagnetisches Kalorimeter (ECAL)	10
3.2.3	Hadronisches Kalorimeter (HCAL)	11
3.2.4	Solenoidmagnet	11
3.2.5	Myonensystem	12
4	Modellunspezifische Suche in CMS	13
4.1	Konzept der modellunspezifischen Suche in CMS	13
4.1.1	Fokussierung auf leptonische Ereignisse	13
4.1.2	Ereignisklassen	14
4.1.3	Suche nach Abweichungen	14
4.2	Der Suchalgorithmus	15
4.2.1	Schritt 1: Bestimmung der Interessanten Region	15
4.2.2	Schritt 2: Bestimmung der Signifikanz	16
4.3	Selektion der Objekte	17
4.3.1	Myonen	18
4.3.2	Elektronen	18
4.3.3	Photonen	19
4.3.4	Jets	20
4.3.5	Fehlende transversale Energie (MET)	20
4.3.6	Bereinigung der Ereignisse	21
4.4	Systematische Messunsicherheiten	21
5	Erkennung von B-Jets	23
5.1	Das Bottom-Quark	23
5.2	Algorithmen zur Erkennung von B-Jets	24
5.2.1	Stoßparameterbasierende Algorithmen	24
5.2.2	Leptonenbasierende Algorithmen	25
5.2.3	Sekundärvertex basierende Algorithmen	25
5.2.4	Kombinierte Algorithmen	26
5.2.5	Messunsicherheiten	26

6 B-Erkennung in MUSiC	27
6.1 Sensitivitätsstudie von B-Erkennung in MUSiC	27
6.1.1 Supersymmetrie	27
6.1.2 Single-Top	29
6.2 Anwendung von B-Erkennung auf Daten	29
6.2.1 Combined Secondary Vertex MVA	29
6.2.2 Track Counting High Purity tight	30
6.2.3 Track Counting High Purity loose	32
6.2.4 Track Counting High Efficiency tight	33
6.2.5 Track Counting High Efficiency loose	35
6.2.6 Vergleich der Ergebnisse	36
7 Zusammenfassung	37
Literatur	I

1 Einleitung

Im Jahr 2009 wurde der Large Hadron Collider (LHC) am CERN bei Genf in Betrieb genommen. Mit seiner Hilfe sollen viele offene Fragen, die das Standardmodell der Elementarteilchenphysik nicht klären kann, gelöst werden. Eines der Experimente am LHC ist der Compact Muon Solenoid (CMS), ein Allzweckdetektor, der bei der Suche nach neuer Physik helfen soll. Dabei sollen auch Theorien überprüft werden, in denen Teilchen vorkommen, die in Bottom-Quarks zerfallen und anschließend als B-Jets detektiert werden. Diese B-Jets spielen bei der Suche nach neuer Physik eine wichtige Rolle.

Bei der modellunspezifischen Suche in CMS (MUSiC) werden die gesammelten Daten bei einer Schwerpunktsenergie von $\sqrt{s} = 7$ TeV mit der Erwartung des Standardmodells, die mithilfe von Monte-Carlo-Simulationen dargestellt wird, verglichen, ohne einen modellbezogenen Ansatz zu verfolgen. So kann man allgemeine Abweichungen erkennen und die modellspezifischen Analysen bei ihrer Suche unterstützen. Zu diesem Zweck wurde die Erkennung von B-Jets in MUSiC implementiert. Die Anwendung von B-Erkennung auf Daten aus dem Jahr 2010 in MUSiC soll im Nachfolgenden untersucht werden.

2 Theoretische Grundlagen

Seit Jahrhunderten versuchen die Menschen, den Aufbau der für uns sichtbaren Welt zu erklären. Bereits im alten Griechenland um ca 400 v. Chr. wurde dabei die Idee von unteilbaren, endlich ausgedehnten Einheiten, den Atomen, entwickelt [1]. Nach heutigem Wissensstand handelt es sich bei den kleinsten Teilchen um punktförmige Teilchen, also ohne Ausdehnung. Bei der Entdeckung der Atome dachte man, diese kleinsten, ununterteilbaren Teilchen gefunden zu haben. Es stellte sich jedoch heraus, dass die vermeintlich ununterteilbaren Atome aus weiteren Teilchen, den Elektronen, Protonen und Neutronen, bestehen, die teilweise ebenfalls eine weitere Substruktur aufweisen. So entwickelte sich die Suche nach den kleinsten Teilchen immer weiter, bis man ein Modell fand, welches sich bis heute bewährt hat: Das Standardmodell der Elementarteilchenphysik.

2.1 Das Standardmodell der Elementarteilchenphysik

Das Standardmodell der Elementarteilchenphysik ist das bis heute gängigste Modell zur Beschreibung der bekannten Wechselwirkungen im Zusammenhang mit Elementarteilchen. Diese Teilchen werden als punktförmig angenommen (bewiesen ist allerdings nur, dass sie kleiner als 10^{-18} m sind). Das Standardmodell der Elementarteilchenphysik beschreibt die nach heutigem Wissensstand bekannten Elementarteilchen und die fundamentalen Wechselwirkungen mit einer großen Präzision [2]. Dennoch wird erwartet, dass das Standardmodell erweitert werden muss, da es einige Phänomene nicht erklären kann.

Das Standardmodell der Elementarteilchenphysik baut auf zwei Arten von Teilchen auf. Zum einen gibt es die Fermionen, Teilchen mit halbzahligem Spin. Diese werden im Standardmodell weiter unterteilt in Quarks und Leptonen. Zum anderen gibt es die Bosonen, Teilchen mit ganzzahligem Spin. Während sich die meiste uns bekannte Materie aus elementaren Fermionen zusammensetzt¹, dienen die Bosonen im Standardmodell als Austauschteilchen der fundamentalen Wechselwirkungen.

Die Fermionen können, wie bereits angesprochen, nochmals in Leptonen und Quarks unterteilt werden. Sowohl Leptonen, als auch Quarks lassen sich in drei Generationen gliedern. Die bekannten Leptonen sind [1]:

Generation	Lepton		Masse /MeV
I.	e^- (Elektron)	ν_e (Elektronneutrino)	$m_e = 0.511, m_{\nu_e} < 10^{-6}$
II.	μ^- (Myon)	ν_μ (Myonneutrino)	$m_\mu = 105.7, m_{\nu_\mu} < 0.25$
III.	τ^- (Tauon)	ν_τ (Tauneutrino)	$m_\tau = 1777, m_{\nu_\tau} < 35$

Das Elektron, das Myon und das Tauon haben eine Ladung von $-e^2$, während die Neutrinos ungeladen sind. Hinzu kommen noch die jeweiligen Antiteilchen, welche die selbe Ruhemasse, jedoch entgegengesetzte Vorzeichen bei den Quantenzahlen, wie z.B. der Ladung

¹Das Proton z.B. besteht aus zwei up- und einem down-Quark.

² e steht für die Elementarladung, die $1.602 \cdot 10^{-19} \text{C}$ beträgt

oder der Leptonenzahl³, besitzen. Vor einigen Jahren ist man noch davon ausgegangen, dass Neutrinos keine Masse haben, jedoch kann man anhand von Neutrinooszillationen eine nicht verschwindende Massendifferenz bei Neutrinos indirekt nachweisen.

Die bekannten Quarks lassen sich wie folgt in ihre Generationen unterteilen:

Generation	Quark		Masse / GeV
I.	u (up)	d (down)	$m_u \approx 3 \cdot 10^{-3}, m_d \approx 6 \cdot 10^{-3}$
II.	c (charm)	s (strange)	$m_c \approx 1.2, m_s \approx 0.1$
III.	t (top)	b (bottom)	$m_t \approx 175, m_b \approx 4.5$

Das Up-, das Charm- und das Top-Quark haben eine Ladung von $+\frac{2}{3}e$, das Down-, das Strange- und das Bottom-Quark haben eine Ladung von $-\frac{1}{3}e$. Wie bei den Leptonen gibt es auch bei den Quarks zu jedem Teilchen ein Antiteilchen. Quarks bilden Bindungszustände zu sogenannten Baryonen (zusammengesetzt aus drei Quarks oder drei Anti-Quarks) und Mesonen (zusammengesetzt aus einem Quark und einem Anti-Quark). Sowohl Baryonen als auch Mesonen gehören zur Gruppe der Hadronen. Alle Quarks tragen eine der drei Farbladungen „rot“, „grün“ oder „blau“ bzw. entsprechende Antifarben. In der Summe müssen Bindungszustände immer farbneutral sein. Dies wird erreicht, indem die Bindungszustände aus einem Quark jeder Farbe (oder Antifarbe) oder aus einem Quark mit beliebiger Farbe und einem anderen Quark mit der entsprechenden Antifarbe zusammengesetzt werden. Das ist auch der Grund für das sogenannte Confinement⁴, also für die fehlende Existenz bzw. die fehlende Beobachtung von freien Quarks und Gluonen, da diese Teilchen immer eine Farbladung tragen.

Die gerade angesprochenen Gluonen gehören jedoch nicht zu den Fermionen, sondern zu den Bosonen. Sie sind Teilchen mit Spin 1 und bilden im Standardmodell der Elementarteilchenphysik die Austauschteilchen der starken Wechselwirkung. Aus gruppentheoretischen Gründen gibt es 8 Gluonen, die an alle Teilchen mit Farbladung und damit auch an sich selbst koppeln, wobei die Farbe der Teilchen durch die starke Wechselwirkung geändert wird. Dadurch fällt das Potenzial der starken Wechselwirkung mit zunehmendem Abstand nicht ab, sondern es steigt an:

$$U_{QCD}(r) = -\frac{4}{3} \frac{\alpha_s}{r} + kr \quad (1)$$

Die Reichweite der starken Wechselwirkung beträgt $\sim 10^{-15}$ m.

Neben der starken Wechselwirkung gibt es noch die elektroschwache Wechselwirkung, welche aus der elektromagnetischen und der schwachen Wechselwirkung besteht. Auch hier sind die Austauschteilchen Bosonen. Das Austauschteilchen der elektromagnetischen Wechselwirkung ist das masselose Photon. Ein typisches Beispiel für die elektromagnetische Wechselwirkung ist die Abstoßung zweier Teilchen aufgrund gleicher Ladung. Die Reichweite der elektromagnetischen Wechselwirkung ist unendlich.

³Die Leptonenzahl ist eine Größe, die bei Zerfällen mit Leptonen erhalten bleibt. Es gilt z.B. $L_{e^-} = +1, L_{e^+} = -1$ und $L_{\nu_e} = 1$.

⁴frei übersetzt bedeutet confinement Einsperrung.

Die schwache Wechselwirkung wird durch die W^\pm -Bosonen und Z^0 -Bosonen übertragen. Da diese Bosonen eine nicht verschwindende Ruhemasse besitzen und nicht stabil sind, hat die schwache Wechselwirkung nur eine sehr geringe Stärke⁵ und Reichweite ($\sim 10^{-18}$ m). Außerdem sind die W^\pm -Bosonen die einzigen Eichbosonen im Standardmodell, die elektrisch geladen sind. Jede der Wechselwirkungen lässt sich durch eine Symmetriegruppe darstellen. Die elektroschwache Wechselwirkung wird durch die Gruppe $SU2_L \otimes U1$ und die starke Wechselwirkung durch die Gruppe $SU3_C$ dargestellt. Das Standardmodell gilt daher als $SU3_C \otimes SU2_L \otimes U1$ -Modell [2]. Zusammengefasst hat man im Standardmodell folgende Bosonen, die die fundamentalen Wechselwirkungen übertragen:

Wechselwirkung	Austauschteilchen	Masse /GeV
elektromagnetische	γ (Photon)	$m_\gamma = 0$
schwache	W^\pm, Z^0 - Boson	$m_{W^{+/-}} \approx 80, m_{Z^0} \approx 92$
starke	8 g (Gluonen)	$m_g = 0$

Neben diesen bereits beobachteten Bosonen wird im Standardmodell noch das Higgs-Boson vorhergesagt. Es ist das einzige Teilchen des Standardmodells, welches noch nicht experimentell nachgewiesen wurde. Im Standardmodell hat das Higgs-Boson die Aufgabe, den Teilchen die träge Masse zu übertragen. Dieser Mechanismus wird als Higgs-Mechanismus bezeichnet. Es wird erwartet, dass das Higgs-Boson ungefähr eine Masse zwischen 114 GeV und 158 GeV besitzt [3]. Aus diesem Grund erwartet man, dass das Higgs-Boson am LHC beobachtet wird, da dieser den erwarteten Energiebereich komplett abdeckt.

⁵Die schwache WW ist um $\sim 10^{13}$ schwächer als die starke WW und um $\sim 10^{11}$ schwächer als die elektromagnetische WW.

3 Das CMS Experiment am CERN

Auf der Suche nach Erweiterungen des Standardmodells stoßen die älteren Teilchenbeschleuniger an ihre Grenzen. Die von vielen Theorien gelieferten Vorhersagen und Lösungsansätze liegen weit außerhalb der Energiebereiche, in denen die bisherigen Teilchenbeschleuniger gearbeitet haben. Aus diesem Grund wurde der Large Hadron Collider am CERN⁶ entwickelt.

Der LHC wurde in den Tunnel des LEP-Experiments⁷ gebaut und ist mit einer Länge von ungefähr 27 km bei einer geplanten Schwerpunktenenergie von 14 TeV der längste und leistungsfähigste Teilchenbeschleuniger der Welt. Im LHC sollen zum einen Protonen mit Energien von jeweils bis zu 7 TeV und zum anderen Blei-Ionen mit Energien von bis zu 2.76 TeV kollidieren. Eine weitere Besonderheit des LHC ist seine hohe Luminosität. Die Luminosität \mathcal{L} ist ein Maß für die Anzahl der Teilchenkollisionen in einem bestimmten Zeitintervall. Der LHC hat eine Design-Luminosität von $\mathcal{L}_{Design} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, soll aber nach einer weiteren Umbauphase in einigen Jahren eine Luminosität von $\mathcal{L} = 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ erreichen. Im Juni 2011 betrug die Luminosität des LHC $1.25 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ [4]. Ebenfalls im Juni 2011 hat die zeitlich integrierte Luminosität $\mathcal{L}_{int} = \int \mathcal{L} dt$, welche ein Maß für die insgesamt erhaltene Datenmenge ist, bei den Experimenten CMS und ATLAS die Grenze von $\mathcal{L}_{int} = 1 \text{ fb}^{-1}$ (ein inverses Femtobarn) überschritten. Dies entspricht ungefähr einer Datenmenge von 70 Milliarden Proton-Proton Kollisionen [5].

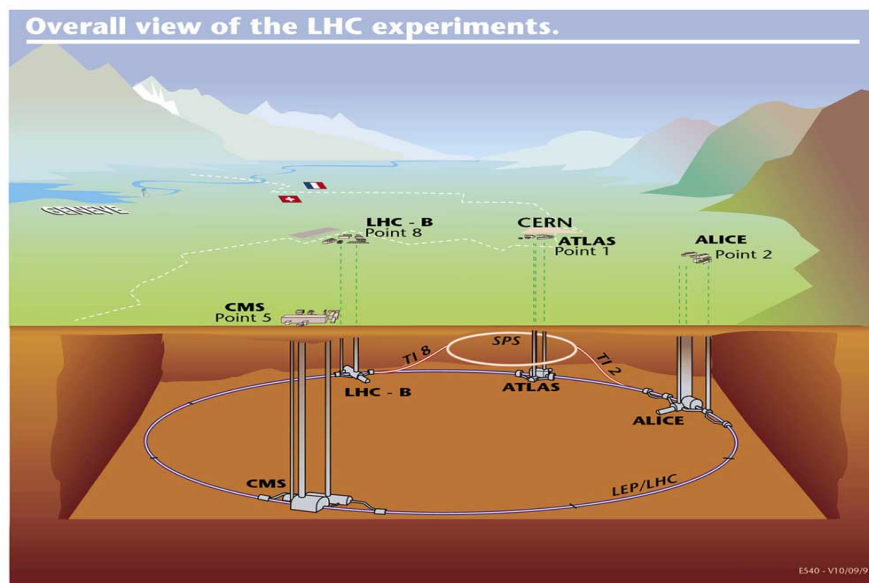


Abbildung 1: Der LHC mit der Lage der dazugehörigen Experimente [6].

⁶früher: Conseil Européen pour la Recherche Nucléaire, übersetzt: Europäische Organisation für Kernforschung

⁷Large Electron-Positron Collider

Neben den bereits erwähnten Experimenten CMS und ATLAS gibt es am LHC noch die Experimente ALICE und LHCb (Abb. 1). Die Hauptaufgaben des CMS- und ATLAS-Experiments liegen im Nachweis des Higgs-Bosons sowie der Suche nach einer eventuellen Substruktur von Leptonen und Quarks, während sich die Experimente ALICE und LHCb mit der Beobachtung bzw. dem Nachweis von Quark-Gluon-Plasma und dem Ungleichgewicht zwischen Materie und Antimaterie im Universum, der CP-Verletzung, beschäftigen [4].

3.1 Wichtige Größen des CMS-Detektors

Um im Detektor die Eigenschaften und die Art der Teilchen zuverlässig rekonstruieren zu können, wird ein einheitliches Größen- und Koordinatensystem verwendet. Dabei liegt die Mitte des CMS-Detektors im Ursprung. Von dort aus zeigt die x-Achse zum Mittelpunkt des LHC, die y-Achse nach oben zur Erdoberfläche und die z-Achse liegt auf der Strahlachse. Da der Detektor zylindrisch ist, ist er ϕ -symmetrisch, wobei ϕ der Polarwinkel in der x-y-Ebene ist. Der Winkel θ wird von der z-Achse in Richtung x-y-Ebene gemessen. Dieser Winkel wird im CMS Experiment durch die Pseudorapidität [2]

$$\eta = \frac{1}{2} \ln \left(\frac{p + p_L}{p - p_L} \right) = - \ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (2)$$

ausgedrückt, da die Differenz $\Delta\eta$ zweier Pseudorapiditäten näherungsweise invariant gegenüber Lorentztransformationen entlang der z-Achse ist. Mithilfe des Polarwinkels und der Pseudorapidität kann man den Abstand zweier Punkte im Detektor, z.B. von zwei Teilchen, wie folgt ausdrücken:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \quad (3)$$

Neben diesen Größen zur Bestimmung der Koordinaten von Teilchen spielen noch andere Größen eine Rolle. Durch die Lorentzkraft werden bewegte Teilchen in einem Magnetfeld abgelenkt, wobei man aus der Ablenkung und der Stärke des Magnetfeldes den Impuls transversal zum Magnetfeld bestimmen kann. Der Transversalimpuls ist die einzige Impulskomponente, die im Detektor rekonstruiert werden kann, weswegen er eine wichtige Rolle spielt. In grober Näherung gilt die Formel

$$p_T / \text{GeV} = 0.3 \cdot B/T \cdot R/m. \quad (4)$$

Da das Magnetfeld im Detektor in Richtung der z-Achse ausgerichtet ist, lässt sich die transversale Komponente des Impulses einfach durch

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin \theta \quad (5)$$

bestimmen. Weitere wichtige Größen sind die transversale Energie E_T und die fehlende transversale Energie E_T^{miss} , wobei man anhand letzterer z.B. erkennen kann, ob Neutrinos oder andere, unbekannt, nicht wechselwirkende Teilchen bei einem Zerfall im Detektor beteiligt waren.

3.2 Der CMS-Detektor

CMS⁸ ist ein 21 m langer, 15 m breiter und 18 m hoher Teilchendetektor am LHC. Der Name „Compact Myon Solenoid“ beruht auf den Eigenschaften des Detektors [5]:

- „Compact“: Trotz eines Gewichts von 12500 Tonnen ist der Detektor relativ klein.
- „Myon“: Der Detektor ist speziell auf die Detektion und Vermessung von Myonen ausgelegt.
- „Solenoid“: Bei der Detektion wird ein starker Solenoidmagnet benutzt.

Den CMS-Detektor kann man in drei Abschnitte gliedern. Die innerste Schicht, welche das Strahlrohr umgibt, besteht aus Silizium-Spurdetektoren und dient der genauen Spurvermessung von geladenen Teilchen. Umgeben wird diese Schicht vom elektromagnetischen Kalorimeter, das für die Messung der Energie von elektromagnetisch wechselwirkenden Teilchen zuständig ist (hier werden auch Photonen absorbiert und ihre Energie gemessen). Außerhalb des elektromagnetischen Kalorimeters befindet sich das hadronische Kalorimeter. Das hadronische Kalorimeter soll die Energie der Teilchen detektieren, die das elektromagnetische Kalorimeter verlassen, wie z.B. Protonen und andere Hadronen. Um den Kalorimeterbereich wurde der supraleitende Solonoidmagnet mit einer magnetischen Feldstärke von fast 4 T angebracht. Durch dieses Magnetfeld wird die Bahn der geladenen Teilchen gekrümmt, sodass man aus der Krümmung den Transversalimpuls bestimmen kann. Den dritten Bereich bilden die Myonendetektoren, die zur Rekonstruktion der Spuren von Myonen dienen. Myonen sollten die einzigen geladenen Teilchen sein, die die ersten zwei Bereiche des Detektors verlassen, da sie mehrere Meter Materie durchfliegen können, ohne zu wechselwirken [4, 5]. Eine schematische Darstellung des CMS-Detektors ist in Abb. 2 gegeben.

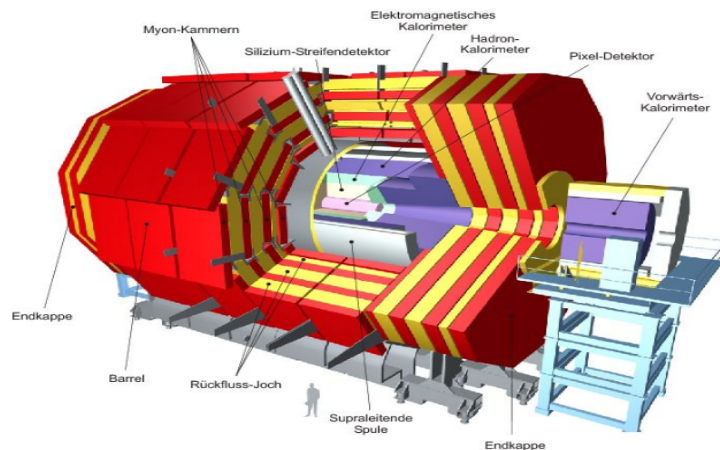


Abbildung 2: Schematische Darstellung des CMS Detektors [7].

⁸Compact Myon Solenoid

Das CMS-Experiment verfolgt mehrere Ziele:

- Nachweis und Untersuchung des vom Standardmodell vorhergesagten Higgs-Bosons
- Suche und Nachweis von supersymmetrischen Teilchen (SUSY)
- Suche nach Hinweisen auf zusätzliche Dimensionen (Stringtheorie)
- Suche nach neuen, nicht vom Standardmodell vorhergesagten Teilchen, z.B. einer Substruktur von Leptonen und Quarks
- Hinweise auf Dunkle Materie

Das vermutlich wichtigste Ziel dabei ist der Nachweis des Higgs-Bosons. Aus diesem Grund sucht neben CMS auch ATLAS nach dem Higgs-Boson, wobei beide Detektoren unterschiedliche Methoden nutzen [8].

3.2.1 Silizium-Spurdetektor

Der Silizium-Spurdetektor besteht aus zwei wesentlichen Teilen: Dem Pixeldetektor, welcher direkt um das Strahlrohr angebracht ist und dem Streifendetektor, der den Pixeldetektor umgibt. Die Abmessungen des Silizium-Spurdetektors betragen 5.8 m in der Länge und 2.5 m im Durchmesser. Der Spurdetektor kann alle Teilchen mit einer Pseudorapidität von $|\eta| \leq 2.5$ detektieren. Wenn der LHC seine Design-Luminosität von $\mathcal{L}_{Design} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ erreicht, werden ungefähr alle 25 ns 1000 Teilchen von bis zu 20 überlappenden Proton-Proton Kollisionen den Spurdetektor durchqueren [9]. Durch die hohe Luminosität und die direkte Nähe zum Strahlrohr ist der Silizium-Spurdetektor einem hohen Verschleiß ausgesetzt, welcher durch Abkühlung auf -10°C bis -27°C verringert werden soll.

Der Pixeldetektor besteht aus drei Lagen mit den Radien 4.4 cm, 7.3 cm und 10.2 cm. Seine Pixelgröße beträgt $100 \mu\text{m} \times 150 \mu\text{m}$ und er erreicht eine Ortsauflösung von $10 \mu\text{m}$ in radialer Richtung und $20 \mu\text{m}$ in Strahlrohrrichtung [4]. Der Streifendetektor besteht aus 10 zylindrischen Lagen, jeweils neun Scheiben in den Endkappen und jeweils drei Scheiben im inneren Bereich. Im ganzen Bereich des Silizium-Spurdetektors beträgt die magnetische Feldstärke ca. 4 T. Dadurch, dass sich im Detektor auch noch Elektronik und Kühlung befinden, weist dieser einige nicht sensitive Regionen auf. Mit dem Silizium-Spurdetektor kann man besonders gut die Spur von hochenergetischen Leptonen rekonstruieren.

3.2.2 Elektromagnetisches Kalorimeter (ECAL)

Das ECAL besteht aus 61200 PbWO_4 -Kristallen im zentralen Bereich und 7324 Kristallen in den beiden Endkappen als Szintillatormaterial des Detektors. Sie haben eine Querschnittsfläche von ungefähr $2.2 \text{ cm} \times 2.2 \text{ cm}$ bis $2.6 \text{ cm} \times 2.6 \text{ cm}$, wobei die Kristalle von innen nach außen hin größer werden. Die Kristalle werden jeweils mit Photodioden ausgelesen [9].

Im zentralen Bereich kann das ECAL Teilchen mit einer Pseudorapidität von $|\eta| \leq 1.479$ und in den Endkappen Teilchen mit einer Pseudorapidität von $1.479 \leq |\eta| \leq 3.0$ detektieren. Die Energie von Photonen kann durch das ECAL auf bis zu einem Prozent genau bestimmt werden.

3.2.3 Hadronisches Kalorimeter (HCAL)

Das HCAL dient in erster Linie zur Messung von Jet-Energien⁹ und der fehlenden transversalen Energie E_T^{miss} , für die z.B. Neutrinos verantwortlich sind. Im Zentralbereich können Teilchen detektiert werden, bei denen $|\eta| \leq 1.3$ ist, im Endkappenbereich werden Teilchen mit $1.3 \leq |\eta| \leq 3.0$ detektiert. Davor gibt es nochmal einen Bereich in dem Teilchen bis $|\eta| \leq 5.2$ erkannt werden. Außerdem befindet sich noch ein kleiner Teil des HCAL außerhalb des Solenoidmagneten. Dieser soll besonders hochenergetische Hadronen detektieren [9].

Das HCAL besteht aus abwechselnden Schichten von Messing und Szintillatormaterial, wobei das Messing als Absorbermaterial dient. Zum Auslesen der Szintillatoren werden Wellenlängenschiefefasern verwendet. Für den Bereich außerhalb des Solenoidmagneten dient der Magnet als zusätzlicher Absorber.

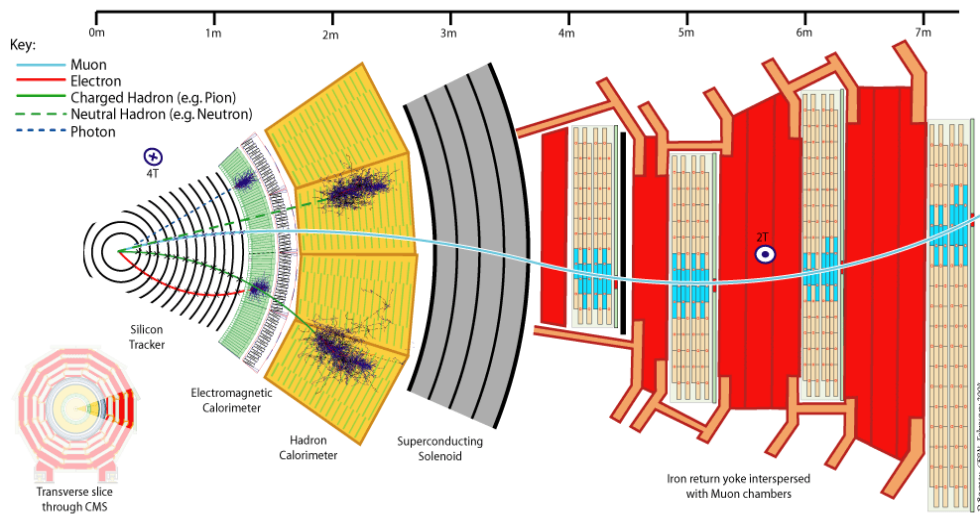


Abbildung 3: Querschnitt des CMS-Detektors [10].

3.2.4 Solenoidmagnet

Da der CMS-Detektor relativ kompakt ist, wird ein entsprechend starker, supraleitender Magnet benötigt, um die Bahn der geladenen Teilchen zu krümmen. Dies dient der Bestimmung des transversalen Impulses. Der Solenoidmagnet hat eine Länge von 13 m und

⁹ Jets sind Schauer aus Hadronen, die durch die Wechselwirkungsenergie zweier gebundener, sich voneinander weg bewegter Quarks entstehen.

einen Durchmesser von 6 m. Der Magnet wird bei einer Temperatur von -270°C betrieben und erreicht eine magnetische Feldstärke von bis zu 4 T [4].

3.2.5 Myonensystem

Das Myonensystem befindet sich im äußersten Bereich des Detektors. Es dient zur Bestimmung der Myonen mittels Bahnkrümmung. Die Myonen sind die einzigen geladenen Teilchen, die diesen Bereich des Detektor erreichen, andere Teilchen zerfallen vorher oder werden auf eine andere Art und Weise absorbiert. Neben den Myonen erreichen noch die ungeladenen Neutrinos und Neutronen diesen Bereich des Detektors. Während Neutrinos eine vernachlässigbar große Wechselwirkung mit dem Myonensystem haben, werden die Spuren der Neutronen rausgefiltert.

Im Zentralbereich werden die Myonen durch Driftkammern detektiert. Diese befinden sich zwischen den Eisenjochen in zwölf Segmenten angeordnet, in denen sich je vier Myonenkammern befinden. In diesem Bereich werden Myonen mit einer Pseudorapidität von $|\eta| \leq 1.2$ gemessen.

Im Endkappen-Bereich befinden sich Kathodenstreifenkammern. Diese haben eine sehr grobe Ortsauflösung von $200\ \mu\text{m}$. Im Bereich der Endkappen werden Teilchen mit einer Pseudorapidität von $0.9 \leq |\eta| \leq 2.4$ detektiert [9, 4]. Abb. 3 zeigt den Querschnitt des CMS-Detektors mit der Flugbahn einiger Teilchen. Gut zu Erkennen ist die Bahn der Myonen, die bis in die äußersten Bereiche des Detektors reicht und sogar darüber hinaus geht.

4 Modellunspezifische Suche in CMS

Durch den Start vom LHC im November 2009 wurde ein Energiebereich jenseits der 3 TeV eröffnet. Es gibt viele Theorien für die Existenz von neuer Physik in diesem Energiebereich, dennoch kann man am LHC eventuell neue Physik finden, die von keiner Theorie vorhergesagt wurde. Dazu ist es hilfreich, eine Analyse zu haben, die alle Daten mit dem Standardmodell vergleicht und die einem mitteilt, wo es eine Abweichung gibt, unabhängig von einem vorher entwickelten Modell. Diesen Zweck erfüllt die modellunspezifische Suche in CMS (MUSiC). Die meisten Informationen des folgenden Abschnitts stammen aus der Analysis Note zur modellunspezifischen Suche in CMS [11].

4.1 Konzept der modellunspezifischen Suche in CMS

Um Hinweise auf neue Physik zu finden, durchsucht MUSiC systematisch Daten und vergleicht diese auf Abweichungen von den Vorhersagen des Standardmodells. Bei diesen Abweichungen kann es sich auch um Detektoreffekte handeln, sodass MUSiC als Kontrollinstrument des CMS-Detektors dienen kann. Ein weiterer Grund für Abweichungen können Fehler in den Monte-Carlo-Simulationen sein. Die modellunspezifische Suche soll modellspezifische Suchen ergänzen und unterstützen, jedoch nicht ersetzen. Eine in MUSiC entdeckte Abweichung des Standardmodells muss immer mithilfe von spezifischen Modellen überprüft werden, da diese im Allgemeinen sensitiver bei der Suche nach neuer Physik sind. MUSiC alleine kann also keine neue Physik entdecken. Aus diesem Grund wird die Datenmenge durch Standardselektionskriterien reduziert, die nicht für irgendeinen modellspezifischen Ansatz optimiert wurden. Dennoch müssen die durch die Selektionskriterien ausgewählten Objekte bekannt und gut studiert sein, wie z.B. hochenergetische Leptonen. Durch die Standardschnitte können jedoch Signale, die sich hinter einem großen Untergrund verbergen, nicht detektiert werden. Die Ziele von MUSiC können also wie folgt zusammengefasst werden:

- Suche nach Abweichungen vom Standardmodell
- Suche nach unbekanntem Detektoreffekten
- Suche nach Fehlern in den Monte-Carlo-Simulationen

Der Grund für eine modellunspezifische Suche am LHC ist relativ offensichtlich: Zwar gibt es eine große Zahl von Theorien, allerdings kann niemand garantieren, dass es in den bislang unbekanntem Energiebereichen nicht auch neue, unbekanntem Effekte gibt. Die modellunspezifische Suche kann also sowohl die entwickelten Modelle unterstützen, als auch nicht erwartete Abweichungen finden, was den Spielraum für neue Modelle vergrößert. Die Software arbeitet in drei wichtigen Schritten, um die Kriterien der modellunspezifischen Suche zu erfüllen.

4.1.1 Fokussierung auf leptonische Ereignisse

Obwohl bei der Selektierung der Datenauswahl modellunspezifische Standardschnitte gemacht werden, müssen dennoch sehr bekannte und gut verstandene Objekte als Grund-

lage für die modellunspezifische Suche verwendet werden. Aus diesem Grund werden nur Ereignisse betrachtet, die mindestens ein geladenes Lepton, ein Elektron oder ein Myon, enthalten.

4.1.2 Ereignisklassen

Nach der Selektierung der Daten werden diese bestimmten Ereignisklassen zugeteilt. Dabei werden die Daten nach ihren Objekten in den Endzuständen kategorisiert. Bei der Einteilung werden folgende physikalische Objekte unterschieden:

- Myonen (μ)
- Elektronen (e)
- Photonen (γ)
- Hadronischen Jets (jet)
- Fehlende transversale Energie (E_T^{miss})

Außerdem wurden noch B-Jets (b) implementiert, die eine Teilmenge der hadronischen Jets sind.

Sollten nach einer Teilchenkollision z.B. ein Myon, ein Elektron und zwei Jets identifiziert werden, so werden diese in die Ereignisklasse „1e+1 μ +2jet“ sortiert. Anhand der Ereignisklassen kann man jedoch nicht erkennen, ob es sich um positiv oder negativ geladene Leptonen handelt. Bei der Einteilung in die Ereignisklassen werden zwei Klassen unterschieden: Exklusive Klassen, in denen genau die aufgeführten Teilchen vorkommen, jedoch keine anderen und inklusive Klassen, die mindestens die dargestellten Teilchen enthalten, aber noch weitere, die durch ein „+X“ gekennzeichnet werden. Die inklusive Klasse „1e+1 μ +X“ beinhaltet also ein Elektron und Myon, sowie eine unbekannte Menge weiterer Objekte. Abbildung 4 zeigt die Klasseneinteilung anhand eines „1e+1 μ +2jet“ Ereignisses.

4.1.3 Suche nach Abweichungen

Um neue Physik zu finden müssen Ereignisse mit der Monte-Carlo-Vorhersage verglichen werden. Dies geschieht anhand von Variablen, die sehr sensitiv für neue Physik sein sollten. Aus diesem Grund erzeugt MUSiC Verteilungen zu den folgenden Variablen:

- $\sum p_T$: Die skalare Summe der transversalen Impulse der Objekte.
- M_{inv} : Die invariante Masse (oder M_T : Die transversale invariante Masse für Klassen mit E_T^{miss}).
- E_T^{miss} : Die fehlende transversale Energie für Klassen, in denen diese größer als 30 GeV ist.

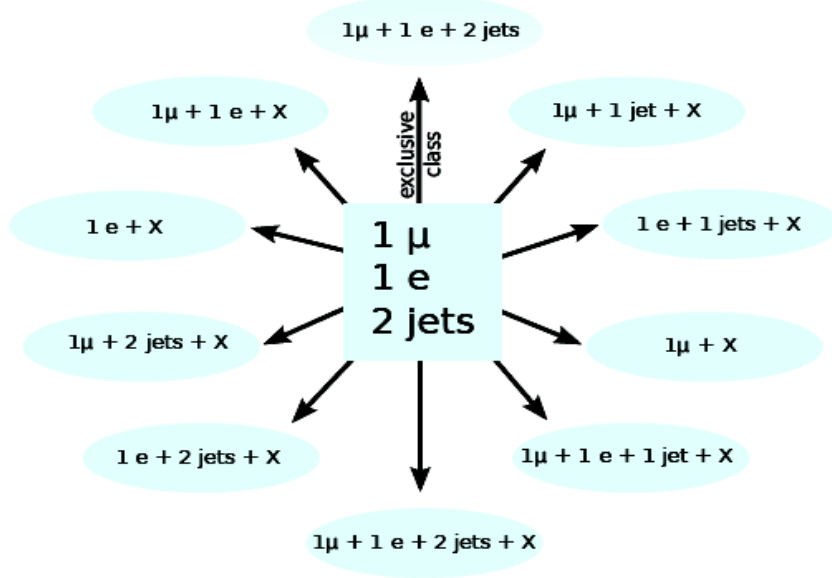


Abbildung 4: Klassifizierung für ein Event mit einem Myon, einem Elektron und zwei Jets [11].

4.2 Der Suchalgorithmus

Wie bereits erwähnt, werden die selektierten physikalischen Objekte in Ereignisklassen eingeteilt, zu denen die drei Verteilungen $\sum p_T$, M_{inv}/M_T und E_T^{miss} generiert werden. Diese werden anschließend mithilfe eines Algorithmus mit den Monte-Carlo-Simulationen verglichen und so die Abweichungen der Daten zu den Erwartungen des Standardmodells quantifiziert. Dies geschieht in mehreren Schritten.

4.2.1 Schritt 1: Bestimmung der Interessanten Region

Damit man die Abweichungen vom Standardmodell berechnen kann, werden die drei Verteilungen in mehrere Bereiche bzw. Regionen unterteilt. Eine Region besteht aus einem oder mehreren Bins. Dabei können die Bins mehreren Regionen angehören, d.h. die Regionen können sich gegenseitig überlappen. Für jede Region wird ein sogenannter p-Wert ermittelt, der die Wahrscheinlichkeit dafür angibt, dass ein Ergebnis nur von den statistischen Schwankungen des Untergrundes abhängt. Der p-Wert gibt also an, dass die Monte-Carlos ähnlich wie Daten fluktuieren können.

Durch einen Vergleich der Summen der Monte-Carlo-Daten mit den gemessenen Daten wird der p-Wert bestimmt. Bei den Monte-Carlo-Daten ergibt sich als Wahrscheinlichkeitsfunktion eine Poisson-Verteilung, als Fehler für den Mittelwert der Monte-Carlo-Simulation wird eine Gaußverteilung mit der Breite

$$\sigma_B = \sqrt{\sigma_{stat}^2 + \sum_i \sigma_{i,syst}^2} \quad (6)$$

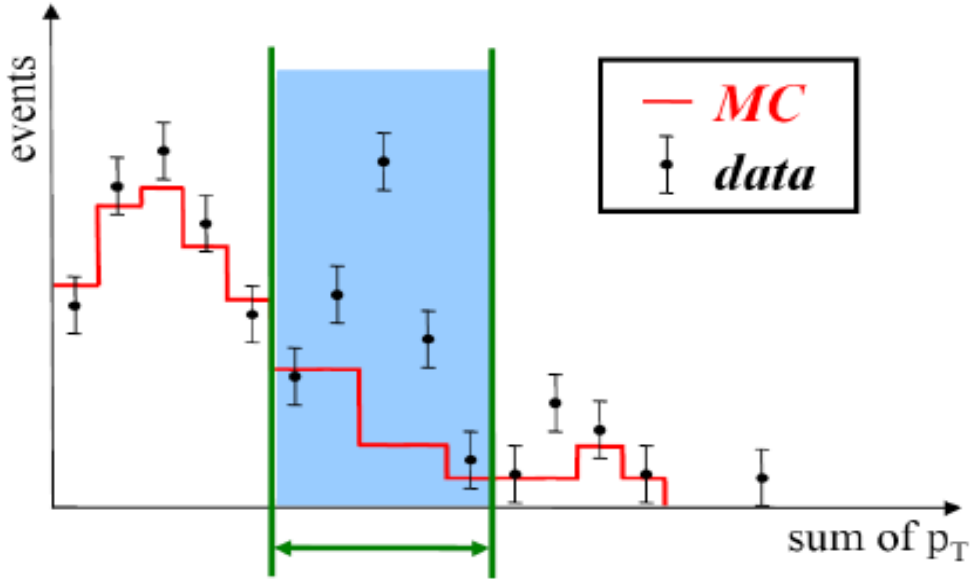


Abbildung 5: Dargestellung der interessanten Region [12].

angenommen, wobei σ_{stat} der statistische Fehler und $\sigma_{i,syst}$ die systematischen Fehler sind. Durch eine Faltung der Poisson-Verteilung und der Gaussverteilung lässt sich der p-Wert darstellen durch

$$p = \begin{cases} \sum_{i=N_{data}}^{\infty} A \cdot \int_0^{\infty} dx \exp\left(\frac{-(x-B)^2}{2\sigma_B^2}\right) \cdot \frac{e^{-x} x^i}{i!} & \text{für } N_{data} \geq B \\ \sum_{i=0}^{N_{data}} A \cdot \int_0^{\infty} dx \exp\left(\frac{-(x-B)^2}{2\sigma_B^2}\right) \cdot \frac{e^{-x} x^i}{i!} & \text{für } N_{data} < B \end{cases}, \quad (7)$$

wobei N_{data} die Anzahl der gemessenen Ereignisse, B die Anzahl der Monte-Carlo-Ereignisse und A ein Normierungsfaktor ist. Die Region mit dem kleinsten p-Wert p_{min}^{data} wird als interessante Region (Region of Interest) bezeichnet (Abb. 5), da sich in dieser Region die größte Abweichung der gemessenen Daten von der Monte-Carlo-Simulation des Standardmodells befindet.

4.2.2 Schritt 2: Bestimmung der Signifikanz

Durch den p-Wert alleine lassen sich keine Aussagen über die Signifikanz einer Region treffen. Der Grund ist, dass man mit einer gewissen Wahrscheinlichkeit ohnehin eine Abweichung findet. Dabei ist entscheidend, wieviele Regionen untersucht wurden, denn bei vielen Regionen ist die Wahrscheinlichkeit, eine Abweichung zu finden, entsprechend größer. Dieser Effekt wird als „look elsewhere effect“ bezeichnet. Die Signifikanz \tilde{P} wird wie folgt ermittelt:

MUSIC führt sogenannte Pseudoexperimente durch, bei denen Daten zufällig erzeugt werden. Dabei wird angenommen, dass nur der Monte-Carlo-Untergrund mit seinen statistischen und systematischen Fehlern zu der Ermittlung des p-Wertes beiträgt, die Monte-Carlo

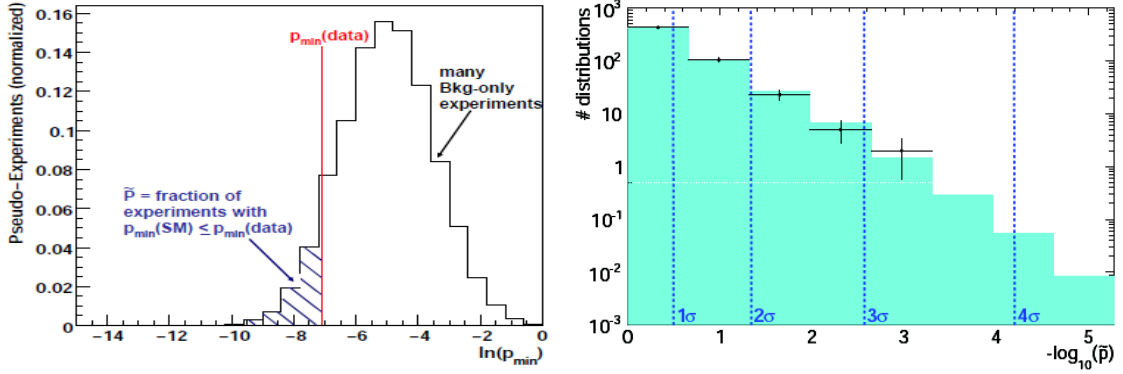


Abbildung 6: links: Bestimmung des \tilde{P} -Wertes [12], rechts: Typisches \tilde{P} -Histogramm [11].

Erwartung also die Verteilung ist, aus der zufällig gezogen wird. Diese Annahme wird Null-Hypothese genannt. Mit den zufällig erzeugten Daten wird Schritt 1 wiederholt, es wird also ein Wert p_{min}^{SM} für jedes der Pseudoexperimente berechnet. Dadurch berechnet sich \tilde{P} durch

$$\tilde{P} = \frac{\text{Anzahl der Pseudoexperimente mit } p_{min}^{SM} \leq p_{min}^{data}}{\text{Anzahl der Pseudoexperimente}}. \quad (8)$$

Die Signifikanz \tilde{P} berechnet sich also aus dem Quotient der Pseudoexperimente mit einer signifikanteren Abweichung als bei den eigentlichen Daten und der Gesamtzahl der Pseudoexperimente. Dies wird in Abbildung 6 in der Grafik auf der linken Seite dargestellt.

Neben dem „look elsewhere effect“ innerhalb einer Klasse, der durch den \tilde{P} -Wert korrigiert wird, kann es noch zu einem globalen „look elsewhere effect“ kommen. Dabei wird angenommen, dass von den untersuchten Klassen eine bestimmte Anzahl abweicht, abhängig von der Gesamtzahl der untersuchten Klassen. MUSiC erstellt zu diesem Zweck ein Histogramm, mit dem man die ermittelten \tilde{P} -Werte mit den erwarteten Werten vergleichen kann (Abb. 6 rechts).

4.3 Selektion der Objekte

Wie bereits erwähnt werden bei einer modellunspezifischen Suche keine modellabhängigen Optimierungen der Schnitte durchgeführt, um ein bestimmtes Signal im Untergrund zu finden. Schnitte werden nur an den Ereignissen vollzogen, die überhaupt die „High-Level-Trigger“ passieren (Tabelle 1).

Außerdem muss mindestens ein Myon bzw. mindestens ein Elektron einen transversalen Impuls $p_T > 25$ GeV bzw. $p_T > 30$ GeV haben. Hinzu kommen Schnitte, um „schlechte“ Ereignisse auszusortieren. Die Rekonstruktion des Primärvertex muss folgende Kriterien erfüllen:

- Mindestens vier Spuren werden zur Rekonstruktion genutzt.
- Der Vertex ist maximal 24 cm in z-Richtung vom Kollisionspunkt entfernt.

Trigger	Erster Lauf	Letzter Lauf	$\int \mathcal{L} dt$
HLT_Mu9	135808	147116	8.3 pb ⁻¹
HLT_Mu11	147117	148818	9.5 pb ⁻¹
HLT_Mu15_v1	148819	149063	18.5 pb ⁻¹
HLT_Ele20_LW_L1R	135808	141949	0.3 pb ⁻¹
HLT_Ele20_SW_L1R	141950	144114	2.9 pb ⁻¹
HLT_Ele17_SW_EleId_L1R	144115	147116	5.1 pb ⁻¹
HLT_Ele17_TightEleId_L1R	147117	148818	9.5 pb ⁻¹
HLT_Ele22_SW_TightEleId_L1R_v2	148819	149063	10.3 pb ⁻¹
HLT_Ele22_SW_TightEleId_L1R_v3	149964	149442	8.1 pb ⁻¹

Tabelle 1: Tabelle der verwendeten High-Level-Trigger für Myonen und Elektronen.

- Der Vertex ist maximal 2 cm senkrecht zur z-Richtung vom Kollisionspunkt entfernt.

Es werden außerdem noch jene Ereignisse aussortiert, die mehr als 10 Spuren haben und von denen 25% schlecht rekonstruiert wurden. Neben diesen globalen Schnitten werden noch einzelne Selektionskriterien für die einzelnen Objekte durchgeführt.

4.3.1 Myonen

Als Myonen in MUSiC werden sogenannte „globale Myonen“ verwendet, die aufgrund des Detektoraufbaus und der Trigger-Akzeptanz eine Pseudorapidität von $|\eta| < 2.1$ besitzen. Die gesamten Selektionskriterien sind:

- Das Myon muss sowohl im Spurdetektor, als auch global detektiert werden.
- Der transversale Impuls beträgt $p_T > 18$ GeV.
- Die Pseudorapidität beträgt $|\eta| < 2.1$.
- Abstand vom Primärvertex: $|d_{xy}| < 0.2$ cm.
- Für die Spurrekonstruktion muss das globale $\chi^2/N_{DOF} < 10$ sein.
- Mindestens 11 Treffer im Spurdetektor, davon mindestens einer im Pixel-Detektor.
- Mindestens 2 getroffene Segmente im Myonendetektor, wovon mindestens eines zur Spurrekonstruktion passt.
- Spurisolation $\sum p_T < 3$ GeV, $\Delta R < 0.3$.

4.3.2 Elektronen

Die genaue Rekonstruktion der Elektronen ist ein wenig komplizierter als bei den Myonen. Während die Elektronen im ECAL abgebremst werden, wo auch deren Energie gemessen wird, entstehen Photonen durch Bremsstrahlung. Diese müssen bei der Rekonstruktion der Elektronenenergie berücksichtigt werden. Zu diesem Zweck werden die Elektronencluster

und die Photoncluster zu einem Supercluster kombiniert. Die Ermittlung der Energie von Elektronen und Photonen wird durch den selben Cluster-Algorithmus durchgeführt. Zur genauen Unterscheidung von Elektronen und Photonen müssen die Treffer im Spurdetektor hinzugezogen werden. Während Elektronen dort detektiert werden, sollten die Photonen ohne zu interagieren diesen Bereich passieren.

- Transversaler Impuls $p_T > 25$ GeV.
- Pseudorapidität $|\eta| < 2.5$.
- Verhältnis von HCAL zu ECAL: $H/E = 0.05$.
- $\Delta\phi < 0.09$ zwischen Spurtrajektorie und Supercluster des ECALs.

Für die Elektronen mit $|\eta| < 1.442$ gelten außerdem folgende Bedingungen:

- $\Delta\eta < 0.05$ zwischen der Spurtrajektorie und dem ECAL Supercluster.
- $E_{2x5}/E_{5x5} > 0.94$ oder $E_{1x5}/E_{5x5} > 0.83$.
- Elektromagnetische Isolation addiert mit der hadronischen Isolation von Auslesebereich 1 des HCALs: $E + H1 < 2 + 0.03 \cdot E_T$.
- Spurisolation: $\sum p_T < 7.5$ GeV.
- Energieanteil in den beiden energiereichsten Kristallen: $E_2/E_9 < 0.9$.

Für Elektronen mit $1.56 < |\eta| < 2.5$ gelten die Bedingungen:

- $\Delta\eta < 0.07$ zwischen der Spurtrajektorie und dem ECAL Supercluster.
- $\sigma_{i\eta i\eta} < 0.03$.
- Elektromagnetische Isolation addiert mit der hadronischen Isolation in Auslesebereich 1 des HCALs: $E + H1 \text{ Isolation} < 2.5 \text{ GeV} + 0.03 \cdot (E_T - 50 \text{ GeV})$.
- Hadronische Isolation in Auslesebereich 2 des HCALs: $H2 \text{ Isolation} < 0.5$ GeV.
- Spurisolation $\sum p_T < 15$ GeV.

4.3.3 Photonen

Die Energie von Photonen wird ebenso wie die Energie von Elektronen im ECAL ermittelt. Da sich die Photonen durch Materie, dem Detektormaterial, bewegen, besteht eine bestimmte Wahrscheinlichkeit, dass es zu Elektron-Positron Paarbildung kommt. Die Wahrscheinlichkeit dafür liegt in einer Größenordnung von ca. 10%. Die genaue Rekonstruktion der durch Paarbildung entstandenen Teilchen ist wichtig für die Ermittlung der Energie von Photonen.

- Transversaler Impuls $p_T > 25$ GeV.

- Pseudorapidity $|\eta| < 1.442$.
- Verhältnis von HCAL zu ECAL: $H/E < 0.05$.
- Keine Treffer im Pixeldetektor.
- Spurisolation im Hohlkegel $< 3.5 \text{ GeV} + 0.001 \cdot p_T$.
- ECAL Isolation $< 4.2 \text{ GeV} + 0.006 \cdot p_T$.
- HCAL Isolation $< 2.2 \text{ GeV} + 0.0025 \cdot p_T$.
- Energieanteil in den beiden energiereichsten Kristallen: $E_2/E_9 < 0.9$.

4.3.4 Jets

Bei der Selektion der Jets werden „loose Particle Flow Jets“ mit einem transversalen Impuls von $p_T > 50 \text{ GeV}$ und einem Winkel $|\eta| < 2.5$ verwendet. „Loose“ bedeutet, dass nur die wirklich schlecht rekonstruierten Jets aussortiert werden. Das führt zwar zu einer Großzahl erkannter Jets, aber auch zu einer größeren Anzahl fehlerhaft erkannter Jets. Die einzelnen Selektionskriterien sind:

- Anteil neutraler Hadronen < 0.99 .
- Anteil elektrisch geladener Hadronen < 0.99 .
- Anteil der Bestandteile > 1 , um einzelne Teilchen auszusortieren.

Für eine Pseudorapidity $|\eta| > 2.4$ werden außerdem noch folgende Selektionen durchgeführt:

- Anteil geladener Hadronen > 0 .
- Anzahl der Ladungen > 0 .
- Elektromagnetisch geladener Anteil < 0.99 .

4.3.5 Fehlende transversale Energie (MET)

Damit Ereignisse als fehlende transversale Energie besitzend akzeptiert werden, muss durch den „Particle Flow Jet“ Algorithmus mindestens eine MET¹⁰ von $E_T^{miss} > 30 \text{ GeV}$ ermittelt werden.

¹⁰missing energy transversal (fehlende transversale Energie).

4.3.6 Bereinigung der Ereignisse

Manchmal werden Teilchen nicht gut genug und eindeutig rekonstruiert, sodass diese doppelt gezählt werden. Im ECAL können z.B. Elektronen ebenso Photonen sein. Aus diesem Grund werden mit den folgenden Schnitten solche Ereignisse aussortiert:

- Bei Myonen, die näher als $\Delta R < 0.2$ beieinander liegen, wird nur das Myon mit dem kleineren χ^2/N_{DOF} bei der Spurrekonstruktion verwendet.
- Wenn zwei mögliche Elektronen näher als $\Delta R < 0.2$ beieinander liegen und dieselbe Spur besitzen, wird nur das höherenergetische Elektron ausgewählt.
- Bei Photonen gilt dasselbe Kriterium, allerdings werden alle Photonen, die eine Spur mit Elektronen teilen, aussortiert.
- Jets, die nur $\Delta R < 0.2$ von ausgewählten Photonen oder Elektronen entfernt liegen, werden aussortiert.

4.4 Systematische Messunsicherheiten

Eine ausreichende Kenntnis der systematischen Messunsicherheiten ist notwendig, um richtige, unerwartete Abweichungen von den erwarteten, aber nicht unbedingt genau bekannten Dektoreffekten unterscheiden zu können. Die bekannten und in MUSiC einbezogenen Messunsicherheiten sind [11, 13]:

- Luminosität: 4%.
- Wirkungsquerschnitt der Ereignisse.
- Unsicherheiten innerhalb der „Parton Verteilungsfunktionen“ (PDF).
- Jet-Energie-Skalierung: Abhängig von p_T und η , normalerweise im Bereich von ein paar Prozent.
- Rekonstruktionseffizienz: 3% für Elektronen und Photonen, 4% für Myonen und 1% für Jets.
- Wahrscheinlichkeit für vorgetäuschte Objekte: 100% für Elektronen, 50% für Myonen und 30% für Photonen.
- Fehler innerhalb der Monte-Carlo-Simulationen, abhängig von der Größe der Monte-Carlos.

Für eine konsistente Handhabung der verschiedenen Messunsicherheiten beachtet der Algorithmus die Korrelation zwischen den einzelnen systematischen Messunsicherheiten.

5 Erkennung von B-Jets

Eine wichtige Rolle bei der Suche nach neuer Physik spielt die Erkennung von B-Jets. Bei B-Jets handelt es sich um Jets, die ihren Ursprung in Bottom-Quarks haben. Diese sind deswegen so interessant, weil es die schwersten hadronisierenden Quarks sind. Das deutlich schwerere Top-Quark hat eine zu kurze Halbwertszeit, um zu hadronisieren. Allerdings zerfallen Top-Quarks häufig in Bottom-Quarks. Dies ist ein weiterer Grund, warum die Analyse von B-Jets so interessant ist. Viele bekannte, aber nur bedingt erforschte Zerfallskanäle bringen B-Quarks hervor, die in Jets hadronisieren. Außerdem sagen viele Theorien für Zerfallskanäle, in denen neue Physik erwartet wird, die Entstehung von B-Quarks und damit von B-Jets voraus. Im folgenden soll ein Einblick in die Physik von B-Quarks und in die verschiedenen Algorithmen zur Erkennung von B-Jets gegeben werden.

5.1 Das Bottom-Quark

Bevor die verschiedenen Möglichkeiten zur Erkennung von B-Jets diskutiert werden, soll ein genereller Einblick in die Physik des B-Quarks gegeben werden. Wie bereits in Abschnitt 2.1 erwähnt, hat das B-Quark eine Ladung von $-\frac{1}{3}e$ und eine Masse von $m_b = 4.5$ GeV. Außerdem gehört es der 3. Generation der Quarks an, sein Partner ist das besonders schwere Top-Quark. Das B-Quark ist das schwerste hadronisierende Quark, was seine Jets sehr interessant bei der Suche nach neuer Physik macht, da neue Teilchen mit relativ hoher Wahrscheinlichkeit in B-Quarks zerfallen. Aus diesem Grund wurde der Spurdetektor des CMS auch auf die Suche nach B-Jets ausgerichtet. Ein weiterer Vorteil von B-Jets ist, dass die entstehenden B-Hadronen aufgrund ihrer hohen Masse, ihrer großen Halbwertszeit und ihrer großen Anzahl an semileptonischen Zerfällen leicht von anderen Jets unterschieden werden können. In erster Linie kann es zu Verwechslungen mit C-Jets¹¹ kommen [14]. Durch ihre Lebensdauer erreichen die B-Hadronen genau den Spurdetektor, wodurch man die Jet-Vertizes relativ gut rekonstruieren kann. Interessante, teilweise noch nicht beobachtete Zerfallskanäle sind [15, 16]:

- Der Zerfall des Top-Paares: $t\bar{t} \rightarrow b\bar{b}W^+W^-$, dieser Zerfall wurde bereits beobachtet.
- Der Zerfall einer hypothetischen 4. Generation Quarks:
 $t'\bar{t}' \rightarrow b\bar{b}W^+W^-$, $b'\bar{b}' \rightarrow t\bar{t}W^+W^- \rightarrow b\bar{b}W^+W^-W^+W^-$
- Der Zerfall des Single-Top, eines einzelnen Top-Quarks: $t \rightarrow bW^+$
- Der Zerfall von schweren Eichbosonen, die ebenfalls noch nicht experimentell beobachtet wurden:
 $Z' \rightarrow t\bar{t} \rightarrow b\bar{b}W^+W^-$, $W'^{\pm} \rightarrow t\bar{b}/\bar{t}b \rightarrow b\bar{b}W^{\pm}$

Besonders die letzten drei Beispiele sind auf der Suche nach neuer Physik interessant. Außerdem zerfällt das Z^0 -Boson häufig in B-Quarks.

¹¹Jets, die aus Charm-Quarks bestehen.

5.2 Algorithmen zur Erkennung von B-Jets

Um B-Jets möglichst gut erkennen zu können, gibt es eine Vielzahl von B-Jet Erkennungsalgorithmen („taggers“). Diese variieren teilweise stark in der Art und Weise, wie die B-Jets erkannt werden sollen. Dabei unterscheidet man vier Arten von Algorithmen. Diese werden im nachfolgenden aufgeführt und erklärt.

5.2.1 Stoßparameterbasierende Algorithmen

Als Stoßparameter¹² bezeichnet man den kürzesten Abstand vom Primärvertex zu einer rekonstruierten Spur. Abbildung 7 zeigt den Impact Parameter einer Spur, die vom Sekundärvertex ausgeht. Impact Parameter basierende Algorithmen sind relativ einfache Algorithmen, insgesamt gibt es vier Stück, die sich in zwei Gruppen unterordnen lassen. Zum einen gibt es die spurzählenden¹³ Algorithmen und zum anderen die Jet-Wahrscheinlichkeitsalgorithmen¹⁴.

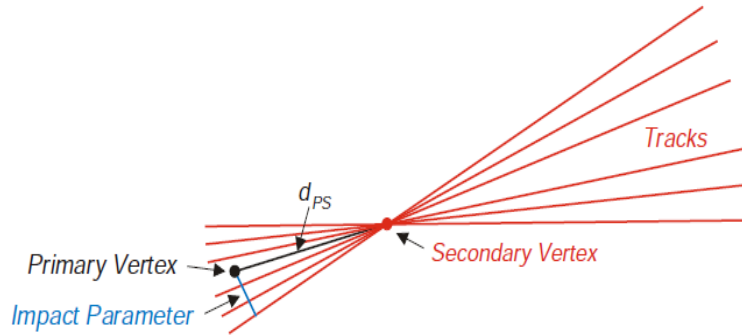


Abbildung 7: Darstellung des Impact Parameters [15].

Die spurzählenden Algorithmen sind die einfacheren Algorithmen. Dabei werden die Spuren vom Primärvertex der Größe des Stoßparameters d_{IP} nach durchgezählt und anschließend die IP-Signifikanz (Impact Parameter Signifikanz) der n-ten Spur als Diskriminator genutzt. Die IP-Signifikanz [17] ist definiert durch

$$S = \frac{d_{IP}}{\sigma_{d_{IP}}}. \quad (9)$$

Für $n=2$, also bei Betrachtung der zweiten Spur, handelt es sich um die Version des Algorithmus für eine hohe Effektivität. Bei Betrachtung der dritten Spur, also $n=3$, handelt es sich um die Version für hohe Reinheit mit der geringsten Fehlerquote bei der Identifizierung der B-Jets. Der Nachteil ist allerdings, dass viele B-Jets nicht als solche erkannt werden und normalen Jets zugeteilt werden.

¹²engl.: Impact Parameter (IP)

¹³engl.: Track Counting (TC)

¹⁴engl.: Jet Probability

Der Jet Probability Algorithmus berechnet abhängig vom IP die Wahrscheinlichkeit, ob die entsprechende Spur vom Primärvertex stammt. Aus den einzelnen Wahrscheinlichkeiten wird ermittelt, wie hoch die Wahrscheinlichkeit ist, dass alle Spuren ihren Ursprung im Primärvertex haben. Bei einer geringen Wahrscheinlichkeit handelt es sich eher um einen B-Jet. Der „Jet B Probability“ Algorithmus, der eine leichte Abwandlung ist, ermittelt die vier Spuren, die am ehesten zu einem B-Jet passen und gewichtet diese verstärkt bei der Berechnung der globalen Wahrscheinlichkeit. Der Vorteil der Jet Probability Algorithmen gegenüber den Track Counting Algorithmen ist, dass alle Spuren mit in die Berechnung des Diskriminators eingehen [18].

5.2.2 Leptonenbasierende Algorithmen

Die zweite Art von Algorithmen zur B-Erkennung sind die auf Leptonen basierenden Algorithmen. Dort gibt es zwei Unterscheidungen: Zum einen den „Soft Electron“ und zum anderen den „Soft Muon“ Algorithmus. Diese Algorithmen verwenden die Eigenschaften der semileptonischen B-Jet Zerfälle, um einen globalen Diskriminator zu bilden. So sollte z.B. der transversale Impuls einen Großteil der Masse von B-Hadronen beinhalten. Die verwendeten Diskriminatorvariablen sind:

- Der Abstand ΔR des Leptons von der Jetachse.
- Die Signifikanz des IP von der Leptonenspur.
- Der transversale Impuls der Leptonen, er sollte im Bereich von ungefähr 5 GeV liegen. Dies ist die durchschnittliche Masse von B-Hadronen, da die Masse eines B-Quarks ungefähr 4.5 GeV beträgt.
- Das Verhältnis des Leptonenimpulses und der gemessenen Jetenergie.

Diese Variablen werden sowohl für Elektronen, als auch für Myonen verwendet. Für Myonen gibt es noch zwei weitere Variablen, um Detektoreffekte bei der Rekonstruktion von Myonen zu beachten:

- Die Pseudorapidität des Jets.
- Die kalibrierte Jet Energie.

Die auf Leptonen basierenden Algorithmen haben eine relativ geringe Fehlerquote, d.h. es werden wenig Jets fälschlicherweise als B-Jets erkannt. Da aber nur 30% der B-Hadronen semileptonisch zerfallen, werden entsprechend viele B-Jets nicht als solche erkannt, d.h. die Algorithmen sind gut für eine hohe Reinheit, aber nicht für eine hohe Effektivität [18].

5.2.3 Sekundärvertex basierende Algorithmen

Bei sekundärvertexbasierenden Algorithmen werden Rekonstruktionen der Spuren, die von Sekundärvertices ausgehen, genutzt, um B-Jets zu erkennen. Kann man den Sekundärvertex genau bestimmen, so wird die Distanz zwischen dem Primär- und dem Sekundärvertex

als Diskriminator verwendet. Da diese Distanz von der Zerfallslänge abhängt, kann man so B-Jets leicht von anderen Jets, z.B. C-Jets, unterscheiden. Der Nachteil dieses Algorithmus ist die Schwierigkeit, den Sekundärvertex hinreichend genau zu rekonstruieren. Dafür verspricht dieser Algorithmus eine hohe Effizienz. Es gibt zwei Algorithmen, die auf Sekundärvertexrekonstruktion aufbauen: Den „Simple Secondary Vertex“ (SSV) Algorithmus und den „Combined Secondary Vertex“ (CSV) Algorithmus.

Während der SSV-Algorithmus nur Vertices von einer erfolgreichen Rekonstruktion des Sekundärvertex verwendet, nutzt der CSV-Algorithmus Variablen von sogenannten Pseudoverteices. Ein Pseudoverteice wird verwendet, wenn kein Sekundärvertex rekonstruiert werden konnte, aber mindestens zwei Spuren eine IP-Signifikanz von 2.0 haben [18].

5.2.4 Kombinierte Algorithmen

Neben den drei einzelnen Methoden gibt es noch einen kombinierten Algorithmus, der als Grundlage den CSV-Algorithmus verwendet: Der „Combined MVA“ Algorithmus. Dieser kann jegliche Information aus den anderen Algorithmen nutzen, um den Diskriminator zu bilden. Dabei kann eingestellt werden, wie groß der jeweilige Anteil der entsprechenden Variablen sein soll, also ob der Diskriminator des Track Counting Algorithmus gegenüber den anderen Algorithmen überwiegen soll. Dadurch lassen sich Effektivität und Reinheit entsprechen anpassen. Der Nachteil liegt in der großen Zahl von Informationen, die einfließen. Diese steigert auch die Anzahl und Größe der Fehler bei der B-Jet-Identifikation [18].

5.2.5 Messunsicherheiten

Bei der Suche nach B-Jets gibt es einige Gründe für Messunsicherheiten, die z.B. durch Teilchen mit ähnlicher Zerfallslänge oder Masse entstehen können. Typische Messunsicherheiten sind [18]:

- Eine der größten Fehlerquellen ist Fehlidentifikation von B-Jets, bei denen es sich eigentlich um C-Jets handelt. Jets mit C-Hadronen haben Massen in der Größenordnung von B-Hadronen, wodurch sich ihre Spuren sehr stark ähneln. Das macht eine Fehlidentifikation relativ wahrscheinlich. Die Unsicherheit beträgt ungefähr $\pm 20\%$.
- Der Anteil an Gluonen, welcher abhängig von der Partondichte und der Hadronisierung ist. Dieser Fehler wirkt sich auch auf MC-Simulationsdaten aus. Der Fehler beträgt ebenfalls ungefähr $\pm 20\%$.
- Kernwechselwirkungen und Mehrfachmessungen von Photonen im Pixeldetektor ($\pm 5\%$).
- Falsch rekonstruierte Spuren, der Fehler beträgt ungefähr $\pm 50\%$.

6 B-Erkennung in MUSiC

Um die Anwendung von B-Erkennung auf Daten in MUSiC zu testen, werden fünf Kombinationen von B-Erkennungsalgorithmen und Diskriminatoren verwendet (Tabelle 2). Der Combined Secondary Vertex MVA mit einem Diskriminator von 0.8 ist auf hohe Effektivität abgestimmt [15]. Bei den Track Counting Algorithmen wurde jeweils die Variante „loose“ und „tight“ gewählt. Loose bedeutet, dass der Anteil der fehlerhaft als B-Jet identifizierten Jets 10% beträgt, während tight für eine Fehlerkennungsrate von 0.1% steht. Demnach hat sowohl der Track Counting High Purity bei einem Diskriminator von 3.41 als auch der Track Counting High Efficiency bei einem Diskriminator von 10.2 eine Fehlerquote von nur 0.1%. Die Vergleiche sollen zeigen, wie sich Unterschiede in den Algorithmen in MUSiC auswirken. In allen Analysen wurden nur Klassen analysiert, bei denen der p-Wert kleiner als 3σ ist. Außerdem werden in der Analyse ausschließlich exklusive Klassen betrachtet, also Klassen, die nur die angegebenen Komponenten beinhalten.

Algorithmus	Diskriminator
Combined Secondary Vertex MVA (CSVMVA)	0.8
Track Counting High Purity ¹⁵ (TCHP)	1.19
	3.41
Track Counting High Efficiency ¹⁶ (TCHE)	1.7
	10.2

Tabelle 2: Verwendete Algorithmen mit Diskriminatoren zur B-Erkennung.

6.1 Sensitivitätsstudie von B-Erkennung in MUSiC

Bevor mit der eigentlichen Analyse der Daten begonnen wird, soll eine Sensitivitätsstudie zeigen, dass die Nutzung von B-Erkennung in MUSiC die Anzahl der signifikanten Ereignisklassen vergrößert bzw. Klassen mit hoher Signifikanz einfacher entdeckt werden. Der verwendete Algorithmus für diese Untersuchung ist der Track Counting High Purity tight (TCHPt). Dadurch kann man ziemlich sicher sein, dass die als B-Jet erkannten Jets wirklich B-Hadronen beinhalten. Der Nachteil ist jedoch, dass viele B-Jets aussortiert werden und in die Klassen mit normalen Jets fallen.

Um die Sensitivität zu messen, werden Modellprozesse, die nicht dem Standardmodell entsprechen, auf den Untergrund der Standardmodell Monte-Carlos addiert. Dann werden zufällige Pseudodaten der sich ergebenden Verteilung entsprechend erzeugt.

Als Untergrundprozesse werden LM-Punkte für Supersymmetrie und Zerfälle des Single-Top gewählt. In beiden Studien werden nur Klassen betrachtet, die fehlende transversale Energie beinhalten.

6.1.1 Supersymmetrie

Bei der Sensitivitätsstudie mithilfe von Supersymmetrie wurden vier LM-Punkte verwendet (Tabelle 3). Dabei wurde eine maximale integrierte Luminosität von 10 fb^{-1} verwendet, da sehr viel höhere Werte nicht zu erwarten sind. Die LM-Punkte LM2 und LM3 zeigen bei

einer integrierten Luminosität von 10 fb^{-1} noch keine auffälligen Signifikanzen, unabhängig davon, ob B-Erkennung eingeschaltet ist oder nicht. Beim Punkt LM0 ergeben sich erste Signifikanzen bei ungefähr 0.3 fb^{-1} . Durch den hohen Wirkungsquerschnitt kann man allerdings keine Unterschiede zwischen den Signifikanzen der Klassen mit B-Erkennung zu den Klassen ohne B-Erkennung finden. Einzig der Punkt LM9 lohnt sich für eine genauere Betrachtung.

LM-Punkt	Wirkungsquerschnitt (pb^{-1})
LM0	38.93
LM2	0.6027
LM3	3.438
LM9	7.134

Tabelle 3: LM-Punkte mit dazugehörigem Wirkungsquerschnitt.

Beim LM9 Punkt zeigt sich der Vorteil von B-Erkennung in MUSiC bereits, wenn man alle Klassen mit einer integrierten Luminosität von 10 inversen Femtobarn betrachtet. Während mit eingeschalteter B-Erkennung 17 Klassen mit einer Signifikanz von mehr als 4σ (ca. $6 \cdot 10^{-5}$) auftreten, finden sich bei ausgeschalteter B-Erkennung nur 6 Klassen mit einer solchen Signifikanz (Tabelle 4). Sucht man nun die integrierte Luminosität, bei der mindestens eine Klasse immer ungefähr eine Signifikanz von 4σ hat, so erhält man mit B-Jets eine integrierte Luminosität von 3.2 fb^{-1} und ohne B-Jets 4.0 fb^{-1} , d.h. mit B-Erkennung braucht man ungefähr 20% weniger Daten, um eine entsprechende Signifikanz zu finden.

mit b-Jets	ohne b-Jets
$1e+3\text{jet}+3b+E_T^{\text{miss}}$	$1\mu+9\text{jet}+E_T^{\text{miss}}$
$1e+4\text{jet}+3b+E_T^{\text{miss}}$	$2\mu+7\text{jet}+E_T^{\text{miss}}$
$1e+5\text{jet}+2b+E_T^{\text{miss}}$	$1\mu+7\text{jet}+E_T^{\text{miss}}$
$1e+7\text{jet}+1b+E_T^{\text{miss}}$	$1e+7\text{jet}+E_T^{\text{miss}}$
$1\mu+4\text{jet}+3b+E_T^{\text{miss}}$	$2\mu+8\text{jet}+E_T^{\text{miss}}$
$1\mu+3\text{jet}+3b+E_T^{\text{miss}}$	$1e+9\text{jet}+E_T^{\text{miss}}$
$1\mu+6\text{jet}+2b+E_T^{\text{miss}}$	
$1\mu+7\text{jet}+1b+E_T^{\text{miss}}$	
$2e+5\text{jet}+1b+E_T^{\text{miss}}$	
$2e+6\text{jet}+E_T^{\text{miss}}$	
$2\mu+6\text{jet}+1b+E_T^{\text{miss}}$	
$2\mu+3\text{jet}+2b+E_T^{\text{miss}}$	
$2\mu+6\text{jet}+E_T^{\text{miss}}$	
$2\mu+4\text{jet}+1b+E_T^{\text{miss}}$	
$2e+7\text{jet}+E_T^{\text{miss}}$	
$1e+9\text{jet}+E_T^{\text{miss}}$	
$2\mu+7\text{jet}+E_T^{\text{miss}}$	

Tabelle 4: Abweichende Klassen des LM9-Punktes für eine integrierte Luminosität von 10 fb^{-1} .

6.1.2 Single-Top

Ein Prozess, bei dem B-Erkennung hilfreich sein könnte, ist die Suche nach dem Single-Top. Diese Form des Top-Quarks zerfällt mit einer hohen Wahrscheinlichkeit in ein B-Quark und ein W-Boson. Die entsprechend genutzten Signale sind TToBLNu_s_MG_PU, TToBLNu_s_MG_PU und TToBLNu_tW_MG_PU.

Auch mit B-Erkennung liefert diese Suche kein positives Ergebnis. Bei einer integrierten Luminosität von 10 fb^{-1} beträgt die maximale Abweichung knapp 3σ . Zwar sind die beiden signifikantesten Klassen mit „1 μ _1b_1met“ und „1e_1b_1met“ diejenigen, die man bei einem Single-Top Zerfall erwarten würde, aber ihre Signifikanz ist nicht ausreichend, um von einer wichtigen Abweichung zu sprechen.

6.2 Anwendung von B-Erkennung auf Daten

6.2.1 Combined Secondary Vertex MVA

Bei Anwendung des CSVMVA-Algorithmus auf Daten in MUSiC ergibt sich keine besonders signifikante Abweichung. Wenn man die einzelnen \tilde{p} -Diagramme (Abb. 8-10) der Verteilungsfunktionen betrachtet, erkennt man, dass die Anzahl der von der Theorie abweichenden Klassen im Bereich des look elsewhere Effektes liegen, diese Klassen also eher zufällig signifikant sind.

Die signifikanteste Klasse ist die „2e+1b“ Klasse der Verteilungsfunktion zur invarianten Masse. Dabei springt besonders der Datenpunkt bei 420 GeV ins Auge. Ein Durchführen der selben Auswertung mit einem Diskriminator von 0.9 hat diesen Punkt allerdings aus der Verteilung fallen lassen. Es handelt sich dementsprechend um einen fehlerhaft als B-Jet erkannte Jet. Die zweitsignifikanteste Klasse bei der invarianten Masse ist die „1 μ +1b+1met“ Klasse, welche auch die signifikanteste Klasse in den beiden anderen Verteilungen ist.

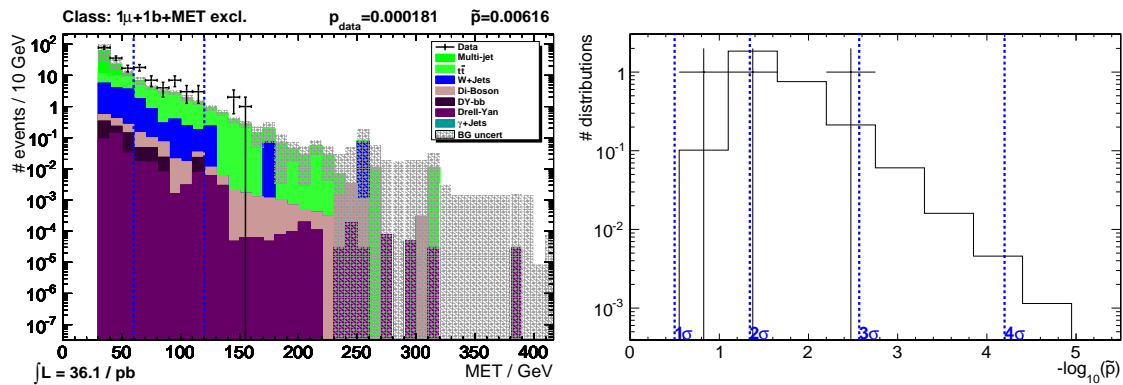


Abbildung 8: Die signifikanteste Klasse und das \tilde{p} -Histogramm der E_T^{miss} -Verteilung des CSVMVA-Algorithmus.

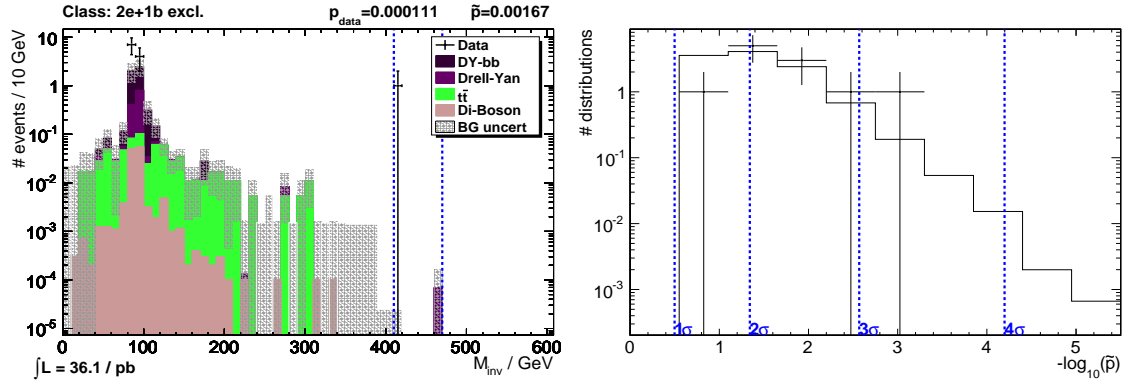


Abbildung 9: Die signifikanteste Klasse und das \tilde{p} -Histogramm der M_{inv} -Verteilung des CSVMVA-Algorithmus.

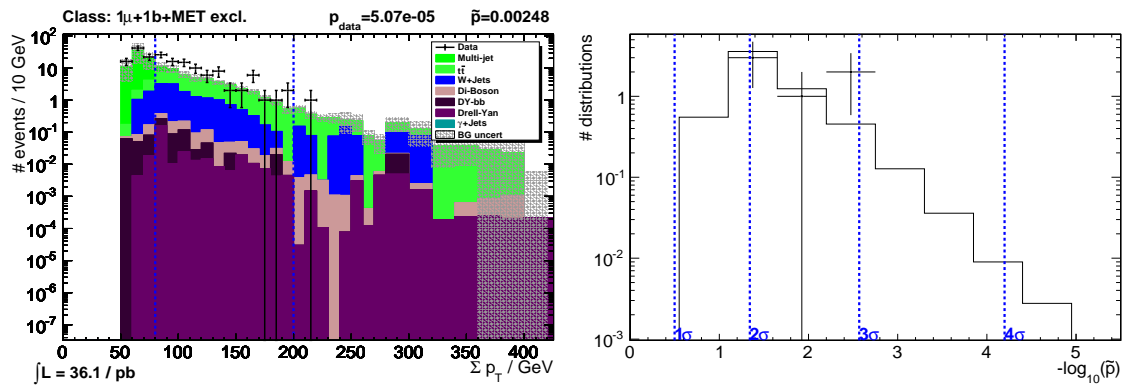


Abbildung 10: Die signifikanteste Klasse und das \tilde{p} -Histogramm der $\sum p_T$ -Verteilung des CSVMVA-Algorithmus.

6.2.2 Track Counting High Purity tight

Der TCHPt-Algorithmus ist auf eine hohe Reinheit der B-Hadronen abgestimmt. Bei Nutzung des TCHPt-Algorithmus erhält man mehr Abweichungen als bei Anwendung des CSVMVA-Algorithmus (Abb. 11-13). Allerdings handelt es sich bei den Abweichungen nur um kleine Signifikanzen im Bereich von 3σ . Eine große Abweichung mit über 4σ sucht man vergebens. Dennoch ist die größere Anzahl der Abweichungen um 3σ auffällig.

Die signifikanteste Klasse in allen drei Verteilungen ist die „1 μ +2b+1met“ Klasse. Man erkennt, dass sich bei der Verteilung des transversalen Impulses und der invarianten Masse mehr Datenpunkte im Massebereich der Eichbosonen (ca. 80 GeV - 90 GeV) befinden als erwartet. Bei der Verteilung der fehlenden transversalen Energie befinden sich mehr Daten im Bereich der halben Masse der Eichbosonen. Außerdem kann man den leichten Ansatz eines Sprunges bei 50 GeV erkennen. Wodurch die Unterschiede von Daten zu Monte-Carlos kommen, kann an dieser Stelle nicht geklärt werden.

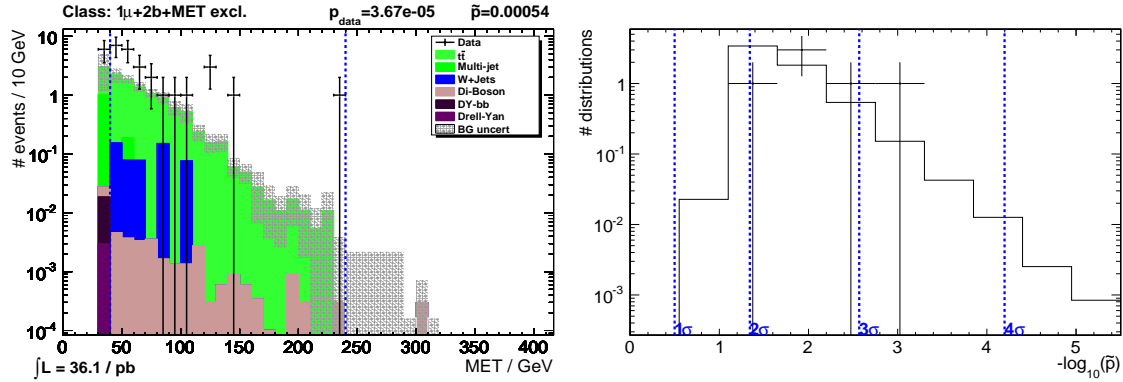


Abbildung 11: Die signifikanteste Klasse und das \tilde{p} -Histogramm der E_T^{miss} -Verteilung des TCHPt-Algorithmus.

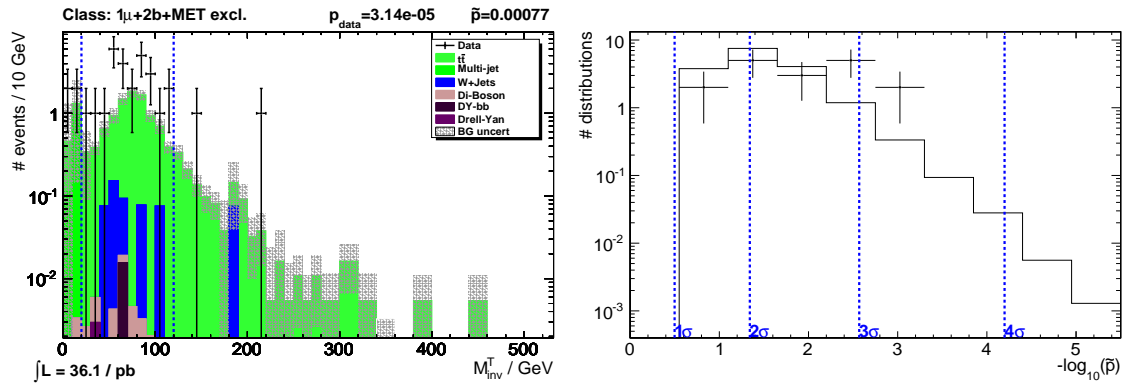


Abbildung 12: Die signifikanteste Klasse und das \tilde{p} -Histogramm der M_{inv} -Verteilung des TCHPt-Algorithmus.

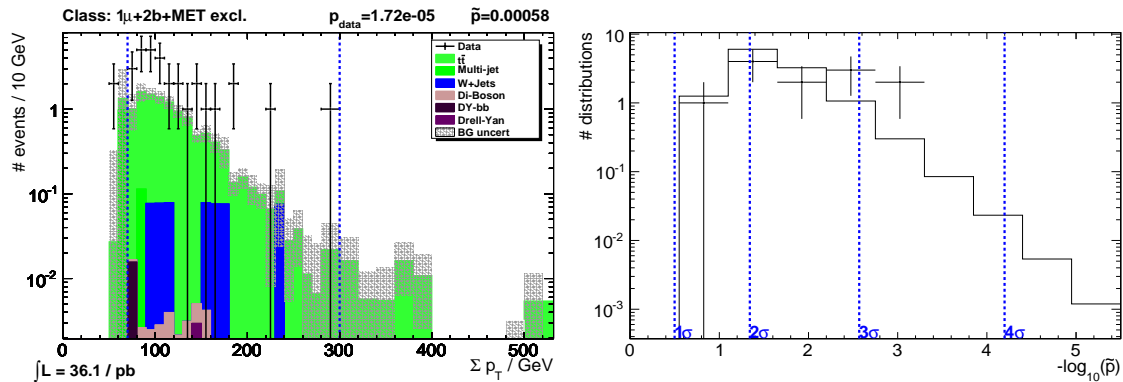


Abbildung 13: Die signifikanteste Klasse und das \tilde{p} -Histogramm der $\sum p_T$ -Verteilung des TCHPt-Algorithmus.

6.2.3 Track Counting High Purity loose

Beim Track Counting High Purity loose (TCHP1) gibt es die größten Abweichungen von der Theorie (Abb. 14-16). Eine Klasse befindet sich immer im Bereich von ungefähr 4σ . Besonders auffällig ist der weit abweichende Datenpunkt der $\sum p_T$ -Verteilung. Diese Klasse hat eine Signifikanz von $\tilde{p} \approx 1 \cdot 10^{-5}$, eine Signifikanz von 4σ erhält man ungefähr bei einem Werte von $\tilde{p} < 6 \cdot 10^{-5}$.

Die so auffällig abweichende Klasse ist wie in Abschnitt 6.2.2 die Klasse „1 μ +2b+1met“. Der Grund für die Abweichung ist wie bei dem TCHPt-Algorithmus eine sehr viel höhere Anzahl Daten als Monte-Carlos im Bereich um 80-90 GeV. Dabei handelt es sich teilweise um einen Faktor zwei Unterschied. Es kann auch hier nicht eindeutig geklärt werden, woher diese Abweichung stammt, allerdings kann es sein, dass sie im Zusammenhang mit dem verwendeten Algorithmus liegt oder dass Monte-Carlos fehlen. Bei den nachfolgenden Algorithmen ist diese Klasse nicht mehr die signifikanteste, obwohl die Fehlerkennungsquote dieselbe sein sollte.

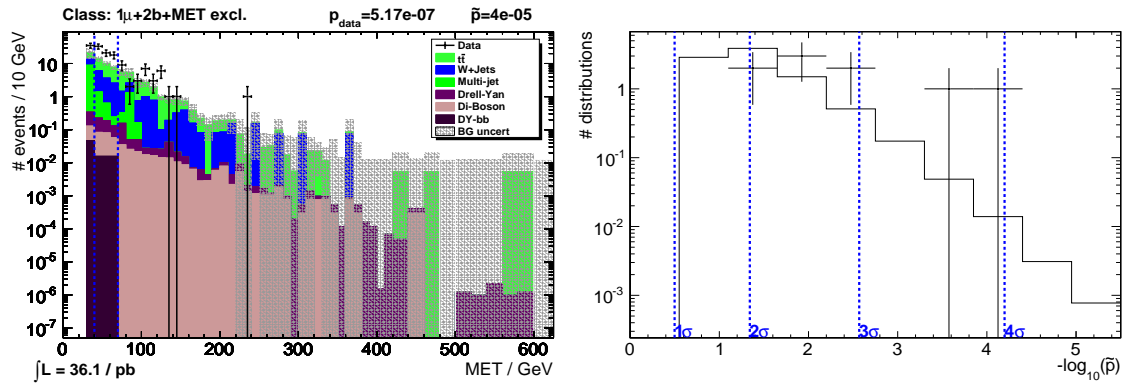


Abbildung 14: Die signifikanteste Klasse und das \tilde{p} -Histogramm der E_T^{miss} -Verteilung des TCHP1-Algorithmus.

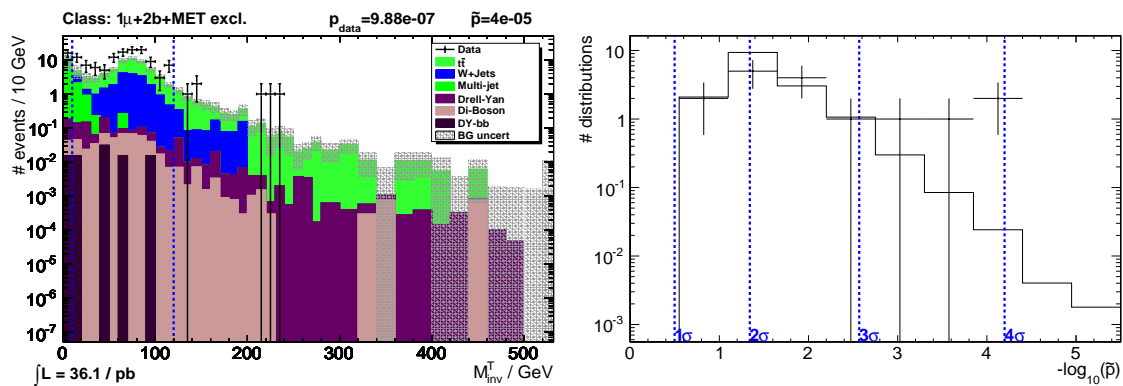


Abbildung 15: Die signifikanteste Klasse und das \tilde{p} -Histogramm der M_{inv} -Verteilung des TCHP1-Algorithmus.

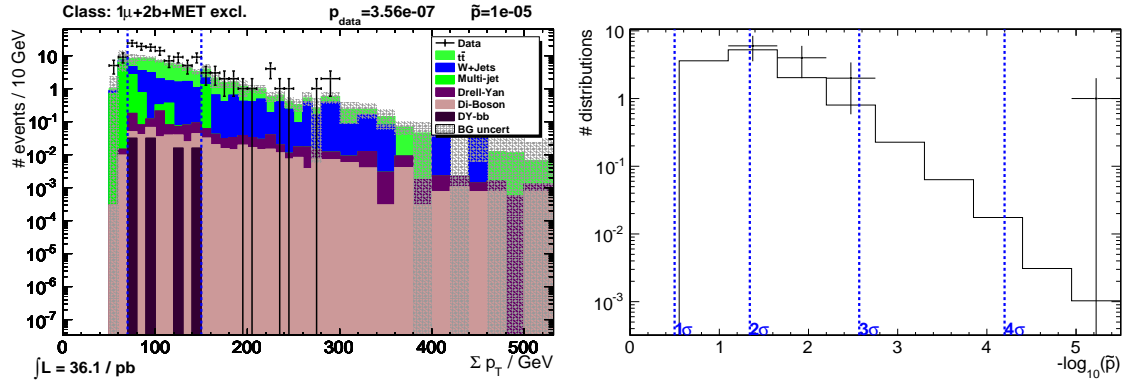


Abbildung 16: Die signifikanteste Klasse und das \hat{p} -Histogramm der $\sum p_T$ -Verteilung des TCHP1-Algorithmus.

6.2.4 Track Counting High Efficiency tight

Der „Track Counting High Efficiency tight“ (TCHet) Algorithmus sollte ähnliche Ergebnisse liefern wie der TCHP_t-Algorithmus. Zwar ist der Algorithmus selbst auf eine hohe Effektivität ausgelegt, aber durch den hoch eingestellten Diskriminator sollte die Fehlerkennungsrate ebenfalls nur 0.1% betragen.

Bei diesem Algorithmus erhält man ebenfalls mehr Klassen mit einer Abweichung zwischen 3σ und 4σ als man erwarten würde (Abb. 17-19). Die signifikanteste Klasse bei Verteilung der fehlenden transversalen Energie und der Summe des transversalen Impulses ist die Klasse „1 μ +1b+1met“. Auch hier handelt es sich wie in Abschnitt 6.2.3 um eine höhere Anzahl Daten als erwartet. Diese liegen aber diesmal nicht in der unmittelbaren Nähe der Massen der W-Eichbosonen, sondern verteilen sich bei der $\sum p_T$ -Verteilung zwischen 100-200 GeV und bei der E_T^{miss} -Verteilung zwischen 50-120 GeV. Die signifikanteste Klasse der M_{inv} -Verteilung ist die „1e+1b+1met“ Klasse. Auch hier hat man mehr Daten als erwartet, diese befinden sich vor der Masse der W-Eichbosonen zwischen 10-80 GeV. Bei der zweitsignifikantesten Klasse handelt es sich wieder um die „1 μ +1b+1met“ Klasse. Die Ergebnisse ähneln tatsächlich denen aus Abschnitt 6.2.4, allerdings beinhaltet die signifikanteste Klasse einen B-Jet weniger.

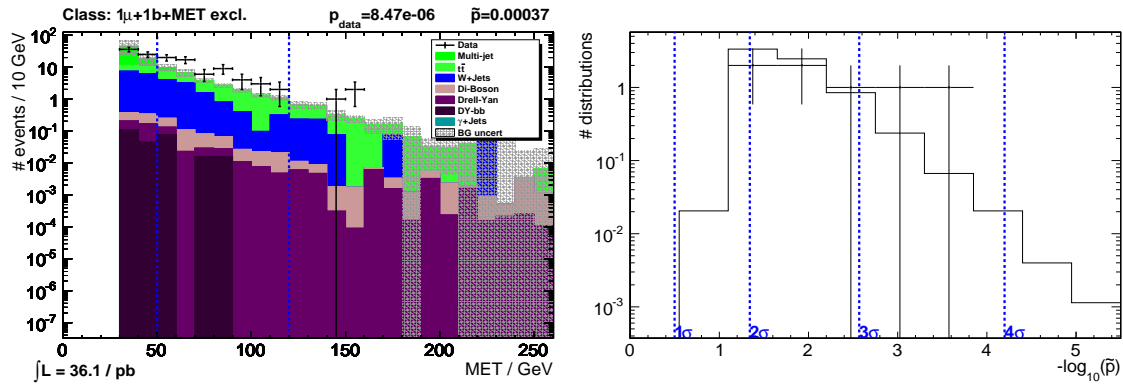


Abbildung 17: Die signifikanteste Klasse und das \tilde{p} -Histogramm der E_T^{miss} -Verteilung des TCHEt-Algorithmus.

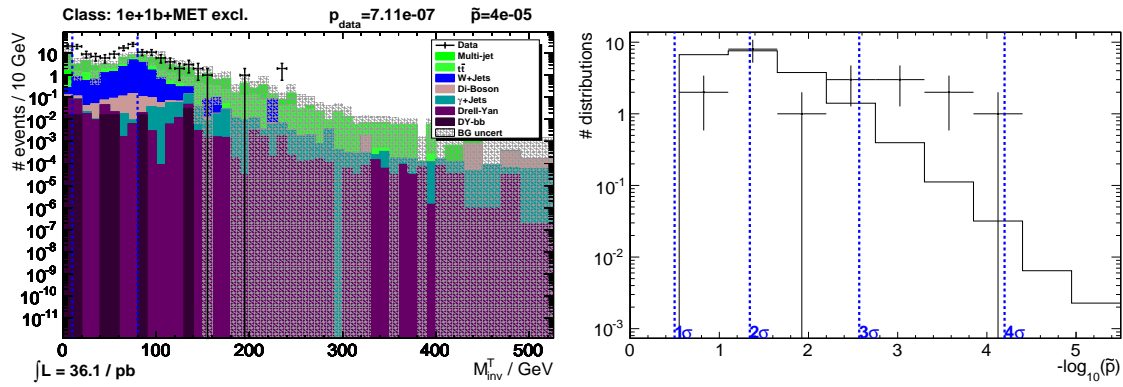


Abbildung 18: Die signifikanteste Klasse und das \tilde{p} -Histogramm der M_{inv} -Verteilung des TCHEt-Algorithmus.

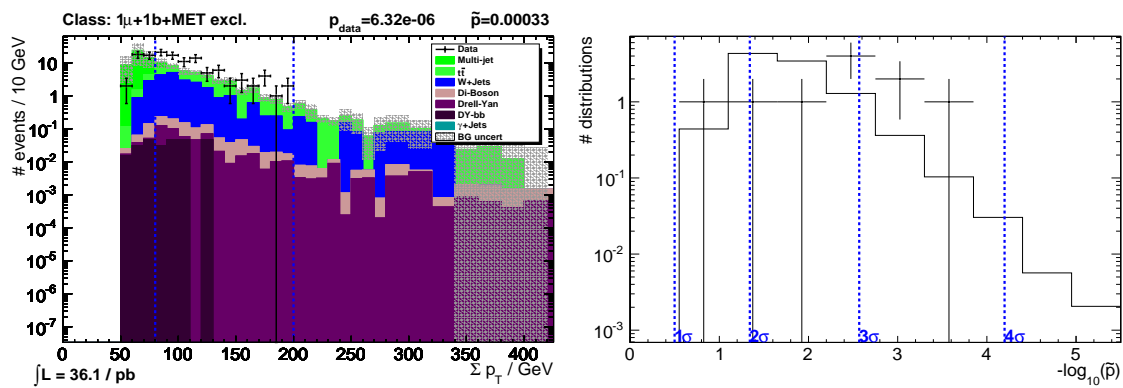


Abbildung 19: Die signifikanteste Klasse und das \tilde{p} -Histogramm der Σp_T -Verteilung des TCHEt-Algorithmus.

6.2.5 Track Counting High Efficiency loose

Der letzte angewendete Algorithmus ist der „Track Counting High Efficiency loose“ (TCHEL). Dieser sollte eine Fehlerkennungsquote von ungefähr 10% haben und ähnliche Ergebnisse wie der Algorithmus aus Abschnitt 6.2.3 liefern. Die \tilde{p} -Diagramme (Abb. 20-22) zeigen, ähnlich wie in den bereits untersuchten Algorithmen, keine großen Abweichungen. Auch hier gibt es ein paar Klassen, die eine Abweichung von ungefähr 3σ haben.

Die signifikanteste Klasse bei der E_T^{miss} -Verteilung und bei der M_{inv} -Verteilung ist wiederum die Klasse „ $1\mu+1b+1met$ “. Interessant ist in diesem Fall die signifikanteste Klasse bei der $\sum p_T$ -Verteilung. Die Klasse „ $1e+1gam+1jet$ “ hat dort die größte Abweichung. Diese Klasse beinhaltet zum einen keinen B-Jet, zum anderen ähnelt sie den sonst signifikantesten Klassen nicht direkt. Auch hier entsteht die Abweichung durch eine Großzahl mehr Daten als Monte-Carlos. Die zweitsignifikanteste Klasse ist wieder die eigentlich als signifikanteste Klasse erwartete Klasse „ $1\mu+1b+1met$ “.

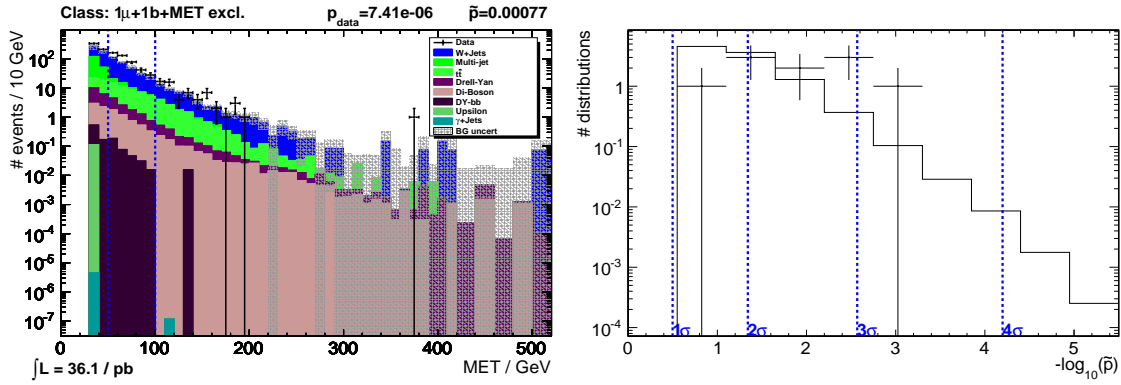


Abbildung 20: Die signifikanteste Klasse und das \tilde{p} -Histogramm der E_T^{miss} -Verteilung des TCHEL-Algorithmus.

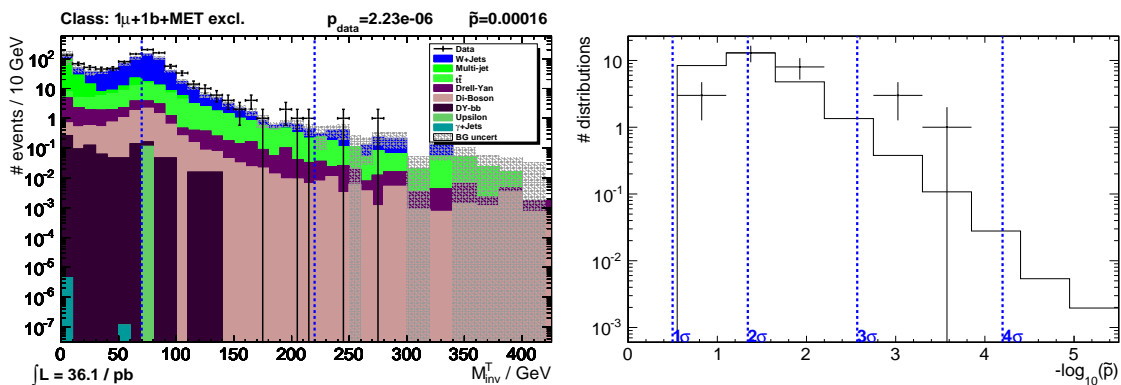


Abbildung 21: Die signifikanteste Klasse und das \tilde{p} -Histogramm der M_{inv} -Verteilung des TCHEL-Algorithmus.

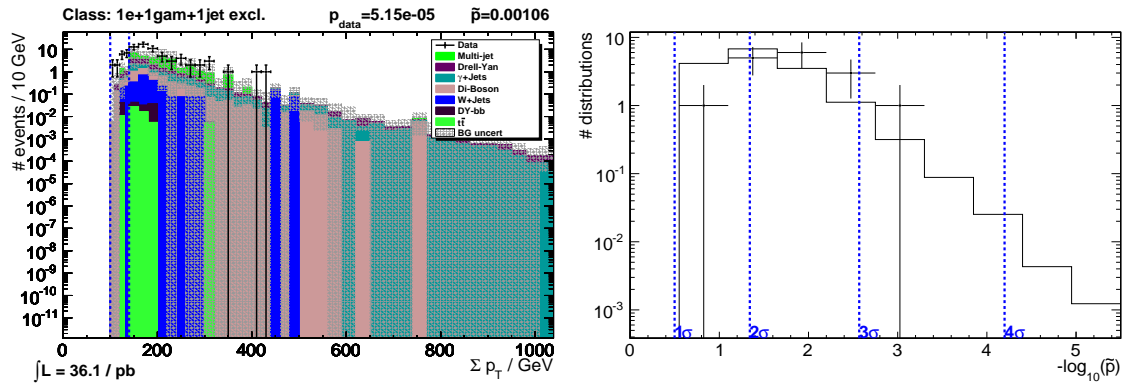


Abbildung 22: Die signifikanteste Klasse und das \tilde{p} -Histogramm der $\sum p_T$ -Verteilung des TCHEL-Algorithmus.

6.2.6 Vergleich der Ergebnisse

Bei der Auswertung der Daten mit den verschiedenen B-Jet Erkennungsalgorithmen hat sich gezeigt, dass die beiden am stärksten abweichenden Klassen die „1 μ +1b+1met“ und die „1 μ +2b+1met“ Klasse sind. Dabei scheint vom verwendeten Algorithmus abhängig zu sein, welche Klasse signifikanter ist. Verwendet man den Track Counting High Purity Algorithmus, so ist die Klasse „1 μ +2b+1met“ signifikanter, bei den anderen Algorithmen ist die Klasse „1 μ +1b+1met“ signifikanter. Dennoch gehören diese beiden Klassen in allen Algorithmen zu den signifikantesten. Da die Daten allerdings in diesen Klassen gleichmäßig zu hoch sind, deutet das auf fehlende Monte-Carlos hin. Diese liegen im Bereich der Massen der W-Eichbosonen, da sich dort die größten Abweichungen ergeben. Ebenfalls auf fehlende Monte-Carlos deutet hin, dass diese beiden Klassen Endzustände von Zerfällen sind, an denen häufig W-Bosonen beteiligt sind. Dennoch kann man festhalten, dass sich verschiedene Algorithmen, auch wenn sie die gleiche Fehlerkennungsquote haben sollten, zu unterschiedlichen Signifikanzen in verschiedenen Klassen führen.

Neben den beiden regelmäßig abweichenden Klassen gibt es keine weitere Klasse, die des Öfteren abweicht. Ohne die beiden signifikanten Klassen „1 μ +2b+1met“ und „1 μ +1b+1met“ sollten die \tilde{p} -Diagramme ungefähr der Theorie entsprechen. Alle weiteren Signifikanzen scheinen also eher zufälliger Natur zu sein. Demnach können mit eingeschalteter B-Erkennung keine großen Signifikanzen in den 2010er Daten festgestellt werden.

7 Zusammenfassung

In einer Sensitivitätsstudie mithilfe von LM-Punkten wurde gezeigt, dass MUSiC mit B-Erkennung für gleiche Signifikanzen nur 80% der Luminosität wie eingeschaltete B-Erkennung benötigt. Dennoch konnte selbst für eine integrierte Luminosität von 10 fb^{-1} bei Signalen für einen Single-Top Zerfall für diesen keine hohe Signifikanz erreicht werden. Wie sich die fehlenden Monte-Carlos auf die Sensitivitätsstudie auswirken, muss noch getestet werden. Bei der Anwendung von B-Erkennung auf Daten in MUSiC wurde gezeigt, dass die Signifikanz der Daten stark von dem gewählten Algorithmus abhängt. Selbst ähnliche Algorithmen mit gleicher Reinheit liefern eine andere Ereignisklasse als signifikanteste Klasse. Die Anwendung auf Daten hat ebenfalls gezeigt, dass es keine wirklich signifikanten Abweichungen in den 2010er Daten gibt. Die Abweichungen in den signifikantesten Klassen „ $1\mu+2b+\text{met}$ “ und „ $1\mu+1b+\text{met}$ “ muss nochmal überprüft werden, da Monte-Carlos fehlen. Klassen, die nicht von den fehlenden Monte-Carlos betroffen sind, weichen eher zufällig ab. Es scheint also in den 2010er Daten keine besonders signifikante Klasse zu geben, die man mit B-Erkennung finden kann. Dennoch sollte das Einschalten von B-Erkennung die Suche nach neuer Physik mit MUSiC vereinfachen, da man weniger Daten für ähnliche Signifikanzen benötigt.

Literatur

- [1] O. Pooth. *Vorlesungsskript Experimentalphysik V*. RWTH Aachen, 2010.
- [2] C. Berger. *Elementarteilchenphysik*. Springer, 2006.
- [3] S. Schael. *Vorlesungsskript Elementarteilchenphysik*. RWTH Aachen, 2011.
- [4] Fakten über den LHC. <http://www.lhc-facts.ch>, Stand: Juni 2011.
- [5] Homepage des CMS-Experiments am CERN. <http://cms.web.cern.ch>, Stand: Juni 2011.
- [6] Homepage des ATLAS-Experiments am CERN. <http://atlas.ch/>, Stand: Juni 2011.
- [7] <http://www.hephy.at/physik/die-technik/detektoren/cms-experiment/>.
- [8] Internetseite des CERN. <http://public.web.cern.ch/public/en/lhc/cms-en.html>.
- [9] The CMS Collaboration. *The CMS Experiment at the CERN LHC*, 2008.
- [10] Internetseite des DESY. <http://www.desy.de/kaefer/vl-hannover/detektoren/>.
- [11] T. Hebbeker, S. Malhotra, A. Meyer, M. Naimuddin, M. Olschewski, H. Pieta, P. Papacz, and S. Schmitz. *MUSiC - A model independent search with 2010 data*, Februar 2011.
- [12] P. Biallass. *Comissioning of the CMS Muon Detector and Development of Generic Search Strategies for New Physics, Doktorarbeit*. RWTH Aachen, 2009.
- [13] T. Hebbeker, S. Malhotra, A. Meyer, M. Naimuddin, M. Olschewski, P. Papacz, H. Pieta, and S. Schmitz. *Model Unspezific Search for New Physics in pp Collisions at $\sqrt{s} = 7$ TeV*, Juni 2011.
- [14] The CMS Collaboration. *Algorithms for b Jet Identification*, Juli 2009.
- [15] E. Dietz-Laursonn. *Model Unspezific Search for New Physics with b-Hadrons in CMS, Diplomarbeit*. RWTH Aachen, Oktober 2010.
- [16] Particle Data Group. *Particle Physics Booklet*, Juli 2010.
- [17] T. Speer, A. Rizzi, C. Saout, F.-P. Schilling, A. Scheurer, T. Boccali, A. Bocci, W. Adam, and A. Schmidt. *Algorithms for b Jet Identification in CMS*, Mai 2009.
- [18] The CMS Collaboration. *Comissioning of b-Jet Identification with pp collisions at $\sqrt{s} = 7$ TeV*, Juli 2010.