

**Studium von t-Kanal-Ereignissen
der Einzel-Top-Quark-Produktion
am CMS-Experiment
mittels neuronaler Netze
zur Ausarbeitung
eines Computerpraktikumversuchs
zur Vorlesung Teilchenphysik**

Bachelorarbeit von

Daniel Hilpert

an der Fakultät für Physik
Institut für Experimentelle Kernphysik (IEKP)

Erstgutachter:	Prof. Dr. Thomas Müller
Zweitgutachter:	Dr. Thorsten Chwalek
Betreuender Mitarbeiter:	Steffen Röcker

08. Mai 2015

Karlsruher Institut für Technologie
Fakultät für Physik
Postfach 6980
76049 Karlsruhe

Ich versichere wahrheitsgemäß, die Arbeit selbstständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Änderungen entnommen wurde.

Karlsruhe, 08. Mai 2015

.....
(Daniel Hilpert)

Inhaltsverzeichnis

1. Einleitung	1
1.1. Der Praktikumsversuch	1
1.2. Das Standardmodell der Teilchenphysik	2
1.3. Das Top-Quark	4
1.3.1. Einzel-Top-Quark-Produktion und Zerfall des Top-Quarks	5
1.3.2. Untergrundereignisse	6
2. Teilchenbeschleuniger und -Detektor	9
2.1. Der Teilchenbeschleuniger Large Hadron Collider	9
2.2. Der Compact-Muon-Solenoid-Detektor	11
3. Neuronale Netze	15
4. Ereignis Selektion und -Modellierung	19
4.1. Jet-Selektion	19
4.2. Myonen-Selektion	19
4.3. Ereignismodellierung	20
4.3.1. Simulierte Ereignisse	20
4.3.2. QCD-Multijet-Untergrund	21
4.4. Reduktion der Untergrundereignisse	21
4.4.1. Reduktion mit kombinierten Schnitten	23
4.4.2. Reduktion mit neuronalem Netz	23
4.4.3. Vergleich der Reduktionsmethoden	26
4.5. Skalierung der simulierten Ereignisse	27
5. Messungen	31
5.1. Wirkungsquerschnitt und CKM-Matrix-Element V_{tb}	31
5.1.1. Wirkungsquerschnitt	31
5.1.2. Matrixelement	32
5.2. Ladungsverhältnis der Leptonen	32
5.3. Masse des Top-Quarks	35
5.3.1. Kalibrierung	35
5.3.2. Rekonstruktion	36
6. Fazit	43
Literatur	45

A. Anhang	49
A.1. Minimierung der Fehlerfunktion des neuronalen Netzes	49
A.1.1. Der Backpropagation-Algorithmus	49
A.2. Verteilungen der Variablen zur weiteren Untergrundreduktion	50
A.3. Übungsblätter für den Praktikumsversuch	66

Abbildungsverzeichnis

1.1.	Parton-Verteilungsfunktion	4
1.2.	Feynman-Diagramme der Einzel-Top-Quark-Produktion im t-Kanal . . .	7
1.3.	Feynman-Diagramme der Einzel-Top-Quark-Produktion	7
1.4.	Feynman-Diagramme des Top-Quark-Zerfalls	7
1.5.	Feynman-Diagramme der Top-Antitop-Quark-Produktion	8
1.6.	Feynman-Diagramme der W-/Z-Boson-Produktion	8
1.7.	Feynman-Diagramme der Diboson-Produktion	8
1.8.	Feynman-Diagramme der Multijet-Produktion	8
2.1.	Einteilung des LHC in Sektoren	10
2.2.	Übersicht über den LHC und die Vorbeschleuniger	11
2.3.	Schematische Darstellung des CMS-Detektors	13
3.1.	Sigmoid- und tanh-Funktion	15
3.2.	Schema eines dreischichtigen neuronalen Netzes	17
4.1.	Verteilung von M_T^W	22
4.2.	Ausgewählte Variablen für Einzelschnitte	23
4.3.	Diskriminator des neuronalen Netzes	24
4.4.	Bestimmung eines Schnittes auf die Ausgabe des neuronale Netzes . . .	26
4.5.	Skalierung der simulierten Ereignisse	28
5.1.	Gemessenes Leptonen-Ladungsverhältnis	34
5.2.	Vergleich Leptonen-Ladungsverhältnis mit PDF-Modellen	34
5.3.	Verteilung der Leptonen-Ladung Q^l	34
5.4.	Verteilung der rekonstruierten Top-Quark-Masse M^{lvb}	38
5.5.	Kalibrierung für kombinierte Schnitte	39
5.6.	Kalibrierung für NN-Schnitte	40
5.7.	Kalibrierung für korrelierten NN-Schnitt	41
A.1.	Verteilung von p_T^l	51
A.2.	Verteilung von η^l	51
A.3.	Verteilung von p_T^W	52
A.4.	Verteilung von η^W	52
A.5.	Verteilung von p_T^ν	53
A.6.	Verteilung von η^ν	53
A.7.	Verteilung von H_T	54
A.8.	Verteilung von E_T^{miss}	54

A.9. Verteilung von M_T^W	55
A.10. Verteilung von $\Delta R^{\text{jet1,jet2}}$	55
A.11. Verteilung von $\cos \Theta^{\text{l,q}}$	56
A.12. Verteilung von $\cos \Theta^{\text{l*,W}}$	56
A.13. Verteilung von $\cos \Theta^{\text{l*,W**}}$	57
A.14. Verteilung von p_T^{jet1}	57
A.15. Verteilung von η^{jet1}	58
A.16. Verteilung von p_T^{jet2}	58
A.17. Verteilung von η^{jet2}	59
A.18. Verteilung von M^{jet1}	59
A.19. Verteilung von M^{jet2}	60
A.20. Verteilung von p_T^{q}	60
A.21. Verteilung von η^{q}	61
A.22. Verteilung von M^{lvb}	61
A.23. Verteilung von M^{lq}	62
A.24. Verteilung von $Q^{\text{l}}\eta^{\text{l}}$	62
A.25. Verteilung von η^{t}	63
A.26. Verteilung von p_T^{t}	63
A.27. Verteilung von η^{b}	64
A.28. Verteilung von p_T^{b}	64
A.29. Verteilung von Q^{l}	65

Tabellenverzeichnis

1.1.	Die vier fundamentalen Kräfte und zugehörige Austauschteilchen nach dem Standardmodell der Teilchenphysik.	4
1.2.	Fermionen des Standardmodells der Teilchenphysik	5
4.1.	Ereignisanzahlen nach Ereignisvorauswahl	20
4.2.	Berechnete Wirkungsquerschnitte für simulierte Ereignisse	21
4.3.	Verwendete Variablen für das Training des neuronalen Netzes	25
4.4.	Vergleich verschiedener Reduktionsmethoden	26
4.5.	Skalierungsfaktoren für simulierte Ereignisse	29
5.1.	Gemessenes Leptonen-Ladungsverhältnis	34
5.2.	Ladungsasymmetrien für simulierte Prozesse	35
5.3.	Massenrekonstruktion für Top-Quark-Massen-Kalibrierung	38
5.4.	Ergebnisse der Regression für Top-Quark-Massen-Kalibrierung	38
5.5.	Ermittelte Top-Quark-Massen für verschiedene Methoden zur Untergrundreduktion	38

1. Einleitung

1.1. Der Praktikumsversuch

Diese Arbeit basiert auf der Vorarbeit von Christoph Alexander Heidecker, welcher im Rahmen einer Bachelorarbeit[1] begann, einen Computerpraktikumsversuch zum Umgang mit neuronalen Netzen zu entwerfen. Ziel dieser Arbeit ist es, den genannten Praktikumsversuch zu einem fertigen Gesamtpaket aus von den Studierenden zu ergänzenden Computerprogrammen und Übungsblatt mit Erklärungen auszuarbeiten. Des Weiteren werden statt der ehemals verwendeten CMS-Daten beim Betrieb des LHC-Teilchenbeschleunigers mit einer Schwerpunktsenergie von $\sqrt{s} = 7 \text{ TeV}$ und einer integrierten Luminosität von $\mathcal{L} = 1,1 \text{ fb}^{-1}$ nun jene mit $\sqrt{s} = 8 \text{ TeV}$ und $\mathcal{L} = 19,7 \text{ fb}^{-1}$ verwendet. Aufgrund der deutlich höheren Anzahl an Ereignissen ist ein weiteres Ziel, das bestehende Programm mit der höheren Datenmenge zu testen und nötigenfalls anzupassen.

Mit dem Praktikumsversuch soll den Studierenden die Möglichkeit gegeben werden, die Analyse von Daten zu einem aktuellen Forschungsgegenstand, dem Top-Quark, durchzuführen und dabei die gängigen Analysemethoden kennen zu lernen, was die Benutzung eines neuronalen Netzes (siehe Abschnitt 3) einschließt.

Das Programmpaket für die Studierenden besteht aus mehreren ausführbaren Programmmodulen, welche die Programmbibliotheken des Programmpakets zur Datenanalyse ROOT [2] und des neuronalen Netzes NEUROBAYES [3] verwenden. Für diese Arbeit wurde das bestehende Anwendungspaket getestet und anschließend in Bezug auf Stabilität und Bedienfreundlichkeit erweitert sowie einige Programmroutinen in Bezug auf Ausführungsgeschwindigkeit optimiert, um eine dem Rahmen eines Praktikumsversuch angemessene Laufzeit auch bei Verwendung der größeren Datenmenge aus 8-TeV-Kollisionen zu erreichen. Das angepasste Programmpaket läuft unabhängig von der CMS-Software und für die verwenden NEUROBAYES-Bibliotheken in der Version für Studierende wird keine Lizenz benötigt, was eine weitere Rahmenbedingung für den Einsatz im Computerpraktikum ist. Mit Hilfe dieser Module können auf verschiedene Arten (siehe Abschnitte 4.4) Untergrundergebnisse reduziert und Schaubilder ausgegeben werden. Diese fertigen Programmteile werden durch Benutzereingaben und Angaben in Konfigurationsdateien gesteuert, um beispielsweise Schnitte zu bestimmen oder das neuronale Netz zu verwenden. Darüber hinaus wurden weitere kleine Anwendungsprogramme (Makros) entwickelt, welche ohne kompiliert werden zu müssen von ROOT ausgeführt werden können. Diese werden für einige der im Abschnitt 5 vorgestellten Messungen verwendet. In den Versionen für die Studierenden wurde der Quelltext teilweise entfernt, so dass die wesentlichen Schritte der Analyse selbst implementiert werden müssen.

1.2. Das Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik ist ein Modell zur Beschreibung aller bisher bekannten Elementarteilchen und deren Wechselwirkung untereinander und beruht im wesentlichen auf der in der ersten Hälfte des 20. Jahrhunderts entwickelten relativistischen Quantenmechanik und im gleichen Jahrhundert gemachten experimentellen Beobachtungen, welche zu erweiterten (Quantenfeld-)Theorien führten. Dieses Standardmodell wurde seitdem immer wieder durch Messungen vorhergesagter Teilchen und deren Eigenschaften bestätigt. Vor Allem aufgrund von astronomischen Beobachtungen, nach welchen der Großteil des Universums aus bisher nicht entdeckten Teilchen (sogenannter Dunkle Materie) bestehen müsste, wird nach schweren Teilchen und Prozessen jenseits des Standardmodells gesucht.

Im Standardmodell der Teilchenphysik werden die bekannten physikalischen Grundkräfte, die starke, elektromagnetische und die schwache Wechselwirkung, durch den Austausch von Wechselwirkungsteilchen mit ganzzahligem Spin (Bosonen) beschrieben (siehe auch Tabelle 1.1). Die schwächste der Grundkräfte, die Wechselwirkung zwischen Massen, Gravitation, wird im bisherigen Standardmodell der Teilchenphysik aufgrund des vernachlässigbaren Einflusses in (sub)atomaren Größenskalen ignoriert. Die starke Wechselwirkung wird durch die Quantenchromodynamik (QCD) beschrieben und wirkt zwischen Teilchen mit sogenannter Farbladung. Das masselose Austauschboson dieser Wechselwirkung, das Gluon, trägt selbst die Farb- und Anti-Farbladung, weshalb Gluonen auch untereinander wechselwirken können. Die elektromagnetische Wechselwirkung wird durch die Quantenelektrodynamik (QED) beschrieben und wirkt zwischen Teilchen mit elektrischer Ladung; das masselose Austauschboson, das Photon, trägt keine elektrische Ladung. Die Austauschbosonen der schwachen Wechselwirkung sind hingegen massereich und tragen unterschiedliche elektrische Ladungen; das W-Boson kann positive oder negative Ladung übermitteln, das Z-Boson ist hingegen elektrisch ungeladen. Nach der vereinheitlichten Theorie von elektromagnetischer und schwacher Wechselwirkung (elektroschwache Wechselwirkung) ist zu erwarten, dass alle Austauschbosonen masselos seien, was nicht mit experimentellen Beobachtungen übereinstimmt. Um dem zu entgegen wird im Higgs-Brout-Englert-Mechanismus des Standardmodells ein weiteres Boson, das Higgs-Boson, mit Spin 0 vorhergesagt.

Die jüngste Bestätigung einer Vorhersage des Standardmodells ist ein im Jahre 2012 erstmals entdecktes Boson mit der Masse $125,09 \pm 0,21 \pm 0,11 \text{ GeV}/c^2$ [4], das höchstwahrscheinlich das postulierte Higgs-Boson ist.

Die Materie selbst wird im Standardmodell durch Teilchen mit halbzahligem Spin (Fermionen) beschrieben, diese sind in Tabelle 1.2 zusammen mit der zugehörigen elektrischen Ladung und Masse gelistet; aufgrund der starken Massenhierarchie sind diese üblicherweise in drei Generationen unterteilt, man unterscheidet zwischen Farbladung tragenden Quarks q , elektrisch geladenen Leptonen $l^{+/-}$ und den entsprechenden elektrisch neutralen Neutrinos ν_l . Zu jedem dieser Teilchen existiert ein Antiteilchen¹ mit umgekehrten

¹Bei dem Neutrino besteht die Möglichkeit, dass dieses sein eigenes Antiteilchen ist.

Vorzeichen der elektrischen Ladung bzw. Anti-Farbladung. Aufgrund der mit dem Abstand abfallenden Stärke der elektromagnetischen Wechselwirkung existieren die Leptonen als einzelne, freie Teilchen, wohingegen die Stärke der starken Wechselwirkung mit dem Abstand zunimmt, weshalb die Quarks nur als gebundene Zustände (Hadronen) zusammen mit den Gluonen existieren. Hochenergetische Hadronen bilden in Wechselwirkung mit weiterer Materie durch QCD-Prozesse viele weitere Hadronen aus, weshalb diese als ein Hadronenbündel (Jet) in Teilchendetektoren zu finden sind. Das Top-Quark ist in keinem Hadron zu finden, da es schneller zerfällt, als dass es hadronisieren kann, weshalb dessen genauere Untersuchung von besonderem Interesse ist. Die Hadronen sind unterteilt in die aus zwei Valenzquarks bestehenden Mesonen und aus drei Valenzquarks bestehenden Baryonen, erstere bestehen aus einem Quark und einem Antiquark gleicher Anti-/Farbladung, letztere aus drei Quarks unterschiedlicher Farbladung. Das Proton ist ein Baryon bestehend aus zwei u-Quarks und einem d-Quark mit resultierender Gesamtladung 1 e, das Neutron ein Baryon bestehend aus einem u-Quark und zwei d-Quarks und Gesamtladung 0 e. Protonen und Neutronen können sich durch schwache und starke Wechselwirkungsprozesse zu Atomkernen verbinden; zusammen mit elektromagnetisch an ein Atomkern gebundenen Elektronen bilden diese Atome, welche zu Molekülen verbunden sein können und die sichtbare stabile Materie bilden.

Neben den Valenzquarks und den sie bindenden Gluonen besteht ein Hadron aus weiteren Quarks (Seequarks genannt) und Gluonen. In experimentell ermittelten Parton-Verteilungsfunktionen (PDF) wird die Wahrscheinlichkeit, ein Quark oder Gluon mit dem Impulsbruchteil x am Gesamtimpuls des Protons zu erhalten, angegeben. Die Verteilung hängt von der betrachteten Energieskala Q^2 ab (siehe Abbildung 1.1).

Im Gegensatz zur elektromagnetischen und zur starken Wechselwirkung sind die Quark-Eigenzustände der schwachen Wechselwirkung nicht die Masseneigenzustände, sondern eine Linearkombination aus diesen und werden mit d' , s' und b' gekennzeichnet. Der Anteil der jeweiligen Masseneigenzustände wird durch die Cabibbo-Kobayashi-Maskawa-Matrix (CKM-Matrix) [5] beschrieben.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\text{CKM}}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.1)$$

Demnach ist die Wahrscheinlichkeit, dass bei einem schwachen Wechselwirkungsprozess mit geladenem Strom ein u-Quark resultiert, wenn ein d-Quark beteiligt ist, proportional zu $|V_{ud}|^2$.

Mit der Einschränkung einer unitären CKM-Matrix sind die ermittelten Werte [6]:

$$V_{\text{CKM}} = \begin{pmatrix} 0,974\,27 \pm 0,000\,14 & 0,225\,36 \pm 0,000\,61 & 0,003\,55 \pm 0,000\,15 \\ 0,225\,22 \pm 0,000\,61 & 0,973\,43 \pm 0,000\,15 & 0,0414 \pm 0,0012 \\ 0,008\,86^{+0,000\,33}_{-0,000\,32} & 0,0405^{+0,0011}_{-0,0012} & 0,999\,14 \pm 0,000\,05 \end{pmatrix}. \quad (1.2)$$

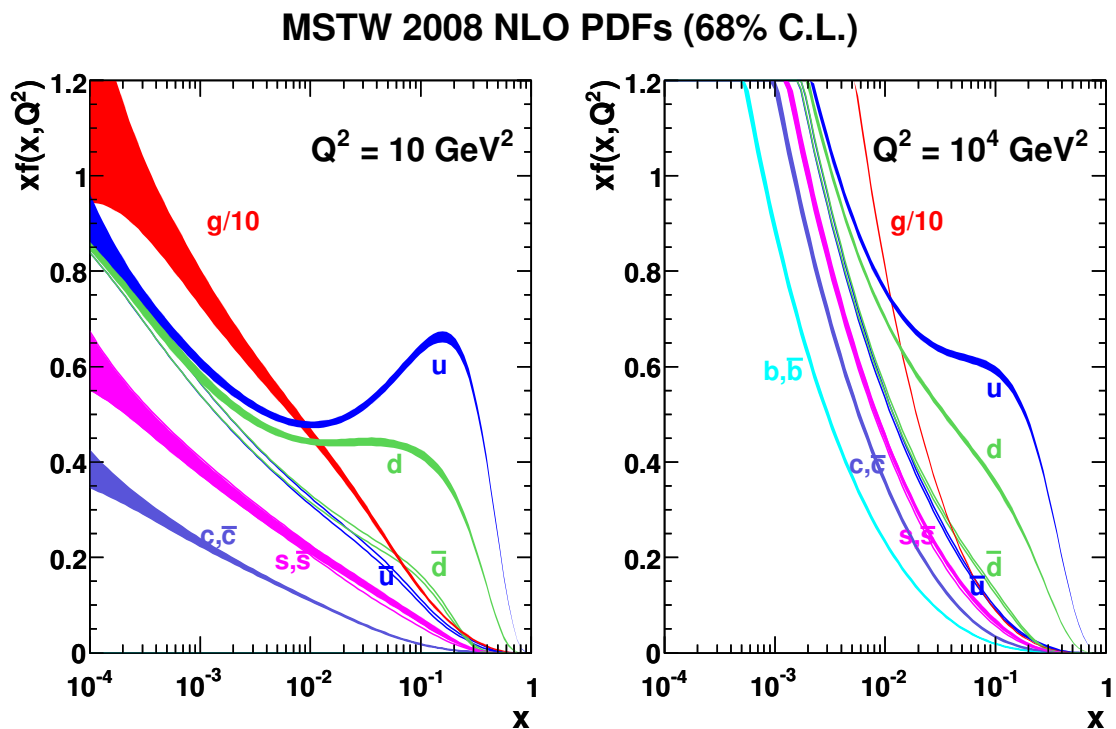


Abbildung 1.1.: Parton-Verteilungsfunktionen [7] für zwei verschiedene Energieskalen Q^2 . Die Wahrscheinlichkeit einer Wechselwirkung mit Gluonen und Seequarks ist für höhere Energien größer.

Anhand der Werte ist zu erkennen, dass eine Umwandlung innerhalb einer Quarkgeneration deutlich wahrscheinlicher als eine Umwandlung zwischen verschiedenen Generationen ist.

1.3. Das Top-Quark

Das Top-Quark wurde 1995 von der CDF- [8] und D0-Kollaboration [9] am Proton-Antiproton-Teilchenbeschleuniger Tevatron des Fermi National Accelerator Laboratorys (Fermilab) entdeckt, nachdem bereits 1977 sein leichter Partner der dritten Teilchengeneration

Tabelle 1.1.: Die vier fundamentalen Kräfte und zugehörige Austauschteilchen nach dem Standardmodell der Teilchenphysik.

Wechselwirkung	Austauschteilchen	relative Stärke
starke	Gluonen (g)	1
elektromagnetische	Photonen (γ)	10^{-2}
schwache	W-/Z-Boson (W^+ , W^- , Z)	10^{-5}

Tabelle 1.2.: Fermionen nach dem Standardmodell der Teilchenphysik unterteilt nach der Teilchengeneration [6].

Name	Symbol		el. Ladung(e)		Masse (MeV/c ²)
	Teilchen	Antiteilchen	Teilchen	Antiteilchen	
Up-Quark	u	\bar{u}	+2/3	-2/3	$2,3^{+0,7}_{-0,5}$
Down-Quark	d	\bar{d}	-1/3	+1/3	$4,8^{+0,5}_{-0,3}$
Elektron	e^-	e^+	-1	1	$0,5 \pm (1,1 \times 10^{-8})$
Elektron-Neutrino	ν_e	$\bar{\nu}_e$	-	-	$< 2 \times 10^{-6}$
Charm-Quark	c	\bar{c}	+2/3	-2/3	1275 ± 25
Strange-Quark	s	\bar{s}	-1/3	+1/3	95 ± 5
Myon	μ^-	μ^+	-1	1	$105,7 \pm (3,5 \times 10^{-6})$
Myon-Neutrino	ν_μ	$\bar{\nu}_\mu$	-	-	$< 2 \times 10^{-6}$
Top-Quark	t	\bar{t}	+2/3	-2/3	$173\,210 \pm 510 \pm 710$
Bottom-Quark	b	\bar{b}	-1/3	+1/3	4180 ± 30
Tauon	τ^-	τ^+	-1	1	$1776,8 \pm 0,16$
Tauon-Neutrino	ν_τ	$\bar{\nu}_\tau$	-	-	$< 2 \times 10^{-6}$

neration, das b-Quark, ebenfalls am Fermilab entdeckt wurde. Nachgewiesen wurde das Top-Quark durch die Erzeugung von Top-Antitop-Quark-Paaren anhand derer Zerfallsprodukte (vergleiche Abbildung 1.5). Im Jahre 2009 wurden die ersten Einzel-Top-Ereignisse vom CDF- und D0-Experiment [10, 11] beobachtet.

1.3.1. Einzel-Top-Quark-Produktion und Zerfall des Top-Quarks

Die Produktion einzelner Top-Quarks ist nach dem Standardmodell der Teilchenphysik für die Kollision zweier Protonen durch schwache Wechselwirkungsprozesse mit geladenen Strömen möglich, die Feynmandiagramme für die Produktionskanäle in führender Ordnung sind in Abbildung 1.3 dargestellt. In der vereinfachten Analyse im Computerpraktikum werden Untersuchungen anhand der t-Kanal-Produktion für die Kollision zweier Protonen am Teilchenbeschleuniger LHC (siehe Abschnitt 2.1) untersucht. Dieser Prozess ist der dominierende für die Kollision zweier Protonen. Das an der Produktion beteiligte b-Quark ist ein Seequark eines Protons und stammt aus einer b-Anti-b-Quarkpaar-Erzeugung aus einem Gluon, jene ist bei den Feynmandiagrammen in Abbildung 1.2 mit berücksichtigt, des Weiteren ist der t-Kanal-Prozess für unterschiedliche Arten von einlaufenden leichten Quarks dargestellt.

Aufgrund der kurzen Zerfallszeit von 5×10^{-25} s [6] zerfällt das Top-Quark in ein W-Boson und in den meisten Fällen ein b-Quark, bevor es hadronisiert. Diese Eigenschaft ermöglicht den direkten Rückschluss auf das Matrixelement $|V_{tb}|$ durch Messung des Wirkungsquerschnittes σ_t bzw. $\sigma_{\bar{t}}$ für die Produktion einzelner Top-Quarks. Die Bestimmung dieser Größe ist Teil des Computerpraktikums und ist im Abschnitt 5.1.2 zu finden.

Berücksichtigt man nur die Valenzquarks eines Protons, so ist aufgrund des darin zweifach vorkommenden u-Valenzquark und nur einfach vorkommenden d-Valenzquark eine im Mittel doppelt so häufige Produktion von Top- wie Antitop-Quarks zu erwarten; jenes Verhältnis wird jedoch durch den Anteil an Seequarks verschoben; das genaue zu erwartende Verhältnis hängt von der Wahl der PDF und der Schwerpunktsenergie ab. Das W-Boson zerfällt entweder hadronisch oder leptonisch, für letzteren Zerfall kann anhand der spezifischen Leptonenladung (l^+ für t , l^- für \bar{t} , siehe Abbildung 1.4) direkt zwischen einem produzierten Top- oder Antitop-Quark unterschieden werden. Das Ergebnis einer Messung dieses Verhältnisses mit der Anwendung für das Computerpraktikum wird in Abschnitt 5.2 diskutiert.

Aus den kinematischen Größen des Leptons, des Neutrinos und des b-Quarks wird die Masse (M^{lvb}) des produzierten Top-Quarks aus der Summe der entsprechenden Viervektoren rekonstruiert. Die Bestimmung der daraus resultierenden tatsächlichen Top-Quark-Masse ist eine weitere Aufgabe des ausgearbeiteten Computerpraktikums und wird im Abschnitt 5.3 erläutert.

1.3.2. Untergrundereignisse

Bei zahlreichen weiteren Prozessen ähneln die Endprodukte denen der Einzel-Top-Quark-Produktion, weshalb diese als ein Einzel-Top-Quark Ereignis im Sinne der im Abschnitt 4 genannte Vorauswahl erfasst werden können, z. B. durch nicht oder fehlerhaft identifizierte Teilchen oder Jets. Die Reduktion dieser Untergrundereignisse und Auswirkungen durch verschiedene Methoden ist ein wesentlicher Aspekt des ausgearbeiteten Computerpraktikums (siehe Abschnitt 4.4).

Der Untergrundprozess mit dem größten Anteil ist bei einer Schwerpunktsenergie von $\sqrt{s} = 8 \text{ TeV}$ die Produktion eines Top-Antitop-Paares ($t\bar{t}$) (Abbildung 1.5). Einen ebenfalls bedeutenden Beitrag liefert die Produktion eines W- oder Z-Bosons (Abbildung 1.6), welche leptonisch zerfallen. Insbesondere beim Prozess (b) entstehen zwei Quark-Jets gleichzeitig mit einem Lepton und dem entsprechenden Neutrino. Der Prozess (d) gleicht diesem im Falle eines nicht detektierten zweiten Leptons; dem gleicht auch der in Abbildung 1.7 gezeigte Fall zweier produzierter Bosonen, der Wirkungsquerschnitt ist jedoch etwa um die Größenordnung 10 geringer (siehe Tabelle 4.2). Einen weiteren wichtigen Anteil liefern QCD-Multijet-Ereignisse aufgrund der starken Wechselwirkung. In Abbildung 1.8 sind zwei mögliche Prozesse mit der Signatur eines Einzel-Top-Quark-Ereignisses dargestellt. Beim Prozess (a) kommt es ebenfalls zu einem leptonischen Zerfall eines W-Bosons, eines b-Jets und zu mehreren leichten Jets; der Fall (b) kann ebenfalls als ein Einzel-Top-Quark-Ereignis klassifiziert werden, wenn ein hochenergetischer Jet die Myonenkammer des Teilchendetektors erreicht und deshalb als Myon erkannt wird (vergleiche Detektoraufbau im Abschnitt 2.2).

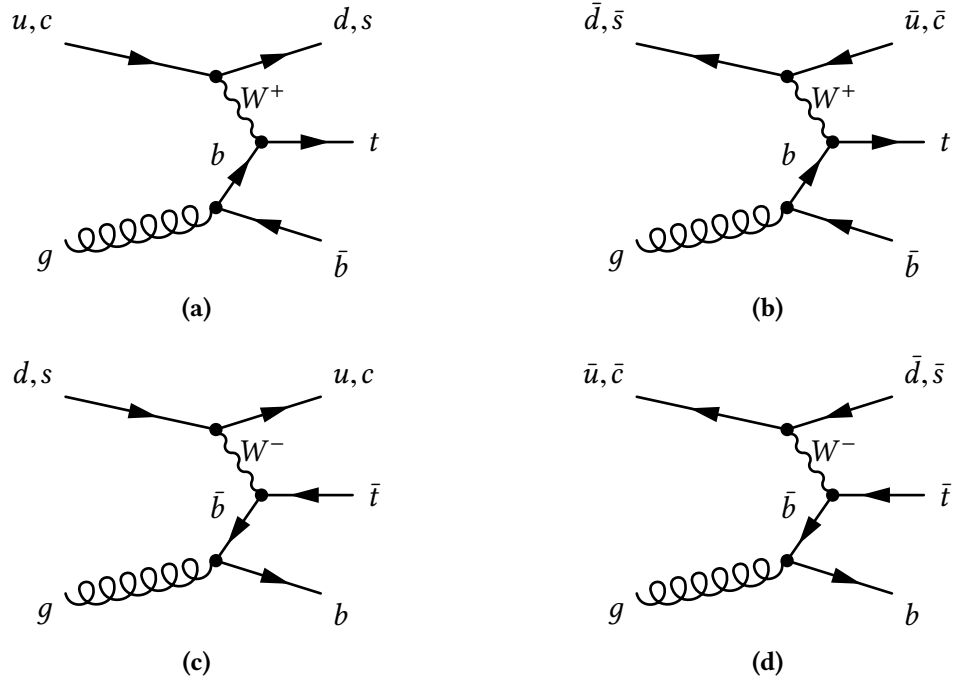


Abbildung 1.2.: Einzel-Top-Quark-Produktion im t-Kanal für ein einlaufendes Quark des Protons (a) und für ein einlaufendes Antiquark (b). Einzel-Antitop-Quark-Produktion im t-Kanal für ein einlaufendes Quark des Protons (c) und für ein einlaufendes Antiquark (d).

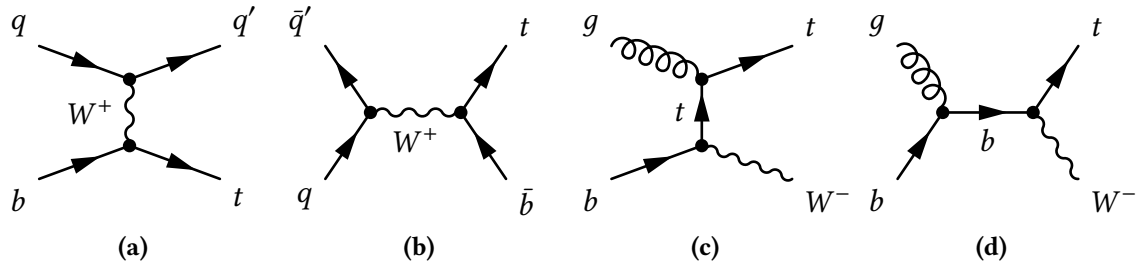


Abbildung 1.3.: Mögliche Einzel-Top-Quark-Produktionskanäle für die Kollision zweier Protonen in führender Ordnung. (a): t-Kanal, (b): s-Kanal, (c) und (d): Einzel-Top-Quark-Produktion mit assoziiertem W-Boson.

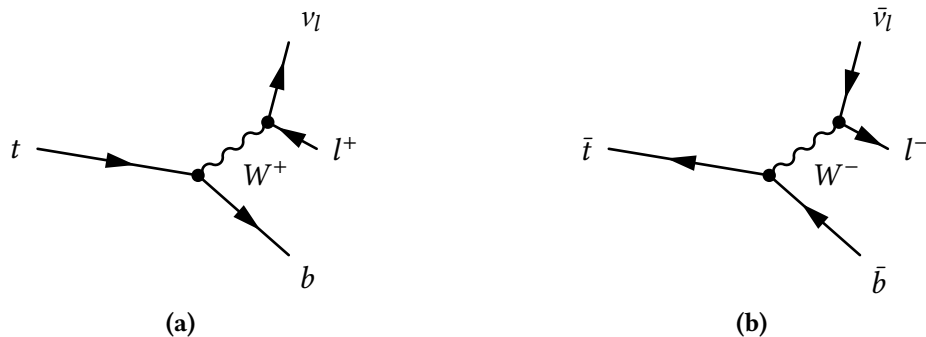


Abbildung 1.4.: Häufigster Zerfall des Top-Quarks (a) und des Antitop-Quarks (b) für einen leptonischen Zerfall des W-Bosons.

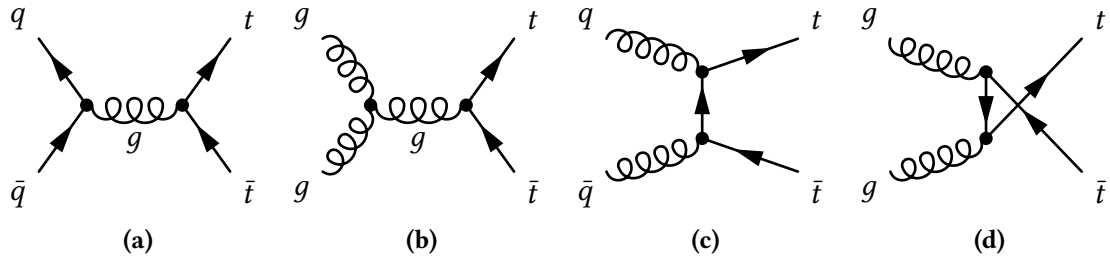


Abbildung 1.5.: $t\bar{t}$ -Produktion

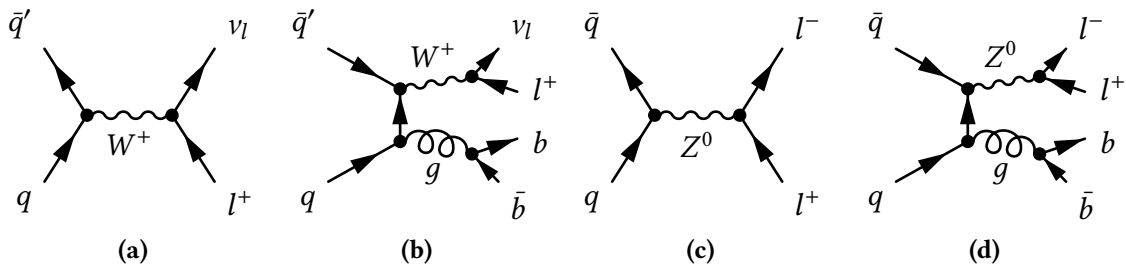


Abbildung 1.6.: W-/Z-Boson-Produktion

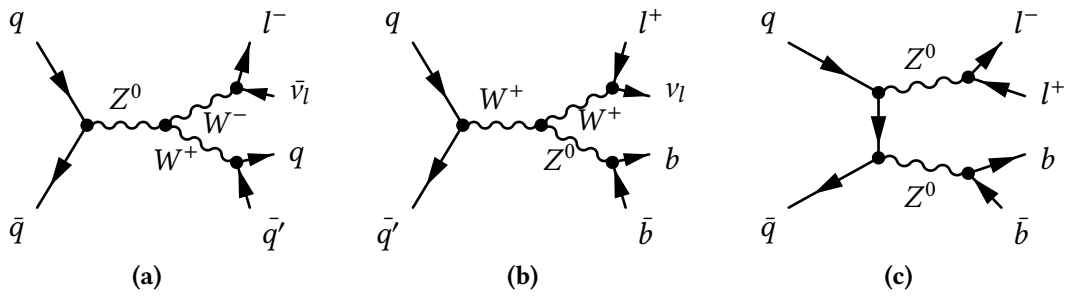


Abbildung 1.7.: Diboson-Produktion

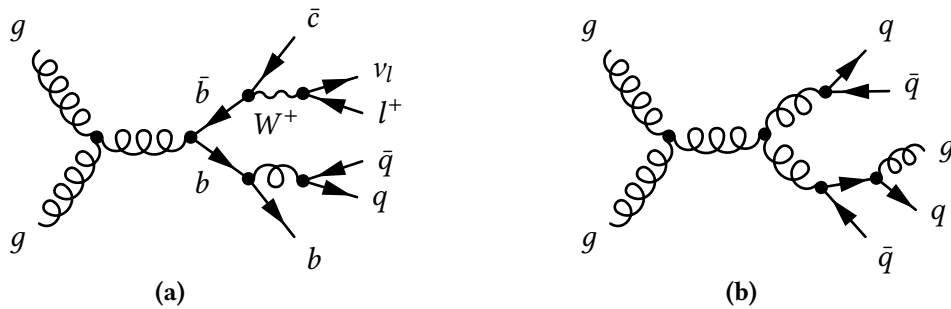


Abbildung 1.8.: Multijet-Produktion: (a) mit gleichen Endprodukten wie bei einem Einzel-Top-Quark-Ereignis, (b) mit hochenergetischen Jet, der Myonenkammer erreicht und fälschlicher Weise als Myon detektiert würde.

2. Teilchenbeschleuniger und -Detektor

Um die nötige Schwerpunktsenergie zu erreichen, bei der die im Abschnitt 1.3 beschriebenen Prozesse erzeugt werden können, werden Protonen in einem Teilchenbeschleuniger beschleunigt und zur Kollision gebracht. Die dabei entstehenden Produkte werden mittels eines am Kollisionspunkt installierten Detektors untersucht.

2.1. Der Teilchenbeschleuniger Large Hadron Collider

Der „Large Hardon Collider“ (LHC) am CERN bei Genf in der Schweiz ist der derzeit leistungsstärkste Teilchenbeschleuniger der Welt; er befindet sich in einem 26,7 km langen künstlichen, kreisförmigen Tunnel in 45 m bis 170 m Tiefe unter der Erdoberfläche. Der Tunnel wurde bereits zwischen 1984 und 1989 für den bis ins Jahr 2000 betriebenen „Large Electron-Positron Collider“ (LEP) angelegt und verläuft sowohl unter schweizerischem als auch französischem Staatsgebiet. In ihm werden sowohl Protonen als auch Blei-Ionen in entgegengesetzte Richtungen beschleunigt und an fest definierten Punkten zur Kollision gebracht.

Der Ring besteht aus insgesamt 8 gebogenen Abschnitten und dazwischen geraden Abschnitten. Ein Oktant wird von der Mitte eines Bogens bis zur Mitte des nächsten Bogens definiert (siehe Abbildung 2.1). Die Mitte eines geraden Abschnitts wird als Punkt mit der entsprechenden Nummer bezeichnet. In den gebogenen Abschnitten halten mit supraflui- dem Helium gekühlte Dipolmagnete die Protonen innerhalb zweier hochvakuumisierter Strahlrohre auf einer Kreisbahn und Quadrupolmagnete fokussieren die umlaufenden Teil- chen. Die geraden Abschnitte dienen verschiedenen Zwecken: Am Punkt 4 beschleunigen acht Hohlraumresonatoren die Teilchen auf die gewünschte Zielenergie, an den Punkten 3 und 7 wird der Strahl durch verschiedene Kollimatoren und weitere Magnete gereinigt. Im LHC befinden sich zwei Strahlen aus mehreren Teilchenbündeln in entgegengesetzten Richtungen im Umlauf. An vier Punkten (Punkt 1, 2, 5 und 8) kreuzen sich die beiden Strahlen; dort befinden sich Detektoren, die die bei der Kollision der Teilchen entstehen- den Produkte untersuchen: „A Toroidal LHC Apparatus“ (ATLAS) und „Compact Muon Solenoid“ (CMS) sind Mehrzweckdetektoren zu Untersuchungen des Standardmodells und von neuer Physik; „A Large Ion Collider Experiment“ (ALICE) ist spezialisiert auf die Untersuchung von Blei-Ionen-Kollisionen zur Untersuchung der starken Wechsel- wirkung bei hohen Energien (Quark-Gluonen-Plasma); „Large Hadron Collider beauty“ (LHCb) ist spezialisiert auf Präzisionsmessungen des Bottom-Quark und der Messung von CP-Verletzung.

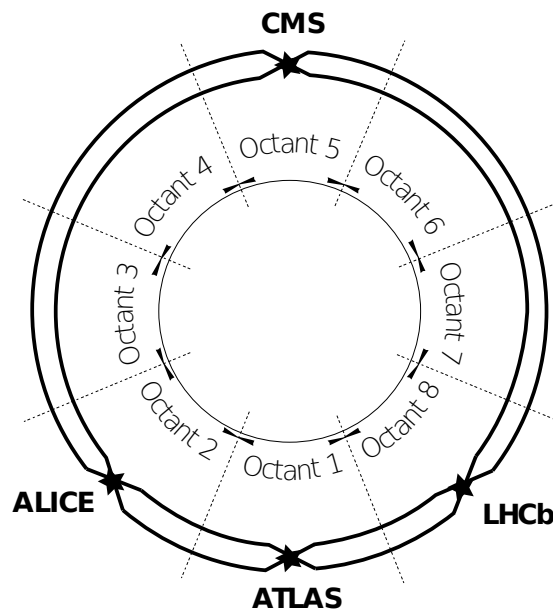


Abbildung 2.1.: Einteilung des LHC in Sektoren (Oktanten). In der Mitte von 4 Abschnitten von insgesamt 8 werden Protonen-Bündel zur Kollision gebracht und die dabei entstehenden Produkte von Detektoren untersucht. Grafik adaptiert von [12].

Da für diese Arbeit Messdaten von Proton-Kollisionen des CMS-Experiments verwendet werden, wird nicht näher auf den Betrieb des LHC mit Blei-Ionen eingegangen. Die Funktionsweise des CMS-Detektors wird im nächsten Abschnitt genauer erläutert.

Die Protonen werden vor der Beschleunigung im LHC auf die Zielenergie in kleineren Beschleunigern vorbeschleunigt, eine schematische Darstellung der Beschleuniger ist in Abbildung 2.2 zu sehen. Erzeugt werden die Protonen durch Ionisation von Wasserstoffatomen mittels eines elektrischen Feldes und anschließend von Hohlraumresonatoren im „Linear Accelerator 2“ (LINAC 2) auf eine Energie von 50 MeV beschleunigt; daraufhin werden sie im ersten Ringbeschleuniger, dem „Proton Synchrotron Booster“ (PSB), auf 1,4 GeV beschleunigt. Eine Beschleunigung auf 25 GeV erfolgt im „Proton Synchrotron“ (PS). Die letzte Vorbeschleunigung auf 450 GeV erfahren sie im knapp 7 km langen „Super Proton Synchrotron“ (SPS), bevor zwei Strahlen in entgegengesetzte Richtungen in den LHC-Beschleuniger entlassen werden. Im LHC werden die Teilchenbündel auf 4 TeV pro Strahl beschleunigt.

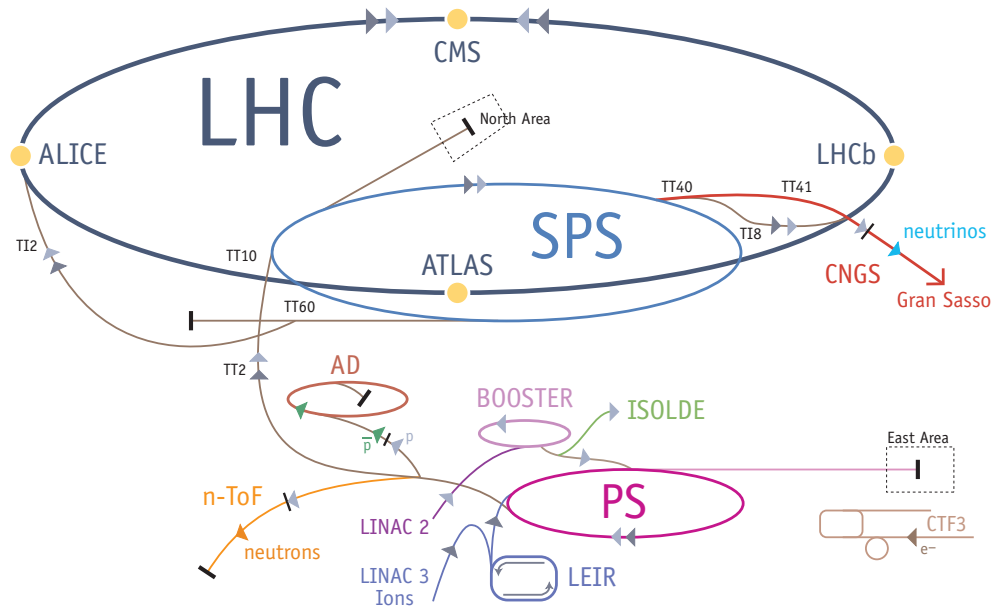


Abbildung 2.2.: Übersicht über den LHC und die Vorbeschleuniger [12].

2.2. Der Compact-Muon-Solenoid-Detektor

Der Compact-Muon-Solenoid-Detektor (CMS-Detektor) (siehe Abbildung 2.3) ist ein Mehrzweckdetektor mit dem Ziel Messungen des Standardmodells der Teilchenphysik betreffend durchzuführen, wie der Nachweis des Higgs-Bosons und neuer Modelle wie Supersymmetrie. Der CMS-Detektor befindet sich am Punkt 5 des LHC-Teilchenbeschleunigers in einer in etwa 100 m Tiefe gelegenen Halle. Zahlreiche Detektorkomponenten werden verwendet, um die Zerfallsprodukte bei der Kollision von im LHC in entgegengesetzter Richtung beschleunigter Protonen-Bündel zu rekonstruieren.

Durch einen supraleitenden Solenoid-Magneten von 6 m innerem Durchmesser wird ein Magnetfeld von 3,8 T Feldstärke im Innern des Magneten und von 2 T im äußeren Bereich, geleitet durch das Eisenjoch, erzeugt. Durch das Magnetfeld werden die bei der Kollision erzeugten elektrisch geladenen Teilchen auf eine gekrümmte Bahn abgelenkt. Silizium-Pixel- und -Streifendetektoren (Tracker) dienen der Rekonstruktion von Teilchenspuren nahe am Kollisionspunkt zur Bestimmung des Impulses aus der Teilchenbahn. Elektromagnetische Kalorimeter (ECAL) dienen der direkten Erfassung der Energie von Photonen. Zusammen mit dem im Spurdetektor bestimmten Impuls wird im ECAL die Energie von Elektronen durch die direkte Messung der im Kalorimeter deponierten Energie in Kombination mit der deponierten Energien assoziierter Photonen aus Bremsstrahlung bestimmt. Durch die Kombination der deponierten Energie im elektromagnetischen Kalorimeter und im Hadronenkalorimeter (HCAL) wird die Energie neutraler Hadronen bestimmt. Für die Energiebestimmung geladener Hadronen wird zusätzlich die mit dem Tracker gewonnene Impulsinformation verwendet. Die Myonen werden mit den Daten der entsprechenden Driftkammern im äußeren Bereich des Detektors rekonstruiert. Durch die Kombination

der Informationen aller Detektorkomponenten werden mittels des Particle-Flow-Event-Algorithmus[13, 14] die Zerfallsprodukte identifiziert und rekonstruiert.

Zur Beschreibung der Ereignispositionen ist ein Koordinatensystem mit Ursprung im Detektor-Mittelpunkt und vertikaler y-Achse, x-Achse in Richtung des Mittelpunkt des LHC und z-Achse tangential zum LHC in Richtung Punkt 4 definiert. Als ϕ ist der Winkel zur x-Achse in der x-y-Ebene definiert, als θ der Winkel zur z-Achse senkrecht zur x-y-Ebene, dieser geht in die Pseudorapidität, definiert als $\eta = -\ln(\tan(\theta/2))$, ein.

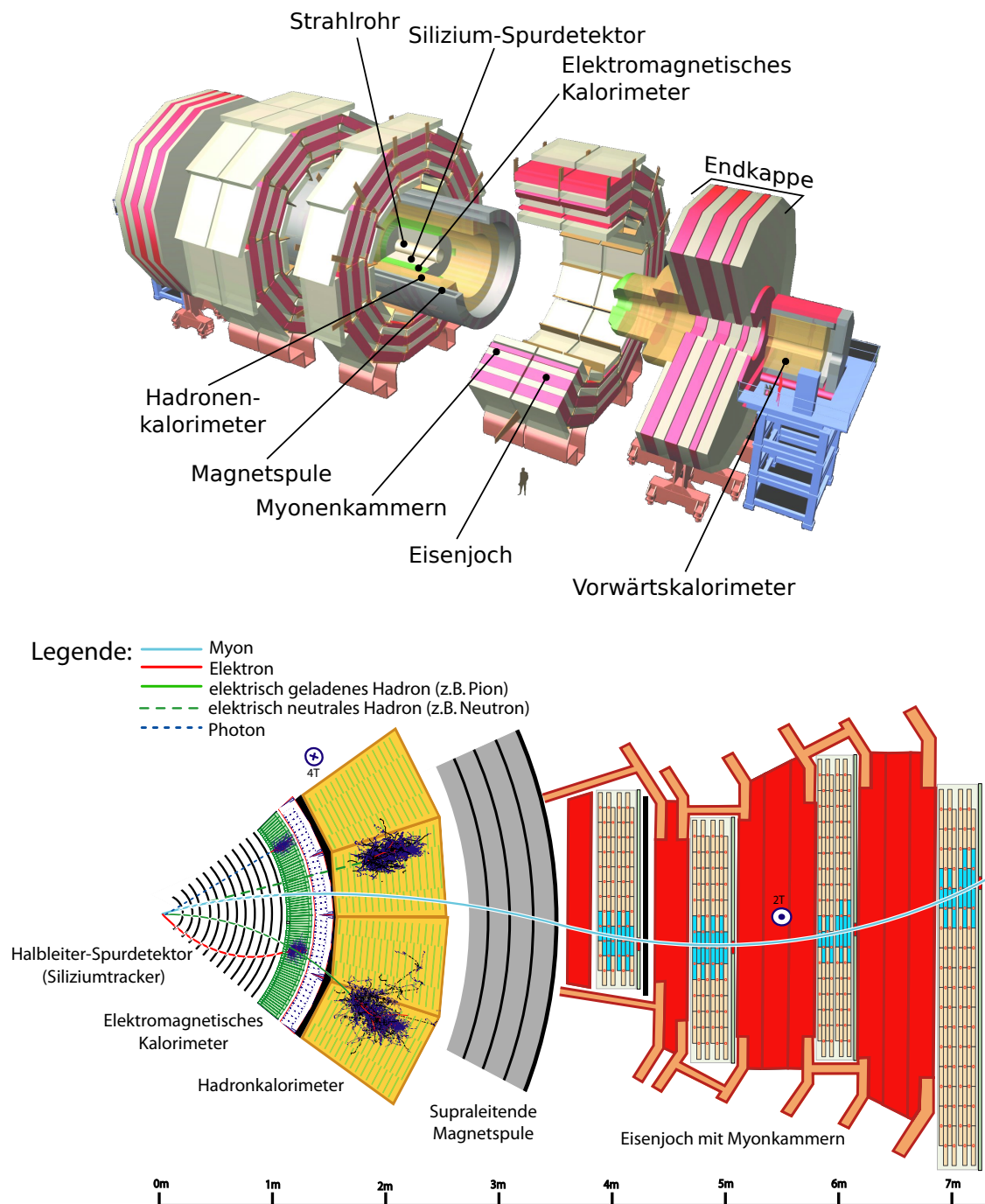


Abbildung 2.3.: Perspektivische Darstellung des CMS-Detektors (oben) und Schnitt senkrecht zur Strahlachse (unten). Im Inneren des Strahlrohrs kollidieren zwei sich entgegengesetzt bewegende Protonen-Bündel. Die geladene Zerfallsprodukte werden durch das mit der Magnetspule erzeugte Magnetfeld abgelenkt und Teilchenspuren mit Hilfe des Siliziumdetektors rekonstruiert. Mit Hilfe der elektromagnetischen und Hadronenkalorimetern wird die Energie der Teilchen rekonstruiert. Die äußeren Myonen-Driftkammern dienen der Spurrekonstruktion speziell von Myonen, das Eisenjoch liefert einen wesentlichen Anteil zur Stabilität der Konstruktion und führt das Magnetfeld zu den äußeren Myonenkammern. Mit Hilfe der Myonenkammern in den Endkappen und der Vorwärtskalorimetern werden auch Teilchen mit Flugrichtung im verhältnismäßig kleinem Winkel zur Strahlrichtung erfasst. (Abbildungen entnommen und teilweise adaptiert von [15] und [16].)

3. Neuronale Netze

Künstliche neuronale Netze verhalten sich ähnlich wie ihre biologischen Vorbilder; sie werden aus vielen einzelnen künstlichen Neuronen aufgebaut, welche abhängig von Eingangssignalen und weiteren Einflüssen ein Ausgangssignal erzeugen. Durch eine Vernetzung der Neuronen können logische Operationen ausgeführt werden. Ein künstliches neuronales Netz kann mit Hilfe von historischen oder simulierten Ereignissen lernen, Aussagen über neue Daten zu treffen. Im Rahmen des in dieser Arbeit erweiterten Anwendungspakets wird ein neuronales Netz verwendet, um nach Training mit simulierten Ereignissen Signalereignisse von Untergrundereignissen in gemessenen Daten zu unterscheiden.

Eine Art von künstlichen Neuronen sind die nach der Sigmoid-Funktion benannten Sigmoid-Neuronen. Für ein Neuron i dieser Art wird der zugehörige Ausgabewert,

$$a_i = s(\underbrace{\vec{w}_i \cdot \vec{x}_i + b_i}_{=z_i}), \quad (3.1)$$

anhand gegebener Eingabewerte \vec{x}_i und Gewichtsparametern \vec{w}_i sowie einem weiteren Parameter, dem Schwellenwert b_i , berechnet. Die Wahl der Funktion $s(z)$ ist je nach Implementierung unterschiedlich; typische Varianten sind die Sigmoid-Funktion oder die tanh-Funktion (siehe Abbildung 3.1).

Durch die Verknüpfung der Neuronen zu einem Netz kann der Ausgabewert eines Neurons als Eingabewert weiteres Neuron verwendet werden. Bei einer speziellen Form dieser Netze (feedforward-Netze) sind die Neuronen in Schichten angeordnet, wobei Neuronen einer Schicht nur mit denen der nächsten Schicht verknüpft sind, jedoch keine Verknüpfung innerhalb der Schicht oder zu vorherigen Schichten vorhanden ist (siehe Abbildung 3.2).

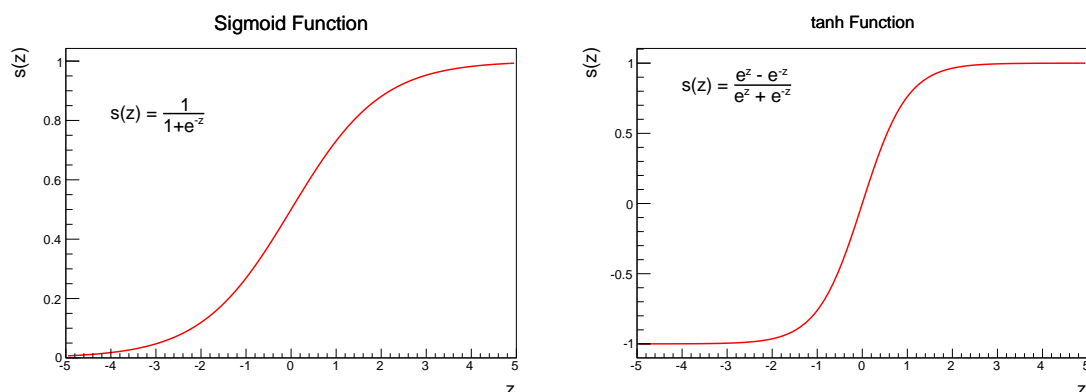


Abbildung 3.1.: Schaubild der Sigmoid- (links) und der tanh-Funktion (rechts).

Im Allgemeinen sind mehrere Neuronen in der Ausgabeschicht und mehrere verborgene Schichten möglich. Im folgenden werden nur Netze mit einer verborgenen Schicht betrachtet ($L = 3$). In einem Netz dieser Art hängen die Ausgabewerte \vec{a}^L des Netzes für einen gegebenen Satz von Eingabewerten \vec{x}^1 von den Gewichtsparametern der Neuronen in der verborgenen Schicht und der Neuronen in der Ausgabeschicht ab. Lernen oder Trainieren des neuronalen Netzes bezeichnet die Bestimmung der Parameter \vec{w}_i und b_i aller Neuronen eines neuronalen Netzes. Die Bestimmung erfolgt durch Minimierung einer (Fehler-)Funktion C , welche von der Ausgabe des Netzes und von einer bekannten zu erwartenden Ausgabe abhängig ist. Die Berechnungsvorschrift für C hängt von der Wahl der Fehlerfunktion sowie von der jeweiligen Implementierung des neuronalen Netzes ab. Übliche Fehlerfunktionen sind beispielsweise die quadratische Fehlerfunktion, welche bevorzugt zur Regression verwendet wird.

$$C = \frac{1}{2N} \sum_{i=1}^N \sum_{j=1}^{N_3} \|y_{ij} - a_j^3\|^2, \quad (3.2)$$

$$C > 0. \quad (3.3)$$

Die Entropie-Fehlerfunktion wird bevorzugt zur Klassifikation verwendet und lautet für die hier betrachteten Netze:

$$C = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_3} (y_{ji} \ln a_j^3 + (1 - y_{ji}) \ln (1 - a_j^3)), \quad (3.4)$$

$$C > 0. \quad (3.5)$$

N ist die Anzahl an Eingabedaten, N_3 die Anzahl der Neuronen in der Ausgabeschicht, y_{ij} ist die zu erwartende Ausgabe des Neurons j der Ausgabeschicht für die Eingabe i und entsprechend ist a_j^3 die tatsächlich erhaltende Ausgabe.

Die Bestimmung der den Wert C minimierenden Gewichtsparameter und Schwellenwerte ist im Allgemeinen analytisch nicht lösbar und wird numerisch ermittelt, beispielsweise durch ein Gradientenverfahren mit mehreren Iterationsschritten n , wobei im ersten Iterationsschritt $n = 1$ (zufällig gewählte) Gewichte und Schwellenwerte vorgegeben werden (siehe auch Anhang A.1).

Ein für die Datenanalyse in der Teilchenphysik verwendetes kommerzielles Anwendungspaket mit dreischichtigem neuronalen Netz ist NEUROBAYES [3, 17]; darin können erweiterte Vorverarbeitungsfunktionen verwendet werden, die das Training des Netzes weiter beschleunigen und stabilisieren sollen, um beispielsweise das Lernen statistischer Fluktuationen zu verhindern und zuverlässig das globale Minimum der Fehlerfunktion zu bestimmen. Für das Training des Netzes werden nicht die Verteilungen selbst als Eingabewerte verwendet, stattdessen wird der Wert einer an der Verteilung durchgeführten Spline-Interpolation verwendet. NEUROBAYES wird ebenfalls im Anwendungspaket für das Computerpraktikum verwendet, genaueres dazu ist im Abschnitt 4.4.2) zu finden.

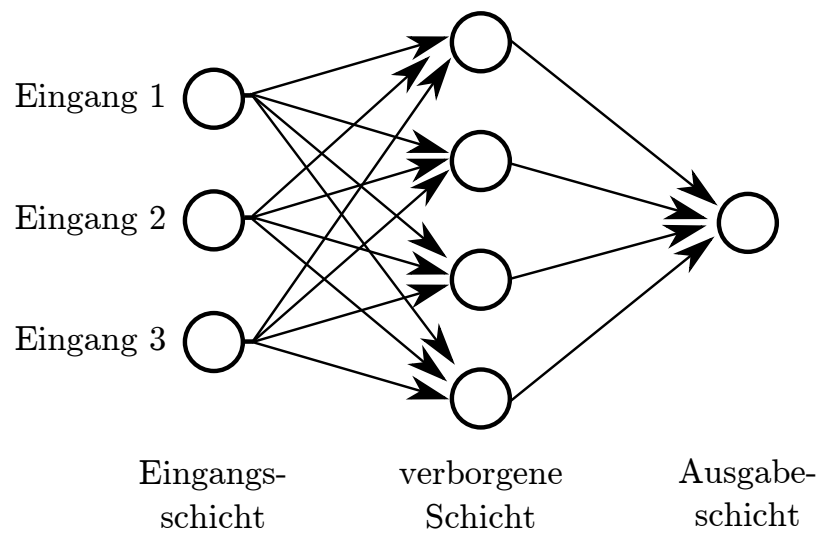


Abbildung 3.2.: Schema eines dreischichtigen neuronalen Netzes mit drei Eingabewerten, vier Neuronen in der verborgenen Schicht und einem Neuron in der Ausgabeschicht.

4. Ereignisselektion und -Modellierung

Es werden nur Ereignisse betrachtet, die im Detektor die Signatur der Endprodukte eines Einzel-Top-Quark-Prozesses aufweisen (vergleiche Abschnitt 1.3). Verlangt werden ein isoliertes geladenes Myon, ein b-Jet sowie mindestens ein weiterer Jet, verursacht durch das leichte Quark.

4.1. Jet-Selektion

Durch eine Jet-Selektion erfolgt die Einschränkung auf die aus der Produktion eines einzelnen Top-Quark und dessen Zerfall hervorgehenden Hadronen.

Selektiert werden nur Ereignisse, bei denen mindestens zwei Jets, jeweils mit einem Transversalimpuls von $p_T^{\text{jet}} > 40 \text{ GeV}/c$ sowie einer Pseudorapidität von $|\eta| < 4,5$ [18] detektiert wurden, wovon einer als Jet eines b-Quarks markiert wurde.

Die Endprodukte sind ein leichtes Quark, ein b-Quark, das in einer Gluon-Aufspaltung entsteht und ein weiteres b-Quark aus dem Zerfall des Top-Quarks (vergleiche Abbildung 1.2 und 1.4); diese bilden durch Hadronisierung Jets aus, welche durch gruppierte Teilchen mittels des anti- k_T -Algorithmus [19] mit einem Abstandsparameter von 0,5 definiert werden. Der Jet des b-Quarks, welches aus der Gluonen-Verzweigung hervorgeht, propagiert vorwiegend in Strahlrichtung und der Transversalimpuls ist oft unterhalb der verlangten p_T -Schwelle, deshalb kann er im Detektor nur selten detektiert werden.

Zur Unterscheidung eines Jets, der einem Bottom-Quark entspringt, von anderen erfolgt die Markierung durch den Combined-Secondary-Vertex-Algorithmus (CSV) [20]; dieser kombiniert Informationen über den Stoßparameter der Jet-Spur und weitere Vertices innerhalb des Jets und berechnet daraus einen Wahrscheinlichkeitsparameter zur Unterscheidung von Bottom-Quark-Jets und Jets anderen Ursprungs. Es wird ein Arbeitspunkt gewählt, der einer Missmarkierungsrate von 0,1% entspricht [18].

4.2. Myonen-Selektion

Es werden zudem nur Ereignisse ausgewählt, bei denen genau ein Myon detektiert wird, welches dem Lepton aus dem Top-Quark-Zerfall entspricht. Für die Myonen wird ein Transversalimpuls von $p_T^l > 26 \text{ GeV}/c$ und eine Pseudorapidität von $|\eta^l| < 2,1$ verlangt, zudem

Tabelle 4.1.: Ereignisanzahlen nach der Jet-, Myonen-Selektion und QCD-Reduktion für die gemessenen Werte und die modellierten Prozesse, normiert auf die gemessene Luminosität, letztere sind zusätzlich separat gelistet.

Ereignisse	Anzahl
gemessene	90260
modellerte	80682
Einzel-Top t	11421
Einzel-Top s	654
tW	4244
t \bar{t}	37247
W+Jets	19467
Z+Jets	1492
Diboson (WW/WZ/ZZ)	602
QCD	5555

wird die relative Isolierung der Myonen auf $I^{\text{rel}} < 0,12$ eingeschränkt, die folgendermaßen berechnet wird:

$$I^{\text{rel}} = \frac{I^{\text{ch.h}} + \max(I^\gamma + I^{\text{n.h}} - I^{\text{PU}}, 0)}{p_{\text{T}}^{\text{l}}} . \quad (4.1)$$

Die darin eingehenden Größen sind die in einem Kegel der Größe $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0,4$ um die Richtung der Myonen deponierte Transversalenergie stabiler geladener ($I^{\text{ch.h}}$) und neutraler ($I^{\text{n.h}}$) Hadronen sowie von Photonen (I^γ) und die Summe von Transversalimpulsen ($I^{\text{PU}} = 0,5 \cdot \sum p_{\text{T}}^{\text{PU}}$) neutraler Teilchen aus Ereignissen, die innerhalb der Totzeit des Detektors auftreten (zeitliche Pile-Up-Ereignisse) [18].

4.3. Ereignismodellierung

Die zu erwartenden Signal- und Untergrundprozesse (vergleiche Abschnitt 1.3) werden durch Simulation mit Monte-Carlo-Generatoren, im Falle des QCD-Multijet-Untergrundes durch eine Datennahme in einem speziellen Seitenband ermittelt. Der QCD-Untergrund wird durch einen Einzel-Schnitt auf die Variable M_{t}^{W} reduziert. Die Ereignisanzahlen für die modellierten Prozesse nach der bisher erfolgten Vorauswahl sind in Tabelle 4.1 zu finden.

4.3.1. Simulierte Ereignisse

Mit Hilfe von Simulationen können die Ereignisanzahlen der meisten Prozesse anhand theoretischer Rechnungen vorhergesagt werden; die zu erwartenden Prozesse werden mit Monte-Carlo-Generatoren simuliert. Hadronisierung und Schauerbildung wird mittels

Tabelle 4.2.: Die berechneten Wirkungsquerschnitte [18] für eine Proton-Proton-Kollision am LHC bei einer Schwerpunktsenergie von $\sqrt{s} = 8$ TeV.

Prozess	Wirkungsquerschnitt (pb)
Einzel-Top t-Kanal	87,1
Einzel-Top s-Kanal	5,6
tW	22,2
t \bar{t}	737,4
W+Jets	3063,6
Z+Jets	3503,7
Diboson (WW/WZ/ZZ)	96,1

PYTHIA 6.4 [21] für jeden Prozess simuliert. Für Einzel-Top-Produktion im t-Kanal, s-Kanal und mit assoziiertem W-Boson wird POWHEG 1.0 [22] verwendet, für t \bar{t} -Produktion, W/Z+Jets-Ereignisse und Diboson-Ereignisse MADGRAPH 5.148 [23]. Die Eigenschaften der generierten Ereignisse im Detektor werden anschließend mittels GEANT 4 [24] simuliert. Die erzeugten Ereignisse sind auf den entsprechenden berechneten Wirkungsquerschnitt (siehe Tabelle 4.2) normiert und werden mit der gemessenen Luminosität multipliziert.

4.3.2. QCD-Multijet-Untergrund

Um den QCD-Multijet-Untergrund zu modellieren, wird dieser aus Daten in einem speziellen Seitenband ermittelt, da bisher keine zuverlässigen Monte-Carlo-Simulationen dieses Untergrundes zur Verfügung stehen. Die Daten werden durch Umkehr der Bedingung für die relative Isolierung ermittelt. Da der somit gewonnene QCD-Untergrund nicht wie die simulierten Daten auf theoretische Wirkungsquerschnitte normiert ist, wird dieser direkt durch eine Likelihood-Anpassung in der von den QCD-Ereignissen dominierten Region $M_T^W < 50 \text{ GeV}/c^2$ auf die Messdaten normiert, dies erfolgt durch Verwendung des THETA-Programmpakets[25]. Der somit ermittelte Faktor wird auf die gesamte M_T^W -Region extrapoliert.

Der Anteil von QCD-Ereignissen ist im Bereich $M_T^W < 50 \text{ GeV}/c^2$ (siehe Abbildung 4.1) besonders hoch, durch einen Schnitt $M_T^W > 50 \text{ GeV}/c^2$ wird dieser Untergrund verringert.

4.4. Reduktion der Untergrundereignisse

Mit den folgenden zusätzlichen Einschränkungen wird angestrebt, die Untergrundergebnisse mit den gleichen Endprodukten wie die eines Einzel-Top-Quark Ereignisses zu reduzieren. Für die weitere Selektion der Ereignisse werden die Pseudorapidität η , der Transversalimpuls p_T und die Massen bzw. Transversalmassen M oder M_T für diverse erzeugte Teilchen oder Jets verwendet sowie die Ladung der Leptonen (Myonen nach Vorslektion) Q^l . Zwischen zwei Teilchen wird der Kosinus des Winkels Θ betrachtet. Der

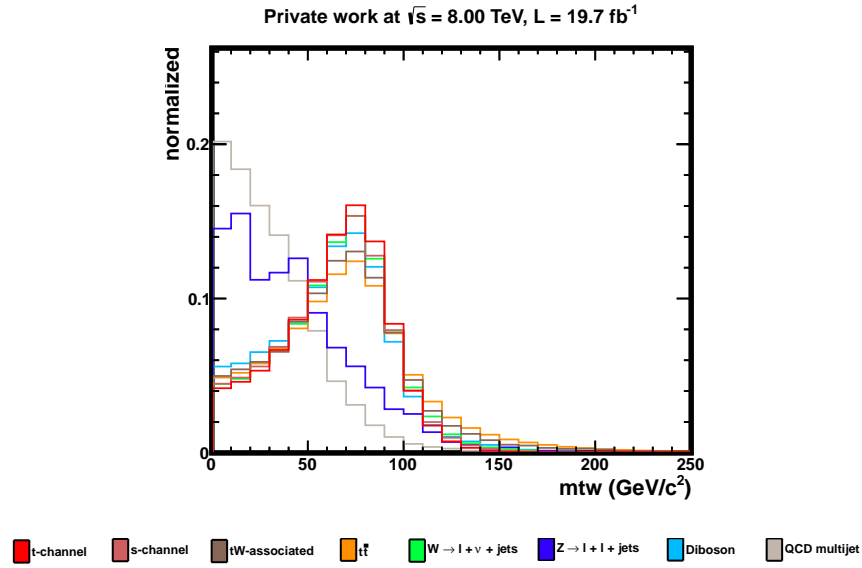


Abbildung 4.1.: Verteilung der Transversalmasse des W-Bosons M_T^W ohne angewandten Schnitt zur Reduktion des QCD-Untergrundes normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses.

Abstand zwischen zwei Jets ist als $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ definiert. Die Größen werden anhand der Daten aus dem Detektor rekonstruiert und die Rekonstruktion gleichermaßen für simulierte Ereignisse angewandt. Da Neutrinos nicht direkt nachgewiesen werden können, werden diese anhand von fehlender Transversalenergie E_T^{miss} registriert und rekonstruiert. Des Weiteren wird die Größe H_T , die betragsmäßige Summe aller Transversalimpulse und der fehlenden Transversalenergie, betrachtet.

In einer multivariaten Analyse mittels neuronalem Netz werden alle Variablen gleichzeitig zur Reduktion des Untergrundes verwendet, indem auf die Ausgabe des neuronalen Netzes geschnitten wird, welches mit modellierten Ereignissen trainiert wurde. Zum Vergleich wird eine Reduktion des Untergrundes durch die Kombination verschiedener Schnitte auf einzelne Variablen untersucht.

Für eine quantitative Einstufung der Untergrundreduktion werden verschiedene Größen anhand der modellierten Ereignisanzahlen berechnet, darin gehen die Gesamtzahl an Signal-Ereignissen s (simulierte t-Kanal-Ereignisse) und die Gesamtzahl an Untergrund-Ereignissen b (übrige simulierte Prozesse und gemessener QCD-Untergrund) ein.

Betrachtet werden die Werte von

$$R_{s,b} = \frac{s}{b},$$

dem direkten Verhältnis beider Größen (Signal-Untergrund-Verhältnis), von

$$p_s = \frac{s}{s+b},$$

der Reinheit des Signals und von

$$\alpha_s = \frac{s}{\sqrt{b}},$$

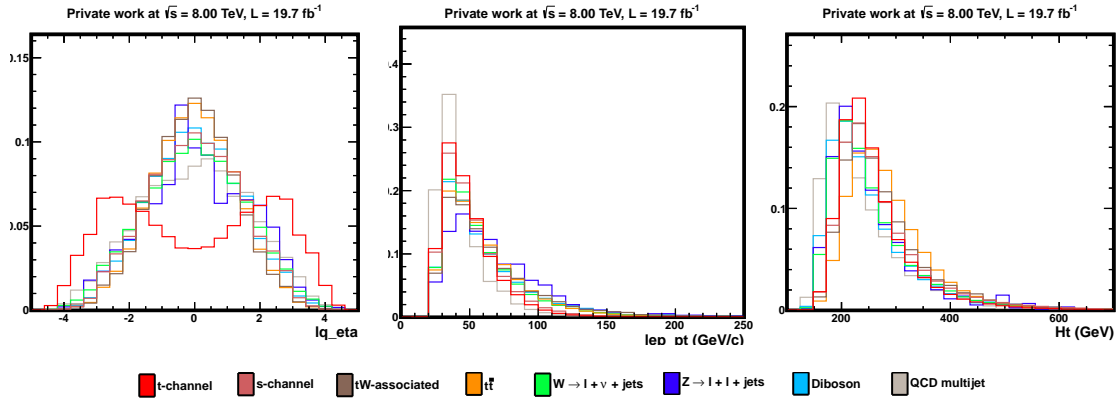


Abbildung 4.2.: Verteilungen der modellierten Prozesse für die kombinierte Einzelschnitte, normiert auf die Ereigniszahl des entsprechenden Prozesses. **Links:** Pseudorapidity des leichten Quarks η^q . **Mitte:** Transversalimpuls des Myons p_T^l . **Rechts:** Betragsmäßige Summe aller Transversalimpulse und der fehlenden Transversalenergie: H_T .

der Signifikanz.

Das Signal-Untergrund-Verhältnis und die Reinheit des Signals sollten durch die zusätzlichen Einschränkungen ansteigen. Ein Anstieg der Signifikanz ist ein Indiz dafür, dass Einzel-Top-Quark-Produktionen im Datensatz vorhanden sind. Die anhand dieser Größen ermittelten Schnitte werden zur späteren Bestimmung der Top-Quark-Masse und dem Verhältnis von im t-Kanal produzierten Einzel-Top-Quarks zu produzierten Einzel-Antitop-Quarks gleichermaßen auf die gemessenen Daten angewandt.

4.4.1. Reduktion mit kombinierten Schnitten

Zunächst wird eine Kombination an Einzelschnitten untersucht, welche auf die Variablen p_T^l , H_T und η^q angewandt werden (siehe Abbildung 4.2).

Für die Schnittkombination

$$\begin{aligned} |\eta^q| &< 1,8, \\ p_T^l &< 90 \text{ GeV}/c, \\ H_T &> 150 \text{ GeV}, \end{aligned}$$

verdreifacht sich der Wert des Signal-Untergrund-Verhältnisses von 0,165 auf 0,513, der Wert der Reinheit des Signals steigt dabei von 0,142 auf 0,339, auch der Wert der Signifikanz wird von 43,4 auf 57,3 erhöht.

4.4.2. Reduktion mit neuronalem Netz

Durch Verwendung eines neuronalen Netzes wird eine Trennung von Signal- und Untergrundereignissen mit gleichzeitiger Berücksichtigung aller verwendeten Variablen erzielt.

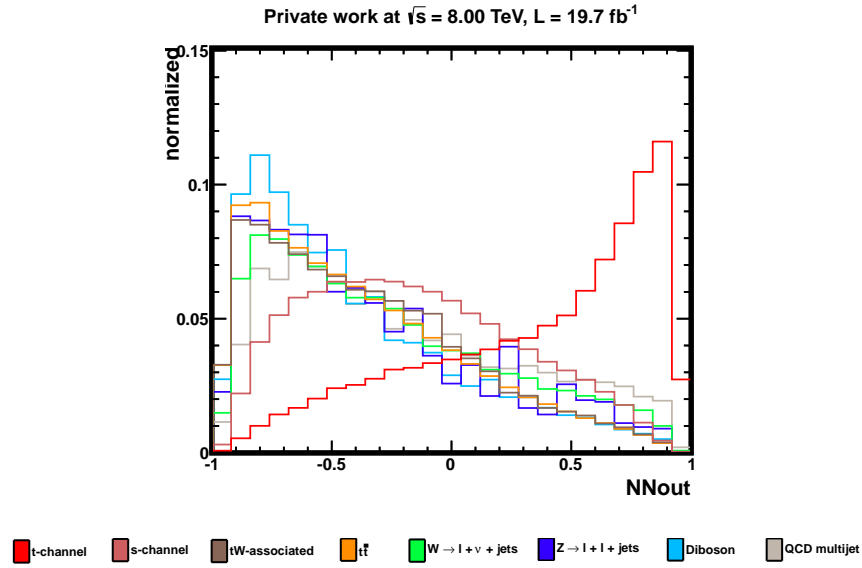


Abbildung 4.3.: Diskriminator des neuronalen Netzes: Die Verteilungen der Untergrundereignisse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses.

Das neuronale Netz wird mit den modellierten Ereignissen trainiert, um zwischen Signal- und Untergrundereignissen zu unterscheiden. Als neuronales Netz wird die Anwendung NEUROBAYES verwendet. Die für das Training verwendeten Variablen sind in Tabelle 4.3 zusammen mit der zusätzlichen Signifikanz aufgelistet, die durch das Training mit diesen Variablen erzielt wurde. Das Programm ist so konfiguriert, dass nur Variablen berücksichtigt werden, die eine zusätzliche Signifikanz von über 3σ im Training liefern. Die nicht verwendeten Variablen sind in der Tabelle gekennzeichnet. Schaubilder der Verteilungen der Variablen befinden sich im Anhang A.2.

Die Ausgabe des neuronalen Netzes (NNout) ist in Abbildung 4.3 dargestellt. Die meisten Untergrundprozesse dominieren für $NNout < 0$, der s-Kanal für Werte bis $NNout = 0,2$. Zur Ermittlung eines geeigneten Schnittes auf diese Variable wird das Verhältnis $\frac{s}{\sqrt{s+b}}$ für verschiedene Schnitte betrachtet, wobei s die Anzahl aller Signalereignisse und b die Anzahl aller Untergrundereignisse ist. In Abbildung 4.4 ist der Wert $\frac{s}{\sqrt{s+b}}$ für verschiedene Schnitte mit der Einschränkung $NNout > NNSchnitt$ aufgetragen. Das Maximum wird für den Schnitt $NNout > 0,3$ erreicht. Der Wert von $R_{s,b}$ wird durch diesen Schnitt von 0,165 auf 0,822 in etwa verfünffacht. Der Wert der Reinheit des Signals steigt dabei von 0,142 auf 0,451. Auch der Wert der Signifikanz wird von 43,4 auf 74,9 erhöht.

Für die Analyse der rekonstruierten Top-Quark-Masse M^{lvb} und des Landungsverhältnisses R_Q der Leptonen aus dem Top-Quark-Zerfall wird jeweils zusätzlich ein Netzwerk ohne die zu untersuchende Variable trainiert, um mögliche Einflüsse durch das Schneiden auf die Ausgabe des neuronalen Netzes erkennen zu können.

Tabelle 4.3.: Eingabevariablen für das Training des neuronalen Netzes und deren zusätzliche Signifikanz. Der Wert N entspricht der in der im Programm verwendeten internen Variablennummer. Variablen, die im Training eine zusätzliche Signifikanz von unter 3σ liefern, werden nicht für die Ausgabe des neuronalen Netzes berücksichtigt.

Variable	Zusätzliche Signifikanz	Bedeutung
η^q	187,48	Pseudorapidität des leichten Quarks
M^{lvb}	98,29	Top-Quark-Masse aus Addition der l -, ν - und b -Vierervektoren
$\Delta R^{\text{jet1,jet2}}$	66,49	Abstand zwischen erstem und zweitem Jet.
M_T^W	58,20	Transversalmasse des W-Bosons
$\cos \Theta^{l,q}$	35,73	Kosinus des Winkels zwischen Lepton und leichtem Quark
Q^l	30,30	Myonen-Ladung
η^W	29,94	Pseudorapidität des W-Bosons
p_T^q	25,77	Transversalimpuls des leichten Quarks
$\cos \Theta^{l*,W}$	22,89	Kosinus des Winkels zwischen Lepton und W-Boson im Laborsystem
M^{jet1}	20,35	Masse des ersten Jets
p_T^t	13,15	Transversalimpuls des Top-Quarks
p_T^W	13,14	Transversalimpuls des W-Bosons
E_T^{miss}	12,72	fehlende Transversalenergie
η^t	12,11	Pseudorapidität des Top-Quarks
p_T^{jet2}	11,02	Transversalimpuls des zweiten Jets
M^{jet2}	10,86	Masse des zweiten Jets
$M^{q,l}$	9,16	Masse des leichten Quarks und des Myons
p_T^ν	8,84	Transversalimpuls des Neutrinos
p_T^l	8,79	Transversalimpuls des Myons
H_T	8,57	Betrag der Summe aller Transversalimpulse und der fehlenden -Energie
p_T^b	8,24	Transversalimpuls des b-Quarks
η^l	7,90	Pseudorapidität des Myons
η^{jet2}	6,98	Pseudorapidität des zweiten Jets
p_T^{jet1}	6,51	Transversalimpuls des ersten Jets
η^b	3,67	Pseudorapidität des b-Quarks.
$\cos \Theta^{l*,W**}$	2,19	Kosinus des Winkels zwischen Lepton und W-Boson im t-Ruhsystem
η^{jet1}	1,92	Pseudorapidität des ersten Jets
η^ν	0,12	Pseudorapidität des Neutrinos
$Q^l \eta^l$	0,00	Produkt $Q^l \eta^l$

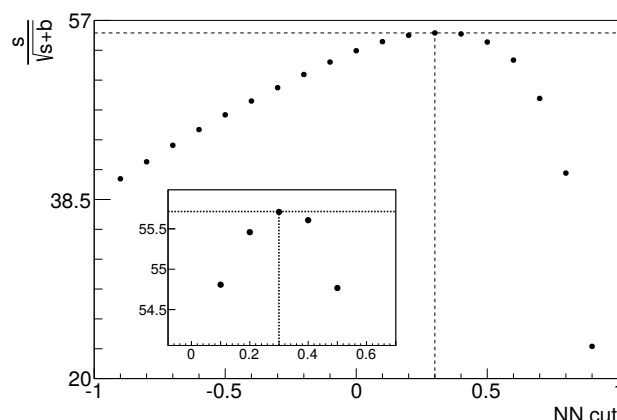


Abbildung 4.4.: Das Verhältnis $\frac{s}{s+b}$ (s : Signal, b : Untergrund) für verschiedene Schnitte der Form $NN_{out} > NN_{Schnitt}$ zur Ermittlung des am besten geeigneten Schnittes. Das Maximum wird für den Schnitt $NN_{out} > 0,3$ erreicht und ist durch eine horizontale und vertikale Linie im Schaubild gekennzeichnet.

4.4.3. Vergleich der Reduktionsmethoden

Die Werte von $R_{s,b}$, p_s und α_s für die Reduktion mittels kombinierten Schnitten, neuronalem Netz und ohne Reduktion sind in Tabelle 4.4 zusammengefasst. Alle Werte steigen durch Anwenden der weiteren Maßnahmen zur Untergrundreduktion an, mit Verwendung eines neuronalen Netzes liegen diese deutlich über den Werten einer Reduktion mit kombinierten Einzelschnitten. Das Signal-Untergrund-Verhältnis $R_{s,b}$ ist etwa das 1,6-fache der kombinierten Einzelschnitte für eine Diskrimination mit neuronalem Netz, in der Arbeit von C. Heidecker [1] war es für die 7-TeV-Messungen noch etwa nur das 1,04-fache. Die bessere Trennbarkeit von Signal- und Untergrundereignissen mit Hilfe neuronaler Netze ist ein wesentlicher Aspekt, der den Studierenden in dem ausgearbeiteten Praktikumsversuch verdeutlicht werden soll. Im Praktikum wird dadurch, dass die Studierenden mögliche Variablen für kombinierte Schnitte selbst wählen müssen, verdeutlicht, dass durch die Verwendung eines neuronalen Netzes sehr viel effektiver einer Untergrundreduktion erfolgen kann als durch das Ermitteln von möglichen Schnittkombinationen.

Tabelle 4.4.: Vergleich der Untergrundreduktion für kombinierte Einzelschnitte, mit neuronalem Netz und keiner Untergrundreduktion.

Methode	$R_{s,b}$	p_s	α_s
Keine Reduktion	0,165	0,142	43,4
Kombinierte Einzelschnitte	0,513	0,339	57,3
Neuronales Netz	0,822	0,451	74,9

4.5. Skalierung der simulierten Ereignisse

Die in Tabelle 4.1 erkennbare Abweichung der modellierten Ereignisanzahlen zu den Ereignisanzahlen aus Messdaten wird durch eine Maximum-Likelihood-Methode korrigiert. In die Anzahl der Ereignisse eines Prozesses $N^{P,\text{Theorie}}$ gehen die integrierte Luminosität \mathcal{L} , der Wirkungsquerschnitt für den Prozess σ_P sowie die Effizienz ϵ_P für diesen Prozess und die Akzeptanz A ein:

$$N^{P,\text{Theorie}} = \sigma_P \cdot \epsilon_P \cdot A \cdot \mathcal{L}. \quad (4.2)$$

Zur Anpassung der Ereignisanzahlen werden Faktoren zur Skalierung der Prozesse, β_P , mit einer Maximum-Likelihood-Methode bestimmt,

$$N^P = \beta_P \cdot N^{P,\text{Theorie}}. \quad (4.3)$$

Die Anpassung erfolgt über den gesamten Bereich der Ausgabe des neuronalen Netzes, welches mit den im Abschnitt 4.4.2 genannten Variablen trainiert wird. Die resultierenden Faktoren sind in Tabelle 4.5 aufgelistet und werden beim rechten Schaubild in Abbildung 4.5 angewandt, um die auf dem linken Schaubild zu erkennende Differenz zwischen den gemessenen Ereignissen (Data) und den aufaddierten modellierten Ereignissen zu korrigieren.

Die verwendete Likelihood-Funktion ist für eine angenommene Poisson-Verteilung definiert als

$$\tilde{L}(\beta_P) = \prod_{i=1}^{N_{\text{KI}}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!}, \quad (4.4)$$

wobei N_{KI} die Anzahl der Bins eines Histogramms, n_i die Anzahl an Ereignissen der gemessenen Daten in Bin i und μ_i die Anzahl an angepassten vorhergesagten Ereignissen in diesem Bin ist.

$$\mu_i = \sum_P^{N_P} \beta_P \cdot N_i^{P,\text{Theorie}}. \quad (4.5)$$

Summiert wird über alle modellierten Prozesse N_P , dabei ist β_P der Skalierungsfaktor für Prozess P und $N_i^{P,\text{Theorie}}$ die Anzahl der vorhergesagten Ereignisse für Prozess P in Bin i . Demnach entspricht für den Wert $\beta_P = 1$ die Anzahl der Ereignisse für den Prozess P der theoretisch vorhergesagten Anzahl, bei einer Abweichung vom Wert 1 wird der Prozess skaliert.

Durch die Multiplikation der Likelihood-Funktion mit einem Faktor C_P wird die mögliche Abweichung vom Wert 1 des Skalierungsfaktors für den Prozess P einschränkt. N_P^E sind alle Untergrundprozesse für die der Parameter eingeschränkt wird.

$$L(\beta_P) = \prod_{i=1}^{N_{\text{KI}}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!} \cdot \prod_P^{N_P^E} C_P. \quad (4.6)$$

Die Faktoren sind als Gaußverteilungen in der Form

$$C_P = \frac{1}{\sqrt{2\pi} \cdot \Delta_P} \cdot e^{-\frac{(\beta_P - 1)^2}{2\Delta_P^2}} \quad (4.7)$$

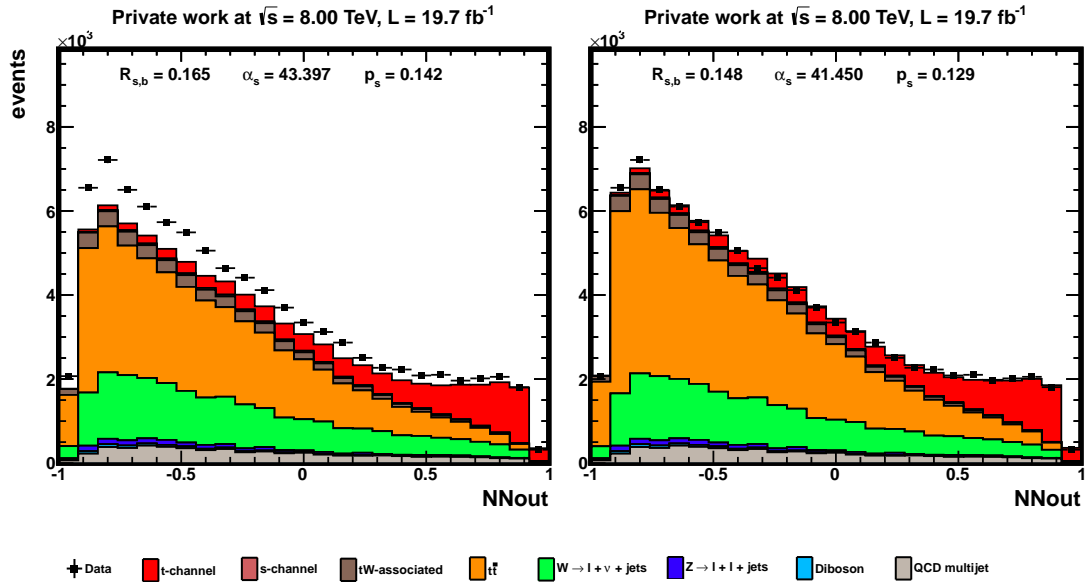


Abbildung 4.5.: Links: Ausgabe des neuronalen Netzes ohne Skalierungsfaktoren. Die Anzahl an Ereignissen der simulierten Ereignisse weicht von den gemessenen Daten ab. **Rechts:** Durch eine Maximum-Likelihood-Methode korrigierte Ausgabe.

realisiert, wobei Δ_P die Standardabweichung des Parameters β_P vom Wert 1 ist.

Die Bestimmung der die Likelihood-Funktion maximierenden Parameter erfolgt im Rahmen der für das Computer-Praktikum entwickelten Anwendung durch ein ROOT-Makro, in welchem das Minimum des negativen Logarithmus der Likelihood-Funktion bestimmt wird. Der beim Logarithmieren auftretende Summand $\sum_{i=1}^{N_{kl}} \ln(n_i!)$ hängt nicht vom Parameter β_P ab, die zu minimierende Funktion lautet somit:

$$\begin{aligned}
 f(\beta_P) &= -\ln(L(\beta_P)) - \sum_{i=1}^{N_{kl}} \ln(n_i!) \\
 &= \sum_{i=1}^{N_{kl}} \mu_i - n_i \ln(\mu_i) - \sum_P^{N_P^E} \left(\frac{(\beta_P - 1)^2}{2\Delta_P^2} + \ln(\sqrt{2\pi} \cdot \Delta_P) \right).
 \end{aligned} \tag{4.8}$$

Die Minimierung erfolgt durch Verwendung des im Minuit-Paket [26] implementierten MIGRAD-Algorithmus [27]. Der Parameter β_{QCD} für den QCD-Multijet Untergrund ist dabei auf 1 fixiert.

Tabelle 4.5.: Durch eine Maximum-Likelihood-Methode bestimmte Skalierungsfaktoren β_{Prozess} zur Anpassungen der vorhergesagten Ereignisanzahlen für simulierte Prozesse an tatsächliche Bedingungen. Für einige Prozesse wurde die mögliche Abweichung Δ_P des Parameters β_P vom Wert 1 eingeschränkt.

Prozess	Δ_P	β_{Prozess}
Einzel-Top t-Kanal	-	$1,02 \pm 0,03$
Einzel-Top s-Kanal	0,3	$0,99 \pm 0,08$
tW	0,3	$0,99 \pm 0,08$
t \bar{t}	-	$1,26 \pm 0,06$
W+jets	-	$0,98 \pm 0,13$
Z+jets	0,3	$1,00 \pm 0,08$
Diboson	0,3	$1,00 \pm 0,08$

5. Messungen

Verschiedene Eigenschaften des Top-Quarks können mit dem im Rahmen der Bachelorarbeit erweiterten Anwendungspaket für das Praktikum berechnet werden. Bei Messungen, bei denen Untergrundprozesse zu reduzieren sind, werden die Resultate der Untergrundreduktion mit den im Abschnitt 4.4 ermittelten kombinierten Einzelschnitten verglichen mit dem Schnitt auf die Ausgabe von neuronalen Netzen (NNout). Es ist zu erwarten, dass durch das Schneiden auf die Netz-Ausgabe nach Training mit der zu untersuchenden Variable das Ergebnis beeinflusst wird, daher wird jeweils das Resultat für ein Training ohne die zu untersuchende Variable und mit dieser betrachtet.

Systematische Unsicherheiten werden für diese Analyse nicht berücksichtigt, bei den angegebenen Messunsicherheiten handelt es sich um rein statistische Größen.

5.1. Wirkungsquerschnitt und CKM-Matrix-Element V_{tb}

5.1.1. Wirkungsquerschnitt

Nach Gleichung 4.3 gilt für den Wirkungsquerschnitt σ_t und die Unsicherheit $\Delta\sigma_t$ unter Verwendung der Proportionalität zwischen Wirkungsquerschnitt und der Anzahl der Ereignisse aus Gleichung 4.2:

$$\sigma_t = \beta_{t\text{-Kanal}} \cdot \sigma_t^{\text{Theorie}} \quad (5.1)$$

$$\Delta\sigma_t = \sqrt{\left(\sigma_t^{\text{Theorie}} \cdot \Delta\beta_{t\text{-Kanal}}\right)^2}. \quad (5.2)$$

Unter Verwendung vom in Tabelle 4.5 angegebenen Wert für $\beta_{t\text{-Kanal}}$ und dem für die simulierten Prozesse verwendeten theoretischen Wirkungsquerschnitt bei einer Schwerpunktsenergie von $\sqrt{s} = 8 \text{ TeV}$, $\sigma_t^{\text{Theorie}} = 87,2^{+4,8}_{-3,2} \text{ pb}$ [28], ist der gemessene Wirkungsquerschnitt

$$\sigma_t = (88,8 \pm 5,5) \text{ pb}. \quad (5.3)$$

Das Ergebnis stimmt innerhalb der Unsicherheit mit dem gemessenen Ergebnis der CMS-Kollaboration von $\sigma_t = (83,6 \pm 2,3) \text{ pb}$ [29] und dem theoretischen Wert überein.

5.1.2. Matricelement

Zieht man anomale Kopplungen jenseits des Standardmodells am Vertex Wtb beim t -Kanal Prozess in Betracht, so kann dies als ein zusätzlicher Faktor f_{Lv} für linkshändige Vektorkopplung am Matricelement als $|f_{Lv}V_{tb}|$ parametrisiert werden[30–32]. Unter der Annahme $|V_{td}|, |V_{ts}| \ll |V_{tb}|$ ist das Verzweigungsverhältnis \mathcal{B} beim Zerfall des Top-Quark in Wb nahezu 1, somit ergibt sich für das Matricelement und die Unsicherheit mit dem Skalierungsfaktor aus Tabelle 4.5:

$$|f_{Lv}V_{tb}| = \sqrt{\beta_{t\text{-Kanal}}}, \quad (5.4)$$

$$\Delta|f_{Lv}V_{tb}| = \sqrt{\left(\frac{\Delta\beta_{t\text{-Kanal}}}{2\sqrt{\beta_{t\text{-Kanal}}}}\right)^2}, \quad (5.5)$$

$$\Rightarrow |f_{Lv}V_{tb}| = 1,009 \pm 0,017. \quad (5.6)$$

Dieser Wert stimmt im Rahmen der Unsicherheit mit dem veröffentlichten Wert [6] von $0,999\,14 \pm 0,000\,05$ überein.

5.2. Ladungsverhältnis der Leptonen

Aus dem Verhältnis R_Q von Ereignissen mit detektierten positiv geladenen Leptonen zu solchen mit negativ geladenen, lässt sich auf das Verhältnis von erzeugten Top-Quarks und Top-Antiquarks schließen (siehe Abschnitt 1.3.1) und mit Modellrechnungen für verschiedene PDFs vergleichen.

Es werden nur solche Ereignisse betrachtet, bei denen das Lepton als ein Myon detektiert wird (siehe Abschnitt 4.2). Die resultierenden Histogramme für die verschiedenen Schnittmethoden sind in Abbildung 5.3 zu finden.

Das Verhältnis R_Q entspricht nach Abzug der modellierten Untergrundereignisse von den gemessenen Daten dem Verhältnis der Anzahl an Ereignissen mit positiven Myonen N_{μ^+} zu der Anzahl an Ereignissen mit negativen Myonen N_{μ^-} :

$$R_Q = \frac{N_{\mu^+}}{N_{\mu^-}} \quad (5.7)$$

$$\Delta R_Q = \sqrt{\left(\frac{\Delta N_{\mu^+}}{N_{\mu^-}}\right)^2 + \left(\frac{N_{\mu^+}}{N_{\mu^-}^2} \cdot \Delta N_{\mu^-}\right)^2}. \quad (5.8)$$

Untersucht wird der Einfluss verschiedener Schnittmethoden auf das Ladungsverhältnis (Tabelle 5.1 & Abbildung 5.1).

Die sich ergebenden Verhältnisse unterscheiden sich für ein Training des neuronalen Netzes ohne die Variable Q^1 und mit dieser Variablen, sowie das letztere Ergebnis mit dem Ergebnis bei kombinierten Einzelschnitten.

Ein Vergleich der ermittelten Ladungsverhältnisse mit simulierten Ereignissen für verschiedene PDF-Modelle ist in Abbildung 5.2 dargestellt. Es ist zu erkennen, dass alle bis auf eines der betrachteten PDF-Modelle mit dem neuronalen Netzwerk Schnitt ohne Q^1 als Trainingseingabe übereinstimmen. Zwei der Modelle stimmen mit den kombinierten Einzelschnitten überein. Kein Modell stimmt mit dem Schnitt auf die Ausgabe des neuronalen Netzes überein, wenn die Variable Q^1 als Eingabe für das Training verwendet wird.

Die Werte der vereinfachten Analyse stimmen für kombinierte Einzelschnitte und Verwendung des neuronalen Netzes ohne die zu untersuchende Variable im Rahmen der Unsicherheit auch mit dem ermittelten Wert der CMS-Kollaboration von $1,95 \pm 0,10$ [29] überein. Das Ergebnis für ein neuronales Netz mit der zu untersuchenden Variablen stimmt auch knapp im Rahmen der Unsicherheit nicht mit diesem Ergebnis überein.

Anhand der ermittelten Asymmetrien der simulierten Prozesse in Tabelle 5.2 und den Histogrammen in Abbildung 5.3 wird deutlich, dass die vorhergesagten Asymmetrien insbesondere für den Einzel-Top-Quark-Prozess im t-Kanal und für den Gesamtuntergrund in etwa gleich ist für kombinierte Einzelschnitte und für ein neuronales Netz ohne Q^1 als Trainingseingabe. Das unterschiedliche Ergebnis für diese beiden Untersuchungen lässt sich demnach nur anhand der Abweichung der gemessenen Daten von den simulierten Ereignissen erklären. Anhand der Histogramme erkennt man unterschiedlich starke Abweichungen für beide Methoden; für kombinierte Einzelschnitte ist die Abweichung für positiv geladene Myonen größer als für negative, das Gegenteil ist für die Reduktion mit neuronalen Netzen der Fall.

Für ein neuronales Netz mit der zu untersuchenden Variable als Trainingseingabe ist bereits eine stärkere Asymmetrie als die zu erwartende doppelt so häufige Einzel-Top-Quark-Produktion im t-Kanal bei alleiniger Berücksichtigung von Valenzquarks des Protons (siehe Abschnitt 1.3.1) zu erkennen, des Weiteren weisen alle Untergrundprozesse eine starke Asymmetrie hin zu mehr positiv geladenen Myonen auf, insbesondere für die gleichzeitige Produktion eines Top- und Antitop-Quarks $t\bar{t}$ ist eine solche Asymmetrie nicht zu erwarten.

Anhand dessen wird deutlich, dass für die Diskrimination von Untergrund und Signal durch Verwendung eines neuronalen Netzes die Wahl der Trainingsvariablen das Ergebnis beeinflussen kann. Neben der Implementierung der Analysemethode durch Ergänzung des Programmcodes ist dies ein wesentlicher Aspekt, der den Studierenden bei der Durchführung dieser Analyse deutlich gemacht werden soll. Durch die Verwendung von Q^1 als Trainingseingabe werden mehr Ereignisse positiver Myonen-Ladung als mit negativer ausgewählt, da der Signalanteil für positive Myonen größer ist, wie anhand des Histogramms in Abbildung A.29 zu erkennen ist, wodurch sich die nicht zu erwartende starke Untergrundasymmetrie ergibt. Eine mögliche Erklärung für die unterschiedliche Abweichung der simulierten Ereignisse zu den Daten lässt sich auf ähnliche Weise finden; durch die kombinierten Einzelschnitte werden die vorhandenen Abweichungen in der ausgewählten Region auf der geschnittenen Variable ausgewählt, wohingegen durch Verwendung eines neuronalen Netzes die Abweichungen aller Variablen gleichzeitig mit berücksichtigt wird. Eine genauere Untersuchung dieser Abweichung ist im Rahmen dieser Bachelorarbeit und im Praktikumsversuch nicht möglich.

Untergrund-Reduktion mit:	Verhältnis R_Q
Kombinierten Schnitten	$2,02 \pm 0,11$
NN-Schnitt ohne Q^1 im Training	$1,81 \pm 0,08$
NN-Schnitt	$2,16 \pm 0,10$

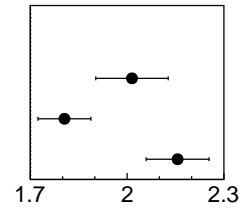


Tabelle 5.1 & Abbildung 5.1: Gemessenes Leptonen-Ladungsverhältnis für verschiedene Schnitte zur Reduktion von Untergrundereignissen.

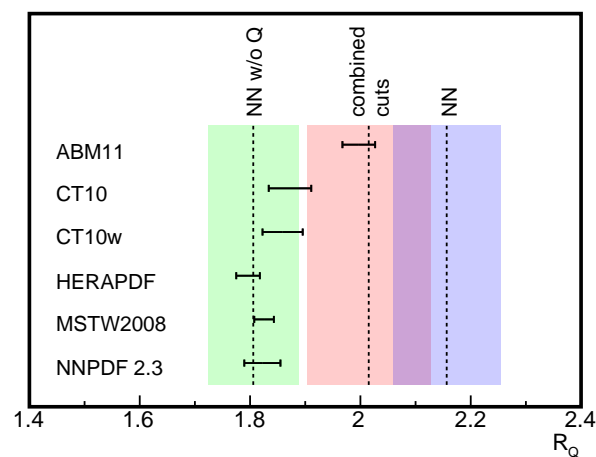


Abbildung 5.2.: Vergleich der verschiedenen Schnittmethoden mit simulierten Werten verschiedener PDF. Die simulierten Werte wurden aus der CMS-Veröffentlichung [29] entnommen.

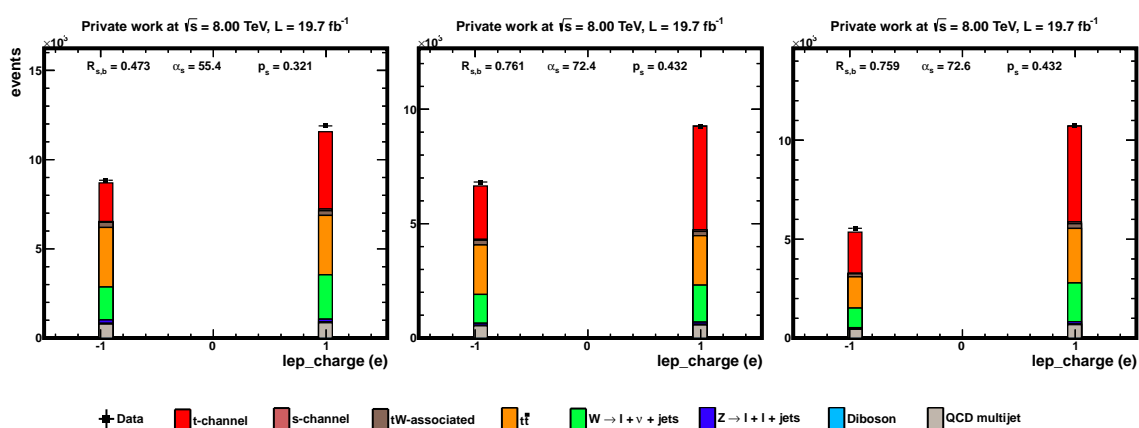


Abbildung 5.3.: Verteilung der Leptonen-Ladung Q^1 für verschiedene Schnitte. **Links:** kombinierte Einzelschnitte. **Mitte:** Schnitt auf neuronales Netzwerk ohne Q^1 als Trainingseingabe. **Rechts:** Schnitt auf neuronales Netzwerk mit Q^1 als Trainingseingabe.

Prozess	kombinierte Schnitte		neuronales Netz			
	R_Q	Anteil (%)	ohne Q^1 -Training		mit Q^1 -Training	
	R_Q	Anteil (%)	R_Q	Anteil (%)	R_Q	Anteil (%)
Einzel-Top t	2,00	32,10	1,95	43,21	2,33	43,38
Untergrund	1,11	67,90	1,10	56,79	1,82	56,62
$t\bar{t}$	1,00	32,92	1,00	27,14	1,78	27,22
W+jets	1,35	21,35	1,28	18,06	2,02	18,11
QCD	1,09	8,09	1,07	6,94	1,57	6,66
tW	0,90	2,71	0,91	2,48	1,60	2,41
Z+jets	0,72	1,62	1,04	1,17	1,50	1,16
Einzel-Top s	2,18	0,68	1,79	0,67	3,01	0,72
Diboson	1,45	0,54	1,41	0,33	2,50	0,34

Tabelle 5.2.: Die Ladungsasymmetrien R_Q aller simulierter Prozesse für verschiedene Methoden zur Reduktion des Untergrundes und der prozentuale Anteil des entsprechenden Prozesses an allen simulierten Prozessen. Die Asymmetrie der Untergrundprozesse zur Einzel-Top-Quark-Produktion im t-Kanal (Einzel-Top t) ist zusammengefasst zu einem Gesamtuntergrund aufgelistet und zusätzlich jeder einzelne Beitrag sortiert nach dessen Anteil.

5.3. Masse des Top-Quarks

Mit Hilfe des ausgearbeiteten Anwendungspakets für das Computerpraktikum kann die Masse des Top-Quarks in einer vereinfachten Analyse bestimmt werden. Es wird das Ergebnis für eine Untergrundreduktion durch Anwenden von mehreren Einzelschnitten und von neuronalen Netzen (siehe Abschnitt 4.4) verglichen. Es wird zudem, wie bereits im Abschnitt 5.2, das Ergebnis bei Verwendung eines Netzes mit der zu untersuchenden Variable als Trainingseingabe mit dem Ergebnis ohne diese Variable im Training verglichen.

Zur eigentlichen Massenbestimmung muss zunächst eine Kalibrierung durch Verwendung verschiedener modellierter Top-Quark-Massen durchgeführt werden.

5.3.1. Kalibrierung

Aufgrund von Detektoreinflüssen stimmen die rekonstruierten Massenwerte bei den Maxima der Verteilung nicht bei der tatsächlichen Masse überein und müssen durch eine Kalibrierung korrigiert werden.

Durch Untersuchungen von Monte-Carlo-Simulationen mit verschiedenen angenommenen tatsächlichen Top-Quark-Massen M_t^T wurde ein linearer Zusammenhang zwischen der tatsächlichen Top-Quark-Masse und der rekonstruierten Top-Quark-Masse M_t^R festgestellt:

$$M_t^T = c + m \cdot M_t^R \quad (5.9)$$

Um die Parameter m und c in Gleichung 5.9 zu bestimmen, werden zusätzlich zu den simulierten t-Kanal-Ereignissen mit einer angenommenen tatsächlichen Top-Quark-Masse M_t^T von 172,5 GeV/c² Simulationen einer tatsächlichen Top-Quark-Masse von 171,5 GeV/c² (−1 GeV/c²) und 173,5 GeV/c² (+1 GeV/c²) verwendet.

Die verschiedenen Schnittmethoden werden gleichermaßen für alle drei simulierten Massen durchgeführt. Durch Gauß-Kurven-Anpassungen wird die rekonstruierte Masse für die simulierten Ereignisse bestimmt (Abbildungen 5.5 - 5.7 (a) - (c)). Die Ergebnisse der Gauß-Kurven-Anpassungen für die verschiedenen Schnittmethoden zur Reduktion der Untergrundereignisse sind in Tabelle 5.3 dargestellt.

Eine lineare Regression an die Datenpunkte (M_t^R, M_t^T) liefert die Werte für die Parameter m und c sowie die Kovarianzmatrix-Elemente $\sigma_{i,j}$. Der Ursprung des Koordinatensystems wurde für die Durchführung der Regression in den Punkt ($M_t^R(172,5 \text{ GeV/c}^2), M_t^T = 172,5 \text{ GeV/c}^2$) gelegt. Die Ergebnisse der Regression sind in Tabelle 5.3 zu finden.

Die Unsicherheit für die tatsächliche Masse wird anhand der ermittelten Werte wie folgt berechnet:

$$\Delta M_t^T = \sqrt{\sigma_{c,c}^2 + (M_t^R \cdot \sigma_{m,m})^2 + (m \cdot \Delta M_t^R)^2 + 2 (M_t^R \cdot \sigma_{c,m})}. \quad (5.10)$$

5.3.2. Rekonstruktion

Durch eine Gauß-Kurven-Anpassung an die Daten nach Abzug der modellierten Untergrundprozesse wird die rekonstruierte Masse der Messdaten $M_t^{R,DATA}$ bestimmt (Abbildung 5.5 - 5.7 (b)). Mit Hilfe der ermittelten Kalibrierung in Abschnitt 5.3.1 wird aus der rekonstruierten Masse die korrespondierende tatsächliche Top-Quark-Masse bestimmt (siehe auch Abbildung 5.5 - 5.7 (d)).

Die rekonstruierten Massen und die ermittelten tatsächlichen Massen sind in Tabelle 5.5 zu finden, die ermittelten Top-Quark-Massen stimmen innerhalb der Unsicherheiten mit dem veröffentlichten Wert [33] von $173,34 \pm 0,27 \pm 0,71 \text{ GeV/c}^2$ überein. Die statistische Unsicherheit ist für beide Untergrundreduktionen mit neuronalem Netz gleich und etwas geringer als bei der Reduktion mittels kombinierter Schnitte. Zu einer nicht vernachlässigbaren Vergrößerung der ermittelten statistischen Unsicherheit trägt die Unsicherheit auf der ermittelten Kalibrierung bei.

Im Gegensatz zur Ladungsasymmetrie stimmen die in dieser Analyse ermittelten Massenwerte im Rahmen der statistischen Unsicherheiten überein, bei der Betrachtung der Histogramme lässt sich jedoch ein ähnlicher Effekt auf die Untergründe wie schon bei der Ladungsasymmetrie ausmachen: Für die kombinierten Einzelschnitte und den Schnitt mittels neuronalem Netz, das ohne M^{lvb} trainiert wurde, ist die Form der Untergrundereignisse in etwa gleich, die Verteilungen sind in etwa gleich breit und die Position des Maximums für die Untergründe ist hin zu kleineren Massen als für die Einzel-Top-Quark-Produktion verschoben, wohingegen für ein neuronales Netz mit dieser Variablen als Eingabe die

Form der Untergründe sichtbar verändert und hin zum Maximum des Signalprozesses verschoben ist.

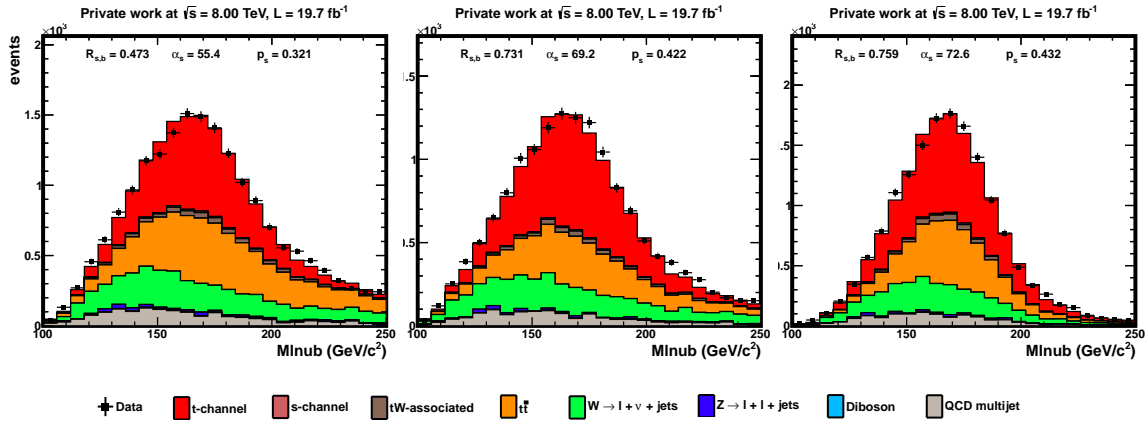


Abbildung 5.4.: Verteilung der rekonstruierten Top-Quark-Masse M^{lvb} für verschiedene Schnitte. **Links:** kombinierte Einzelschnitte. **Mitte:** Schnitt auf neuronales Netzwerk ohne M^{lvb} als Trainingseingabe. **Rechts:** Schnitt auf neuronales Netzwerk mit M^{lvb} als Trainingseingabe.

Tabelle 5.3.: Durch Gauß-Kurven-Anpassung an simulierte t-Kanal-Ereignisse bestimmte Top-Quark-Massen für verschiedene angenommene tatsächliche Massen zur Bestimmung der Kalibrierung für die Korrektur der aus den Daten rekonstruierten Top-Quark-Masse. Durchgeführt für verschiedene Verfahren zur Reduktion des Untergrundes. Werte in GeV/c^2 .

Untergrund-Reduktion mit:	Simulierte rekonstruierte Masse		
Kombinierten Schnitten	$167,7 \pm 0,1$	$168,4 \pm 0,1$	$169,0 \pm 0,1$
NN-Schnitt ohne M^{lvb} -Training	$167,4 \pm 0,1$	$168,2 \pm 0,1$	$168,8 \pm 0,1$
NN-Schnitt mit M^{lvb} -Training	$167,6 \pm 0,1$	$168,2 \pm 0,1$	$168,7 \pm 0,1$
Simulierte „wahre“ Masse	171,5	172,5	173,5

Tabelle 5.4.: Ergebnisse der lineare Regression an die rekonstruierten simulierten Top-Quark-Massen (Tabelle 5.3) für Rückgewinnung der „wahren“ Top-Quark Masse aus der rekonstruierten.

Untergrund-Reduktion mit:	c	m	$\sigma_{c,c}$	$\sigma_{m,m}$	$\sigma_{m,c}$
Kombinierten Schnitten	$0,10 \pm 0,10$	$1,57 \pm 0,20$	0,010	0,042	0,003
NN-Schnitt ohne M^{lvb} -Training	$0,08 \pm 0,10$	$1,49 \pm 0,18$	0,009	0,033	0,002
NN-Schnitt mit M^{lvb} -Training	$0,01 \pm 0,10$	$1,80 \pm 0,23$	0,009	0,051	0,001

Tabelle 5.5.: Gemessene Top-Quark-Massen für verschiedene Schnitte und mit Hilfe der Kalibrierung korrigierte Massen. Werte in GeV/c^2 .

Untergrund-Reduktion mit:	Rekonstruierte Masse	Korrigierte Masse
Kombinierten Schnitten	$168,1 \pm 0,7$	$172,1 \pm 1,1$
NN-Schnitt ohne M^{lvb} -Training	$168,0 \pm 0,6$	$172,3 \pm 0,9$
NN-Schnitt mit M^{lvb} -Training	$168,1 \pm 0,5$	$172,4 \pm 0,9$

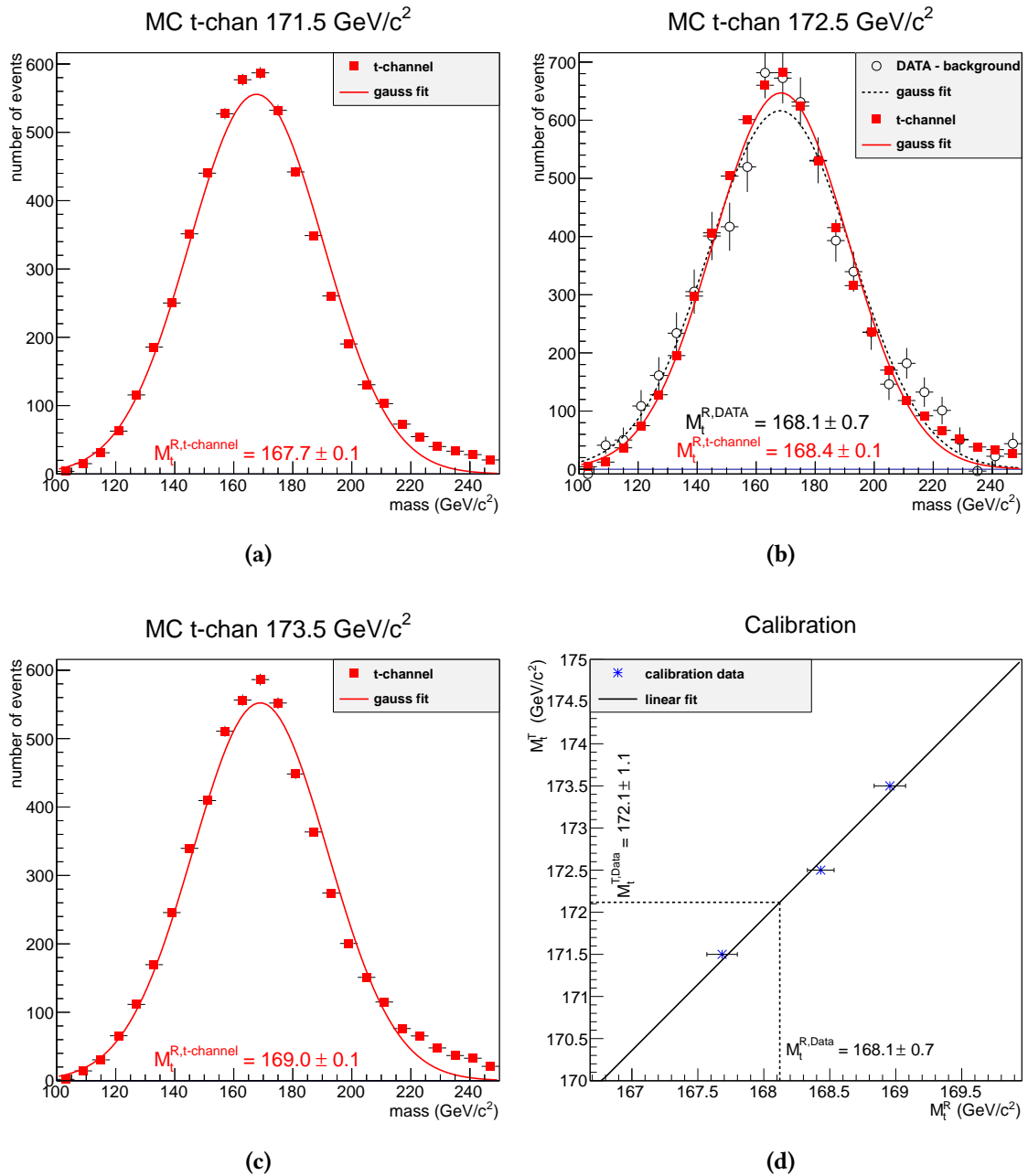


Abbildung 5.5.: Kalibrierung für kombinierte Einzelschnitte:

- (a) - (c): Simulierte t-Kanal-Ereignisse für verschiedene simulierte Top-Quark-Massen und Gauß-Kurven-Anpassung für die Bestimmung der rekonstruierten Masse zur Kalibrierung.
- (b): Zusätzlich gemessene Ereignisse mit abgezogenen Untergrundprozessen und Gauß-Kurven-Anpassung zur Bestimmung der rekonstruierten Top-Quark Masse.
- (d): Kalibrierung mit den simulierten Anpassungen aus (a) - (c) zur Rückgewinnung der Top-Quark-Masse aus der rekonstruierten.

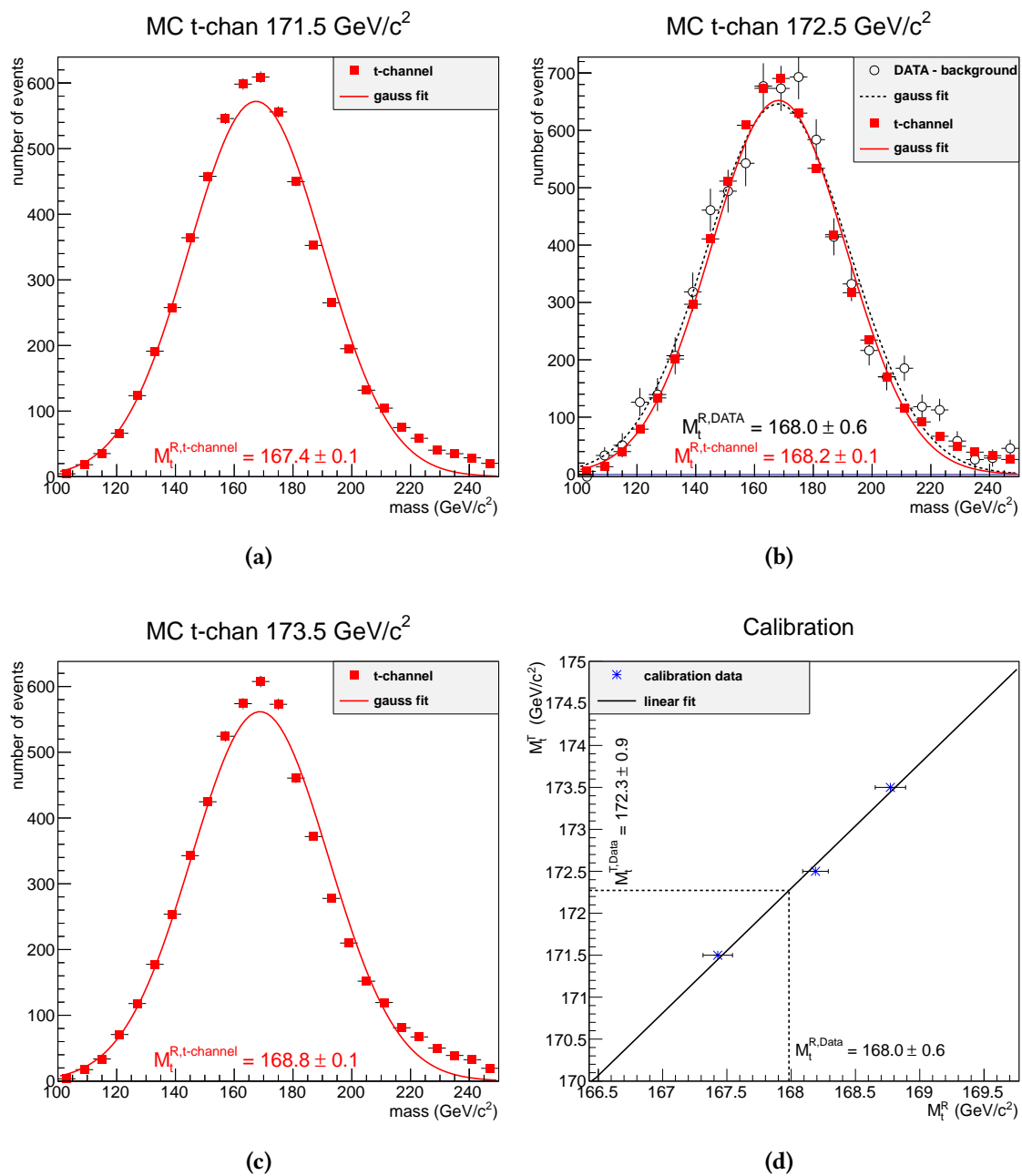


Abbildung 5.6.: Kalibrierung für NN-Schnitt (ohne $M^{l\nu b}$ als Trainingseingabe):

- (a) - (c): Simulierte t-Kanal-Ereignisse für verschiedene simulierte Top-Quark-Massen und Gauß-Kurven-Anpassung für die Bestimmung der rekonstruierten Masse zur Kalibrierung.
- (b): Zusätzlich gemessene Ereignisse mit abgezogenen Untergrundprozessen und Gauß-Kurven-Anpassung zur Bestimmung der rekonstruierten Top-Quark Masse.
- (d): Kalibrierung mit den simulierten Anpassungen aus (a) - (c) zur Rückgewinnung der Top-Quark-Masse aus der rekonstruierten.

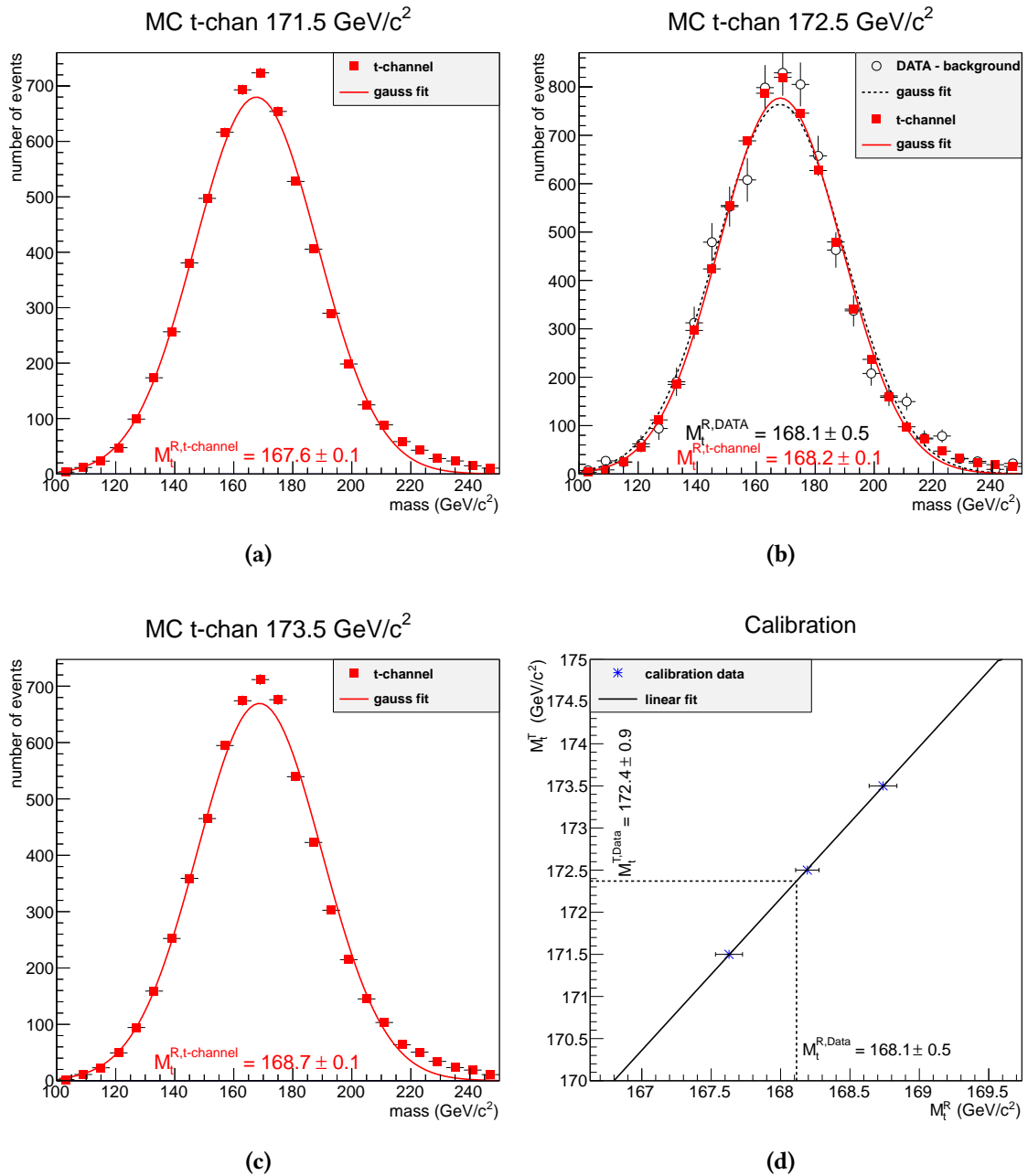


Abbildung 5.7.: Kalibrierung für NN-Schnitt (mit M_t^{lvb} als Trainingseingabe):

- (a) - (c): Simulierte t-Kanal-Ereignisse für verschiedene simulierte Top-Quark-Massen und Gauß-Kurven-Anpassung für die Bestimmung der rekonstruierten Masse zur Kalibrierung.
- (b): Zusätzlich gemessene Ereignisse mit abgezogenen Untergrundprozessen und Gauß-Kurven-Anpassung zur Bestimmung der rekonstruierten Top-Quark Masse.
- (d): Kalibrierung mit den simulierten Anpassungen aus (a) - (c) zur Rückgewinnung der Top-Quark-Masse aus der rekonstruierten.

6. Fazit

Das auf der Vorarbeit von C. Heidecker[1] basierende und mit dieser Arbeit ausgearbeitete Programmpaket lässt sich mit einigen optimierten ProgrammROUTINEN auch auf die mit dem CMS-Detektor gewonnenen Daten aus dem 8-TeV-Betrieb des LHC und zugehörigen modellierten Prozessen anwenden.

Durch Verwendung des neuronalen Netzes kann eine stärkere Reduktion von Untergrundereignissen erzielt werden als durch kombinierte Schnitte auf einzelne Variablen. Das Signal-Untergrund-Verhältnis ist für die 8-TeV-Ereignisse bei Verwendung des neuronalen Netzes etwa 1,6-fach so hoch wie das Verhältnis bei Verwendung von kombinierten Einzelschnitten, daher wird für die 8-TeV-Ereignisse die Fähigkeit des neuronalen Netzes zur Trennung von Signal- und Untergrundereignissen deutlicher als bei Verwendung der 7-TeV-Ereignissen, bei welchen die Reduktion unter Verwendung eines neuronalen Netzes nur etwa das 1,04-fache der kombinierten Schnitte betrug. Die in den vereinfachten Analysen bestimmten Größen σ_t , $|f_{Lv} V_{tb}|$ und M^{lvb} stimmen im Rahmen der statistischen Unsicherheit mit den theoretisch vorhergesagten bzw. in vorherigen Experimenten gemessenen Werten überein. Eine Abweichung von Q^l von allen betrachteten theoretischen Vorhersagen ist nur für ein Training des neuronalen Netzes mit dieser Variablen zu beobachten, was jedoch den Erwartungen entspricht und im Praktikumsversuch mögliche Beeinflussungen von Prozessen bei unsorgfältiger Auswahl der Eingabevariablen verdeutlicht. Die gemessene Abweichung für kombinierte Schnitte und für ein neuronales Netz ohne Q^l als Trainingseingabe kann durch die ungleichmäßige Abweichung der simulierten Prozesse von den Daten auf den geschnittenen Variablen erklärt werden. Eine genauere Untersuchung und mögliche Verbesserungen würde über den Rahmen dieser Arbeit hinausgehen, ebenso wie die Analyse der systematischen Unsicherheiten. Die ermittelten statistischen Unsicherheiten aller Messungen sind unter Verwendung von neuronalen Netzen zur Untergrundreduktion etwas geringer als bei Verwendung von kombinierten Einzelschnitten.

Durch die Verwendung von 8-TeV-Daten wird demnach eines der Hauptziele des Praktikumsversuches, die Fähigkeit zur stärkeren Trennung von Untergrund- und Signalereignissen durch den Einsatz neuronaler Netze, deutlicher erkennbar als mit den 7-TeV-Daten, bei welchem der Datensatz nur etwa ein Zwanzigstel des in dieser Arbeit verwendeten betrug. Durch die Analyse der Ladungsasymmetrie wird erkennbar, dass die Wahl der Eingabeparameter für das neuronale Netz einen Einfluss haben kann. Eine mögliche Erweiterung des Praktikumsversuches wäre dahingehend denkbar, dass auch der Einfluss der Gestalt des Netzes selbst z. B. durch die Anzahl der Neuronen in der verborgenen Schicht oder die Wahl der Fehlerfunktion untersucht werden kann.

Die für den Praktikumsversuch entworfenen Übungsblätter für die Studierenden sind im Anhang A.3 beigelegt. Durch die Bearbeitung der Übungsblätter soll den Studierenden die in dieser Arbeit aufgezeigte Möglichkeit zur besseren Trennung von Signal- und Untergrundereignissen bei Verwendung von neuronalen Netzen gegenüber von Einzelschnittkombinationen verdeutlicht werden und in der anschließenden Analyse auch unerwünschte Nebeneffekte durch unvorsichtige Verwendung von neuronalen Netzen aufgezeigt werden. Des Weiteren sind gebräuchliche Methoden in der Datenanalyse wie die Anpassung von simulierten Ereignissen mittels einer Maximum-Likelihood-Methoden und die Möglichkeit einer Messkorrektur mit simulierter Ereignissen von den Studierenden mit Hilfestellungen selbst anzuwenden.

Literatur

- [1] C. A. Heidecker. “Studium von t-Kanal Ereignissen der Einzel-Top-Quark-Produktion am CMS Experiment mittels Neuronaler Netzwerke. Vorbereitung eines Computerpraktikumsversuchs zur Vorlesung Teilchenphysik für Fortgeschrittene II”. Bachelorarbeit. Fakultät für Physik, Karlsruher Institut für Technologie (KIT), 2. Aug. 2013.
- [2] *ROOT A Data Analysis Framework*. CERN. URL: <https://root.cern.ch> (besucht am 17. 04. 2015).
- [3] *NeuroBayes*. phi-t. URL: <http://neurobayes.phi-t.de/> (besucht am 26. 03. 2015).
- [4] G. Aad u. a. *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*. 2015. arXiv: 1503.07589 [hep-ex].
- [5] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 1963), S. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.10.531>.
- [6] K. Olive u. a. “Review of Particle Physics”. In: *Chin.Phys.* C38 (2014), S. 090001. DOI: 10.1088/1674-1137/38/9/090001.
- [7] A. Martin u. a. “Parton distributions for the LHC”. In: *Eur.Phys.J.* C63 (2009), S. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].
- [8] F. Abe u. a. “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Phys. Rev. Lett.* 74 (14 1995), S. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [9] S. Abachi u. a. “Observation of the Top Quark”. In: *Phys. Rev. Lett.* 74 (14 1995), S. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [10] V. M. Abazov u. a. “Observation of Single Top-Quark Production”. In: *Phys. Rev. Lett.* 103 (9 2009), S. 092001. DOI: 10.1103/PhysRevLett.103.092001. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.103.092001>.
- [11] T. Aaltonen u. a. “Observation of Electroweak Single Top-Quark Production”. In: *Phys. Rev. Lett.* 103 (9 2009), S. 092002. DOI: 10.1103/PhysRevLett.103.092002. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.103.092002>.
- [12] *CERN faq LHC the guide*. Feb. 2009. URL: <http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf> (besucht am 10. 03. 2015).

- [13] CMS Collaboration. *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*. Techn. Ber. CMS-PAS-PFT-09-001. 2009. Geneva: CERN, 2009. URL: <http://cds.cern.ch/record/1194487>.
- [14] CMS Collaboration. *Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector*. Techn. Ber. CMS-PAS-PFT-10-001. 2010. URL: <http://cds.cern.ch/record/1247373>.
- [15] S. Boreham u. a. “Photos from the CMS Photo Book”. CMS Collection. 2008. URL: <http://cds.cern.ch/record/1344500>.
- [16] D. Barney. *CMS-doc-5697-v2*. Feb. 2012. URL: https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/RetrieveFile?docid=5697&version=2&filename=slice_white_v3_Deutsch.pdf (besucht am 10. 03. 2015).
- [17] M. Feindt und U. Kerzel. “The NeuroBayes neural network package”. In: *Nucl. Instrum.Meth.* A559 (2006), S. 190–194. DOI: 10.1016/j.nima.2005.11.166.
- [18] T. Chwalek, W. A. Khan, T. Müller, S. Röcker, J. Wagner-Kuhr. *Measurements of the differential cross section of single top-uquark production in the t channel at 8 TeV*. CMS AN-13-321. CMS internal, 2013.
- [19] M. Cacciari, G. P. Salam und G. Soyez. “The Anti-k(t) jet clustering algorithm”. In: *JHEP* 0804 (2008), S. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].
- [20] C. Weiser. *A Combined Secondary Vertex Based B-Tagging Algorithm in CMS*. Techn. Ber. CMS-NOTE-2006-014. Geneva: CERN, 2006.
- [21] T. Sjostrand, S. Mrenna und P. Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 0605 (2006), S. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].
- [22] S. Alioli u. a. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 1006 (2010), S. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].
- [23] J. Alwall u. a. “MadGraph 5 : Going Beyond”. In: *JHEP* 1106 (2011), S. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].
- [24] S. Agostinelli u. a. “GEANT4: A Simulation toolkit”. In: *Nucl.Instrum.Meth.* A506 (2003), S. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [25] J. Ott. URL: <http://www-ekp.physik.uni-karlsruhe.de/~ott/theta/theta-auto/> (besucht am 26. 03. 2015).
- [26] *Minuit2 Package*. CERN. URL: https://root.cern.ch/root/html/MATH_MINUIT2_Index.html (besucht am 10. 03. 2015).
- [27] F. James und M. Roos. “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”. In: *Comput.Phys.Commun.* 10 (1975), S. 343–367. DOI: 10.1016/0010-4655(75)90039-9.
- [28] N. Kidonakis. *Differential and total cross sections for top pair and single top production*. 2012. DOI: 10.3204/DESY-PROC-2012-02/251. arXiv: 1205.3453 [hep-ph].

-
- [29] V. Khachatryan u. a. “Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s}=8$ TeV”. In: *JHEP* 1406 (2014), S. 090. DOI: 10.1007/JHEP06(2014)090. arXiv: 1403.7366 [hep-ex].
- [30] J. A. Aguilar-Saavedra. “A minimal set of top anomalous couplings”. In: *Nucl. Phys. B* 812 (2009), S. 181. DOI: 10.1016/j.nuclphysb.2008.12.012. arXiv: 0811.3842 [hep-ph].
- [31] G. L. Kane, G. A. Ladinsky und C. P. Yuan. “Using the top quark for testing standard model polarization and CP predictions”. In: *Phys. Rev. D* 45 (1992), S. 124. DOI: 10.1103/PhysRevD.45.124.
- [32] T. G. Rizzo. “Single top quark production as a probe for anomalous moments at hadron colliders”. In: *Phys. Rev. D* 53 (1996), S. 6218. DOI: 10.1103/PhysRevD.53.6218. arXiv: hep-ph/9506351 [hep-ph].
- [33] ATLAS, CDF, CMS, D0 Collaborations. *First combination of Tevatron and LHC measurements of the top-quark mass*. 2014. arXiv: 1403.4427 [hep-ex].
- [34] D. E. Rumelhart, G. E. Hinton und R. J. Williams. “Learning representations by back-propagating errors”. In: *Nature* 323.6088 (Okt. 1986), S. 533–536. DOI: 10.1038/323533a0.

A. Anhang

A.1. Minimierung der Fehlerfunktion des neuronalen Netzes

Die Bestimmung des Minimums erfolgt aufgrund der Eigenschaft (3.5) einer Fehlerfunktion durch Anwenden der Regeln

$$\vec{w}^n \rightarrow \vec{w}^{n+1} = \vec{w}^n - \eta \nabla_{\vec{w}^n} C, \quad (\text{A.1})$$

$$\vec{b}^n \rightarrow \vec{b}^{n+1} = \vec{b}^n - \eta \nabla_{\vec{b}^n} C. \quad (\text{A.2})$$

Die Wahl von η muss dabei genügend klein sein, damit gilt

$$\Delta C \approx \nabla_{\vec{w}} C \cdot \Delta \vec{w} + \nabla_{\vec{b}} C \cdot \Delta \vec{b}, \quad (\text{A.3})$$

wobei hier die Dimension des Vektors \vec{w} der Anzahl aller Gewichte des neuronalen Netzes und die Dimension des Vektors \vec{b} der Anzahl aller Schwellenwerte des neuronalen Netzes entspricht. Aufgrund der meist hohen Anzahl an Gewichten und Schwellenwerten genügt oft eine stochastische Gradientenbestimmung anhand einer zufällig gewählten Untermenge aller Gewichte und Schwellenwerte.

A.1.1. Der Backpropagation-Algorithmus

Ein Algorithmus zur effizienten Bestimmung der in Gleichung (A.1) und (A.2) vorkommenden Gradienten ist der Backpropagation-Algorithmus [34]. Zunächst wird für alle Neuronen (Index j) einer Schicht l der Wert \vec{z}^l und \vec{a}^l berechnet, wobei (vergleiche mit Gleichung (3.1))

$$\vec{z}^l = \begin{pmatrix} z_1^l \\ z_2^l \\ \vdots \\ z_{n^l}^l \end{pmatrix}, \vec{a}^l = \begin{pmatrix} s(z_1^l) \\ s(z_2^l) \\ \vdots \\ s(z_{n^l}^l) \end{pmatrix}, z_j^l = \vec{w}_j^{l-1} \cdot \vec{x}_j^{l-1} + b_j^l. \quad (\text{A.4})$$

.

Durch eine wie in Abbildung 3.2 dargestellte Verknüpfung ergibt sich

$$\vec{x}_j^l = \vec{a}^{l-1}; \quad (1 < l \leq L), \quad (\text{A.5})$$

die Werte können demnach rekursiv aus dem jeweiligen Satz an Eingabewerten \vec{x}^1 berechnet werden.

Für die Gradienten gilt

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (\text{A.6})$$

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l \quad (\text{A.7})$$

wobei folgende rekursive Beziehung für δ_j^l gilt

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} s'(z_j^L) \quad (\text{A.8})$$

$$\delta_j^l = \sum_k \left(w_{kj}^{l+1} \delta_k^{l+1} s'(z_j^l) \right); \quad (0 < l < L) . \quad (\text{A.9})$$

Die rekursive Gestalt der Gleichung (A.9) ist die namensgebende Eigenschaft des Backpropagation-Algorithmus. Aus dem ermittelten Gradienten der Ausgabeschicht L ergeben sich die Gradienten der vorherigen Schichten.

A.2. Verteilungen der Variablen zur weiteren Untergrundreduktion

Die nachfolgenden Abbildungen enthalten die Histogramme der verfügbaren Variablen zur weiteren Auswahl von t-Kanal-Ereignissen wie in Abschnitt 4.4 beschrieben.

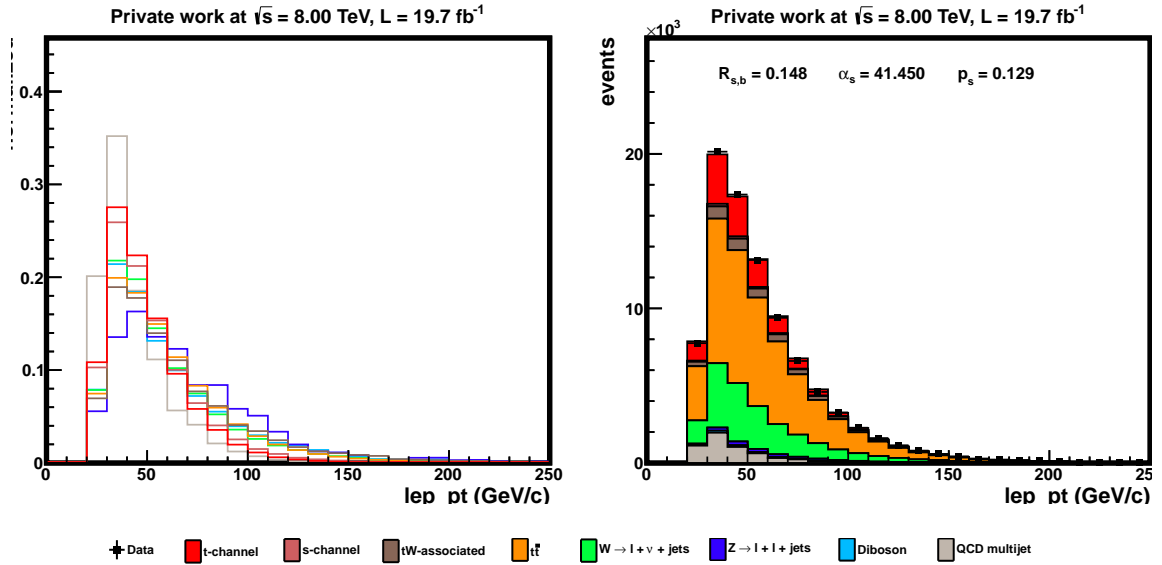


Abbildung A.1.: Verteilung des Leptonen-Transversalimpulses p_T^l . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

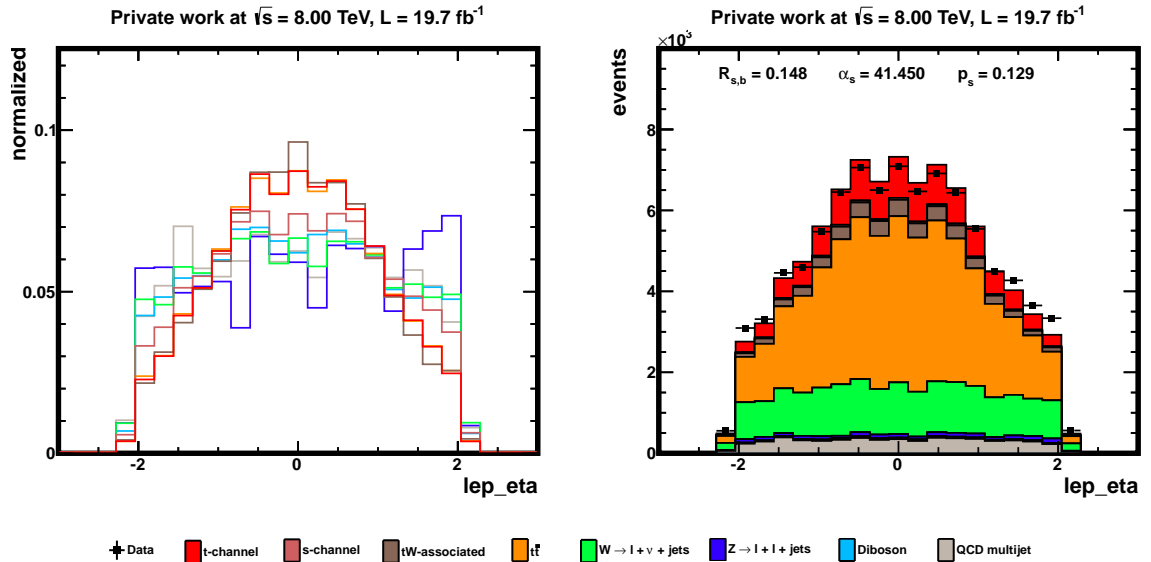


Abbildung A.2.: Verteilung der Leptonen-Pseudorapidität η^l . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

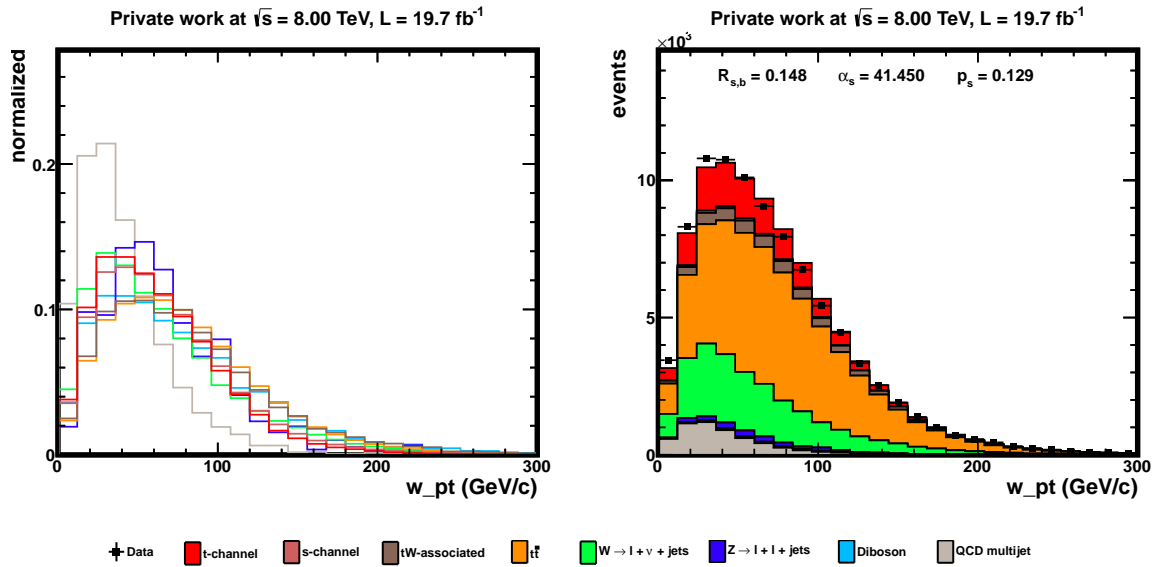


Abbildung A.3.: Verteilung des W-Boson-Transversalimpulses p_T^W . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

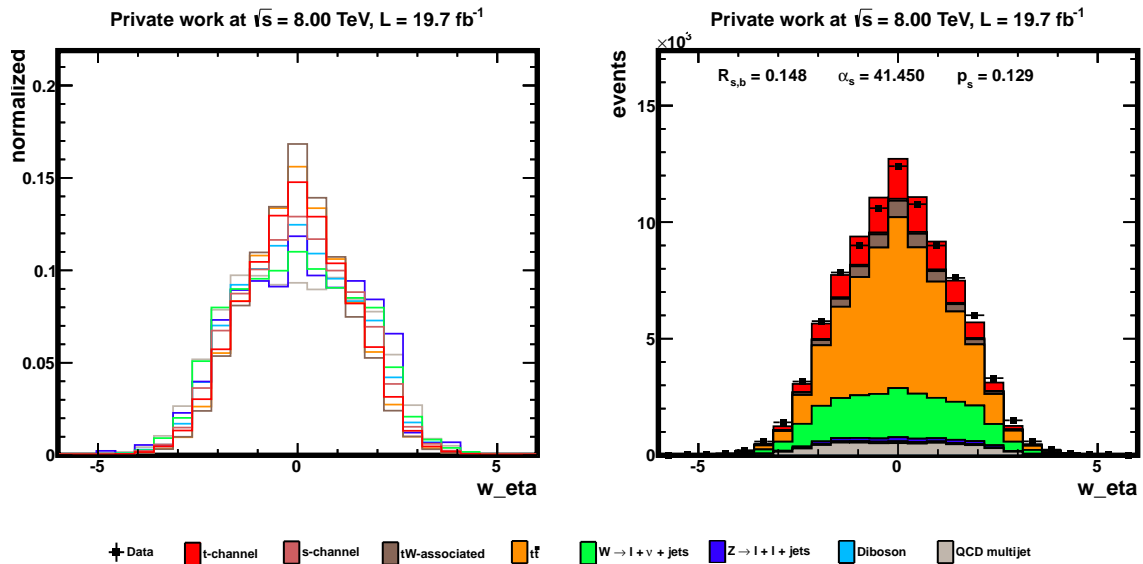


Abbildung A.4.: Verteilung der W-Boson-Pseudorapidität η^W . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

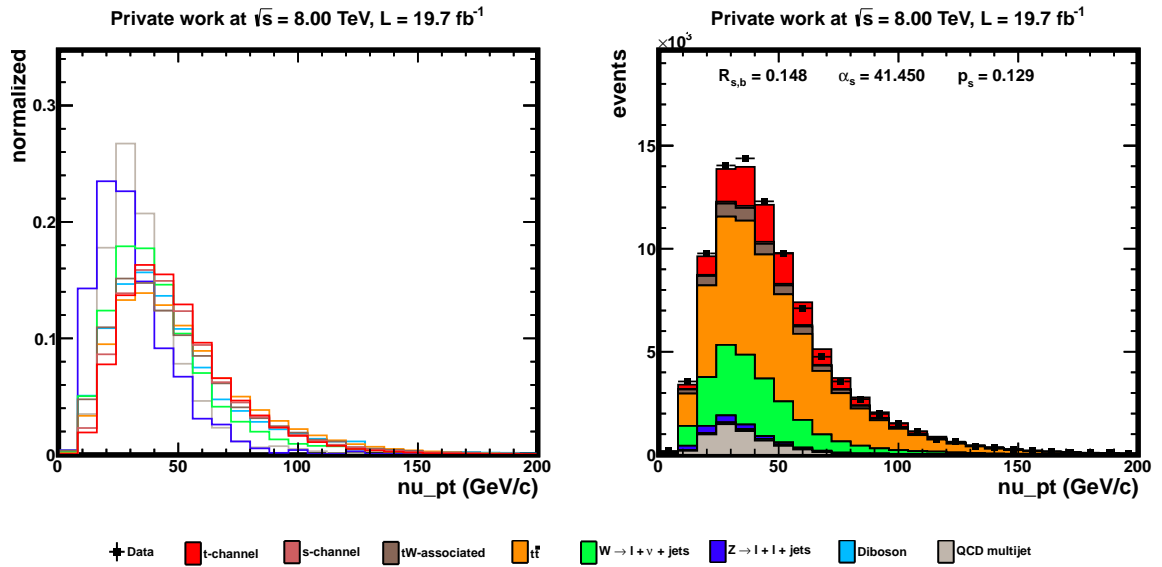


Abbildung A.5.: Verteilung des Neutrino-Transversalimpulses p_T^ν . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

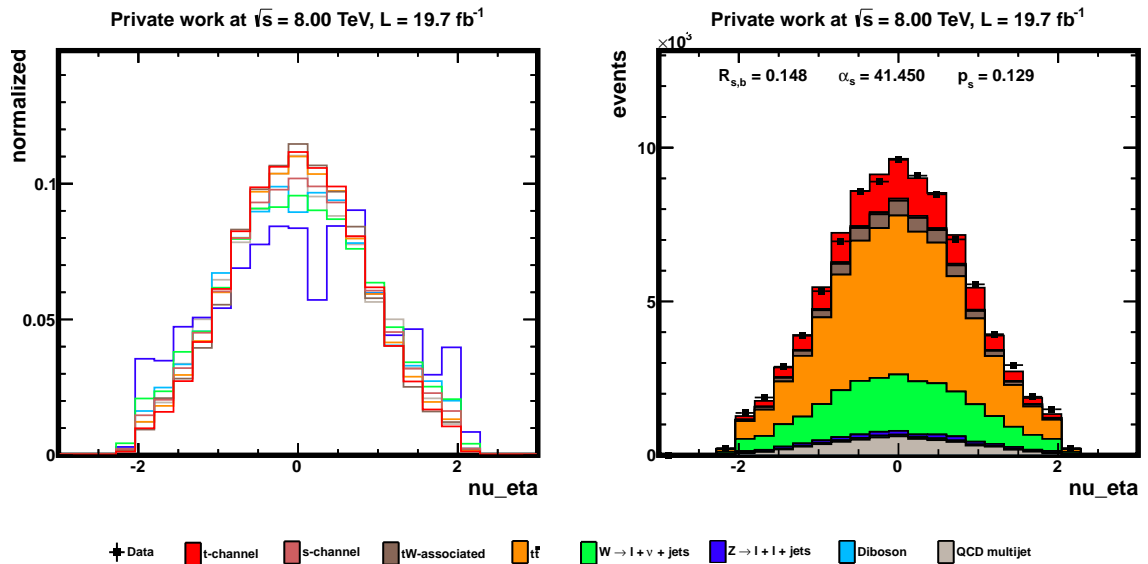


Abbildung A.6.: Verteilung der Neutrino-Pseudorapidität η^ν . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

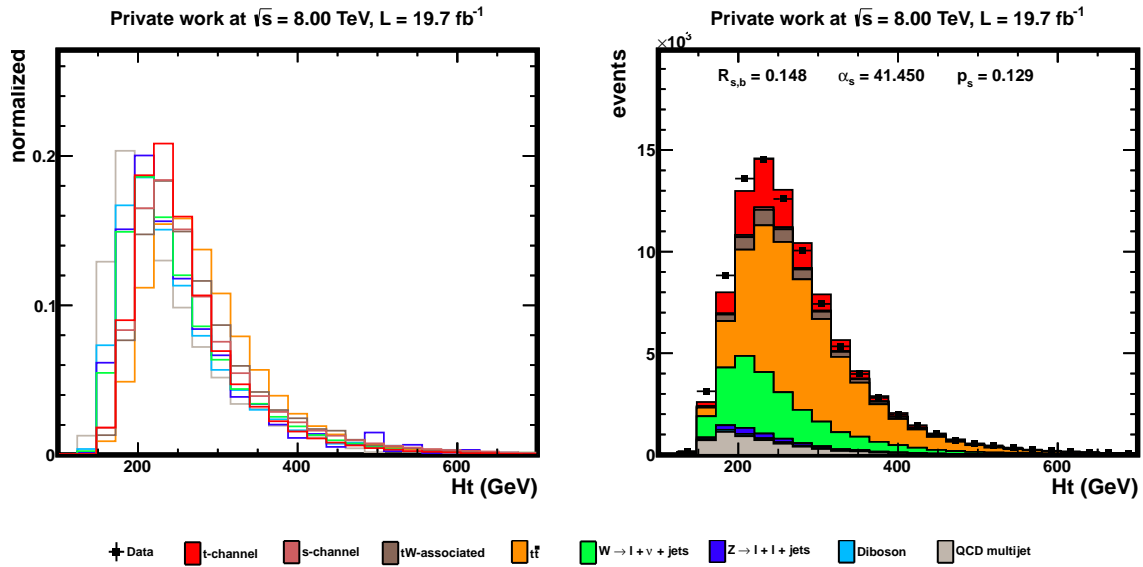


Abbildung A.7.: Verteilung der betragsmäßigen Summe aller Transversalimpulse und der fehlenden Transversalenergie H_T . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

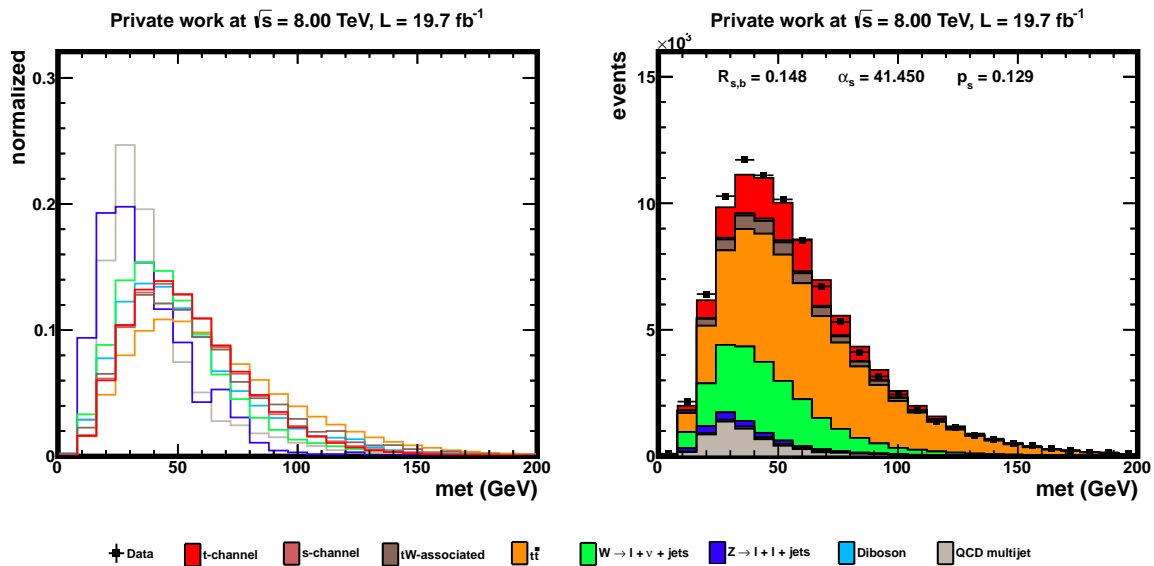


Abbildung A.8.: Verteilung der fehlenden Transversalenergie E_T^{miss} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

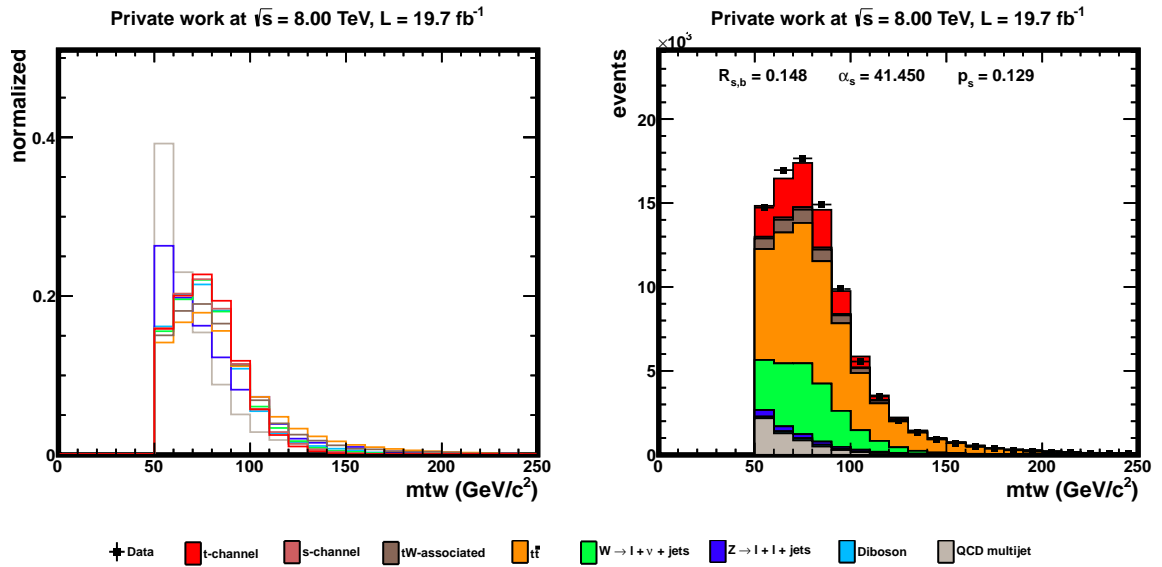


Abbildung A.9.: Verteilung der rekonstruierten W-Boson-Transversalmasse M_T^W . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

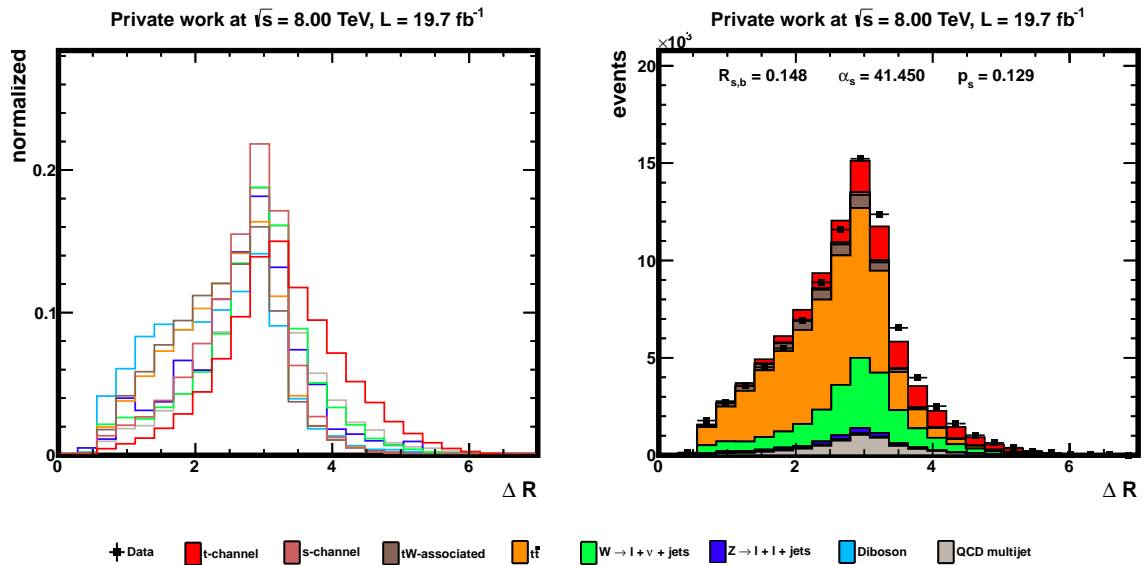


Abbildung A.10.: Verteilung des Abstandsparameters des ersten und zweiten Jet $\Delta R^{\text{jet1,jet2}}$. **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

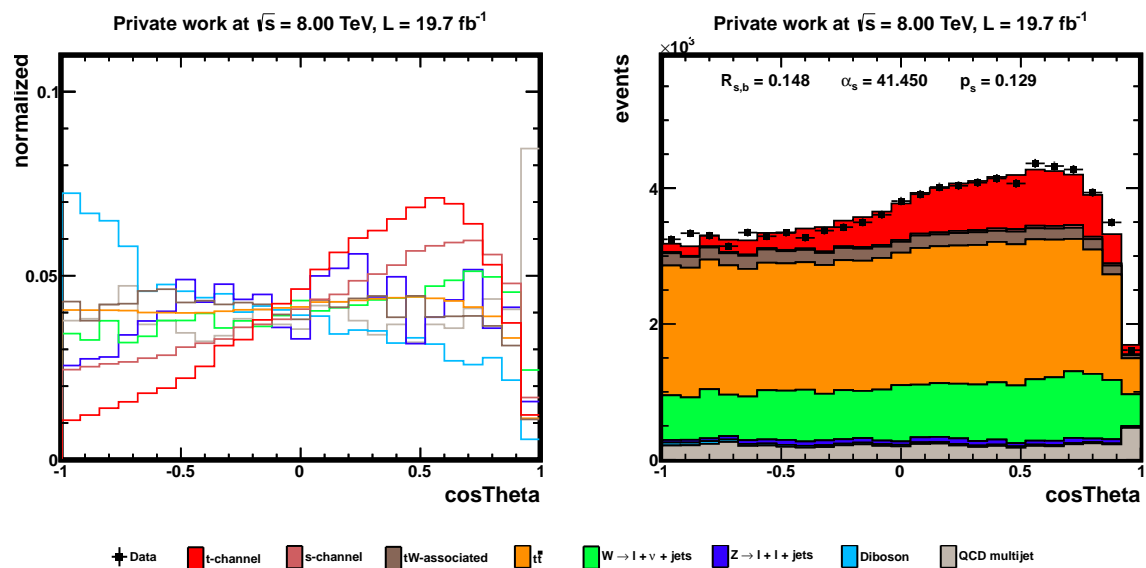


Abbildung A.11.: Verteilung des Kosinus des Winkels zwischen dem Lepton und dem leichten Quark $\cos \Theta^{l,q}$ im Laborsystem. **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

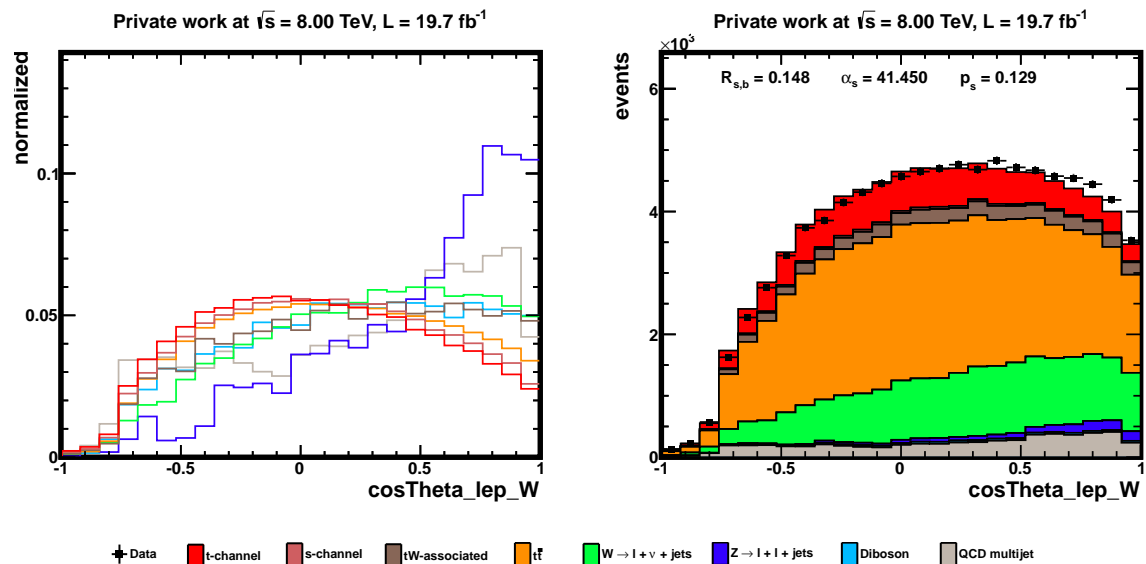


Abbildung A.12.: Verteilung des Kosinus des Winkels zwischen dem Lepton im Ruhesystem des W-Bosons und dem W-Boson im Laborsystem $\cos \Theta^{l*,W}$. **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

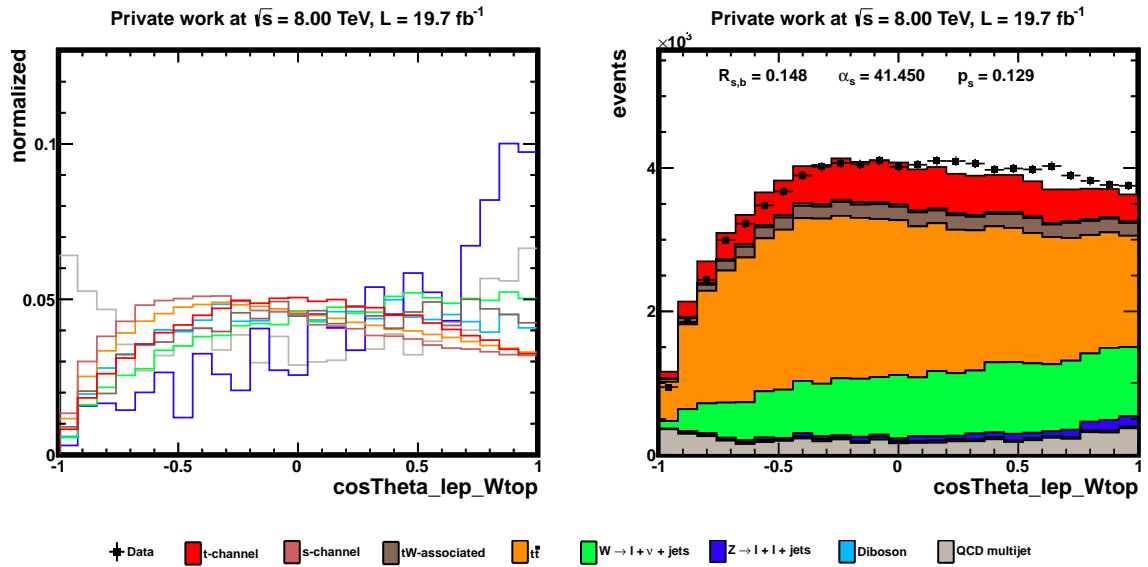


Abbildung A.13.: Verteilung des Kosinus des Winkels zwischen dem Lepton im Ruhesystem des W-Bosons und dem W-Boson im Ruhesystem des Top-Quarks $\cos \Theta^{l*, W**}$. **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

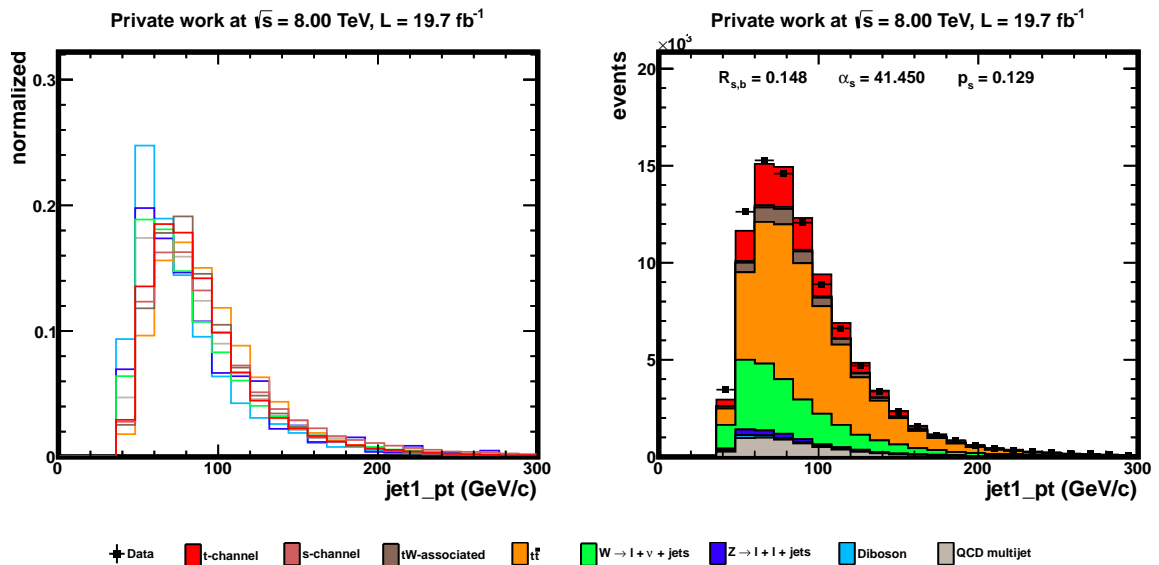


Abbildung A.14.: Verteilung des Jet1-Transversalimpulses p_T^{jet1} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

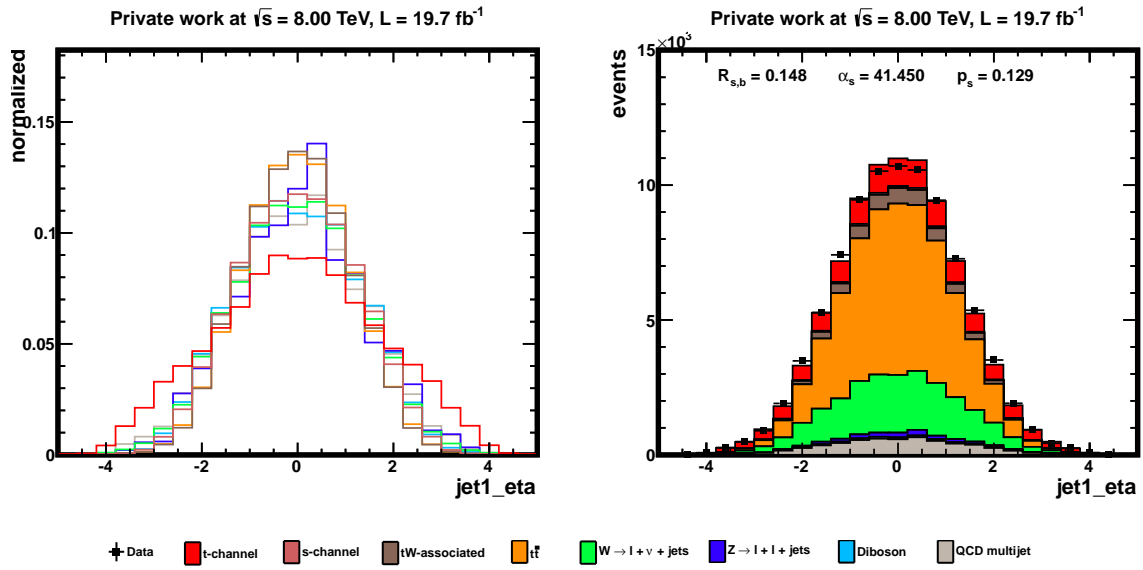


Abbildung A.15.: Verteilung der Jet1-Pseudorapidität η^{jet1} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

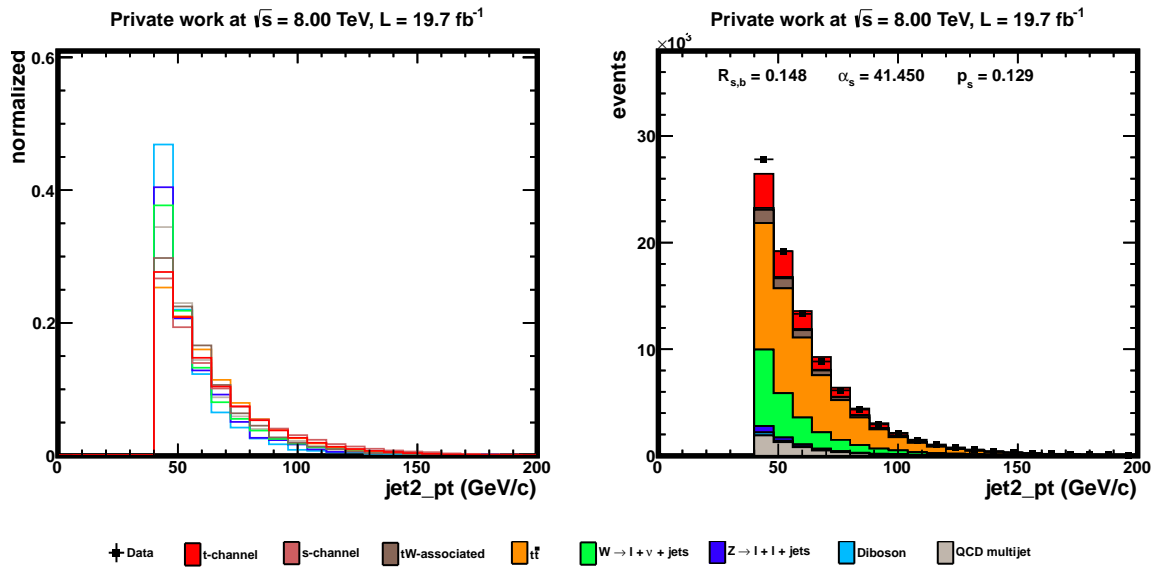


Abbildung A.16.: Verteilung des Jet2-Transversalimpulses p_T^{jet2} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

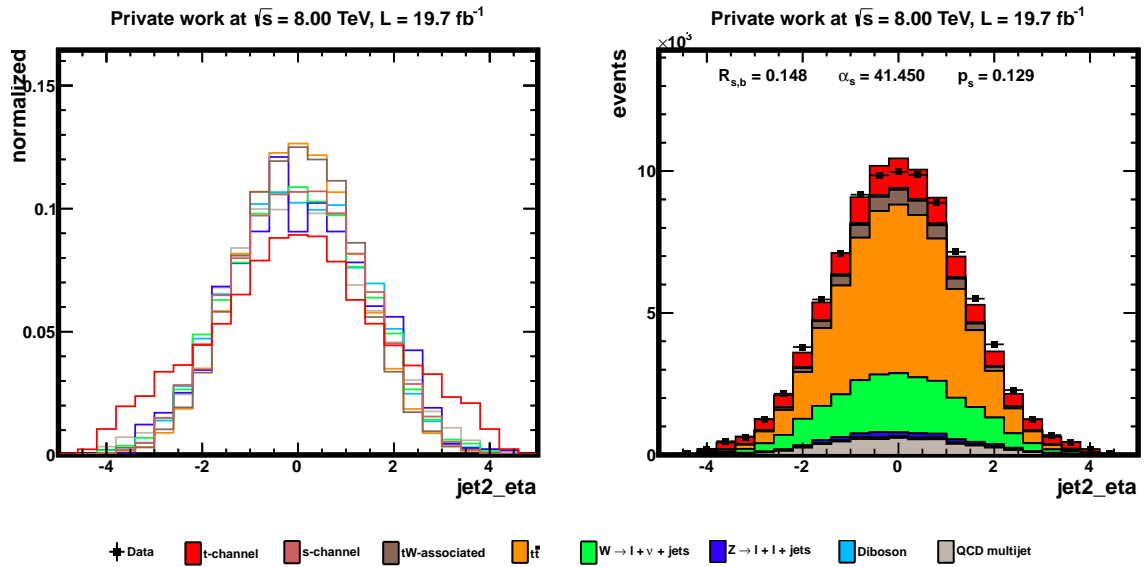


Abbildung A.17.: Verteilung der Jet2-Pseudorapidität η^{jet2} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

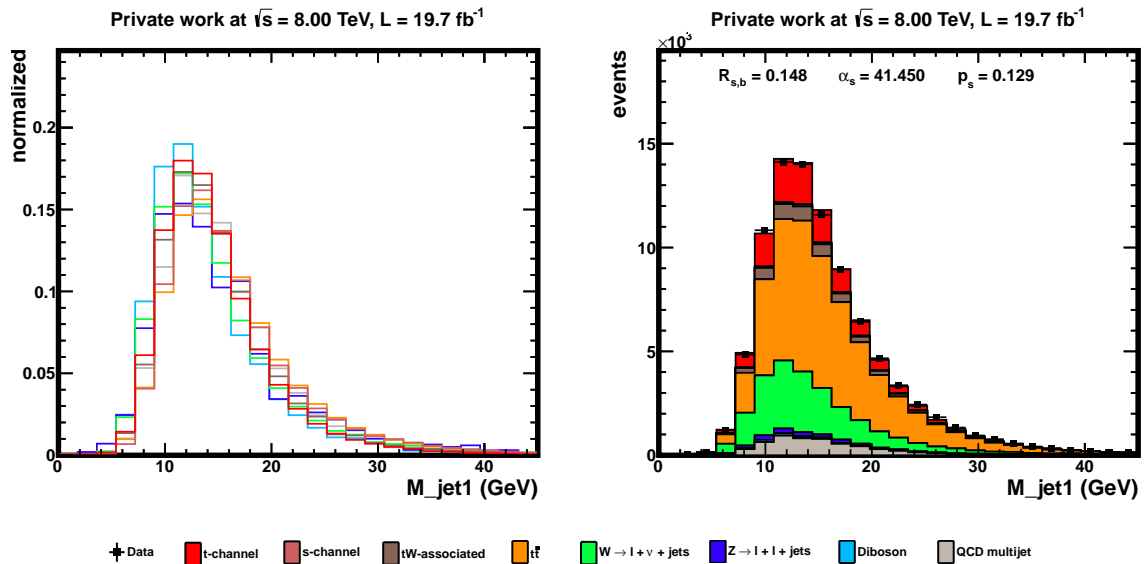


Abbildung A.18.: Verteilung der Jet1-Masse M^{jet1} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

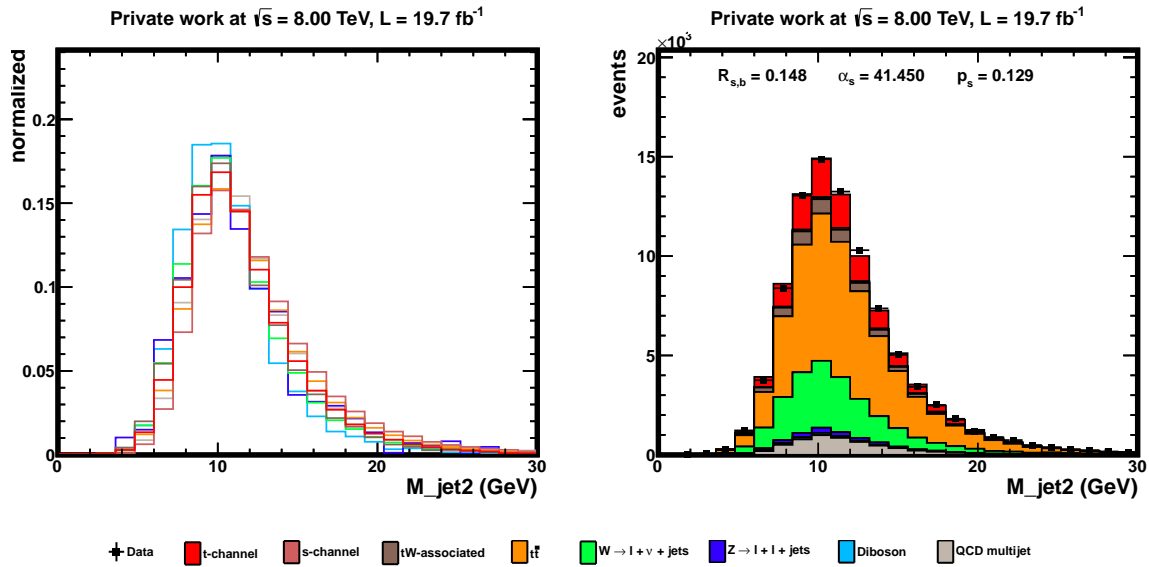


Abbildung A.19.: Verteilung der Jet2-Masse M_{jet2}^{jet2} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

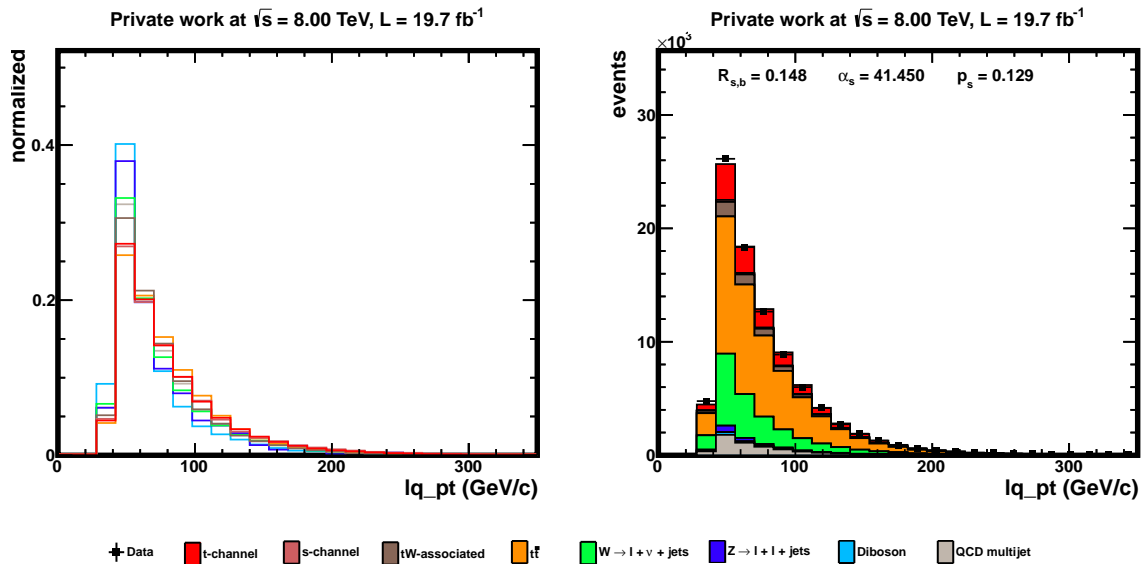


Abbildung A.20.: Verteilung des Jet2-Transversalimpulses p_T^q . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

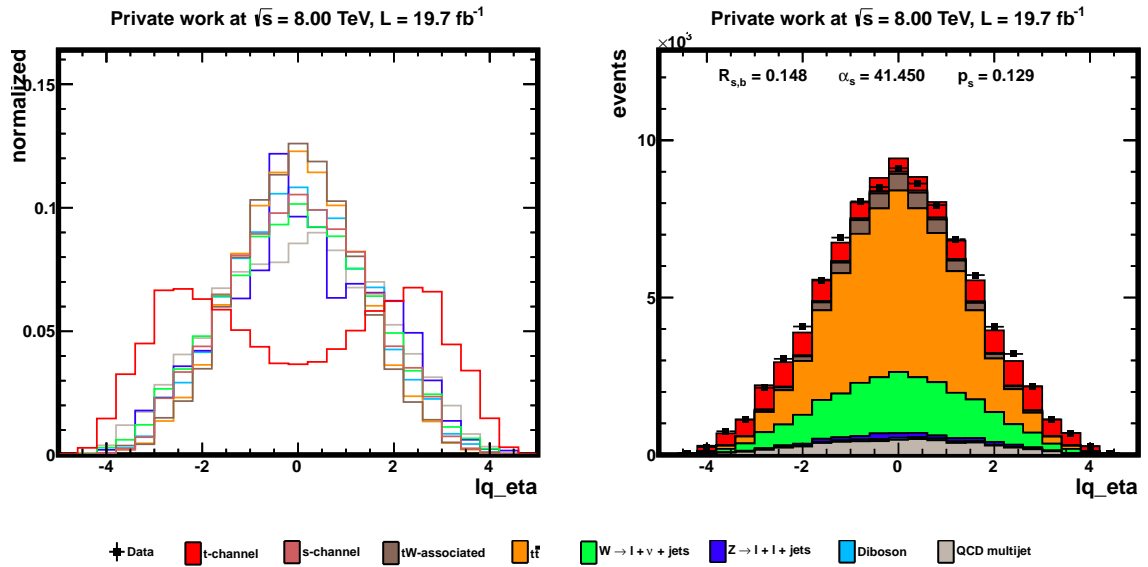


Abbildung A.21.: Verteilung der Jet2-Pseudorapidity η^q . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

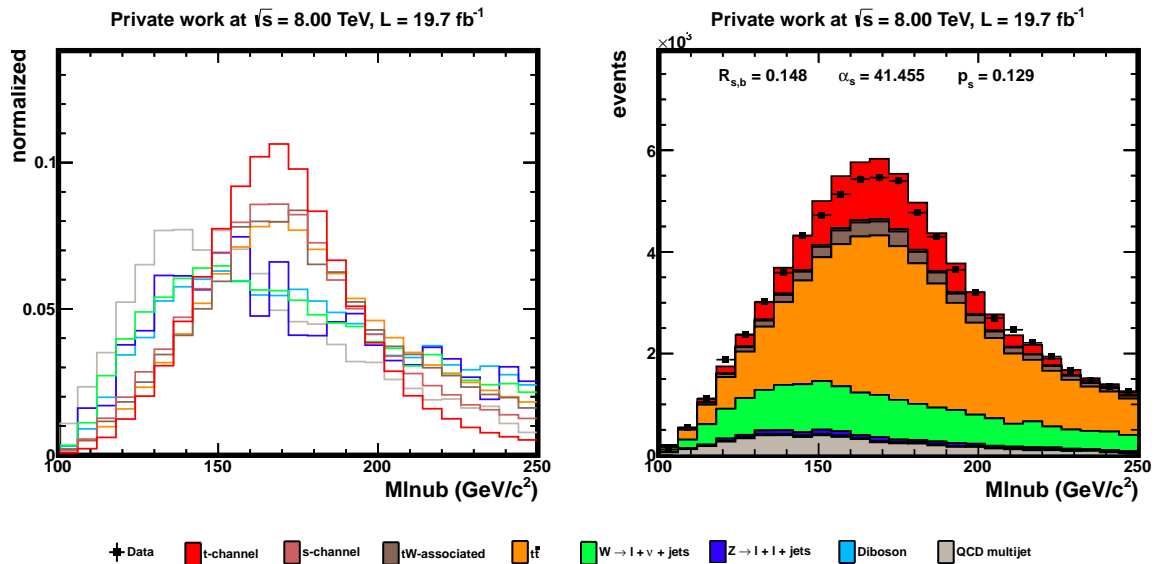


Abbildung A.22.: Verteilung der rekonstruierten Top-Quark-Masse M^{lvb} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

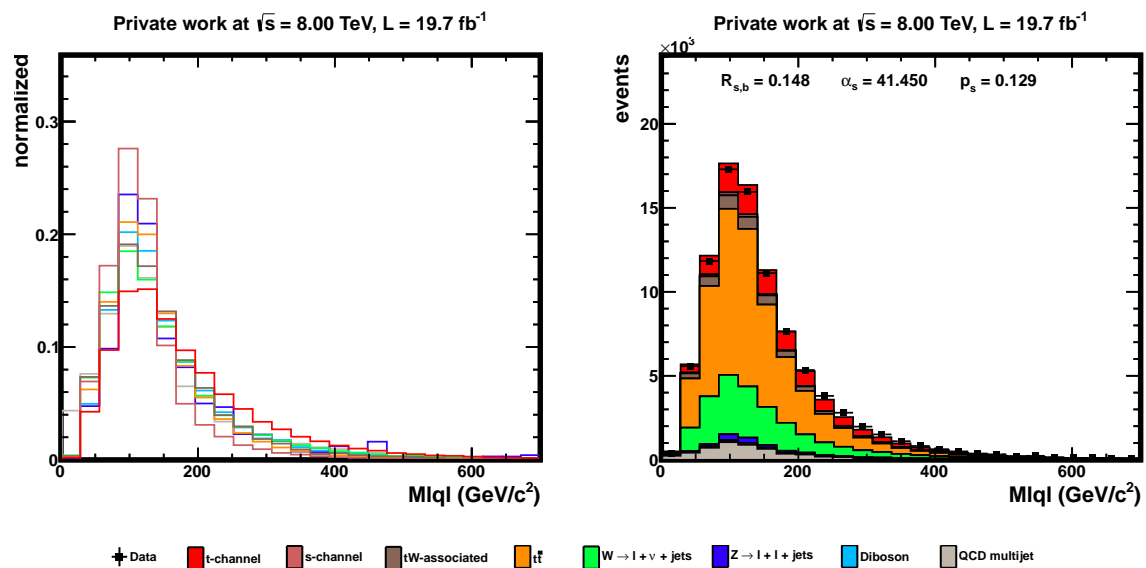


Abbildung A.23.: Verteilung der invarianten Masse des leichten Quark und des Leptons M^{lq} . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

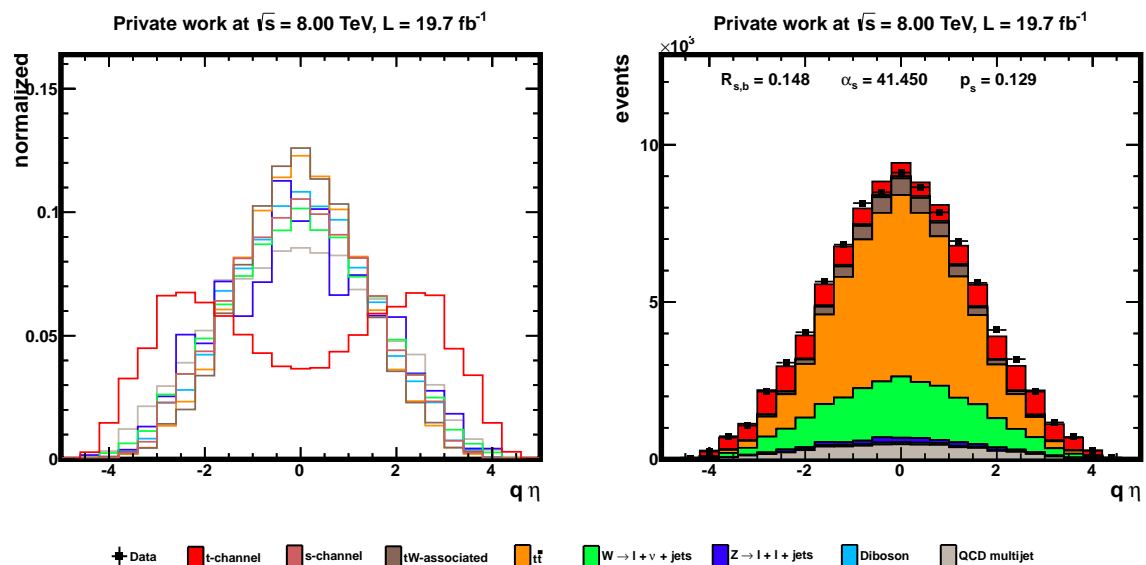


Abbildung A.24.: Verteilung des Produktes $Q^{l\eta}$. **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

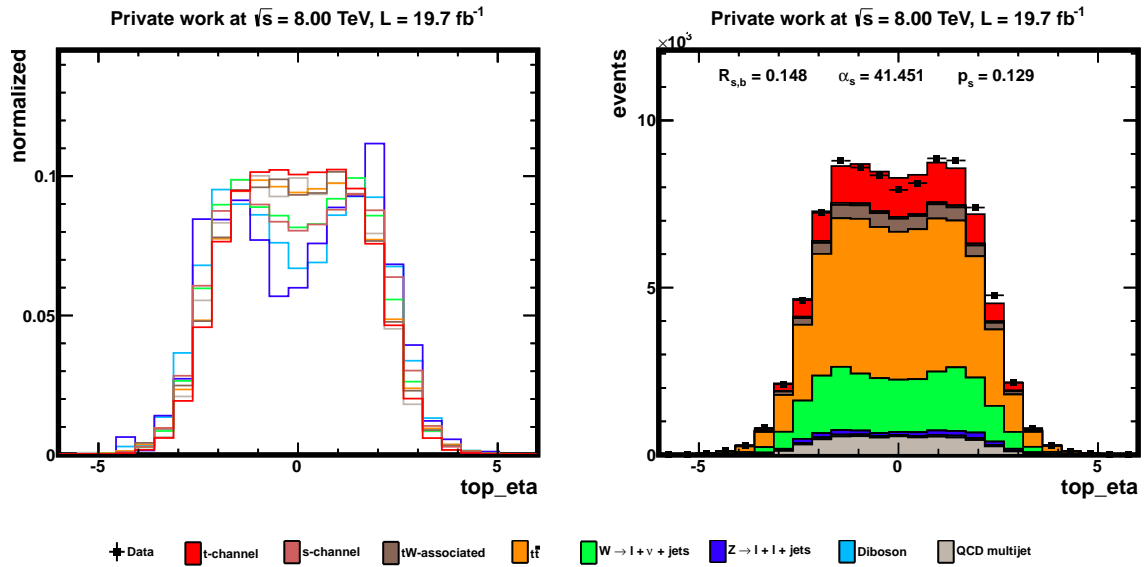


Abbildung A.25.: Verteilung der t-Quark-Pseudorapidität η^t . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

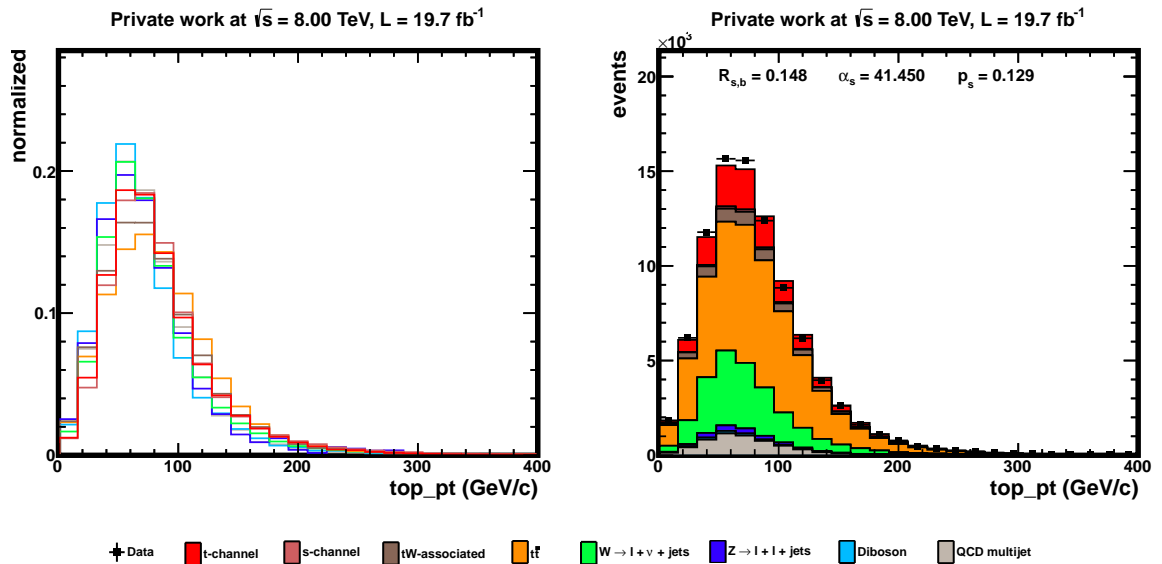


Abbildung A.26.: Verteilung des t-Quark-Transversalimpulses p_T^t . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

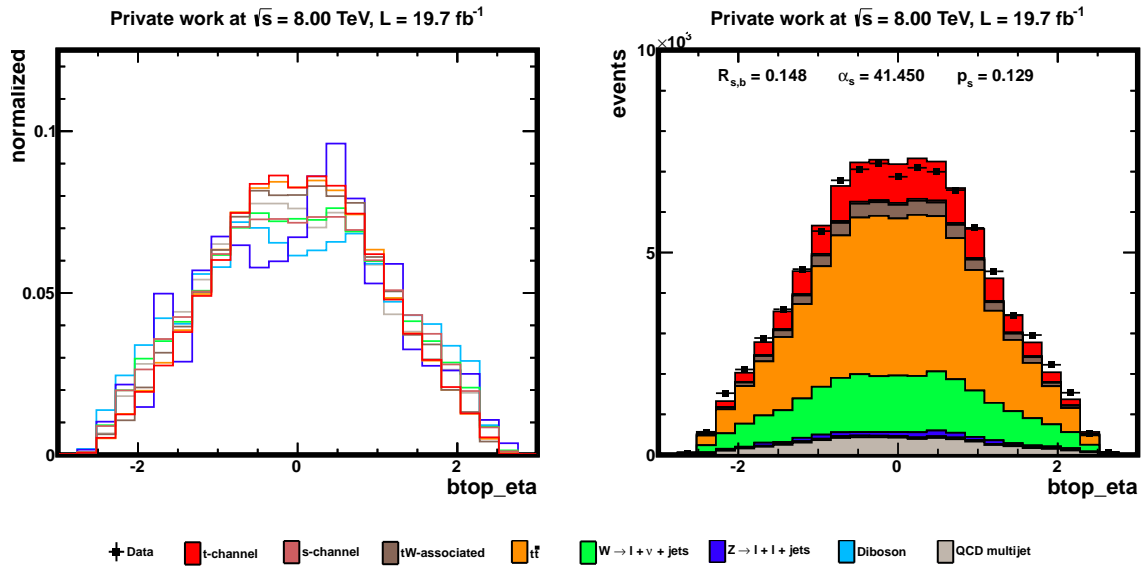


Abbildung A.27.: Verteilung der b -Quark-Pseudorapidity η^b . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

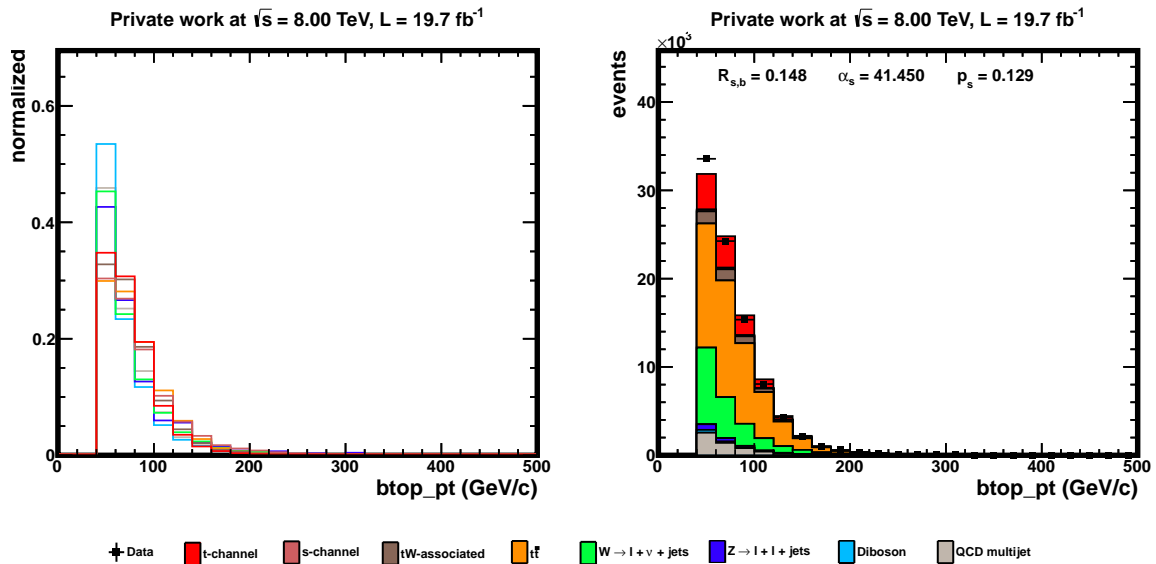


Abbildung A.28.: Verteilung des b -Quark-Transversalimpulses p_T^b . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignisse des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

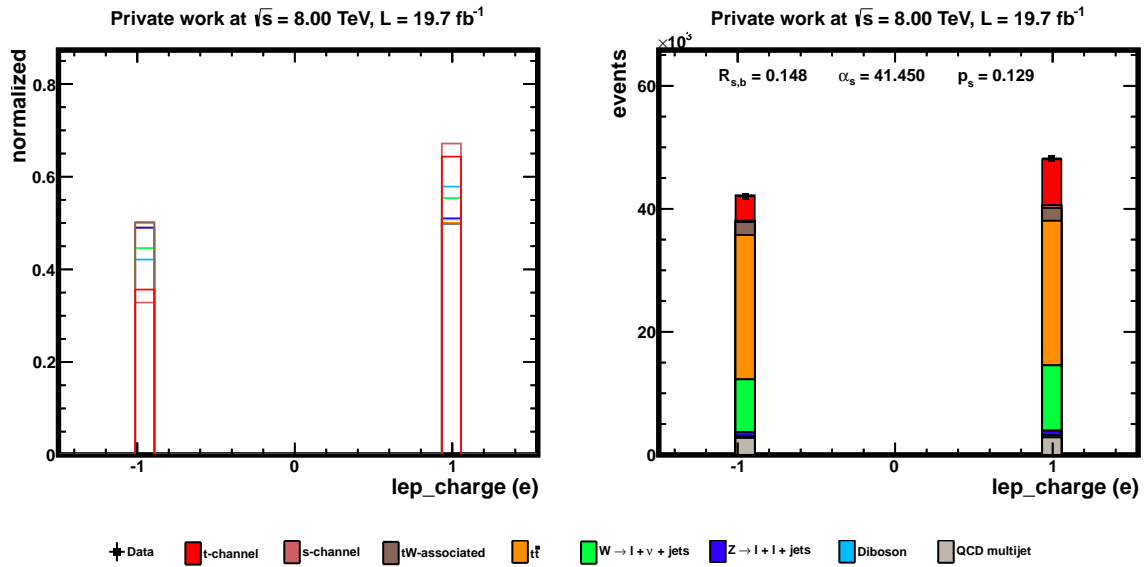


Abbildung A.29.: Verteilung der Leptonen-Ladung Q^l . **Links:** Verteilung der simulierten Prozesse normiert auf die Gesamtanzahl an Ereignissen des entsprechenden Prozesses. **Rechts:** Verteilung normiert auf die integrierte Luminosität mit angewandten Skalierungsfaktoren.

A.3. Übungsblätter für den Praktikumsversuch

Die entworfenen Übungsblätter geben den Studierenden eine knappe Einleitung in die relevante Top-Quark-Physik und dienen des Weiteren als Anleitung für die Handhabung des Programmpakets zur Reduktion der Untergrundereignisse. Für die Programmieraufgaben ist eine kurze Quelltext-Sammlung beigelegt, die alle relevanten Funktionsaufrufe beinhaltet.

Computerpraktikum zur Vorlesung Teilchenphysik

Using Neural Networks for Single Top Analysis at CMS

Introduction

By doing this exercise sheet you are about to learn how to use a neuronal network in data analysis, the benefits and the problems that can arise and how to avoid them by studying the single top quark production at the LHC using CMS detector data.

There are two main parts of exercises, you will:

- learn how to use a neural network for data analysis to reduce background processes and compare this multivariate analysis with an analysis based on cuts on various variables.
- use the results for top quark analysis (cross section, CKM matrix element V_{tb} , charge asymmetry and top mass reconstruction).

Brief introduction to single top physics

At the *proton-proton* collider LHC there are several production channels that create top quarks, some of them create a *single top* quark, the most important is the t-channel process (figure 1 left), where the quark q is one of the proton's quarks (most probable the up valence quark). The gluon is one of the other proton's gluons which splits into b \bar{b} . In a weak interaction scattering process between the b quark and the q quark, a *light quark* q' and a top quark t are produced.

This is the signal process and all the other production channels are considered as backgrounds in *the scope of this exercise sheet*.

Top quark decay

Due to its very short lifetime, the top quark can be only detected and analysed by its decay products. Almost all top quarks decay into a b quark and a W boson. The latter

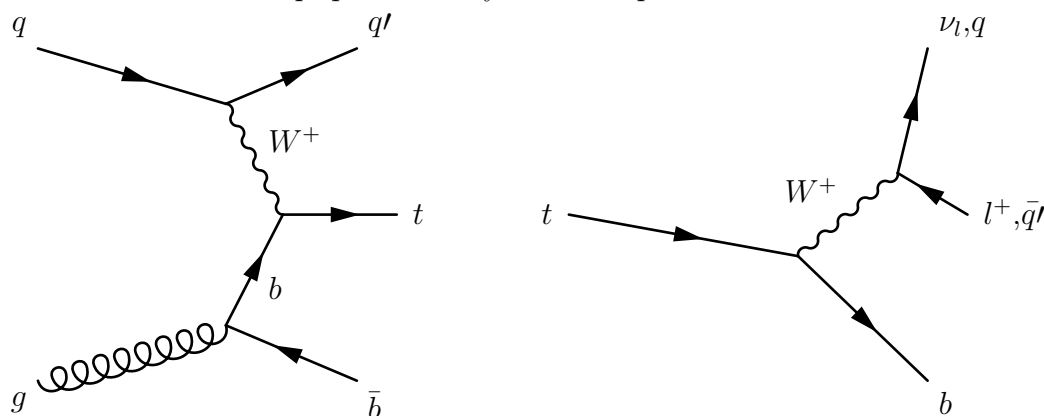


Figure 1: **Left:** t-channel single top production. **Right:** Top quark decay.

decays either hadronically into a pair of quarks or leptonically into a charged lepton and a neutrino (figure 1 right).

Event selection

In the data used for the following exercises, only events are selected which contain exactly one isolated muon (from the leptonic top quark decay). Furthermore only events are selected that contain two jets from which one is b tagged¹. One of these jets originates from the light quark of the t-channel production process and the other from the b quark (b tagged) of the top decay.

Simulated events and QCD data side band

In addition to single top production other processes can mimic the final state topology of one muon and two jets. To distinguish between signal and background processes we make use of samples of simulated events for each process. The only background contribution that is modelled directly from data is the QCD multijet production.

Questions:

How does the Feynman diagram of an anti top event in the t-channel process look like and how can a top event be distinguished experimentally from an anti top event?

Are there as many single top events as single anti top events expected in proton-proton collisions at the LHC?

Brief description of neural networks

Artificial neural networks behave similar to biological neural networks. They are built from many neurons, where each neuron has several input connections and one² output connection. Dependent on the strength of all inputs together, the strength of the output is calculated by comparing the input value with a threshold parameter of the neuron. Additionally the connections are weighted, that means that one connection has more influence than another. Dependent on the weights and threshold parameters for a set of input values an output value of the overall network is calculated.

The determination of the weights and threshold parameters is called training or learning of the neural network. To train the neural network it is required that for each set of input values the correct output values must be known. The network output is then compared with the correct output by a function called cost function which is chosen to be always positive. Since the neural network output depends on all the threshold and weight parameters the same holds for the cost function, thus this function is a high dimensional one.

This function reaches its global minimum when the network output matches the correct output, therefore by minimizing this function, the weights and threshold parameters that lead to an output close to the desired output for a set of input values are determined.

¹There is also a third jet that originates from the \bar{b} jet of the gluon splitting in the t-channel process, which often escapes detection by the CMS detector for this goes mainly into the beam direction and its transverse momentum is below the threshold value.

²The output of one neuron can be connected with several neurons though.

A special class of neural networks is the feed-forward neural network, this means that the neurons are grouped into different layers and the neurons of one layer are only connected with the neurons of the next layer (figure 2).

Such a neural network is trained using the modelled events to distinguish between signal and background events, the output can be used to increase the signal-to-background ratio and is also applied to the measured data.

The neural network used for these exercises is NEUROBAYES³

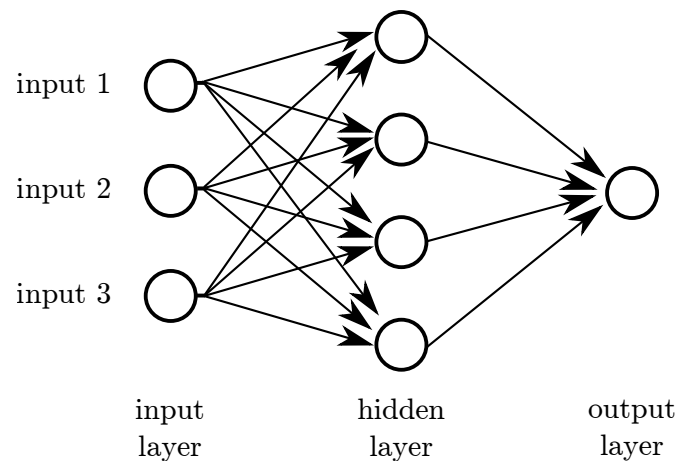


Figure 2: Scheme of a three layered feed-forward neural network with three input nodes

³<http://neurobayes.phi-t.de>

Overview of the Program

A program package for the single top analysis can be found inside the directory `program/`. There are several subdirectories where actions have to be done in the exercises:

- `current-save/` - contains output files and internal files of the neural network.
- `current-save/trainings/` - contains previously created training files and expertises.
- `data-files/` - contains general configuration files that specify parameters of the program tools such as input variables and cut values.
- `histos/` - all histograms generated by the program tool are stored there as `.root` files.
- `pics/` - all the plots that are generated by the program tool are stored there as `.png` files.

It might be useful to understand the full program workflow during the step by step instructions in the exercises, this is shown in figure 3.

There are basically two main programs that are used in this exercises: `./plotmain` is used to create the histograms and plots with user defined settings (e. g. variables to plot, applied cuts) and `./nbmain` to start the neural network expert tool.

When executed both programs ask for user input that specify what action has to be done:

- `./plotmain` will ask you to select whether during the creation process cuts shall be applied.

It will also ask you to enter a name suffix that will be added to the histograms and pictures: This is useful if you want to compare different results. If you skip entering a suffix or you are using the same suffix again the **existing histograms and plots will be overwritten!**

- The program `./nbmain` will ask you whether you want to start the neural network trainer (not available for this exercises) or the expert. The trainer creates a file with the training results for the expert and also an analysis file with information of the training performance. The expert uses the training file to create a single variable `NNout` which can be used for discrimination. **The created files will be overwritten with each training and expert cycle!**

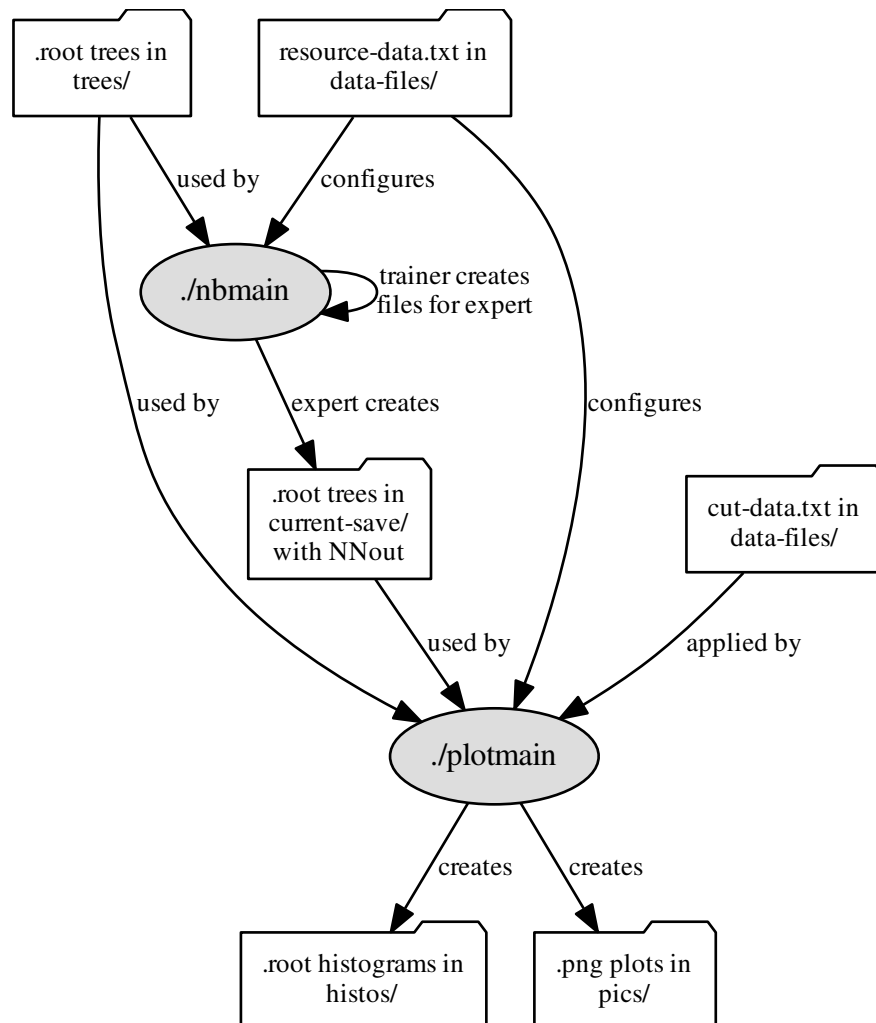


Figure 3: Overview over the main program tools

Exercises

Exercise 1: Use multiple cuts to reduce background processes

First use the method of multiple single cuts to reduce the background processes. The results will be compared with the results from a single cut on the neural network output for signal and background separation (exercise 2). It is desired to achieve a better signal to background ratio $R_{s,b} = \frac{s}{b}$, the significance $\alpha_s = \frac{s}{\sqrt{s+b}}$ and the signal purity $p_s = \frac{s}{s+b}$ should increase, too. Look for up to three variables that are suitable for background and signal separation based on the modelled signal and underground processes, optimize the values of $R_{s,b}$, α_s and p_s by calculating them for different ranges of each variable. Apply a combination of the determined cut ranges to get the values for the combined cuts.

Your Tasks:

- Change to the program's root directory and execute `./plotmain`, the program will ask you to choose an option (select option (1)) and to enter a name for the created files, you can skip a name addition (by entering `xx`) here or choose one if desired.
- Study the created plots in `pics/` and look for up to three variables that are suitable for reducing the background by applying a single cut. You can find a short description of the variables in the appendix. 0.4.
- Have a look at the created `.root` files in `histos/` of the selected variables by using the `TBrowser` of ROOT. Scan the selected variables by using these ROOT histograms for ranges with the highest α_s value and select the range with the highest value of $R_{s,b}$ and p_s for this α_s value. It might be helpful to write a ROOT macro for this, you can use the code reference in appendix 0.3 for useful commands.
- Use `data-files/cut-data.txt` to define the determined single cut for the first of the chosen variables first and then create plots by using `./plotmain` ((2), an unique name is recommended) to check that the cut is correctly applied, repeat the task for each chosen variable with the determined cut range.
- Define combinations of the chosen single cuts using `data-files/cut-data.txt` to reduce the background. Create plots with these cuts applied by using `./plotmain`, ((2), choose an unique name is recommended). The values of $R_{s,b}$, p_s and α_s for the cut combination can be found on the created plots.

Exercise 2: Use the neural network to reduce background processes

Use the neural network to get the variable `NNout` for discrimination. Because the training of the neural network is a time consuming task and runs automatically, training files for the expert have already been created. Optimise the values of $R_{s,b}$, α_s and p_s by selecting the optimal cut range as in exercise 1. Compare the values with those of the combined single cuts that have been applied in exercise 1.

Your Tasks:

- Select the first training (`t0-full`) by copying *the content* of the directory `current-save/trainings/t0-full/` to `current-save/`.
The training contains the file:
 - `analysis.pdf` with informations on how the training has performed and the correlation between the input variables.
 - `input-vars.txt` that lists the input variables and a number by which it is represented in the analysis file.
- Execute `./nbmain` and select option (2) to start the expertise, the expertise can take up to a few minutes. The program will create `.root` files (with a `NNout_` prefix) in `current-save/`.
Optionally have a look at the newly created `.root` files, they should contain the single variable `NNout`.
- Use `./plotmain` ((1), an unique name is recommended), this produces plots of all activated variables again with the selected name suffix and this time also of `NNout` in `pics/`, furthermore histogram outputs in `histos/` will be created. The created `NNout` histogram will be used for the following steps.
- Choose the ranges of `NNout` with the highest α_s value and select the range with the highest value of $R_{s,b}$ and p_s for this α_s value.
- Use `data-files/cut-data.txt` to define the determined single `NNout` cut and plot the results using `./plotmain` ((2), choose an unique name) to check that the cut is correctly applied and
- Compare the values $R_{s,b}$, p_s and α_s with those of the combined single cuts.
- *Optionally:* Use the other training that is available which only uses a subset of all variables that serve a low additional significance, determine the optimal cut range and compare the results with the full training and the results from exercise 1.

Exercise 3: Single top cross section and CKM matrix element

To measure the cross section and for further analysis the amount of modelled processes has to be adjusted to fit to the number of measured data events. This is done by fitting the MC generated samples using a binned maximum likelihood method which are normalised to the cross section and already scaled by the luminosity to the amount of data events of the neural network output in the whole region from -1 to 1.

The parameters β_P will be determined by maximizing the likelihood function. The scale factor of the t-channel β_{tchan} can be used to calculate the cross section of the single top production in this channel:

$$\sigma_{t-channel}^{\text{measured}} = \beta_{t-channel} \cdot \sigma_{t-channel}^{\text{theoretical}}, \quad (1)$$

where the theoretical cross section is

$$\sigma_{tchan}^{\text{theoretical}} = 87.2^{+4.8}_{-3.2} \text{ pb}. \quad (2)$$

Your Tasks:

- Implement the likelihood fit by completing the code of the `LikelihoodFit.C` ROOT macro, the fit factors will be determined by minimising the *negative logarithm* of constrained likelihood function. The configuration of the minimisation tools is already implemented in this macro using the MINUIT2-package⁴ of ROOT:
- Calculate the negative logarithm of the constrained likelihood function $L(\beta_P)$ and neglect terms of the sum that are independent of the parameter β_P .

The likelihood function is

$$\tilde{L}(\beta_P) = \prod_{i=1}^{\mathcal{N}_{\text{bins}}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!}, \quad (3)$$

where i is the bin index, n_i is the number of measured events in bin i .

$$\mu_i = \sum_P \beta_P \cdot N_i^P, \quad (4)$$

where \mathcal{N}_P is the number of processes P , N_i^P is the predicted number of events in bin i for the process P and β_P is the corresponding scale factor, where a value of $\beta_P = 1$ means that the number of events of a process is the same as the predicted theoretical value.

By multiplying the Likelihood function with gauss distributed factors (C_P), the values of the parameter β_P for process P can be constrained. This factor is defined as

$$C_P = \frac{1}{\sqrt{2\pi} \cdot \Delta_P} \cdot e^{-\frac{(\beta_P - 1)^2}{2\Delta_P^2}}, \quad (5)$$

where Δ_P is the deviation of the parameter β_P from the value 1. The resulting constrained likelihood function is

$$L(\beta_P) = \prod_{i=1}^{\mathcal{N}_{\text{kl}}} \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!} \cdot \prod_P^{\mathcal{N}_P^E} C_P. \quad (6)$$

\mathcal{N}_P^E is the number of processes for which the parameter β_P is constrained (since not all parameters require a constrained factor).

- Implement this function as instructed by the comments in `LikelihoodFit.C`
- Execute the ROOT macro by using
`root -q -b 'LikelihoodFit.C("NNout","THE_NAME")'` where `THE_NAME` is the name you might have entered as you used `./plotmain` for the histogram, choose

⁴<https://root.cern.ch/root/html/TMinuit.html>

the `NNout` histogram on which no cut has been applied! The fit scales will be written to `data-files/fit-results-tchan.txt`. When this file is present the scales will be applied automatically for plotting. Use `./plotmain` ((1), choose an unique name) and have a look at the created plots to check the applied scale. The modelled processes should fit to the measured data now.

- Calculate the cross section.
- Calculate $f|V_{tb}|$:

The cross section is proportional to $|V_{tb}|^2$, where the CKM-matrix element $V_{tb} \approx 1$. Considering physics beyond the standard model it is predicted that anomalous coupling occurs, this is parameterized by a factor $f \Rightarrow f \cdot |V_{tb}|$. By considering that $|V_{td}|, |V_{ts}| \ll |V_{tb}|$, the branching fraction of the top decay into Wb can be considered as 1, therefore $|fV_{tb}| = \sqrt{\beta_{t\text{-channel}}}$.

Exercise 4: Top quark charge asymmetry

In the proton-proton collision at the LHC, more top quark events than anti top quark events are expected, those can be distinguished by the lepton charge Q^l of the top quark decay (compare with figure 1 right), thus there are not as many events with a positive charged lepton as with a negative charged expected.

In this exercise this charge asymmetry will be measured. It can be determined by calculating the ratio of the events with positively charged muons and events with negatively charged muon after background subtraction.

$$R = \frac{N_{\text{pos}}}{N_{\text{neg}}}$$

For this exercise a training independent of the variable Q^l is used.

Your Tasks:

- Copy the `.root` files from `current-save/trainings/expert/full-lep_charge/` to `current-save/` these were created using a training without the variable Q^l .
- Create histograms in `histos/` with the `NNout` cut determined in exercise 2 applied by defining these in `data-files/cut-data.txt` (should be still present) and by executing `./plotmain` ((2), choose an unique name) and check the results.
- Copy the newly created Q^l histograms from `histos/` to `analysis-tools/lep_charge/`.
- Change to `analysis-tools/lep_charge/` and complete the code of the ROOT macro `ChargeAsymmetry.C` to calculate the charge ratio where the comments tell you. Take notice of the reference in appendix 0.3.
- Execute this macro using `root -q -b 'ChargeAsym.C("lep_charge",THE_NAME)'` where `THE_NAME` is the name you might have entered as you used `./plotmain` for the histograms.

- *Optional:* Use a neural network training that includes Q^1 (`current-save/trainings/expert/full/`) and compare both results.

Exercise 5: Top mass (*optional*)

In this exercise the top quark mass is derived from data. The top mass is retrieved by analysing the peak of $M^{\nu b}$ which will be called the *reconstructed* top mass M_t^R . This variable is reconstructed by adding the four vectors of the reconstructed muon, neutrino and b-quark. A training independent from $M^{\nu b}$ is used.

Your Tasks:

- Execute `./massvar.sh` and select option 0 to copy the .root files (preserving symlinks) from `current-save/trainings/expert/full-Mlnub-massvar/` to `current-save/`.
- Create histograms in `histos/` with an applied `NNout` cut using the `data-files/cut-data.txt` and by executing `./plotmain ((2),` an unique name is recommended: *memorize this name for later use*) and check the results.
- Copy the newly created $M^{\nu b}$ histogram from `histos/` to `analysis-tools/Mlnub/`, where the analysis will be done.

Due to detector effects the *reconstructed* top mass M_t^R is not equal to the *true* top mass M_t^T , thus a calibration has to be performed to get the relationship between those (compare also with figure 4):

The calibration is based on the simulated t-channel samples, since the *true* top mass for these is known and the detector influences can be simulated to get the corresponding *reconstructed* top mass.

Here additionally to the default true top mass t-channel MC sample of $172.5 \text{ GeV}/c^2$, samples with shifted true top masses of $\pm 1 \text{ GeV}/c^2$ are used. A linear dependency between M_t^T and M_t^R is assumed:

$$M_t^T = c + m \cdot M_t^R. \quad (8)$$

The pairs of $M_t^{T,MC}$ and $M_t^{R,MC}$ are used to determine the values m and c in (8) using a linear fit.

After this calibration one can measure the peak value of **DATA** (after background subtraction) to get $M_t^{R,DATA}$ and use the relationship (8) to get the corresponding true value $M_t^{T,DATA}$. This will be done by the ROOT macro `TopMass.C` in `analysis-tools/Mlnub/`

Your Tasks:

- Change to `analysis-tools/Mlnub/` and complete the code of the ROOT macro `TopMass.C` to implement the calibration.

Calculate the propagation of the uncertainty for the correlated values m and c using this formula

$$\sigma_Y^2 = \sum_i \left(\frac{\partial f}{\partial X_i} \right)^2 \sigma_{i,i}^2 + \sum_{i \neq j} \left(\frac{\partial f}{\partial X_i} \right) \left(\frac{\partial f}{\partial X_j} \right) \sigma_{i,j}$$

, where $Y = f(X_1, X_2, \dots, X_n)$ and where X_i are input random variables. $\sigma_{i,j}$ are the covariance matrix elements you get from the linear fit, and $\sigma_{k,k} = \Delta X_k$ (uncertainty) when the input variable is not correlated to other input variables.

- Execute `./massvar.sh` and select option 1 to use the $+1 \text{ GeV}/c^2$ (massup) sample.
- Execute `./plotmain` ((2), use the name a specified by your code in the macro to distinguish it from the normal mass sample).
- Copy the newly created $M^{\text{t}\nu\text{b}}$ histograms from `histos/` to `analysis-tools/Mlnub/`.
- Repeat these steps using the massdown ($-1 \text{ GeV}/c^2$) sample.
- Execute the `TopMass.C` ROOT macro using `root -q -b 'TopMass.C("Mlnub",THE_NAME)'`. The syntax is the same as for the `ChargeAsymmetry.C` exercise macro.
- *Optional:* Create the expertise of a full neural network training using different mass variations (set the symbolic link to the correct sample in `trees/` as done by `massvar.sh` before executing `./nbmain` and choosing the expert). You will also have to rename the created t-channel `NNout` histograms to distinguish between the three mass samples, it is also recommended to use symbolic links for that. Compare the full training result with the one that does not include the $M^{\text{t}\nu\text{b}}$.

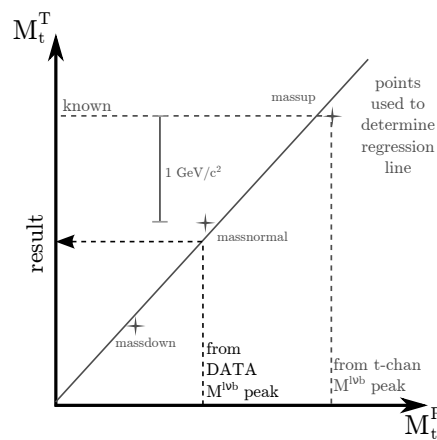


Figure 4: Scheme of the calibration method to determine the true top mass from the reconstructed.

Appendix

0.1 Resource file format

The file uses the following structure:

```
number plot train varname min max
```

where:

- **number** is a sequential number for internal use (do not change this)
- **plot** specifies whether this variable should be plotted when using `./plotmain` (1 = plot, 0 = don't plot)
- **train** specifies this variable should be used for training when using the NeuroBayesTM trainer (1 = use, 0 = don't use)
- **varname** is the name of the variable (do not change this). Compare with appendix 0.4 and table 1.
- **min** is the minimum value of the plot range (do not change this)
- **max** is the maximum value of the plot range (do not change this)

0.2 Cut file format

The file uses the following structure:

```
varname min max
```

where:

- **varname** is the name of the variables (the name must be the same as the variable names in `resource-data.txt`)
- **min** is the minimum value of the cut range
- **max** is the maximum value of the cut range

the area between min and max will be cut away.

You can define multiple cuts on the same variable, but *do not* set up a cut on another variable in between, this will cause plotting issues!

Working example:

mtw	0	50	keep only events with $M_T^W > 50 \text{ GeV}/c^2$
Ht	0	160	keep only events with $H_T > 160 \text{ GeV}$
Ht	450	999	keep only events with $H_T < 450 \text{ GeV}$

Not working example:

Ht	0	160
mtw	0	50
Ht	450	999

0.3 Reference for ROOT macros

This is a *pseudo reference* of ROOT commands that are useful to know to complete the code in the ROOT macros, you can find the full reference at <https://root.cern.ch/root/html/index.html>, especially for the histogram and graph commands look at <https://root.cern.ch/root/html/TH1.html> and <https://root.cern.ch/root/html/TGraph.html>.

```
Double_t TH1F->Integral(Int_t min, Int_t max)
```

Calculates the sum of all bin contents within the bin range between min and max.

```
Double_t TH1F->GetBinContent(Int_t bin)
```

Get the bin content using the bin number of a histogram.

```
Int_t TH1F->GetXaxis()->FindBin(Double_t x)}
```

Get the bin number using the x-value of a histogram.

```
Double_t TH1F->GetXaxis()->GetBinLowEdge(Int_t bin)
Double_t TH1F->GetXaxis()->GetBinUpEdge(Int_t bin)
```

Get the lower / upper value of the bin.

```
Double_t TH1F->GetBinError(Int_t bin)
```

Get the bin uncertainty (error) using the bin number of a histogram.

```
TH1F->Add(TH1F, Double_t c)
```

Adds another histogram to a histogram, c is a factor which will be used before each bin is added (set c=-1 to subtract histograms for instance).

```
TGraphErrors* gr = new TGraphErrors(Int_t n, Double_t* x, Double_t* y, Double_t* ex, Double_t* ey)
```

Creates a new graph. n is the size of the input arrays x, y, ex, ey which hold the values and their uncertainties.

```
TFitResultPtr TH1F->Fit(char* formula, Option_t* option)
TFitResultPtr TGraphErrors->Fit(char* formula, Option_t* option)
```

Performs a fit on the histogram / graph (similar behaviour in this context but not fully equal - look at the ROOT reference for details).

- **formula**⁵ specifies which formula you want to use, you can define an own function or use an internal expression for instance.
 - "gaus" for a gauss function, this is an alias for the input "[0]*exp(-0.5*((x-[1])/[2])**2)", where [0] is the parameter of this function with index 0, [1] with index 1 and so on.
 - "polN" for a polynomial function of degree N (for example "pol1" is an alias for the input "[0]+[1]*x").

⁵look at <https://root.cern.ch/root/html/TFormula.html> for a more details if required

- `option` defines options ⁶ for the fit.

For instance "S" must be used to get the `TFitResultPtr` for further access, it can be useful to also to use "E" for an improved uncertainty calculation. To combine two options set this to "SE" for example.

```
double TFitResultPtr->Parameter(int i)
```

Get the parameter value of the fit function that has been applied using the parameter index i.

```
double TFitResultPtr->CovMatrix(int i, int j)
```

Get the covariance matrix value of the fit that has been applied using the parameter index i and j.

0.4 Brief description of the available variables

The corresponding names of the input variables in the .root files and all the program tools can be found in table 1.

0.4.1 Pseudorapidity: η

The pseudorapidity $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$ of several particles is measured (l, q_1) or reconstructed ($\nu_1, W, t, \text{jet 1, jet 2, } b$).

0.4.2 Transverse momentum: p_T

The transverse of several particles is measured (l, q_1) or reconstructed ($\nu_1, W, t, \text{jet 1, jet 2, } b$).

0.4.3 Invariant Mass: M

The invariant mass is calculated for several variables (jet 1, jet 2, W) and combination of those (q_1, l). Especially the invariant mass $M^{\nu_1 b}$ is reconstructed by adding the four vectors of the components, which corresponds to the top mass.

0.4.4 Distance of jet 1 and jet 2 in η - ϕ plane: ΔR

$\Delta R^{\text{jet1, jet2}} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ describes the distance in the η - ϕ plane between the areas of jet 1 and jet 2.

0.4.5 Lepton charge: Q^l

The electric charge of the isolated lepton (muon in the data used for these exercises).

⁶look at <https://root.cern.ch/root/html/TGraph.html#TGraph:Fit> for more details if required

0.4.6 Cosine of angle between particles: $\cos \Theta$

The angle reconstruction has been done between

- the light quark q_l and lepton l in the laboratory frame.
- the lepton l in the W boson rest frame and the W boson in the laboratory frame.
- the lepton l in the W boson rest frame and the W boson in the top quark rest frame.

0.4.7 Missing transverse energy: E_T^{miss}

The missing transverse energy is reconstructed by vectorially adding up transverse momenta of all measured particles.

0.4.8 Transverse Mass: M_T

The transverse mass of the W boson is reconstructed from the charged lepton four-vector and the E_T^{miss} vector.

0.4.9 Absolute sum of all transverse momenta: H_T

H_T is calculated by the absolute scalar sum of all p_T and E_T^{miss} .

Table 1: Variable names in program files

name in program	variable	name in program	variable
lep_pt	p_T^l	jet2_pt	p_T^{jet1}
lep_eta	η^l	jet2_eta	η^{jet2}
w_pt	p_T^W	M_jet1	M^{jet1}
w_eta	η^W	M_jet2	M^{jet2}
nu_pt	p_T^ν	lq_pt	p_T^{jet1}
nu_eta	η^ν	lq_eta	η^{q1}
Ht	H_T	Mlnub	M^{lq1b}
met	E_T^{miss}	Mlql	$M^{\text{q1,l}}$
mtw	M_T^W	qeta	$Q^{\text{l}\eta^{\text{l}}}$
DeltaR_jet1_jet2	$\Delta R^{\text{jet1,jet2}}$	top_eta	η^{t}
cosTheta	$\cos \Theta^{\text{l,q1}}$	top_pt	p_T^{t}
cosTheta_lep_W	$\cos \Theta^{\text{l*,W}}$	btop_eta	η^{b}
cosTheta_lep_Wtop	$\cos \Theta^{\text{l*,W*}}$	btop_pt	p_T^{b}
jet1_pt	p_T^{jet1}	lep_charge	Q^{l}
jet1_eta	η^{jet1}		

Danksagung

Das Anfertigen dieser Bachelorarbeit wäre ohne die Unterstützung zahlreicher Personen nicht möglich gewesen.

Daher bedanke ich mich zunächst bei Herrn Prof. Dr. Thomas Müller für die Aufnahme in die CMS-Arbeitsgruppe des EKP und der Ermöglichung dieser Bachelorarbeit.

Mein besonderer Dank richtet sich an Dr. Thorsten Chwalek für für das intensive Korrekturlesen der angefertigten Arbeit und der entworfenen Übungsblätter und daraus hervorgehender Verbesserungsvorschläge. Nicht minder gilt mein Dank Steffen Röcker für die unmittelbare Betreuung beim Entwickeln der nötigen Programme, für die Beantwortung von Verständnisfragen und darüber hinaus für Lektürevorschläge zum Erlangen von tieferen Verständnis einzelner Teilbereiche; des Weiteren für das Lesen schon früher Entwürfe dieser Arbeit, der entworfenen Übungsblätter und den daraus hervorgehenden nützlichen Verbesserungsvorschlägen.

Außerdem richtet sich mein Dank an die gesamte restliche EKP-Arbeitsgruppe für die tolle Arbeitsatmosphäre und die zahlreichen kleine Anregungen und Hilfestellungen von allen Seiten, die ebenfalls von großem Nutzen für das Anfertigen dieser Arbeit waren.

Ich möchte des Weiteren mein Dank an all diejenigen Personen richten, die mich während des bisherigen Studiums der Physik begleitet und unterstützt haben. An dieser Stelle möchte ich namentlich noch meine Schwester Deniese Gebele erwähnen, ohne deren Hilfe manch Tippfehler sicherlich unentdeckt geblieben wäre.

Mein Dank gilt überdies auch meiner restlichen Familie, allen Freunden und Bekannten, die mich ebenfalls sowohl während des Studiums als auch bereits zuvor tatkräftig in vielerlei Hinsicht unterstützt haben.